Vinu V Das
Yogesh Chaba (Eds.)

Communications in Computer and Information Science 296

Mobile Communication and Power Engineering

Second International Joint Conference, AIM/CCPE 2012
Bangalore, India, April 2012
Revised Selected Papers

Springer

Communications in Computer and Information Science 296

Vinu V Das Yogesh Chaba (Eds.)

Mobile Communication and Power Engineering

Second International Joint Conference, AIM/CCPE 2012
Bangalore, India, April 27-28, 2012
Revised Selected Papers

Springer

Volume Editors

Vinu V Das
IDES - Institute of Doctors Engineers and Scientists
Amsterdam, The Netherlands
E-mail: prof.vinuvdas@gmail.com

Yogesh Chaba
Guru Jambheshwar University of Science and Technology
Hisar, Haryana, India
E-mail: yogeshchaba@yahoo.com

ISSN 1865-0929 e-ISSN 1865-0937
ISBN 978-3-642-35863-0 e-ISBN 978-3-642-35864-7
DOI 10.1007/978-3-642-35864-7
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012954704

CR Subject Classification (1998): C.2.0-5, B.4.2, J.1, J.2, J.7, D.2, H.3.4-5, I.5.1-4, H.4.1-3, K.6.5, K.4.4

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

It is a great privilege to write this preface for the proceedings of the Second International Joint Conference on AIM 2012 and CCPE 2012, organized by the Association of Computer Electronics and Electrical Engineers (ACEEE) at Bangalore, India, during April 27–28, 2012.

The primary goal of this joint conference is to promote research and development activities in computer science, information technology, computational engineering, mobile communication, control and instrumentation, communication system, power electronics and power engineering. Another goal is to promote scientific information interchange between researchers, developers, engineers, students, and practitioners working in India and abroad. It gives an opportunity to engineering professionals to present their research and development work that has potential for applications in several disciplines of engineering.

The Organizing Committee rigorously invited submissions for many months from researchers, scientists, engineers, students, and practitioners related to the relevant themes and tracks of the conference. This effort guaranteed submissions from an unparalleled number of internationally recognized top IT, network, and power electronic experts. All the submissions underwent a strenuous peer-review process by expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The review process judged the contributions, on their technical content, originality, and clarity. The entire process including the submission, review, and acceptance phases was done electronically. All these efforts undertaken by the Organizing and Technical Committee led to an exciting, rich, and high-quality technical conference, which featured high-impact presentations, for all attendees to enjoy, appreciate, and expand their expertise in the latest developments in the area of IT, mobile communication, and power electronics.

People responded to the conference with great enthusiasm and academicians, researchers, students, and corporate professionals have contributed a great deal. I hope that the professional dialogue among the researchers, scientists, engineers, students, and educators continues beyond the workshop and that the friendships and collaborations forged will linger and prosper for many years to come.

Yogesh Chaba

Organization

Technical Chairs

Hicham Elzabadani	American University in Dubai
Prafulla Kumar Behera	Utkal University, India

Technical Co-chairs

Natarajan Meghanathan	Jackson State University, USA
Gylson Thomas	MES College of Engineering, India

General Chairs

Janahanlal Stephen	Ilahiya College of Engineeing, India
Beno Benhabib	University of Toronto, Canada

Publication Chairs

R. Vijaykumar	MG University, India
Brajesh Kumar Kaoushik	IIT Roorke, India

Organizing Chairs

Vinu V. Das	The IDES
Nessy T.	Electrical Machines Group, ACEEE

Program Committee Chairs

Harry E. Ruda	University of Toronto, Canada
Durga Prasad Mohapatra	NIT Rourkela, India

Program Committee Members

Shu-Ching Chen	Florida International University, USA
T.S.B.Sudarshan	BITS Pilani, India
Habibollah Haro	Universiti Teknologi Malaysia
Derek Molloy	Dublin City University, Ireland

Jagadeesh Pujari	SDM College of Engineering and Technology, India
Animesh Adhikari	S.P. Chowgule College, India
Anirban Mukhopadhyay	University of Kalyani, India
Malabika Basu	Dublin Institute of Technology, Ireland
Tahseen Al-Doori	American University in Dubai
V.K. Bhat	SMVD University, India
Ranjit Abraham	Armia Systems, India
Naomie Salim	Universiti Teknologi Malaysia
Abdullah Ibrahim	Universiti Malaysia Pahang
Charles McCorkell	Dublin City University, Ireland
Neeraj Nehra	SMVD University, India
Muhammad Nubli	Universiti Malaysia Pahang
Zhenyu Y. Angz	Florida International University, USA
Keivan Navi	Shahid Beheshti University, Iran
Rama Shankar Yadav	MNNIT, India
Smriti Agrawal	MNNIT, India
Vandana Bhattacherjee	BITS Mesra, India
R.D. Sudhaker Samuel	S.J. College of Engineering, India
Amitabha Sinha	West Bengal University of Technology, India
Shyam Lal	MIT Moradabad, India
Debasish Jena	Biju Patnaik University of Technology, India
Srinivasa K.G.	M.S. Ramaiah Institute of Technology, India
Bipin Bihari Jayasingh	CVR College of Engineering, India
Seyed-Hassan Mirian-Hosseinabadi	Sharif University of Technology, Iran
Malay K. Pakhira	Kalyani Government Engineering College, India
Sarmistha Neogy	Jadavpur University, India
Sreenath Niladhuri	Podicherry Engineering College, India
Ananta Ojha	ICFAI University, India
A.K. Sharma	YMCA Institute of Engineering, India
Debasis Giri	IIT Kharagpur, India
Suparna Biswas	WBUT, India

Table of Contents

Full Paper

Short Paper

Poster Paper

Cluster Based Dynamic Keying Technique for Authentication in Wireless Sensor Networks

Thiruppathy Kesavan V. and Radhakrishnan S.

Department of Computer Science and Engineering,
Kalasalingam University, Krishnankoil – 626 126, Tamil Nadu, India
{kesavan,srk}@klu.ac.in

Abstract. This paper proposes a Cluster Based Dynamic Keying Technique (CBDKT) for authentication in wireless sensor networks. The Cluster Head (CH) determines a Combined Cost Value (CCV) for each sensor node based on node location, node degree and residual battery power. Since the key to the encryption scheme changes dynamically when there is a change in CCV, separate dynamic keys are generated in each cluster by the CH. While the source CH forwards the data to the sink along different clusters, verification of authenticity is done. When the source CH needs to transmit the aggregated data packet to the sink, it splits the packet into shares by threshold secret sharing algorithm and forwards them to the sink using multi-path dispersal routing technique.

Keywords: Clustering, Authentication, Multi-path routing, Heterogeneous.

1 Introduction

Wireless Sensor Network (WSN) is formed of identical or dissimilar sensor nodes. The two types involved in deploying WSN are the homogeneous deployment and the heterogeneous deployment. The heterogeneous deployment involves dissimilar sensor nodes having different functionalities. In this type of network, some of the nodes are more powerful where as numerous sensors are less powerful. Such a combination of sensor nodes covers a large Heterogeneous Wireless Sensor Networks (H-WSNs) [17], [18]. H-WSNs can be utilized for various real time applications and can function in diverse environments [15]. The network performance is improved in the presence of high potential sensors with a better sensing range and transmission range [1].

Data confidentiality, integrity, authenticity and availability are considered to be the security requirements which the WSN should provide even in the presence of powerful attackers [2]. The threats and challenges which occur in the sensor networks are spoofed, dropped, altered, or replayed routing information, selective forwarding, Sinkhole attacks, Sybil attacks, Wormholes, HELLO flood attacks and Blackhole attack [3].

In WSNs, the three simplest keying models such as network keying, pair-wise keying and group-wise keying are used to compare the relationship among the WSN security and its operational needs [2]. The features of both network and pair-wise

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 1–8, 2013.

keying models are combined together to form the group-wise keying model. Inside a cluster, it uses a single, secret key where as between the clusters it uses pair-wise keying model [2, 20]. The data confidentiality in heterogeneous WSN are safeguarded using group-wise key distribution.

The spending of energy can be reduced by generating keys dynamically and then by avoiding the key distribution. Once the keys are generated dynamically, the packets should be authenticated to defend against the adversaries for security. Routing is one of the factors in WSN to achieve good packet delivery ratio, by reducing packet drops in the network where attackers exist.

This paper proposes a cluster-based design approach for dynamic key estimation that uses location of the nodes (XY-co ordinates), node degree and residual battery power to reduce the energy expenditure of the key distribution. This approach uses multi-path dispersal routing that provides good packet delivery ratio and it consumes less energy even in the presence of attackers.

2 Related Work

To provide stronger protection, some pairing-based schemes [6, 16] have been proposed with the objective to make the scheme smaller and faster. Ying Qui et al. [8] proposed a technique for key establishment that reduces communication cost by using cluster heads as sub-base stations. Probabilistic key pre-distribution schemes require large storage [9] if the key pool size is large. Guohua Ou et al. proposed a key-chain based key management scheme for HSN [26], where the sensor nodes need to be pre-loaded with an initial key which is to be used during pair-wise key establishment. They have included a key renovation process to provide security of pair-wise keys. Chin-Ling Chen and Cheng-Ta Li [19] proposed a dynamic key management mechanism for WSNs, where the keys required in the next round are generated dynamically using the previous two keys that are already preset in each sensor nodes. VEBEK [10] protocol avoids exchanging of messages to dynamically change its key. The keys are updated as a function of residual virtual energy of the sensor node.

A node authentication scheme [25] has been proposed where each sensor nodes has unique clock skew which is different from other sensor node. Since this method is based on time synchronization, it is very difficult to achieve in WSN. Since conventional authentication schemes are insufficient to solve false data injection and path-based DoS attacks, Chia-Mu Yu et al. [11] proposed a constrained function-based message authentication scheme using hash function that supports the en-route filtering function.

3 Secure Cluster Based Multi-path Routing

This paper provides a cluster based dynamic keying technique for wireless sensor networks (WSN). The CHs are assumed to be enabled with GPS system. Firstly the nodes have to be grouped into clusters using their distance [23]. Secondly the CHs

estimate their member locations using the location estimation scheme [13]. Thirdly the cluster member estimates its node degree and the CH maintains it in its member table. Fourthly CH computes the Combined Cost Value (CCV) for each nodes and key is dynamically updated based on this CCV. The sensor nodes also calculate CCV to change its key while sending the data. Finally the data is then transmitted to sink using multi-path dispersal routing technique. The forwarded data is verified for authenticity by different CHs along the path to the sink.

3.1 Cost Computation

First, CH constitutes the member table which includes the details of each member ID, its location, node degree and residual battery power. Afterwards CH determines the CCV for each sensor node based on these parameters. The location of each node (NL) is given by the (x, y) coordinates and is computed as per the location estimation scheme [13]. The node degree (ND) is estimated based on the neighboring nodes information. As location and node degree of a node are constant for a static network, both NL and ND values remain static. The residual battery power available at time *t* is $a(t) = \alpha - \alpha(t)$. The parameter α represents the total charge in the battery and $\alpha(t)$ is a cost function which represents the actual charge lost by the battery at time *t* [21]. When the residual battery power is changed every time, the cost function is updated. The computation of cost function is as follows:

$$CCV = \acute{\alpha}\,(xa + yb) + \beta\, ND + \gamma\, R_N\,(c)\,. \tag{1}$$

where $\acute{\alpha}$, β and γ are normalization factors and a and b are constants

3.2 Dynamic Key Generation

Each sensor node has residual energy during its initial deployment in the network. During the above functional state, there may be depletion of residual energy in the sensor node. The key generation scheme is initiated whenever the data is sensed. Hence no separate methodology is required to update the keys. The dynamic key is computed as follows:

```
At start, i ← 1
            If i = 1
Then
K_i ← F_k (C_i, V_i), i← i+1
Else
K_i ← F_k (K_i-1, C_c), i← i+1
End if
```

where K_i is the current key, F_k is the key generation function, C_i is initial cost of the sensor node, C_c is the current cost of the sensor node, K_{i-1} is the previous key and V_i is the initialization vector.

3.3 Encoding and Authentication

The key to RC5 [12] is generated with the dynamic key generation module as per the section 3.2. The packet includes source ID, packet type and data fields together denoted as [z]. Then the packet is encrypted and an additional copy of the source ID has to be included along with the encoded message which appears to be [ID {k (z)}]. Here k (z) constitutes the message z encoded by key k. When the next cluster along the path to sink receives the packet, it generates the local permutation code to decode the packet. The node computes the message authenticity and integrity code (MAC) which offers hash function over the encoded message and destination address and finally this message will be forwarded to CH. The message forwarded to the CH is in the following format:

$$\text{MAC}\,[\text{ID}\,\{C_p(z)] = H\,[\text{ID}\,\{C_p(z) \parallel Ad_d\}]\,. \tag{2}$$

where ID is the identity of the originating node

$C_p(z)$ constitutes the encoding data with permutation code C_p
Ad_d represents the destination address of the cluster head

When a CH receives the message, it checks for the associated $R_N(c)$ stored in the sending node. Then it extracts $R_N(c)$ to estimate CCV and derives the dynamic key k. CH then computes the MAC over the received message k[z] from the sensor node which is represented as MAC_{ch}. On comparison, if MAC and MAC_{ch} are found to be similar, then the message is said to be authentic. CH repeats this authentication process for every received packet from its member. Afterwards it aggregates the entire data received from the source and forwards them towards the sink through multi-path dispersal routing technique.

4 Multi-path Dispersal Routing

When the cluster head needs to transmit the aggregated data packet to the sink, it splits the packet into q shares according to (t, q) – threshold secret sharing algorithm [14]. The application of the scheme is concerned with the cooperation of group of mutually suspicious nodes with incompatible interests. The properties of this algorithm are:

- The size of the each share does not exceed the size of the original data.
- When the t is kept fixed, shares are dynamically added or deleted without affecting the other shares.

The multi-path dispersal routing is used to transmit the shares to the next CH. Each share holds a Time-to-Live (TTL) field and the source node assigns the initial value of TTL field to control the total number of random transmissions. When the sink gathers minimum of t shares, it will rebuild the original packet.

4.1 Multi-path Routing

Multi-path routing [22] is established for transmission of q shares from CH to the sink. The multi-path routing protocol usually desires the node-disjoint paths owing to the use of the most accessible network resources. In order to perform the multi-path dispersal routing, the following phases are considered:

Initialization Phase. Every CH broadcasts a HELLO message through the network that contains enough information regarding its nearest neighbors. The message includes source ID, hop count and residual battery power. In this phase, CH maintains and updates a buffer table which includes the information about the list of the neighbor CHs.

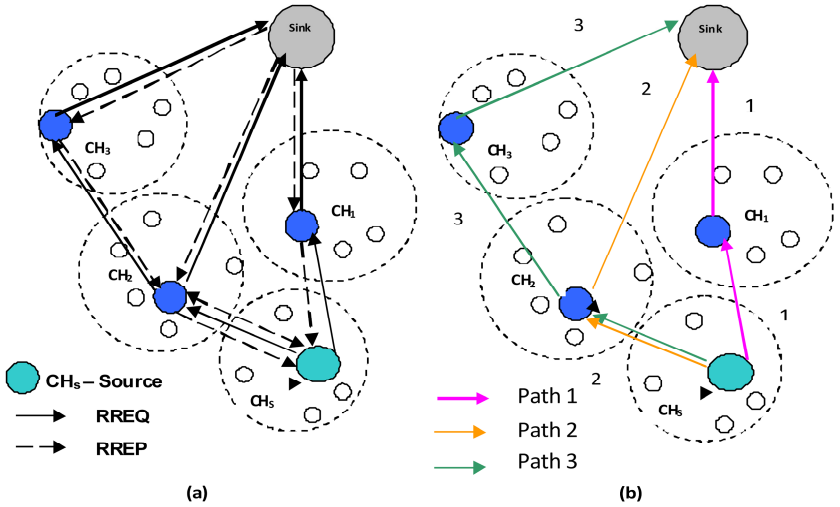


Fig. 1. **(a)** Route Discovery. **(b)** Routing from CHs to Sink.

Route Discovery Phase. After initialization phase, sufficient information is available in every CH for computation of cost function for its neighboring CHs. The source CH (CH_s) computes the preferred next hop CH. Then the route request (RREQ) message has to be broadcasted to the chosen next hop CH (shown in Fig. 1(a)). The RREQ message includes the source ID, destination ID, cluster ID, and residual battery power. After the reception of all RREQ messages, the sink node replies back to CH_s with the route reply (RREP) message via the path traversed by the RREQ messages. Based on the RREP, CH_s discovers the available paths and forwards q shares through these paths towards the sink node. The above process of multi-path dispersal routing from cluster head to the sink is described in Fig. 1(b). Through the selected paths, the CH_s forward the shares towards the sink node.

5 Simulation Results

The proposed CBDKT is evaluated using NS2 [24]. We consider a random network of 100 sensor nodes deployed in an area of 500 X 500m. Initial energy is assumed as 17 J. The number of clusters formed is 9. The performance of CBDKT is compared with the VEBEK-II [10] scheme. The performance is evaluated in terms of Packet Delivery Ratio (PDR) and energy consumed by varying the number of attackers from 1 to 5, from various clusters performing black hole and packet dropping attacks.

Fig. 2(a) shows that when the attackers are increased, the percentage of packet drop increases and hence the PDR is decreased. In VEBEK, the route from source to the sink is fixed during the delivery of data. Since the attackers are increased along the path to the sink, more numbers of packets are dropped and hence it reduces the PDR. But CBDKT has shown good PDR, when compared to VEBEK. This is because of the fact that CBDKT uses the multi-path routing technique to reach the destination. Also due to the dispersal technique, the sink needs only *t* shares to build the original data which further improves PDR.

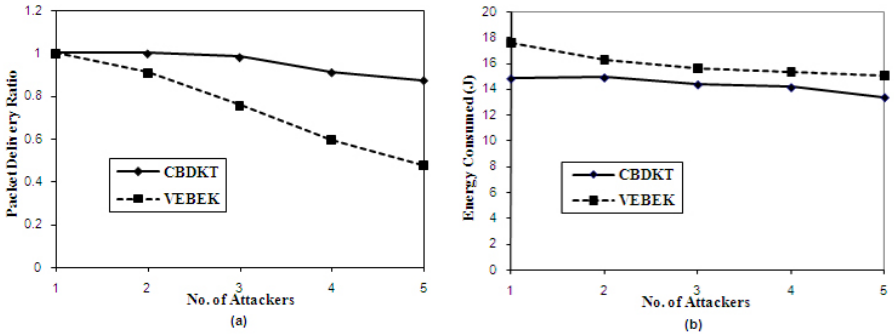


Fig. 2. Network Performance: **(a)** Packet Delivery Ratio, **(b)** Energy Consumption

Fig. 2(b) shows the average energy consumption of both the schemes. From the figure, it is defined that the average energy consumption decreases, when the attackers are increased. This is because the percentage of correctly received data traffic reduces, as the number of dropped packets increases by the attackers in the network. In VEBEK, if the forwarding node is not able to extract the key successfully, it tries for another key to classify the packet as malicious. Since our simulation launches packet drop attack along the path to sink, the forwarding nodes take many attempts to extract the key though the packet is not malicious. So it consumes more energy compared to CBDKT. When compared CBDKT consumes 11% energy less than VEBEK because of the cluster based routing approach.

6 Conclusion

This paper illustrates the key to the encryption scheme which changes dynamically based on the CCV of the sensor and eliminates re-keying. Since separate dynamic

keys are generated in each CH, the data to the sink along different clusters are able to verify the authenticity and integrity and provides non-repudiation. Due to the multi-path dispersal routing technique, this method provides improved PDR compared to VEBEK. Since the authentication happens only at CHs, it reduces the energy consumption compared to VEBEK. If any cluster head becomes an attacker, any misbehaving action will not affect the data, since it can access only a part of the data. Since this proposed method involves less key updating operations, energy cost associated with keying is reduced. This technique can be applied for mobile nodes since node location and node degree are taken into consideration while computing key cost function. Adding up secure routing will further improves the security of this approach.

References

1. Wang, Y.: Intrusion Detection in Gaussian Distributed Heterogeneous Wireless Sensor Networks. In: Global Communications Conference, GLOBECOM 2009, Honolulu, Hawaii, USA, pp. 1–6. IEEE (2009)
2. Bojkovic, Z.S., Bakmaz, B.M., Bakmaz, M.R.: Security Issues in Wireless Sensor Networks. International Journal of Communications 2(1), 106–115 (2008)
3. Clark, J.A., Murdoch, J., McDermid, J.A., Sen, S., Chivers, H.R., Worthington, O., Rohatgi, P.: Threat Modelling for Mobile Ad Hoc and Sensor Networks. In: Annual Conference of ITA, September 25-27 (2007)
4. Vaidya, B., Chen, M., Rodrigues, J.J.P.C.: Improved Robust User Authentication Scheme for Wireless Sensor Networks. In: Fifth IEEE Conference on Wireless Communication and Sensor Networks (WCSN 2009), Allahabad, India, pp. 1–6 (2009)
5. Benenson, Z., Gedicke, N., Raivio, O.: Realizing Robust User Authentication in Sensor Networks. In: Workshop on Real-World Wireless Sensor Networks (REALWSN 2005), Sweden (2005)
6. Xiong, X., Wong, D.S., Deng, X.: TinyPairing: Computing Tate Pairing on Sensor Nodes with Higher Speed and Less Memory. In: Eighth IEEE International Symposium on Network Computing and Applications (NCA 2009), Cambridge, MA, USA, pp. 187–194 (2009)
7. Moon, S.Y., Cho, T.H.: Intrusion Detection Scheme against Sinkhole Attacks in Directed Diffusion Based Sensor Networks. IJCSNS International Journal of Computer Science and Network Security 9(7), 118–122 (2009)
8. Qiu, Y., Zhou, J., Baek, J., Lopez, J.: Authentication and Key Establishment in Dynamic Wireless Sensor Networks. Sensors 10(4), 3718–3731 (2010)
9. Du, X., Guizani, M., Xiao, Y., Chen, H.: A Routing-Driven Elliptic Curve Cryptography Based Key Management Scheme for Heterogeneous Sensor Networks. IEEE Transactions on Wireless Communications 8(3), 1223–1229 (2009)
10. Uluagac, A.S., Beyah, R.A., Li, Y., Copeland, J.A.: VEBEK: Virtual Energy-Based Encryption and Keying for Wireless Sensor Networks. IEEE Transactions on Mobile Computing 9(7), 994–1007 (2010)
11. Yu, C.-M., Tsou, Y.-T., Lu, C.-S., Kuo, S.-Y.: Constrained Function-Based Message Authentication for Sensor Networks. IEEE Transactions on Information Forensics and Security 6(2), 407–425 (2011)

12. Rivest, R.L.: The RC5 Encryption Algorithm. In: Preneel, B. (ed.) FSE 1994. LNCS, vol. 1008, pp. 86–96. Springer, Heidelberg (1995)
13. Stoleru, R., Stankovic, J.A.: Probability Grid: A Location Estimation Scheme for Wireless Sensor Networks. In: 1st Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON), Santa Clara, CA, pp. 430–438 (2004)
14. Shamir, A.: How to share a secret. Communications of the ACM (CACM) 22(11) (November 1979)
15. Han, X., Cao, X., Lloyd, E.L., Shen, C.-C.: Fault-tolerant Relay Node Placement in Heterogeneous Wireless Sensor Networks. In: 26th IEEE International Conference on Computer Communications (INFOCOM), pp. 1667–1675 (2007)
16. Shirase, M., Miyazaki, Y., Takagi, T., Han, D.-G., Choi, D.: Efficient implementation of pairing based cryptography on a sensor node. IEICE Trans. E92-D, 909–917 (2009)
17. Lu, X., Spear, M., Levitt, K., Felix Wu, S.: iBubble: Multi-keyword Routing Protocol for Heterogeneous Wireless Sensor Networks. In: IEEE, 27th Conference on Computer Communications (INFOCOM), pp. 968–976 (2008)
18. Qiu, M., Xue, C., Shao, Z., Liu, M., Sha, E.H.-M.: Energy Minimization for Heterogeneous Wireless Sensor Network. Special Issue of Journal of Embedded Computing (JEC) 3(2), 109–117 (2009)
19. Chen, C.-L., Li, C.-T.: Dynamic Session-Key Generation for Wireless Sensor Networks. EURASIP Journal on Wireless Communications and Networking 2008, Article ID 691571, 10 pages (2008)
20. Du, W., Deng, J., Han, Y.S., Varshney, P.K.: A Pairwise Key Predistribution Scheme for Wireless Sensor Networks. ACM Transactions on Information and System Security 8(2), 228–258 (2005)
21. Dasika, S., Vrudhula, S., Chopra, K., Srinivasan, R.: A framework for battery-aware sensor management. In: The Conference on Design, Automation and Test in Europe, vol. 2, pp. 962–967. IEEE Computer Society, Washington, DC (2004)
22. Yahya, B., Ben-Othman, J.: REER: Robust and Energy Efficient Multipath Routing Protocol for Wireless Sensor Networks. In: Global Telecommunications Conference (IEEE GLOBECOM 2009), Honolulu, HI, pp. 1–7 (2009)
23. Quan, S.G., Kim, Y.Y.: Fast Data Aggregation Algorithm for Minimum Delay in Clustered Ubiquitous Sensor Networks. In: International Conference on Convergence and Hybrid Information Technology (ICHIT 2008), Daejeon, pp. 327–333 (2008)
24. Network Simulator, `http://www.isi.edu/nsnam/ns`
25. Huang, D., Teng, W., Wang, C., Huang, H., Hellerstein, J.M.: Clock Skew Based Node Identification in Wireless Sensor Networks. In: Global Telecommunications Conference (IEEE GLOBECOM 2008), New Orleans, LO, pp. 1–5 (2008)
26. Ou, G., Huang, J., Juan Li, J.: A Key-Chain Based Key Management Scheme for Heterogeneous Sensor Network. In: IEEE International Conference on Information Theory and Information Security (ICITIS), Beijing, pp. 358–361 (2010)

RrMm - A Measurement Model to Quantify the Effect of Reengineering Risk in Quality Perspective of Legacy System

Anand Rajavat and Vrinda Tokekar

Shri Vaishnav Institute of Technology and Science, Indore, M.P., India
Institute of Engineering and Technology (DAVV), Indore, M.P., India
anandrajavat@yahoo.co.in, vrindatokekar@yahoo.com

Abstract. Legacy systems are the backbone of any software organization. To satisfy the essentials of current users and modern business environment of global competition legacy system needs to be modernized. Now a days reengineering offers an approach to migrate a legacy system towards an evolve system. However improving quality of legacy system through reengineering requires examining and analyzing different reengineering risk. Quantifiable risk measures are necessary for the measurement of reengineering risk to take decision about when the evolution of legacy system through reengineering is successful. Existing methods used for evaluating the current state of legacy system mostly rely on user survey which are too subjective and need much processing time and cost. We present a reengineering risk measurement model RrMm to measure comprehensive impact of different reengineering risk in the evolution process of legacy system. We also measure reengineering risk arises from quality perspective of legacy system using proposed RrMm model. The quality perspective RrMm model consists of five reengineering risk component, including Availability risk, Reliability risk, Modularity risk, Usability risk and performance risk .RrMm offers better performance in terms of risk measurement to support the decision-making process.

Keywords: Legacy System, Reengineering, Risk Engineering, Software Quality.

1 Introduction

Older software systems that remain vital to an organization are called legacy systems [1]. Legacy systems are still essential for the normal functioning of the business. Currently, Organizations face with a very high competition and consequently they have to continuously improve their legacy systems. This improvement could require changes of the system component so it is important to evaluate the present state of the legacy system to facilitate their evolution [2]. Variety of options is available for the evolution of legacy system. Continued maintenance, some form of reengineering, or replacement are general evolution strategies, of which one or a combination may be an appropriate way of evolving a legacy system. Currently legacy system reengineering has emerged as a well-known system evolution technique. Reengineering concerns the examination of the design and implementation of an existing legacy system and

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 9–16, 2013.

applying different techniques and methods to redesign and reshape that system into hopefully better and more suitable software. Unfortunately most reengineering projects are only concentrated on functional aspects of reengineering, however, system managerial and technical aspects play an important role in the successful improvement of the legacy system through reengineering. Software reengineering projects is often faced with unanticipated problems which pose risks within the reengineering process. Successful Implementation of reengineering effort requires a systematic measurement of overall impact of reengineering risk from system, managerial, and technical domain of legacy application. [3]

Proposed RrMm model analyze current state of legacy system and desired state of target system to measure different reengineering risk components in the evolution process of legacy system. The behavior of the RrMm system is represented through pentagram model used to compute comprehensive impact of all risk components as part of the decision-making process.

We also use RrMm model for quantifying the impact of reengineering risk arises from quality perspective of legacy system. The quality perspective RrMm model consists of five reengineering risk component, including Availability risk, Reliability risk, Modularity risk, Usability risk and performance risk components. Different Measurement metrics are used to examine and analyze different reengineering risk components.

2 Related Work

Measurement is an important phenomenon in our daily lives. We use measurement all of the time to help us make better decisions based on facts and information. In analyzing the feasibility of a reengineering process for the evolution of legacy software system, it is important to measure the overall impact of reengineering risks emerge from system, managerial and technical domain of legacy system [4].

Yennun huang in [5] present a model for analyzing software renovation in running application and express downtime and costs due to downtime during renovation using some parameters in that model. On the other side Kishor S. Trivedi in [6] presents a stochastic model to measure the effectiveness of proactive fault management in software systems and determine optimal time to perform rejuvenation, for different scenarios. Model develops different methodologies to detect software aging and estimate its effect on various system resources.

Most of the measurement models [7-10] are static and only target specific quality or risk component. Henceforth there is a need of a consolidated reengineering risk measurement model to measure overall impact of all reengineering risk components in the evolution process of legacy system.

3 RrMm - Reengineering Risk Measurement Model

The purpose of RrMm model is to design quantifiable metrics for the evaluation of legacy system by measuring different reengineering risk components materialize in the evolution process of legacy system. The RrMm model is able to design quantifiable

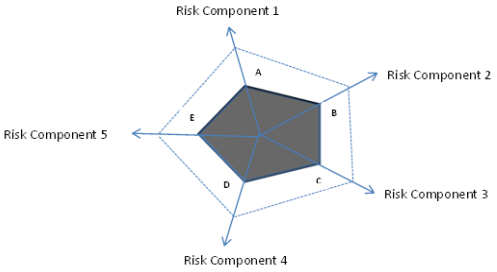


Fig. 1. RrMm Model

metrics to measure comprehensive impact of all reengineering risk from system, managerial, and technical domain of legacy system.

As shown in Figure 1 RrMm model is presented using a pentagram diagram based on the measurement of its five risk components. Each risk component is measured based on the results of their metrics during measurement process. Let us assume: the measurement results of each risk component is a value from 0 to 1. The value "1" indicates the maximum value for each risk component, and "0" indicates the minimum value. The area of the pentagram is used as the measurement of overall impact of five risk components. Clearly, the smallest value of this pentagram area is 0, and the maximum value is approximately 2.4. As the pentagram consists of five triangles, the area of each triangle can be computed 0.5 * L1 * L2 * Sinα where L1, L2 represent the sides of the triangle and α represents the 72-degree angle between the two sides. The letters A, B, C, D and E are used to represent the five risk components of RrMm Model respectively. Since each risk component is measured using different quantifiable measurement metrics the Total Risk Impact (TRI) of all risk components can be computed as below:

$$\begin{aligned} \textit{Total Risk Impact (TRI)} &= 1/2\ Sin\ 72\ (AB + BC + CD + DE + EA) \\ &= \ ½\ X\ 0.9511X(AB + BC + CD + DE + EA) \\ &= 0.48\ X\ (AB + BC + CD + DE + EA) \end{aligned}$$

4 Quality Perspective RrMm Model

Quality perspective RrMm model is used to measure overall impact of five different reengineering risks from technical domain of legacy system. The purpose of quality perspective RrMm is to design a pentagram model used to measure overall impact of different reengineering risk component in quality perspective of legacy system.

5 Measurement Metrics

In this section, we describe the measurement metrics of each side of the Quality perspective RrMm pentagram model.

5.1 Availability Risk Component (A)

Availability is the probability that the system is operating properly when it is requested for use. Availability risk measurement required to thoroughly understand the system's configuration to accurately measure availability of legacy system as experienced by end users. This includes all the components and resources used by the application, both local and remote; and the hardware and software components required to access those resources. The next step is to monitor all these components for outages, then calculate end-to-end availability [11].

The total impact of availability risk component is calculated as follows:

$$\text{Total Impact of Availability Risk (TIAR)} = 1/\text{ Total Availability (TA)} \quad (1)$$

As the total availability can be calculated as

$$\text{Total Availability (TA)} = A(t) + Am(t) + A(\infty) + Ao \quad (2)$$

Where

A (t) represents Point availability, Am (t) represents mean availability

A (∞) represents Steady State availability, Ao represents Operational availability

5.2 Reliability Risk Component (B)

Software reliability is an important facet of legacy software quality. Reliability implies a time factor in that a reliable product is expected to perform correctly over a period of time i.e. reliability is measured as the probability that a system will not fail to perform its intended functions over a specified time interval [12].

The total impact of reliability risk component for a legacy system can be measured using following metrics

$$\text{Total impact of Reliability Risk (TIRR)} = 1/\text{ Total Reliability (TR)} \quad (3)$$

As the total reliability (TR) can be calculated as,

$$Total\ Reliability\ (TR) = \sum_{k=0}^{n} (MTBF + MTTR) \quad (4)$$

Where

MTBF represents Mean time between failures, MTTR represents Mean time to repair and K represents number of modules

5.3 Modularity Risk Component (C)

Modularity is designing a system that is divided into a set of functional units (named modules) that can be composed into a larger application. A module represents a set of related concerns. Coupling and Cohesion is used to measure the quality of a module [12].

The total impact of Modularity risk component can be measured using following metrics

$$\text{Total Impact of Modularity Risk (TIMR)} = 1/\text{Total Modularity(TM)} \quad (5)$$

As total modularity (TM) can be calculated as

$$\text{Total Modularity (TM)} = 1 - \frac{1}{di + 2 * Ci + do + 2*Co + gd + 2*gc + w + r} \quad (6)$$

Where

di: number of input data parameters, ci: number of input control parameters, do: number of output data parameters, co: number of output control parameters , gd: number of global variables used as data , gc: number of global variables used as control , w: number of modules called (fan-out), r: number of modules calling the module under consideration (fan-in).

TM makes the value larger the more coupled the module is. This number ranges from approximately 0.67 (low coupling) to 1.0 (highly coupled).

5.4 Usability Risk Component (D)

Usability is defined as how usable is your software program that means conveniences and practicality of use. From the customers' standpoint, all problems they encounter while using the legacy system, not just the valid defects, are problems with the software. Problems that are not valid defects may be usability problems [13]. The total impact of usability risk component can be measured using following metrics

$$\text{Total Impact of Usability Risk (TIUR)} = 1/\text{ Total Usability (TU)} \quad (7)$$

$$\text{Total Usability (TU)} = \frac{\text{Total problems that customers reported for a time period}}{\text{Total number of license-months during the period}} \quad (8)$$

Where

Number of license-months = Number of install licenses of the software * Number of months in the calculation period.

5.5 Performance Risk Component (E)

Performance of a software system is defined as the fulfillment of system purpose without waste of resources, such as memory, space and processor utilization, network bandwidth, time, etc. Correctness is an important measure for the measurement of the performance risk component. One of the most common measures is Defects per KLOC. KLOC means thousands (Kilo) of Lines of Code.) [14].

The total impact of performance risk component can be measured using following metrics

$$\text{Total impact of Performance Risk (TIPR)} = 1/\text{Total Performance (TP)} \quad (9)$$

As the total performance can be calculated as,

$$\text{Total performance (TP)} = E / (E + D) \quad (10)$$

Where
E represents No. of Errors found before delivery of the software
D represents No. of Errors found after delivery of the software.

Ideal value of TP should be 1 which means no defects found. If TP score low it means to say there is a need to re-look at existing legacy system. It encourages the team to find as many defects before they take decision to evolve legacy system.

6 Experiment and Analysis

The two different legacy library management systems are used to check the correctness of RrMm model. The total impact of five identified risk is calculated for each legacy library management system using different measurement metrics for respective risk components. Applying RrMm pentagram model to five risk components we have the following results:

$$\textit{Total Risk Impact (TRI) (LS1)} = 0.48\ X\ (AB + BC + CD + DE + EA)$$
$$\textit{TRI}=0.48\ X\ [(.45*.30) + (.30*.36) + (.36*.60) + (.60*.50) + (.50*.45)]$$
$$\textit{TRI}=.984$$
$$\textit{Total Risk Impact (TRI) (LS2)} = 0.48\ X\ (AB + BC + CD + DE + EA)$$
$$\textit{TRI}=0.48\ X\ [(.60*.50) + (.50*.70) + (.70*.60) + (.60*.70) + (.70*.60)]$$
$$\textit{TRI}=1.91$$

Where LS1 represents legacy system 1 and LS2 represents legacy system2

Table 1. Results

System	A	B	C	D	E	TRI
LS1	.45	.30	.36	.60	.50	.984
LS2	.60	.50	.70	.60	.70	1.91

Based on the TRI values of two legacy library management system the measurement results for both systems i.e. LS1 and Ls2 tests are shown in Figure 2. It is clear that the TRI of LS2 (Left) is higher than the TRI of LS1 (Right).

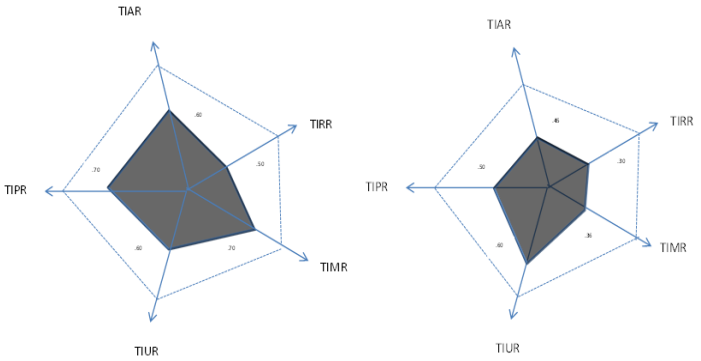


Fig. 2. Comparative Analysis

The most widely-used metric to quantify perceptual quality impairment severity due to total reengineering risk impact is the mean opinion score. We give comparative TRI values for LS1 and LS2 with the mean opinion score. The mean opinion scores and the corresponding RrMm model based impact values are shown on the Table 2. It shows that reengineering is successful if the TRI value is less than or equals to 1 and the values higher than this level required massive risk engineering or tends to reengineering failure.

Table 2. Mean Opinion Score

Level of Satisfaction	Risk Impact	Range of values
Reengineering successful	Low TRI value	0-1
Need Risk Resolution strategy	Average TRI value	1-1.5
Massive Risk engineering /Reengineering failure	High TRI value	1.5-2.4

7 Conclusion

Software reengineering is a modernizing technique used for revolution of any legacy system. Reengineering involves analyzing the current state of the legacy system and reengineers the system state according to the needs of current business environment. Successful reengineering required to measure wide-ranging impact of reengineering risk in the evolution process of legacy system. In this paper, we propose a measurement-based model RrMm to estimate the comprehensive impact of reengineering risk arises from system, managerial and technical domain of legacy system. We also use RrMm model for quantifying the impact of reengineering risk arises from quality perspective of legacy system. Finally to check the correctness of referred reengineering risk measurement model RrMm for quality perspective of legacy system we compare Total risk Impact TRI of two legacy system of the same domain. As a final point founded on the experimental result a mean opinion score board is developed to take decision about when evolution of a legacy system through reengineering is successful.

References

1. Brodie, M.L., Stonebraker, M.: Migrating Legacy Systems: Gateways, Interfaces, & the Incremental Approach. Morgan Kaufmann Publishers, Inc. (1995)
2. Garcia, F., Piattini, M., Ruiz, F., Visaggio, C.: Maintainability of Software Process Models: An Empirical Study. In: Ninth European Conference on Software Maintenance and Reengineering, CSMR 2005, pp. 246–255 (2005), doi:10.1109/CSMR.2005.35
3. Rajavat, A., Tokekar, V.: SysRisk –A Decisional Framework to Measure System Dimensions of Legacy Application for Rejuvenation through Reengineering. Published in International Journal of Computer Applications (IJCA) 16(2), 16–19 (2011), doi:10.5120/1985-2674, ISBN: 978-93-80747-56-8

4. Rajavat, A., Tokekar, V.: ReeRisk –A Decisional Risk Engineering Framework for Legacy System Rejuvenation through Reengineering. In: Das, V.V., Stephen, J., Chaba, Y. (eds.) CNC 2011. CCIS, vol. 142, pp. 152–158. Springer, Heidelberg (2011)
5. Huang, Y., Kintala, C., Kolettis, N., Fulton, N.D.: Software rejuvenation: Analysis, Modeling, and applications. In: Twenty-Fifth International Symposium on Fault-Tolerant Computing, FTCS-25. Digest of Papers, pp. 381–390 (1995), doi:10.1109/FTCS.1995.466961
6. Trivedi, K.S., Vaidyanathan, K., Goseva-Popstojanova, K.: Modeling and analysis of software aging and rejuvenation. In: Proceedings of the 33rd Annual Simulation Symposium, pp. 270–279 (2000), doi:10.1109/SIMSYM.2000.844925
7. Lindroos, J.: Code and Design Metrics for Object-Oriented Systems. Seminar on Quality Models for Software Engineering, Department of Computer Science, pp. 1–10. University of Helsinki (2004)
8. Basili, V.R., Caldiera, G., Dieter Rombach, H.: The Goal Question Metric Approach, technical report, department of computer science, institute for advanced computer studies, University of Maryland
9. Westfall, L.: Kiviat Charts, technical report, the west fall team, partnering for software excellence
10. Gong, Y., Yang, F., Huang, L., Su, S.: Model-Based Approach to Measuring quality of Experience. In: First International Conference on Emerging Network Intelligence, pp. 29–32 (2009), doi:10.1109/EMERGING.2009.17
11. Reliability HotWire eMagazine, Relationship between Availability and Reliability, Reliability Engineering Newsletter (26) (April 2003)
12. Gui, G., Scott, P.D.: New Coupling and Cohesion Metrics for evaluation of software component Reusability. In: The 9th International Conference for Young Computer Scientists, pp. 1181–1186 (2008), doi:10.1109/ICYCS.2008.270
13. Kan, S.H.: Metrics and models in software quality engineering, 2nd edn. Pearson education, ISBN 0201729156
14. Garvey, P.R., Cho, C.-C.: An Index to Measure System's Performance Risk. Tutorial (Spring 2003)

Association Based Image Retrieval: A Survey

R. Priyatharshini* and S. Chitrakala

Easwari Engineering College, Anna University, Chennai
{priya.sneham,ckgops}@gmail.com

Abstract. With advances in the computer technology and the World Wide Web there has been an explosion in the amount and complexity of multimedia data that are generated, stored, transmitted, analyzed, and accessed. In order to extract useful information from this huge amount of data, many content based image retrieval (CBIR) systems have been developed in the last decade. A typical CBIR system captures image features that represent image properties such as color, texture, or shape of objects in the query image and try to retrieve images from the database with similar features. Recent advances in CBIR systems include relevance feedback based interactive systems. In order to make the Image Retrieval more effective, Researchers are moving towards Association-Based Image Retrieval (ABIR) a new direction to CBIR. Image retrieval can be done effectively by associating the textual and visual clues of an image by reducing the semantic gap. This Survey paper focuses on the detailed review of different methods and their evaluation techniques used in the recent works based on ABIR systems. Finally, several recommendations for future research directions in ABIR have been suggested based on the recent technologies.

Keywords: Content Based Image Retrieval, Text Based Image Retrieval, Semantic Gap, Semantic Image Retrieval, Association Based Image Retrieval.

1 Introduction

The dramatic growth of digital media at home, in enterprises, and on the web, has over the last decade spawned great interests in developing methods for effective indexing and searching of desired visual contents to open the value of these contents. Conventional text based search techniques have been widely used in commercial search engines over large content corpora, such as World Wide Web. However, using text-based search techniques on non-textual unstructured content, such as image and video data, is not nearly as mature or effective as on text documents. In fact, this kind of approach works fairly well for retrieving images with text annotation such as named entities, e.g., specific people, objects, or places.

However, it does not work well for generic topics related to general settings of objects, as the text annotation rarely describes the background setting or the visual appearance, such as color, texture, shape, and size of the objects .Because of these and other limitations, it is now apparent that conventional text search techniques on their own are not sufficient for effective image and/or video retrieval, and they need to be

* Corresponding author.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 17–26, 2013.

combined with techniques that consider the visual features of the content. Thus, we need a seamlessly integrated paradigm of text-based and visual content-based query and retrieval in a unified interface. Hierarchy of Image Retrieval System is shown in Fig.1.

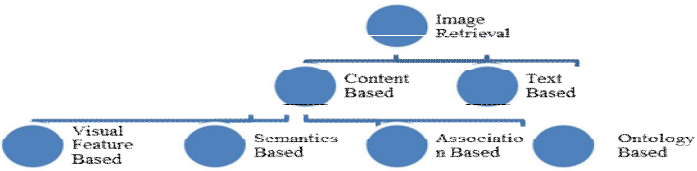


Fig. 1. Hierarchy of Image Retrieval System

2 Text Based Image Retrieval

In conventional image retrieval system, keywords are used as descriptors to index an image however the content of an image is much richer than what any set of keywords can express. Text Based Image Retrieval(TBIR) techniques employ text to describe the content of the image which often causes ambiguity and inadequacy in performing an image database search and query processing. In TBIR systems the first thing to do is providing textual descriptions/annotations for images, which is very tiresome, inefficient, and expensive and has a problem of undesirable mismatch due to annotation impreciseness.

Text annotation is extremely tedious in large image collections. To provide text descriptions or annotations, two approaches can be applied. The first approach acquires annotations *manually* by human annotators. The second approach is to *automatically* annotate images using machine-learning techniques that learn the correlation between image features and textual words. In both cases manual labeling is too expensive while automatic methods are not reliable. In recent years indexing and retrieval approaches based on keywords and visual features together are getting attentions.

3 Content Based Image Retrieval

CBIR has been used as an alternative to TBIR. IBM was the first, who take an initiative by proposing query-by image content (QBIC). QBIC developed at the IBM Almaden Research Center is an open framework and development technology. CBIR can be categorized based on the type of features used for retrieval which could be either low level or high level features. At early years, low level features include color, texture, shape and spatial relations were used. The CBIR researches were done in retrieving the image on the basis of their visual content. Neither a single feature nor a combination of multiple visual features could fully capture high level concept of images. Besides, due to the performance of Image retrieval based on low level

features are not satisfactory, there is a need for the mainstream of the research converges to retrieval based on semantic meaning by trying to extract the cognitive concept of a human to map the low level image features to high level concept (semantic gap).

In practical applications both text based and content based systems have limitations. .Ontology-based image retrieval has two components: semantic image annotation and semantic image retrieval. Semantic annotation describes about the semantic content in images and retrieval queries. The subjectivity of human perception is one of the key motivating reasons to make use of interaction model and specifically relevance feedback in CBIR systems. For instance, people under different circumstances may recognize the same image content in a different way. Continuous learning is other motivating factor to employ these kinds of techniques. In a typical Relevance Feedback (RF) approach, retrieval system provides initial image results in response to the query of user (by image or keyword). Then, user's judgment (relevant/irrelevant) on the retrieved images is employed by specific algorithm for tuning system parameters. These steps are iteratively carried out till user is satisfied with the image results. The RF approach has been used to reduce the semantic gap between high-level semantics and low-level visual features. Even though there are a lot of significant efforts and works on image retrieval research, there are still some spaces which need to be improved besides the challenges that is associated with mapping low level to high level concepts.

3.1 Related Work

Currently, most image retrieval systems use either purely visual features or textual metadata associated with images. Bridging the semantic gap for image retrieval is a very challenging problem yet to be solved. They have advantages and disadvantages respectively. To overcome their drawbacks and improve the performance without sacrificing the efficiency, a multimodal image retrieval scheme has been proposed which is based on annotation keywords and visual content, which can benefit from the strength of text and content-based retrieval.

In [5] a technique for Learning Image-Text Associations has been proposed. An associated text and image pair and an illustration of implicit associations between visual and textual features are shown in Fig.2. In [6] a technique for Web Image Co-Clustering Based o Tag and Image Content Fusion has been proposed. The current error-driven fusion method is more efficient than the traditional flat tag co-clustering methods. In [7] a technique for Multimodal Image Retrieval Based on Annotation Keywords and Visual Content has been proposed. The System architecture is shown in Fig.3 which achieves annotation offline, while performs the image retrieval online.

A technique for Content-Based Geospatial Information Retrieval using Associative Mining Techniques for Automatic learning of geospatial intelligence is proposed in [1] is challenging due to the complexity of articulating knowledge from visual patterns and to the ever-increasing quantities of image data generated on a daily basis. The knowledge-discovery algorithm uses content-based methods to link low-level

Fig. 2. Association of Text and Image pair as in [3]

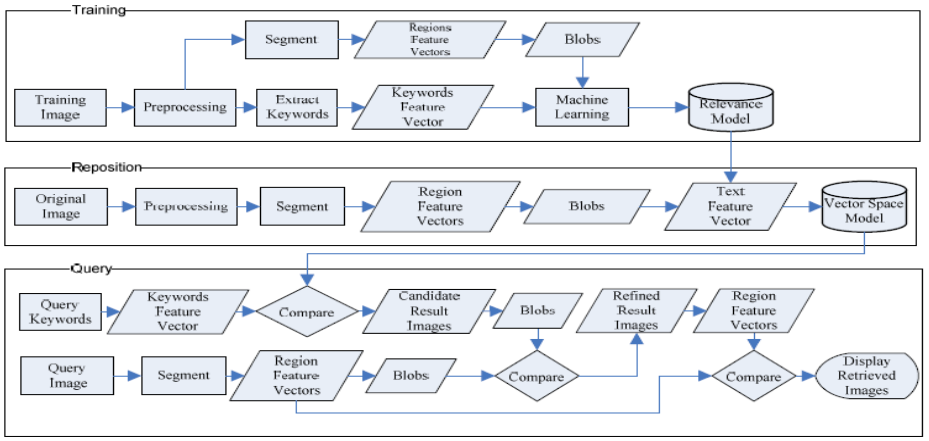


Fig. 3. System Architecture of work proposed in [7]

image features with high-level visual semantics in an effort to automate the process of retrieving semantically similar images. This algorithm represents geospatial images by using a high-dimensional feature vector and generates a set of association rules that correlate semantic terms with visual patterns represented by discrete feature intervals.

A technique for Master Defect Record Retrieval Using Network-Based Feature Association (NBFA) is proposed in [2]. As electronic records (e.g., medical records and technical defect records) accumulate, the retrieval of a record from a past instance with the same or similar circumstances has become extremely valuable. This is because a past record may contain the correct diagnosis or correct solution to the current circumstance. These two records are referred as master and duplicate records. A new paradigm has been proposed for master defect record retrieval using NBFA for training the master record retrieval process by constructing feature associations to limit

the search space. The retrieval paradigm was employed and tested on a real-world large-scale defect record database from a telecommunications company. The empirical results suggest that the NBFA was able to significantly improve the performance of master record retrieval, and should be implemented in practice.

A technique for extracting Spatial Semantics in Association Rules for Weather Forecasting Image has been proposed in [12]. A new approach to apply association rule on weather forecasting image has been proposed by building an image system to store weather forecasting images and retrieve them later for further research, for example, to predict future temperature, relative humidity, rainfall, wind speed and atmospheric pressure. Spatial reference method can be used to get the spatial relationships between objects for a certain image. Spatial association rules are also mined and are subsequently used as a basis for retrieving additional images.

3.2 Issues in Content Based Image Retrieval

The semantic gap is the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation. User seeks semantic similarity, but the database can only provide mathematical similarity by data processing. Semantic gap characterization is a difficult problem and is a challenge to the Image Processing computer scientist. The hierarchy of steps from low level features to semantics involves

- Low level feature extraction
- Segmentation
- Salient regions

Effectiveness of CBIR system depends on the choice of the set of visual features and on the choice of the similarity metric that models the user perception of similarity.

• Level 1: Retrieval by primitive features such as color, texture, shape or the spatial location of image elements. Typical query is query by example, ‘find pictures like this’.
•Level 2: Retrieval of objects of given type identified by derived features, with some degree of logical inference. For example, ‘find a picture of a flower’.
•Level 3: Retrieval by abstract attributes, involving a significant amount of high-level reasoning about the purpose of the objects or scenes depicted. This includes retrieval of named events, of pictures with emotional or religious significance, etc. Query example, ‘find pictures of a joyful crowd’.

Levels 2 and 3 together are referred to as semantic image retrieval, and the gap between Levels 1 and 2/3 as the semantic gap. Some challenges in learning semantics in CBIR are:

- Semantic gap characterization
- Huge amount of objects to search among
- Incomplete query specification
- Incomplete image description.

Some state-of-the-art techniques in reducing the 'semantic gap' as in [11]:

1) Using object ontology to define high-level concepts and generating Semantic Template (ST) to support high level image retrieval
2) Using machine learning tools to associate low level features with query concepts
3) Introducing Relevance Feedback into retrieval loop for continuous learning of users' intention
4) Making use of both the visual content of images and the textual information obtained from the Web for WWW (the Web) image retrieval.

3.3 Analysis of Techniques in ABIR

Most Image Retrieval tasks use both precision and recall to evaluate the performance. In [5] the performance is measured in terms of precision which is defined by Precision= Nc/N. where Nc is the number of correctly identified image-text associations, and N is the total number of images. The vague-transformation based method utilizes an intermediate layer of information categories for capturing indirect image-text associations. The fusion-ART-based method learns direct associations between image and text features by employing a resonance environment of the multimedia objects. The Seven Cross Media Models used are shown in Table 1.The experimental results suggest that both methods are able to efficiently learn image-text associations from a small training data set. Comparison of Cross-Media Models for discovering image-text Associations is shown in Fig.4.

In [6] an error-driven fusion method for effectively co-clustering a large number of web images. The co-clustering method combines both the image's contents as well as the image's tags to yield good clusters. In [7] a multimodal image retrieval scheme, which adopts stepwise refinement approach including keywords, blobs, and region-based retrieval scheme. Two of the most popular evaluation measures are precision and recall. MAP is the arithmetic mean of average precision calculated over a number of different queries, while average precision is given as the mean of all the individual precision scores. Compared with prior image retrieval system, the system proposed in [7] achieved better performance with respect to MAP and recall Example of Multimodal Image Retrieval with Annotation is shown in Fig.5.

In [2] a total of 155 images in this data set were assigned with multiple labels. Using this image collection, we created seven data sets with numbers of classes ranging from two to eight by randomly removing one class at a time as given in Table 1.Getting semantic-query results from 200 geospatial images takes 0.45 s, on average, using a standard personal computer. The search time for ranking is proportional to the number of association rules and ranked images. The validations are carried out by by querying the database with duplicate record vectors and determining if the corresponding master record vector is returned in the top n retrieved document vectors from the database. Table2 reports the Experimental Results for Seven Geospatial Datasets.

Table 1. Seven Cross Media Models in Comparison as in [3]

Model	Descriptions
SDT	Single Direction Vague Transformation
DDT	Dual Direction Vague Transformation
DDT_VP_BG	DDT with visual space projection using bipartie graph based similarity matrix
DDT_VP_CT	DDT with visual space projection using category-to-text transformation based similarity matrix
fusion_ART	The fusion ART based cross media
CMRM	Cross media relevance model
DWH	Dual-wing harmonium model

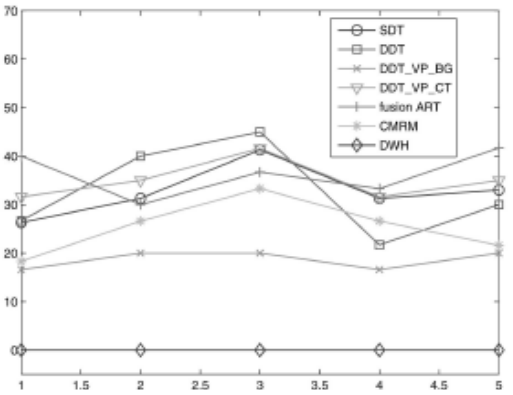


Fig. 4. Comparison of Cross-Media Models for discovering Image-Text Associations as in [3]

Table 2. Experimental Results for Seven Geospatial Datasets as in [1]

Data set	Data set size	Feature number	Ranking Precision		Classification accuracy								
			20% recall	40% recall	Bayes Net	AIRS	LVQ1	SVM	C4.5	CBA	CPAR	ATFPC	Barb & Shyu
Geospatial-2	278	2	100.0	99.5	93.88	94.96	97.12	97.84	97.84	91.73	95.71	90.99	97.48
Geospatial-3	407	3	98.5	98.3	90.9	93.85	95.08	94.59	91.89	84.9	95.07	86.72	94.07
Geospatial-4	529	4	99.1	97.2	84.68	84.12	88.22	92.62	86.57	71.1	90.05	81.09	89.56
Geospatial-5	652	5	92.1	87.5	78.06	72.54	74.69	85.42	75.46	57.36	80.36	66.88	80.82
Geospatial-6	786	6	89.2	84.2	74.04	70.86	70.1	83.33	72.64	49.84	78.1	64.36	75.69
Geospatial-7	915	7	87.9	84.2	68.19	64.91	61.53	78.68	67.97	36.32	68.63	61.53	71.63
Geospatial-8	1032	8	88.0	81.1	58.54	53.39	50.58	66.79	58.25	22.87	54.07	51.35	62.31
Average			93.5	90.2	78.32	76.37	76.9	85.61	78.66	59.16	80.28	71.84	80.63

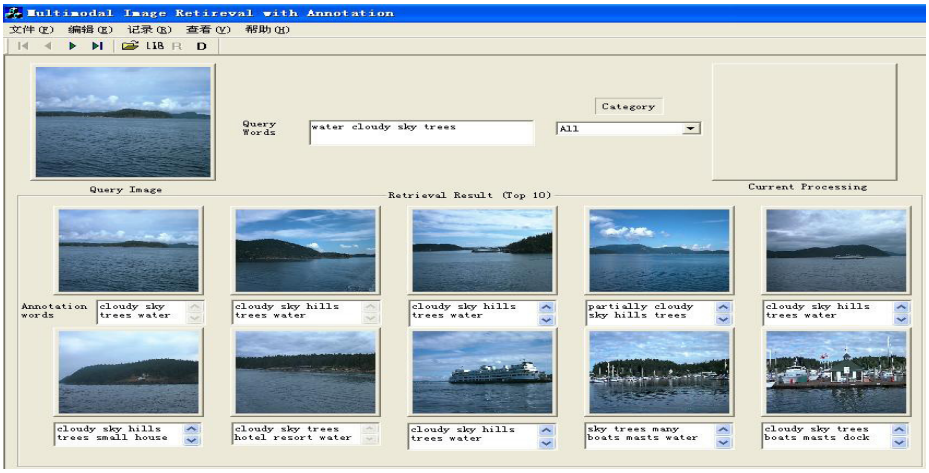


Fig. 5. Example of Multimodal Image Retrieval with Annotation proposed in [7]

4 Inferences Made

- By combining visual and textual features the overall performance of discovering cross media relationships between components of multimedia documents can be improved. Advanced similarity calculation methods with visterm and term taxonomies can be used to provide better performance.
- Multimodal Image Retrieval scheme based on annotation keywords and visual content has been used to improve the performance of Image Retrieval systems.
- The Massive visual data associated annotations have facilitated development of novel techniques for image indexing at the semantic level.
- By exploring the relationship among image content and textual terms in the associated metadata the image indexing and retrieval process can be made effective.

5 Applications

Most traditional and common methods of image retrieval utilize some method of adding metadata such as captioning, keywords, or descriptions to the images so that retrieval can be performed over the annotation words. Manual image annotation is time-consuming, to address this, there has been a large amount of research done on automatic image annotation. Semantic Ontology techniques have been developed for CBIR in order to improve the performance of Image Retrieval systems for Remote Sensing data. Many novel Image Retrieval methods has been proposed for improving the effectiveness of Web Image Retrieval systems by associating visual and textual

clues or by associating semantic and visual features to return the image results with great visual and semantic similarities. Application wise Association of Features done is summarized in Table 3.

Table 3. Association of Features in Various Domains

Domain	Researches	Association of Features
Web based Multimedia document	Ning Zhou Cheung[3]	Image-Text Association
Geo Spatial Information Retrieval	Adrian S. Barb[1]	Semantic term-Visual feature based Association
Master Defect Record Retrieval in Telecommunication	Andrew Rodriguez [2]	Network based feature Association
Weather Forecasting	SenduruSrinivasalu [14]	Association among spatial objects

6 Future Directions in ABIR

By associating visual features and textual keywords we can discover semantic concepts to explicitly exploit synergy among modalities. Association rule mining can be applied to domain of large scale multimedia data to find semantic relationship between keywords and visual features and is explored to improve the search results. Visual feature space can be represented using advanced wavelet transform based techniques. Wavelet coefficients provide approximation of visual features which captures the texture, shape and spatial information in to single signature. Textual terms or the Concept space can be represented using techniques such as Latent Semantic Indexing and Singular Value Decomposition can be used to reduce the concept space in to much lower dimensionality. Multimodality Ontology can be constructed for exploring the association between extracted visual features and textual terms from which semantic association rules can be mined. There is a lack of coincidence between high level semantic concepts and low level image features. Introducing ontology in to multimodal image retrieval will not only help better analyze text but also useful for knowledge matching.

7 Conclusion

A wide variety of researches have been made on image retrieval in multimedia databases. Each work has its own technique, contribution and limitations. As a survey paper, we might not include each and every aspect of individual works, however attempt has been made to deal with a detailed review of the association based image retrieval systems. We review those works mainly based on the methods/approaches they used to come up to an efficient retrieval system together with the limitations/challenges and the evaluation mechanisms used. This paper provides a study of image retrieval work towards narrowing down the Semantic gap. Even

though there are a lot of significant efforts and works on image retrieval research, there are still some spaces which need to be improved besides the challenges that is associated with mapping low level to high level concepts.

References

1. Barb, A.S., Shyu, C.-R.: Visual-Semantic Modeling in Content-Based Geospatial Information Retrieval Using Associative Mining Techniques. IEEE Geo Science and Remote Sensing Letters 7(1) (January 2010)
2. Rodriguez, A., Art Chaovalitwongse, W., Zhe, L., Singhal, H., Pham, H.: Master Defect Record Retrieval Using Network-Based Feature Association. IEEE Transactions on Systems, Man, and Cybernetics— PART C: Applications and Reviews 40(3), 319–329 (2010)
3. Zhou, N., Cheung, W.K., Qiu, G., Xue, X.: A Hybrid Probabilistic Model for Unified Collaborative and Content-Based Image Tagging. IEEE Transactions on Pattern Analysis and Machine Intelligence 33(7), 1281–1294 (2011)
4. Su, J.-H., Chou, C.-L., Lin, C.-Y., Tseng, V.S.: Effective Semantic Annotation by Image-to-Concept Distribution Model. IEEE Transactions on Multimedia 13(3), 530–538 (2011)
5. Jiang, T., Tan, A.-H.: Learning Image-Text Associations. IEEE Transactions on Knowledge and Data Engineering 21(2), 161–177 (2009)
6. Chen, J., Tan, J., Yin, X., Liao, H.: Web Image Co-Clustering based on Tag and Image Content Fusion. In: 2nd IEEE International Conference on Network Infrastructure and Digital Content, pp. 378–382 (December 2010)
7. Song, H., Li, X., Wang, P.: Multimodal Image Retrieval Based on Annotation Keywords and Visual Content. In: IITA International Conference on Control, Automation and Systems Engineering, CASE, pp. 295–298 (July 2009)
8. Rafiee, G., Dlay, S.S., Woo, W.L.: A Review of Content-Based Image Retrieval. In: CSNDSP, pp. 775–779. IEEE (2010)
9. Wang, H., Liu, S., Chia, L.-T.: Image retrieval with multi-modality ontology. Multimedia Systems 13, 79–390 (2008)
10. Setia, L.: Machine Learning Strategies for Content Based Image Retrieval, M.Sc Dissertation, Albert-Ludwigs-Universität, Freiburg im Breisgau (2008)
11. Srinivasulu, S., Sakthivel, P.: Extracting Spatial Semantics in Association Rules for Weather Forecasting Image. Trendz in Information Sciences & Computing (TISC), 54–57 (December 2010)

Development of Wireless Sensor Node to Monitor Poultry Farm

Santoshkumar[1], Kelvin Chelli[2], and Suresh Chavhan[1]

[1] Electrical Engineering Department, SDM College of Engineering and Technology, Dharwad, Karnataka, India
[2] Electronics Engineering Department, Siddaganga Institute of Technology, Tumkur, Karnataka, India
santoshkumar777@yahoo.com,
{kelvin.chelli,suresh.chavhan046}@gmail.com

Abstract. In this paper we have proposed a Wireless Sensor Network (WSN) System and developed sensor node to monitor the Poultry farm and detect the highly pathogenic avian influenza virus in Chickens. WSN's have tremendous potential and play a key role in such monitoring applications. The developed wireless sensor node monitors the body temperature and movement of the Chickens. Based on sudden variation in body temperature and slow movements of the Chicken, early detection of virus is possible. The proposed low cost WSN System was developed with the use of open source technology-ARDUINO (open electronics prototype and open source).

Keywords: Wireless Sensor Network (WSN), ZigBee, XBee, Avian Influenza, Arduino.

1 Introduction

Avian influenza is an infection caused by avian (bird) influenza (flu). This influenza virus occur naturally among birds. Wild birds worldwide get flu A infections in their intestines, but usually do not get sick from flu infections. However, avian influenza is very contagious among birds and some of these viruses can make certain domesticated bird species, including chickens, ducks, and turkeys, very sick and kill them. Infected birds can shed influenza virus in their saliva, nasal secretions. Susceptible birds become infected when they have contact with contaminated secretions or excretions. Domesticated birds may become infected with avian influenza virus through direct contact with infected waterfowl or other infected poultry. Infection with avian influenza viruses in domestic poultry causes two main forms of disease that are distinguished by low and high extremes of virulence. The "low pathogenic" form may go undetected and usually causes only mild symptoms (such as ruffled feathers and a drop in egg production). However, the highly pathogenic form spreads more rapidly through flocks of poultry. This form may cause disease that affects multiple internal organs and has a mortality rate that can reach 90-100% often within 48 hours. One approach to solve this problem would be to employ

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 27–32, 2013.

a Sensor Network that would enable users to monitor the required factors such as body temperature and movement using sensors. The sensed parameters can be viewed at the receiver side in a personal computer. The smart sensor network system is designed using open source technology-ARDUINO. ARDUINO is an open electronics prototype and open source.

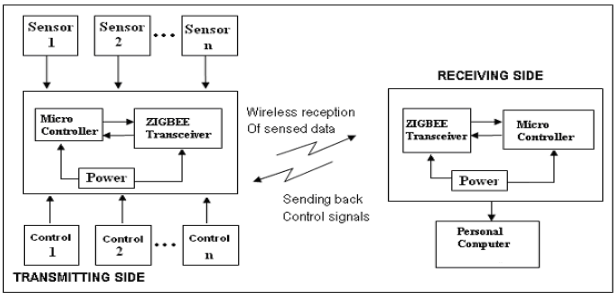


Fig. 1. Simplified block diagram of WSN

2 Related Work

After reviewing many articles, there are very few papers that discuss about monitoring and detection of the avian influenza in Poultry farm. In the available paper, there is mention of body temperature and movements of Chickens in poultry farm [1], but there is less description of wireless technology. However, there are papers which address the societal concerns like agriculture, animal, asset, vehicle monitoring using WSN's.

3 Wireless Sensor Networks

Wireless sensor network is used in special situation for signal collection, processing and transmitting. A wireless sensor network (WSN) is a system consisting of a collection of nodes and a base station. A node is composed by a processor, local memory, sensors, radio and battery and a base station is responsible for receiving and processing data collected by the nodes. They perform collaborative activities due to limited resources, such as battery, processor and memory [2]. Nowadays, the applications of WSN's are many and varied, and the applications in monitoring Poultry farm are still incipient. One interesting WSN application is in poultry farm, where the conditions such as body temperature and movement are to be monitored. To monitor and detect avian influenza, sensors are essential. In such application, the sensors are fixed on the abdomen of the Chicken and WSN should guarantee real time monitoring. A WSN consists of generally two main parts: Data Acquisition Network

and Data Distribution Network. The figures fig.1 show a simplified WSN block diagram [2].The block diagram of wireless sensor network consists of two main parts, transmitting side and receiving side respectively. At the transmitting side, various sensors and controllers are connected with the FREEDUINO and XBEE. At the receiving side the transceiver and the microcontroller are present along with the display device.

4 ZigBee Protocol Overview

ZigBee is an established set of specifications for wireless personal area networking (WPAN). The relationship between IEEE 802.15.4 and ZigBee is similar to that between IEEE 802.11 and the Wi-Fi Alliance.

Coordinator
ZigBee networks always have a single coordinator device. This radio is responsible for forming the network, handing out addresses, and managing the other functions that define the network, secure it, and keep it healthy.
Router
A router is a full-featured ZigBee node. It can join existing networks, send information, receive information, and route information.
End device
There are many situations where the hardware and full-time power of a router are excessive for what a particular radio node needs to do. End devices are essentially stripped-down versions of a router. They can join networks and send and receive information. End devices always need a router or the coordinator to be their parent device.

5 Methodology

A very systematic approach was considered for the overall design of the project, in which two parameters were monitored. The temperature sensor is used to monitor the body temperature and 3-axis accelerometer is used to monitor the activities or movements of the Chicken. There is sudden reduction observed in the body temperature and activities of the infected Chicken. The node is designed for the increased battery life and the ZigBee technology supports the same. The hardware details of a sensor node are shown in figure Fig.2 below.

5.1 Poultry Farm Details

A Poultry farm is a place where Chickens grow and stay. It provides a friendly location for growing Chickens and protect them from atmospheric variation. The temperature sensor is used to monitor temperature and accelerometer is used to monitor movements.

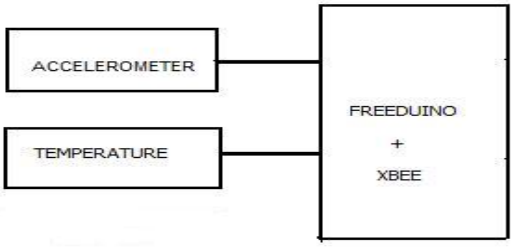


Fig. 2. Hardware details

5.2 Hardware Details

The Arduino/Freeduino board uses the ATMega328 and was selected as the microcontroller. This was an ideal selection, as the processor is extremely strong and cost efficient. An on-board 10-bit analog to digital converter (ADC), aids in the digitization of the analog signal acquired from the sensor.

Temperature Sensor
The LM35 temperature sensor produces an analog voltage directly proportional to temperature with an output of 1 mill volt per 0.1°C (10 mV per degree). The sketch(program) converts the analogRead values into milli volts and divides this by 10 to get degrees.

Accelerometer
The ADXL320 accelerometer is used to detect the motion of the Chicken. The accelerometer is connected to Arduino and corresponding readings are recorded.

6 Design Issues

The prototype wireless sensor node is designed and the weight is kept at minimum for a six weeks old Chicken, to which the sensor node is to be tied to the abdominal surface. The sensor node is designed keeping in mind the battery life. We have considered the temperature range from 38 to 40 degrees centigrade. The figure Fig.3 shows developed wireless sensor node. The Threshold value is set to 36 degrees and if the body temperature of the Chicken is below this the alarm is sent to observing station. The body temperature and the movements of the Chicken are observed and recorded using sensors. The software design consists of two main sections: Transmitting side and Receiving side. At the transmitting side, various sensors are connected to the FREEDUINO. Continuously the values are monitored and the details of distantly located sensors are stored in the base station as data base. The parameters read are temperature and movements of the Chicken.

At the Base Station(receiving side) the transceiver and the FREEDUINO are present along with the display device. The sensed parameters with their precision values are transmitted to the base station through Wireless Communication and details are monitored by the administrator and stored in the data base.

Fig. 3. Developed Wireless Sensor Node

7 Results

The various laboratory based experiments were performed to validate the proposed hardware design. The node was connected to the abdomen of the Chicken. The prototype developed demonstrated promising results. The measured body temperature and movements of the Chicken are shown in figures Fig.4 and Fig.5.

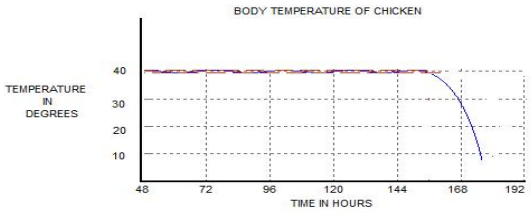


Fig. 4. Body temperature of Chicken

Fig. 5. Movements of the Chicken

8 Conclusion

The developed prototype wireless sensor network system using open source technology-ARDUINO proved to successfully acquire accurate measurements for the above mentioned parameters like body temperature and movement of the Chickens. The alarm is sent to the observing station, if the value received at the observing station is less than the threshold. This helps us to detect the H5N1 virus before the Chickens die. The mortality rate is reduced by using such kind of systems.

References

1. Hironao Okada, S., Suzuki, K., Kenji, T., Itoh, T.: Wireless sensor system for detection of avian influenza outbreak farms at an early stage. In: IEEE SENSORS 2009 Conference (2009)
2. Santoshkumar, Wakde, D.G.: Development of Wireless Sensor Node for Precision Agriculture. In: IEEE International Conference on Computational Intelligence and Computing Research (IEEE-ICCIC-2011) at Kanyakumari, India (2011)
3. Santoshkumar, Rakhee, K.: Design and Development of Wireless Sensor Network System based on ZIGBEE for Greenhouse. In: National Conference on Advances in Wireless Communications (NCACW 2010) Shegaon, India (2010)
4. Akyildiz, F., et al.: A Survey on Sensor Networks. IEEE Communications Magazine 40(8) (2002)
5. Chong, C., Kumar, S.P.: Sensor Networks: Evolution, Opportunities, and Challenges. Proc. IEEE 91(8) (2003)
6. Culler, D., Estrin, D., Srivastava, M.: Overview of Sensor Networks. Computer 37(8) (2004)

Energy Optimization in Wireless Sensor Network

M.M. Chandane[1], S.G. Bhirud[1], and S.V. Bonde[2]

[1] Department of Computer Technology, VJTI, Mumbai, India
[2] Department of Electronics and Telecommunication, SGGSIE&T, Nanded, India
{mmchandane,sgbhirud}@vjti.org.in, svbonde11@yahoo.com

Abstract. Energy consumption is a major challenge in wireless sensor network (WSN). Most of the routing algorithms focus on energy efficient paths, For the analysis of such algorithms at low cost and in less time; we believe that, simulation gives the better approximation. Therefore, in this paper, we are proposing a simulation model for WSN. Literature survey is done on energy-aware routing protocols, in which, it is found that, Minimum Total Transmission Power Routing (MTPR) and Minimum Battery Cost Routing (MBCR) Protocol, most comprehensively captures tradeoffs of energy efficiency and network lifetime respectively. Proposed simulation model is implemented using Qualnet 4.5 and applied to MTPR and MBCR to analyze its performance. Through this study, we lay a foundation for further research work on enhancements in extending the network lifetime of WSN. Experimental result shows that there is always a tradeoff between energy efficiency and network lifetime.

Keywords: Energy Aware Routing Protocol, Wireless Sensor Network, Quality of Service, Power Aware Routing, Qualnet 4.5.

1 Introduction

Traditional routing protocols are not designed as per the specific requirement of WSN. Therefore, energy efficient routing paradigms are an area of active research [1]. Power aware routing protocols make routing decisions based on power, energy related metrics to optimize the performance of routing protocol. The two routing objectives of "minimum total transmission energy" and "total working time of the network" can be mutually contradictory. For example, when several minimum energy routes share a common node, the battery power of this node will quickly run into depletion, shortening the network lifetime. In WSN, excessive energy conservation, neglecting energy consumption at individual nodes, speeds up network partition by draining batteries of the energy critical nodes in the network. In effect, it shortens the network lifetime. On the other hand, not considering overall energy consumption, this commits to paths with large number of hops and longer total distance. Consequently, the total energy dissipated is high and on an average, the battery power decays faster. In effect, it also shortens the network lifetime. The above observations suggest that both battery level and transmission energy needs to be considered while designing energy aware routing protocols.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 33–40, 2013.

The paper is organized as follows: Section 2 presents related work. Proposed simulation model is discussed in section 3. Section 4 presents experimental set-up. Results are discussed in section 5 and finally, conclusion is given in section 6.

2 Related Work

This section gives the brief information of research in routing protocols for WSN. In [1], the authors have said that the Traditional routing algorithms are not optimized for energy conservation; therefore energy efficient routing paradigms are an area of active research. Ian et al. [2] present a survey of sensor networks. W. R. Heinzelman et. al. [3] have proposed LEACH (Low Energy Adaptive Clustering Hierarchy), a clustering based protocol that utilizes random selection of cluster head. Itanagonwiwat et al. [4] have presented Directed Diffusion Protocol [DDP]. In this protocol, a sink requests data by sending interests for named data. Data matching the interest is then drawn toward that node. Intermediate nodes can cache or transform data, and may direct interests based on previously cached data. Ganesan et al. [5] present Multi path Routing algorithm. The focus of this algorithm is to extend the working time of network. A family of protocols called Sensor Protocols for Information via Negotiation (SPIN) is proposed in [6]. SPIN is a source initiated directed diffusion scheme, developed at the Manchester's Institute of Technology (MIT). SPIN uses negotiation and resource adaptation to address the deficiencies of flooding. N. Chilamkurti et. al [7], have exploited cross-layer optimizations technique that extends the DSR to improve its routing efficiency by minimizing the frequency of reforming routes. In [8], Nadeem Ahmed et. al, have proposed a multi hop routing protocol for ZigBee based WSN that considers the interference caused by WiFi networks and uses multiple channels at different frequencies to increase the network throughput. Hui Wang et. al. [9] investigates a cross-layer design approach to minimize energy consumption and maximize network lifetime (NL) of a multiple source and single sink (MSSS) WSN. C.-K. Toh [10] has discussed Minimum Total Transmission Power Routing and Minimum Battery cost routing along with the Network Performance Parameters. In [11], S. Singh, M. Woo and C. S. Raghavendra have presented a new power aware matrics for determing routes in Ad-Hoc network.

2.1 Minimum Total Transmission Power Routing (MTPR) [10]

This protocol focuses on end-to-end energy efficiency. Generally, the route selected for conserving energy is the shortest distance path or minimum hop path, the end-to-end shortest path naturally leads to conservation of energy in transmission. In a non-partitioned network, there exists at least one path for communication with any other node. So theoretically, any node can reach any other node in the network through a random forwarding path. However, the energy consumption along different paths differs, due to its dependence on variations of distance between directly communicating nodes and noise interference levels. The greater the values of these parameters, larger the amount of energy required for transmission. Successfully

transmission of packets requires the Signal-to-Noise Ratio (SNR) at the receiver to be greater than a predetermined threshold ψ_j that is closely related to the Bit Error Rate (BER). Mathematically, this is expressed as:

$$SNR_j = \frac{P_i G_{i,j}}{\sum_{k \neq i} P_k G_{k,j} + \eta_j} > \psi_j(BER) \tag{1}$$

where P_i is the transmission power of node n_i, $G_{i,j}$ is the path gain, inversely proportional to the distance d between nodes n_i and n_j (i.e., $G_{i,j} = 1 / d_{i,j}n$, usually n = 2 for short distance and n = 4 for longer distance) and η_j is the thermal noise at n_j.

Minimization of the power consumption can be obtained by selecting a routing path with minimum total transmission power. The transmission power P (n_i, n_j) between nodes n_i and n_j are used as the metric to construct such routing path. The total transmission power for a possible path l, P_l can be derived from:

$$P_l = \sum_{i=0}^{D-1} P(n_i, n_{i+1}) \qquad \text{For all nodes } n_i \text{ on route l} \tag{2}$$

Where n0 and n_D are the source and destination nodes. Hence, a path k will be selected if it satisfies:

$$P_k = \min_{l \in A} P_l \tag{3}$$

Where A is the set of all possible routing paths.

2.2 Minimum Battery Cost Routing (MBCR)[10]

Though the transmission power is an important metric to consider, if multiple minimum total power paths pass through some common node, then this common node will soon experience battery exhaustion. MTPR has a drawback in violating fair distribution of power consumption among nodes. It does not reflect the lifetime of individual nodes. This indicates that, as an alternative, node's residual energy can be used as a cost metric in route selection. MBCR is such a scheme that minimizes the path battery cost so as to maximize the total network life time. The cost function f in MBCR is defined such that the lower the remaining battery capacity c of a node i, the more reluctant the node is to receive and forward the packet. One possible function f is

$$f_i(c_i) = \frac{1}{c_i} \tag{4}$$

This shows that as a node's battery power decreases, the cost to include the node into the routing path increases.

By using residual power as a cost metric, MBCR avoids excessive usage of network nodes, and attempts to evenly distribute battery capacity over the network to delay network partitioning. However, it has a drawback, again because only the end-to-end consideration is taken. Although the total battery cost achieves minimum, some weak links where nodes have little residual power can still exist in the paths, which may lead to early network partitioning.

3 Proposed Simulation Model

Fig.1. shows that, the proposed simulation model has six major components: Network Scenario Designer, Network Animator, Packet Generator, Static/Mobile Scenario Generator, Routing Protocol Engine, and Statistics Analyzer.

The module Network Scenario Designer states the space boundary, number of network nodes, their location and the maximum transmission range as an input parameter. The network animator is the simulation ground for packet delivery and mobile node movement events. The number of active communicating parameters including mobility speed, pause time can be varied.

The routing protocol engine includes MTPR, MBCR, AODV, and other routing protocols whose performance is to be analyzed. MTPR, MBCR are on top of AODV, in which MTPR, MBCR handles the route selection whereas AODV manages route discovery, route maintenance and route refreshments through in cooperation with

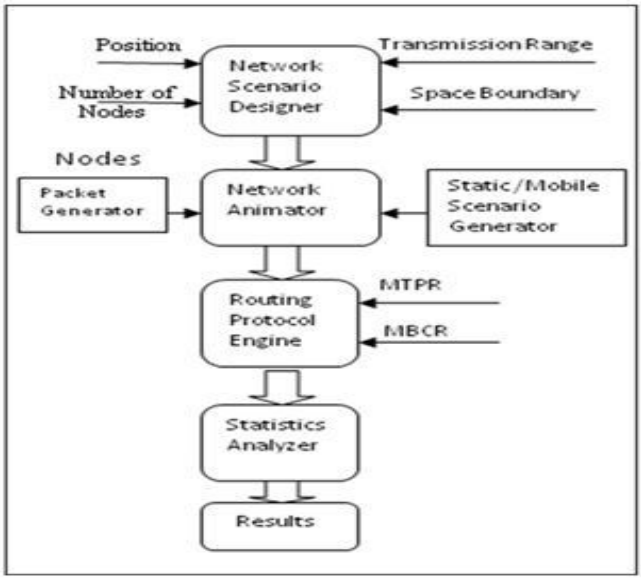


Fig. 1. WSN Simulation Model

MAC and physical layers in the TCP/IP stack. When the routing protocol engine processes packet transmission or node movement events, statistics such as energy consumption, node dead time, and other QoS parameters are recorded. Statistics analyzer examines the recorded data and draw out specified analysis results.

4 Experimental Setup

The proposed simulation model is implemented using Qualnet 4.5[12] and the performance of MTPR and MBCR is tested for the simulation of 16 nodes, forming a grid of 4*4 cells with 200M distance between every pair of nodes. IEEE 802.11 protocol is used at MAC and Physical layer. Constant bit rate (CBR) is used to apply data traffic over User Datagram Protocol (UDP), a connection between source and destination node. The simulations parameters used are as shown in table1.

Table 1. Simulation Parameters

Parameters	Values
No. of Nodes and Area	16 and 1500m*1500m
Simulation time	240 simulation units
Channel frequency	2.4GHz
Transmission range	300 meter
TX-Power	15dBm
Path Loss Model	Two Ray Model
Phy and MAC Model	802.11
Energy Model	MICAZ Mote
Battery Model	Simple Linear,1200 mAhr,
Data Rate	0.1,1,2,5,10
Payload Size	512 bytes

5 Results and Discussion

This section provides the experimental results to compare and validate the effectiveness of MTPR, MBCR. We have implemented the MTPR and MBCR algorithms in “C” and used the environment of Qualnet 4.5 for simulation.

Fig. 2 shows the residual energy of nodes after the completion of simulation using MTPR and MBCR. In MTPR, It is an end to end consideration to selects the shortest path for routing and does not consider the energy critical nodes along the path. MBCR considers node’s residual energy as a cost metric in route selection but it is again a end to end consideration. Results showed in Fig. 3 to 6 shows the effect of path selection metric used in MTPR and MBCR. Fig. 3 shows the comparative analysis of network lifetime of MBCR and MTPR. Experiment is repeated for different data rate values such as 0.1, 1, 2, 5, and 10. Fig. 3 shows that for high data rate, MTPR has low network lifetime than the MBCR. Fig. 4 shows that the amount of data received

(bits/Sec) at destination node. For all data rate values, throughput obtained using MTPR is higher as compared to that of MBCR. It can be easily judged from Fig. 3 and 4 that there is always a tradeoff between energy efficient path and network life time. Fig. 5, and 6 shows end to end delay and jitter observed for the given scenario using MBCR and MTPR respectively.

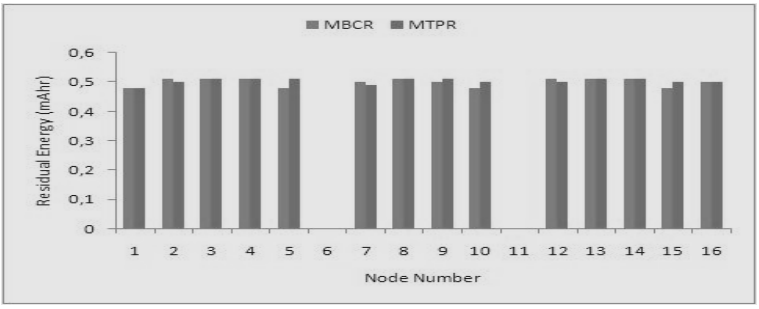


Fig. 2. Residual Energy

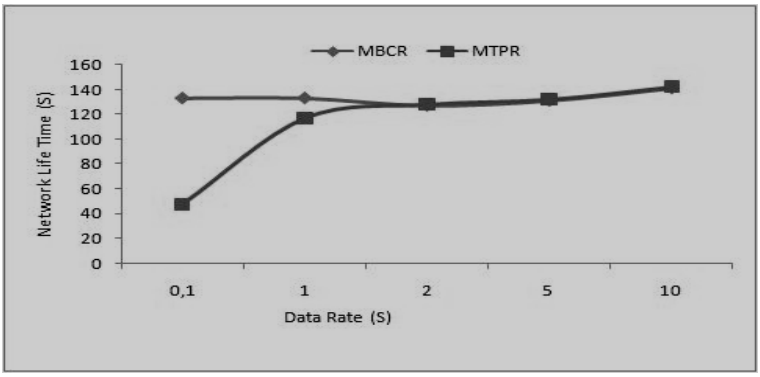


Fig. 3. Network Lifetime

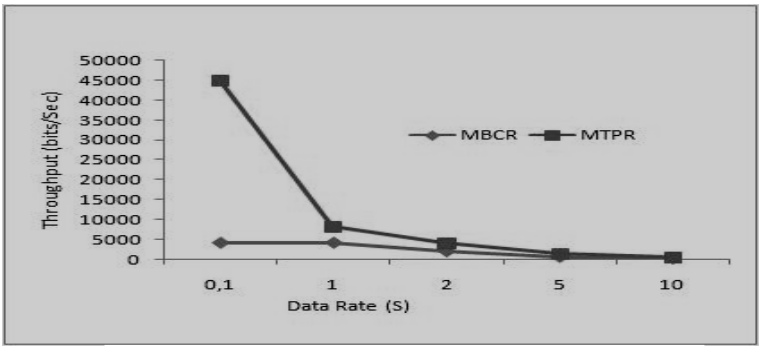


Fig. 4. Throughput

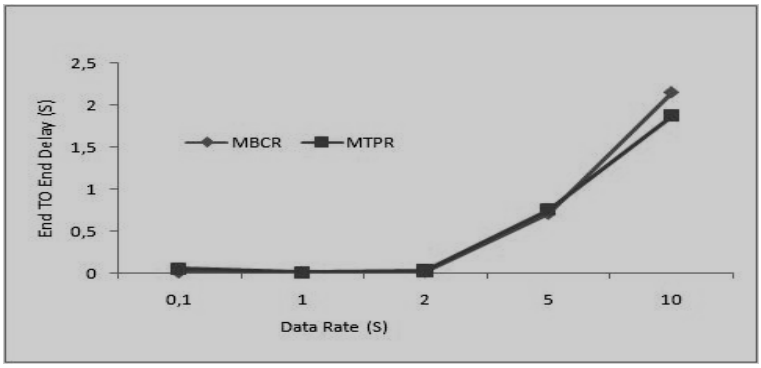


Fig. 5. Comparative Analysis of End TO End Delay

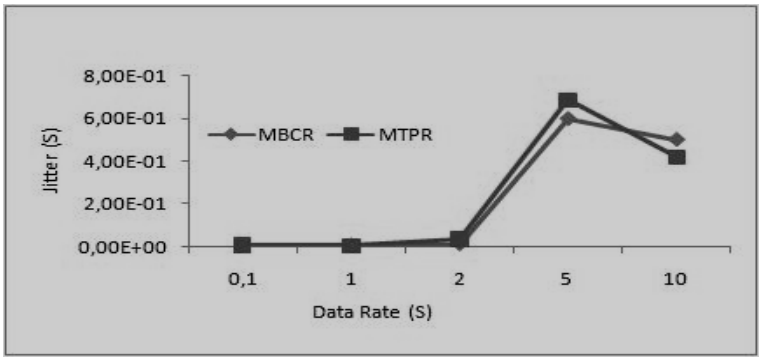


Fig. 6. Comparative Analysis of Jitter

5 Conclusion

In this paper, we have proposed a simulation model for the analysis of routing protocols and also presented some energy aware routing protocols such as Minimum Total Transmission Power Routing (MTPR) and Minimum Battery Cost Routing (MBCR) that are prominent in the research community. Proposed simulation model is applied to MTPR and MBCR, to investigate the performance they provides on energy saving, network lifetime and other Quality of Service (QoS) parameters. It is found that, as MTPR focuses on shortest distance path or minimum hop path, the total battery consumption achieves minimum, However, some weak links where nodes have little residual power can still exist in the paths, which may lead to early network partitioning. MBCR avoids excessive usage of network nodes, and tries to balance battery capacity over the network to delay network partitioning. However, because of end to end energy consideration, some weak links where nodes have little residual energy can still exist in the paths, which may lead to early network partitioning.

Future Work: Energy aware routing protocols are energy-saving strategies designed at the network layer. These are effective in power saving, but are still limited in the ability of making use of strategies or parameters designed at other layers. Through cross layering, Making use of the parameters defined at various layers such as MAC, physical, application layers and network layer are expected to bring improvements. Therefore, our further work will be based on cross layer design for enhancements in extending the energy efficiency and network lifetime of Wireless Sensor Network.

References

1. Tarannum, S., Aravinda, B., Nalini, L., Venugopal, K.R., Patnaik, L.M.: Routing Protocol for lifetime Maximization of Wireless Sensor Networks. IEEE Explore, 401–406 (2006)
2. Akylidiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E.: Wireless Sensor Network: A Survey on Sensor Networks. IEEE Communications Magazine 40(8), 102–114 (2002)
3. Heinzemman, W.R., Chandrakasan, A., Balakrishnan, H.: Energy-Efficient Communication Protocol for Wireless Microsensor Networks. In: IEEE Proceedings of the 33rd Hawai International Conference on System Sciences, pp. 1–10 (2000)
4. Intanagonwiwat, C., Govindan, R., Estrin, D.: Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor networks. In: Proceedings of 6th ACM/IEEE MOBICOM Conference, pp. 56–67 (August 2000)
5. Ganesan, D., Govindan, R., Shenker, S., Estrin, D.: Highly Resilient Energy Efficient Multi-path Routing in Wireless Sensor Networks. ACM SIGMOBILE, Mobile Computing and Communications Review 4(5) (October 2001)
6. Heinzelman, R., Kulik, J., Balakrishnan, H.: Adaptive protocols for Information Dissemination in Wireless Sensor Networks. In: Proceedings of ACM MOBICOM, pp. 174–185 (August 1999)
7. Chilamkurti, N., Zeadally, S., Vasilakos, A., Sharma, V.: Cross Layer Support for Energy Efficient Routing in Wireless Sensor Networks. Journal of Sensors 2009, Article ID 134165, 9 pages (2009)
8. Ahmed, N., Kanhere, S., Jha, S.: Poster Abstract: Experimental Evaluation of Multi-Hop Routing Protocols for Wireless Sensor Networks. In: IPSN 2010, April 12-16. ACM (2010) 978-1-60558-988-6/10/04
9. Wang, H., Yang, Y., Ma, M., He, J., Wang, X.: Network Lifetime Maximization with Cross-Layer Design in Wireless Sensor Networks. IEEE Transactions on Wireless Communications 7(10) (October 2008)
10. Toh, C.-K.: Maximum Battery Life Routing to Support Ubiquitous Mobile Computing in Wireless Ad Hoc Networks. IEEE Communications Magazine 39, 138–147 (2001)
11. Singh, S., Woo, M., Raghavendra, C.S.: Power-Aware Routing in Mobile Ad Hoc Networks. In: Proceedings of the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking (October 1998)
12. Qualnet Simulator, `http://www.scalable-networks.com`

Adaptive Delay and Energy Aware Data Aggregation Technique in Wireless Sensor Networks

S. Sivaranjani[1], S. Radhakrishnan[1], and C. Thangaraj[2]

[1] Department of Computer Science & Engineering, Anand Nagar, Krishnankoil – 626 126
{sivaranjani,srk}@klu.ac.in
[2] Vice-Chancellor, Anna University of Technology, Chennai
vc@annatech.ac.in

Abstract. In wireless sensor networks, data aggregation is performed to eliminate redundant data transmission and to enhance the network life time. The existing data aggregation techniques rarely consider the different types of applications like critical and non-critical. Also they are generally energy aware in nature but not delay aware. To minimize the delay and energy consumption in the data aggregation, this paper proposes an adaptive data aggregation technique. We consider three types of application data namely critical, bursty and regular. When the data type is critical, a shortest path with minimum delay is constructed and the critical data is transmitted directly towards the sink. In case of bursty data, the structure free data aggregation technique is used and it gets fractional aggregation. For regular data, an energy efficient aggregation tree is constructed. By simulation results, we show that the proposed technique significantly reduces the end-to-end delay, energy consumption and overhead.

Keywords: Data Aggregation, Critical Data, Non critical Data, Delay.

1 Introduction

A network having individual nodes that are distributed spatially using the sensors in order to observe the physical or environmental conditions at different locations is known as wireless sensor networks (WSN). In order to fulfill their responsibilities, the wireless communication helps in the collaboration of the nodes [1]. Based upon the sensing mechanism the data is generated by the sensor nodes and the sensed data packet is transmitted to the base station (sink). Efficient techniques which have fewer nodes for sending data to the base station are implemented since more energy is needed for transmitting data over long distances. This process is called data aggregation and the nodes are called the aggregator nodes [2]. Collection of data from multiple sensor nodes of network, aggregating data packet using aggregation function like sum, average, count, max min, sending the aggregate result to the upper aggregator node or querier node are the basic functions of aggregator nodes [3].

An important factor to consider in the data gathering application is the average delay per round. When the data gathering rounds are far apart and the only traffic in

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 41–49, 2013.

the network is due to sensor data. So the data transmissions in each round are completely scheduled. With the direct transmission scheme, nodes have to transmit to the base station one at a time, making huge delay. In order to reduce delay, one needs to perform simultaneous transmissions [4]. Many applications need the fast report with minimum energy consumptions. For such sensitive applications this paper proposes a data aggregation technique which can adaptively work in the critical and non critical situation.

2 Related Work

Mario Macedo et al. [5] have proposed Latency-Energy Minimization Medium Access (LEMMA), a new TDMA based MAC protocol for Wireless Sensor Networks to extern the network lifetime. In this approach the slot allocation is based on the interference of the node. LEMMA includes two phases: initialization phase and interference-free data transmission phase. They achieve the minimum delay with energy efficient way by using the near-optimal cascading TDMA slot allocation with a very low duty-cycle. They assume that all nods are synchronized.

Luisa Gargano et al. [6] have proposed an optimal solution to the collision free path coloring problem in any graph. By using the optimal data gathering algorithm they reduce the time to complete the process. The network is represented as a graph G = (V, E) and they find the path by coloring the edges such that the coloring is collision-free, that is, no two edges incident on a common node are assigned the same color. They use the centralized optimal time gathering algorithm to solve this issue.

Zheng Jie et al. [7] have proposed a dynamic tree based energy equalizing routing scheme (DTEER). It is a dynamic multi-hop route selecting scheme based on weight-value and height-value to form a dynamic tree. They propose the distributed algorithm to organize the nodes for the fast data gathering.

XiaoHua Xu et al. [8] have studied the problem of distributed aggregation scheduling in WSNs and propose a distributed scheduling method with an upper-bound on delay. This is a nearly constant approximate algorithm which significantly reduces the aggregation delay. This paper focuses on data aggregation scheduling problem to minimize the delay (or latency).

3 Adaptive Delay and Energy Aware Data Aggregation

Data aggregation is the common way to save the network life time. But it also raises the overall network delay. In some applications, sensor nodes report standard events periodically and report the dangerous events whenever it occurs in the sensing region. Any dangerous event should be immediately informed to the sink node. The sensed information can be typed as standard event or risky event. So the aggregation is done based on the type of the sensed information. This paper proposes an adaptive minimum Delay data aggregation by considering the following types of data which are Critical data and Non-Critical data. Non-Critical data is further classified into Bursty data and Regular data.

In case of critical data, the shortest path with minimum delay is to be constructed and using that path, the critical data is to be transmitted directly towards the sink. Initially the expected per hop delay is estimated and finds the shortest paths between each sensor node and the sink. In the case of non-critical data, it is further classified into fully aggregated and fractionally aggregated. In fractional aggregation, no aggregation tree will be established and the structure free data aggregation technique will be used for data aggregation. In full aggregation, an energy efficient aggregation tree will be constructed. An aggregation tree is a connected sub graph of G containing the sink, the source S and any other nodes and edges should connect these nodes without creating cycles. Data aggregation takes place through this aggregation tree, towards the sink.

Before sending the data, each sensor node marks the type of data and transmits it to its parent node. On receiving the data, the parent node will check the type of data and adaptively invokes the appropriate scheme. Since minimum delay based shortest path is used for transmitting critical data, the delay involved will be very much reduced. Similarly, since structure free data aggregation is used in case of fractional aggregation, the time taken for tree construction will be eliminated. By using the energy efficient aggregation tree for the full aggregation, we have reduced the energy consumption.

3.1 Finding Shortest Path with Minimum Delay

In order to find the path with minimum delay, consider a network where the transmitter to receiver path is composed of time varying number of hops h. This paper considers the following possible delays:

- The queuing delay, Q_d: The time taken by a packet, which needs to wait in the source node buffer for processing.
- The propagation delay, P_d: The time taken for the packet header to travel from the source node to the next hop neighbor in the link.
- The transmission delay, T_d: The time taken to transmit the end of the packet to the next hop receiver.

Risky information should be reported to the sink without fail. To achieve the reliable delivery it allows multiple reports. The critical data is transferred to the base station without eliminating the duplication without considering the retransmission delay. The total delay D_{ij} to transmit a packet from the i^{th} source to the next hop node j can be expressed as:

$$D_{ij} = Q_d + P_d + T_d \tag{1}$$

After estimating the per hop delay, the shortest path with minimum delay between the sink and the source is constructed by using dijkstra's shortest path algorithm [9]. Given a network represented as a graph $G = (N, E)$ where N is the set of sensor nodes with a positive edge cost D_{ij} (which represents the delay) for all edges $(i, j) \in E$ where

i and *j* are the adjacent nodes, starting node *S,* generally the sink node and a set *P* of permanently labeled nodes, the shortest path from the starting node *S* to every other node *j* in the set *P* as shown below:

```
Algorithm 1

  /* Initialization */
  P = {S};
  C_j = D_Sj  ∀ j∈N
  While ( ( P ⋂ N) == ϕ )
     Find a node i such that C_i=min C_j  ∀ j∈P iff i ∈ N - P
     P = P ⋃ {i}.
     For each j ∈ P
           C_j = min [C_j, C_i + D_ij ]
     End For
  End While
```

3.2 Structure-Free Data Aggregation

Structured aggregation techniques has many drawbacks. It causes high overhead in maintaining the constructed aggregation tree. The aggregation may vary because of the waiting period at the intermediate nodes. Structure free techniques [10] were developed in order to avoid these overheads. In structured aggregation tree, packet transmission and aggregation are done simultaneously and it requires a waiting period. In structure free aggregation, packet location and waiting period are unavailable. Thus spatial convergence or temporal convergence needs to be improved.

The Data-aware anycast protocol (DAA) is used to improve the spatial convergence. A set is created among the sources in DAA for optimal aggregation. Nodes in the independent set represent the aggregation points. Since the packets are forwarded automatically to the sink, maintenance overhead can be reduced. In DAA, any-casting is done in order to find the next-hop. Aggregation ID (AID) represents the associated packets that need to be aggregated. RTS contains the AID and when neighboring packets has similar AID, CTS will be generated. The two packets generated at the same time are likely to be aggregated. The receivers delay the CTS transmissions, in order to avoid CTS collisions. Interference range is high between the neighbors of the sender. The CTS collisions can be prevented by canceling the CTS transmission during packet overhearing. The packets transmitted do not meet at the intermediary node temporarily in the structure free aggregation technique. Thus artificial delays were introduced using the randomized waiting (RW) which increases temporal convergence. Delay is based on the proximity to the sink. Nodes closer to the sink have greater delay. RW sets the maximum delay τ based on the size of the network. Each node in the network sets its transmission delay between 0 and τ.

3.3 Data Aggregation Tree Construction

This proposed dynamic aggregation tree construction algorithm considers the steiner tree problems [11]. Let G (V, E) be a graph with V number of sensor nodes as vertices. Edge (u, v) ϵ E represents the link between the two nodes u and v. The objective is to find the minimum weighted steiner tree for the set of sources S ϵ V. This approach constructs the energy efficient aggregation tree so the edge cost is defined based on the energy requirement. Edge cost of the link between the node u and v is,

$$E_{\cos t}(u,v) = E_r(b,d) + E_t(b,d) + \mu \tag{2}$$

Where E_r is the energy to receive the b bits of data, E_t is the energy to transmit the data, d is the distance between node u and v and μ is the aggregation energy. Algorithm 2

Let us consider a set of source nodes S = {s1, s2, …. sn} in the network and assume that the sink node has the knowledge about sensor network topology. Let T be the aggregation tree. Initially the set T is empty.

```
Node s_i is chosen as the Source node at random.
Find another node s_j in S
If s_j is closer to s_i and E_cost (i, j) = minimum, then
          Choose s_j
End if
s_i is connected to s_j
Build the tree T = T ⋃ (i, j)
For each {s_k in S, k ≠ i}
          Choose next node s_k closer to y, where y ϵ T
          Connect s_k and y
          T = T ⋃ (k, y)
End For
```

Constructed aggregation tree is used to perform the full aggregation when the node sensed the regular event.

3.4 Adaptive Delay-Aware Algorithm for Data Aggregation

The typical problem with data aggregation is the delay. At aggregator node, various causes increase the delay in data aggregation. Waiting time to collect the information from the set of nodes and the time to perform the data aggregation function. These are the major factors for the higher delay in data aggregation. Some critical applications like fire alarms, theft alarms, health status monitoring and nuclear monitoring systems, need to report the information on time. For such applications it adaptively performs the data aggregation in order to reduce the overall delay. The sensor node marks the type of data based on the event detected and transmits to its parent node.

The parent node receives the data and checks the type of data and adaptively invokes the appropriate scheme.

Algorithm 3 explains the type of aggregation to be used according to the data packet type. If the monitored event is a critical event then the node marks its packet type as critical packet. For the regular event it sets the packet type as non critical packet. Depending upon the applications of the data, the packet type is differentiated into critical and non-critical. Initially the generation rates of the sensor node are detected for different time intervals. The difference in the data generation rate is calculated later. The non-critical data is further classified into fractionally aggregate and fully aggregate. Due to bursty packet reception, few data packets require only fractional aggregation. When the data packets have regular intervals, they require full aggregation.

A threshold D_{Th} is considered and when the difference exceeds the threshold level, this requires construction and frequent maintenance. So it causes high overhead in dynamic scenarios. The sensor node transmits the data packet along with the data type. The data type is taken as fractionally aggregated and uses structure free aggregation so that overhead and waiting period can be reduced. When the difference does not exceed the threshold level, the data packet type is taken for full aggregation and the data aggregation tree is used for aggregation process. When a critical data is detected, it is directly transmitted to the sink using the minimum delay shortest path.

Let us denote '*C*' for the critical data, 'F_rA' for the fractional aggregation data, and 'F_uA' for the fully aggregation data. Let D_{Th} denotes the threshold for the difference in the data generation rates. Let α and β denote the data generation rates of the sensor at time τ_1 and τ_2.

```
Algorithm 3

If(N_i = Source_Node)
 If(Sensed_Event = critical) then
  Packet_Type = 'C'
 Else
   /* α & β are data generation rates at time τ_1 and τ_2 */
   diff_αβ  = abs(β - α)
   If (diff_αβ > D_Th) then
     Packet_Type = 'F_rA'
   Else
     Packet_Type = 'F_uA'
   End if
End if
End if
/*Transmission of the data packet */
If (Packet_Type = 'C') then
  /*Data is transmitted directly to the sink */
  Use minimum delay shortest path (described in Algo.1)
Else if ( Packet_Type = 'F_rA') then
```

```
  Use the Structure free data aggregation
Else if ( Packet_Type = 'F_uA'), then
  Use the data aggregation tree (described in Algo.2)
End if
```

The proposed aggregation technique reduces the transmission delay of critical data, by selecting the delay-aware shortest path. For fractionally aggregated data, structure free aggregation is used. Since all the non critical data are aggregated, the redundant data gets eliminated completely and this avoids the heavy traffic in the network. Since the aggregation tree is built based upon the energy cost, it enhances the lifetime of the network. Since data aggregation is adaptive, the performance is not degraded when the network state or topology is changed. Thus the proposed data aggregation technique reduces overhead, minimizes delay and improves the lifetime of the network effectively.

4 Simulation Results

The performance of Adaptive Delay and Energy Aware Data Aggregation **(ADEADA)** protocol is evaluated through NS2 simulation. A random network deployed in an area of 500 X 500 m is considered. The simulated traffic is CBR for critical and full aggregation and exponential traffic for fractional aggregation. There are 5 traffic sources in which one is of critical traffic, 2 fractional aggregation traffic and 2 full aggregation traffic types. The performance of ADEADA is compared with the Distributed Improved Aggregation Scheduling (DIAS) [8] protocol.

When the data sending rate is increased, more packets are generated and so the aggregation latency is also increased resulting in the increase of end-to-end delay of the flows. From Fig. 1 (a) it is identified that the delay is linearly increased when the rate is increased from 50kb to 250kb. Since ADEADA adaptively chooses the aggregation mode depending on the type, the delay is less when compared to DIAS.

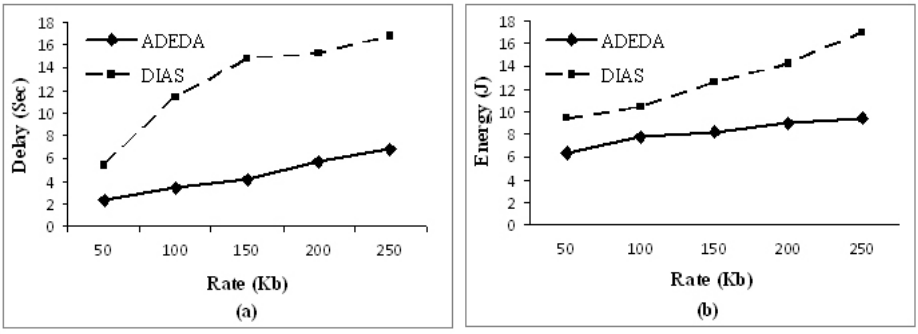


Fig. 1. (a) Rate Vs Delay, **(b)** Rate Vs Energy

Since more packets are generated for aggregation, the energy consumption increases when the data sending rate is increased. But the energy cost of building the aggregation tree is reduced by the use of fractional aggregation tree in ADEADA. From Fig. 1 (b), it is identified that the average energy consumption of ADEADA is less when compared to DIAS.

The next experiment varies the number of nodes as 100, 150, 200 and 250 with rate as 250Kbps. When the number of nodes is increased, the aggregation tree length and routing path length are increased resulting in the increase of end-to-end delay of the flows. From Fig. 2 (a), it is identified that the delay is linearly increased when the nodes are increased from 100 to 250. Since ADEADA adaptively chooses the aggregation mode depending on the type, the delay is less when compared to DIAS.

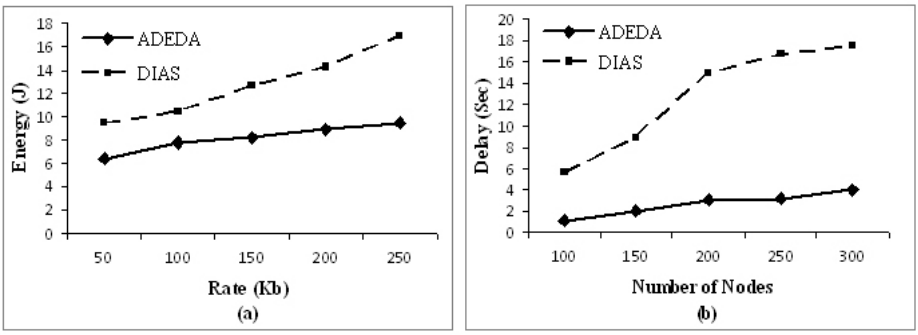


Fig. 2. (a) Nodes Vs Delay, (b) Nodes Vs Energy

Because of the increased aggregation tree depth and routing path length, the energy consumption increases when the number of nodes is increased. But the energy cost of building the aggregation tree is reduced by the use of fractional aggregation tree in ADEADA. Fig. 2 (b) shows that the average energy consumption of ADEADA is less when compared to DIAS.

5 Conclusion

This paper proposes an adaptive data aggregation technique using minimum delay with shortest path strategy. Each sensor marks the type of data: critical data, aggregated data and fractionally aggregated data. The estimation of shortest path with minimum delay is provided. When a data type happens to be critical, the data is transmitted directly to the sink using the shortest path. A structure free data aggregation technique is proposed and when a fractional aggregation data type is received, this technique is used for aggregation. Data aggregation tree is constructed by designing a cost model based on per-hop energy consumption. Depending upon the data type the aggregation process is selected. Delay can be reduced, since the proposed technique is adaptive and involves energy and overhead in the transmission. From the simulation results it is shown that the proposed technique reduces the end-to-end delay, energy consumption and overhead when compared with the existing techniques.

References

1. Nithyanandan, L., Sivarajesh, G., Dananjayan, P.: Modified GPSR Protocol for Wireless Sensor Networks. Int. J. of Computer and Electrical Engineering 2 (2010)
2. Huang, Q., Liu, X., Guo, C.: Reliable Aggregation Routing for Wireless Sensor Networks based on Game Theory. InTech (2010)
3. Maraiya, K., Kant, K., Gupta, N.: Wireless Sensor Network: A Review on Data Aggregation. Int. J. of Scientific & Engineering Research 2 (2011)
4. Lindsey, S., Raghavendra, C., Sivalingam, K.M.: Data Gathering in Sensor Networks using the Energy Metric. In: International Symposium on Parallel and Distributed Processing (2001)
5. Macedo, M., Grilo, A., Nunes, M.: Distributed Latency-Energy Minimization and interference avoidance in TDMA Wireless Sensor Networks. Computer Networks, 569–582 (2009)
6. Gargano, L., Rescigno, A.A.: Collision-free path coloring with application to minimum-delay gathering in sensor Networks. Discrete Applied Mathematics 157(8), 1858–1872 (2009)
7. Jie, Z., Shu-jie, G., Yu-gui, Q.U., Bao-hua, Z.: Energy equalizing routing for fast data gathering in Wireless Sensor Networks. J. of China Universities of Posts and Telecommunications 14(4) (2007)
8. Xu, X., Wang, X., Mao, X., Li, X., Tang, S.: A Delay Efficient Algorithm for Data Aggregation in Multi-hop Wireless Sensor Networks. IEEE Transactions on Parallel and Distributed Systems 22, 163–175 (2010)
9. Thomas, H., Charles, E., Ronald, L., Stein, C.: Introduction to Algorithms, 2nd edn. The MIT Press, Cambridge
10. Fan, K., Liu, S., Sinha, P.: Structure-Free Data Aggregation in Sensor Networks. IEEE Transactions on Mobile Computing 6(8) (2007)
11. Du, D., Hu, X.: Steiner Tree problem in computer communications networks. World Scientific Publishing Co. Pte. Ltd., Singapore

Investigation on the Influence of Packet Sizes on WRR Queue Management in Ethernet Switch

Soumen Sarkar[1], Nikolay Kakanakov[2], and Shyam Sundar Prasad[1]

[1] Dept. of ECE, National Institute of Technology, Jamshedpur, India
[2] Dept. of Comp. System and Tech., Tech University, Plovdiv, Bulgaria

Abstract. The paper investigates the influence of the packet sizes of different traffic flows on the queuing delay in Ethernet switch. Investigation is based on the queue management algorithms most often used in Ethernet switches - weighted round robin and expedite queuing. As a base model the Cisco Catalyst 2950 queue discipline 1p3q is used. It has one expedite queue and three queues serviced using weighted round robin discipline. The experimental scenario is based on the common controller network traffic model - mixture of periodic CBR data flows and randomly occurred configuration and bulk traffic flows. Investigation is made using Network Calculus and Network Simulator.

Keywords: Queue management, Packet delay, controller network.

1 Introduction

In the common case Ethernet do not have real-time characteristics. Layer 2 prioritization in switches gives an opportunity to give to some guaranties against flow delay and priority service. In that case strict priority algorithms cannot provide fair service to other low priority flows. Weighted round robin (WRR) algorithms provide fair service based on predefined weights associated to each flow. As long as the packets are with variable size in every flow and there are differences in the packet sizes between different traffic types, the modeling of queue management is hard to do. In this paper the influence of the packet sizes on the delay in WRR queue is investigated using analytical and simulation models.

2 Backgrounds and Related Work

Investigating the delay for traversing node with SP or WRR queues is not a new task. However, investigations on combined mechanisms are missing. Zhang and Harrison [11] present a deep investigation on such combined strict priority and weighted round robin (called priority-weighted round robin) but their results are statistical due to the mathematical model they use – queuing theory. Queuing theory is often intractable for deterministic models. Classical queuing theory usually deals with average performance of aggregate flows, while for guaranteed performance service bounds need to be derived for worst-case scenario on per-flow basis [7].

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 50–55, 2013.

Georges et al. [10] propose a comparison of the delays of weighted fair queuing and strict priority. They provide a formulation of the service curves for WRR nodes and prove that the delay for each flow through it depends mostly on the assigned weights. Diouri et al. [1] present an investigation of WRR (and WFQ) using network calculus based on the results in [10]. The work provides some mechanisms for determination of the best values for the weights using an iterative method of substitution of weights to achieve the desired delay bounds. These results are not reflecting the physical implementation of WRR schedulers.

S. Floyd and V. Jacobson in their work on Class-based Queuing [11] present some experimental results on the work of packet-by-packet and bit-by-bit WRR schedulers. They work provide a source of data for analysis of the influence of the packet size on round robin scheduler but only in NS CBQ implementation.

2.1 Network Calculus

Network Calculus is a collection of results based on Min-Plus algebra, which applies to deterministic queuing systems found in communication networks. Main input parameters of the Network Calculus theory are: per-flow arrival curves, service curves of network elements, and the route selecting mechanism.

The formula for representing the service curve depends on the queue management policy. Knowing the service curve for the switch and arrival curves for the flows, the worst case delay for queuing can be obtained using the formula [2]:

$$D_i = S_i + \frac{M_i}{R_i} + \left(\frac{b_i - M_i}{p_i - r_i} \right) \times \frac{(p_i - R)^+}{R_i} \tag{1}$$

Expression above $()^+$ means true when it is positive and is equal to zero otherwise. Parameters *M, b, p, r* are TSPEC parameters of the flow and *R, S* are RSPEC parameters of the switch.

2.2 Network Simulator

The network simulator NS2 is a Linux based open source software tool for evaluation of network protocols and topologies. While it represents a discrete, event based simulator, can easily be extended and modified to include additional network elements during simulation of local controller networks.

In NS2 there is no definition of Ethernet switch and queue management is associated with links, not with nodes. To investigate the delays of 1P3Q queue management in switch the assumption shown in [4] is used. For current simulation, we have implemented Class of Service (CoS) by mapping multiple flows through the use of CBQ/WRR by applying different bandwidth and delays (IEEE 802.1p, layer 2&3 respectively).

3 Experimental Scenario

The experimental scenario is built up to cover common automation model in local controller networks with different types of application traffic - monitoring and control, configuration, event driven, and background. In the current paper it is assumed a case with four traffic flows traversing a single Ethernet switch. For the switch it is assumed the Cisco Catalyst 295x Series. It has 1P3Q queue management – one expedite and three WRR queues. The four flows are simulated with four different workload types with different distribution, each representing a common class of traffic type in industrial network: WT1 - Monitoring, Control and diagnostics – TCP/UDP – 75-95%; WT2 - Event driven exception message, alarm (Poisson distribution) – 5-15%; WT3 - Configuration scenario – exponential/CBR - <5%; WT4 - Big messages, files upload etc – <1%.

For reducing the complexity of the experiments, the paths of the packets of each flow are predefined. WT1 packets are generated from Source1 and are targeted to Sink1, all other traffic are generated from Source2 and WT2 packets goes to Sink1 and WT3 and WT4 packets goes to Null - figure 1.

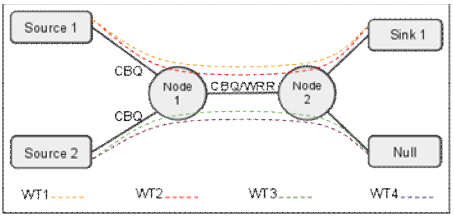


Fig. 1. Experimental Scenario

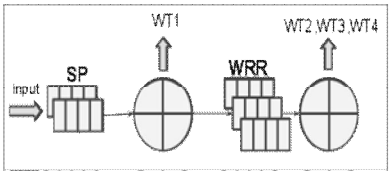


Fig. 2. Two step analytical model

WT1 is served with Strict Priority; WT2, WT3 and WT4 are served with Weighted Round robin with weights, respectively 75%, 25%, 5%.

For the analytical model the switch service is represented with R-SPEC curve in general. It has two main parameters: forwarding rate - R, and latency (slack term) - S. The first traffic flow is assumed as periodic and all others as T-SPEC. Periodic traffic has two parameters: the period - T and the packet size - L. T-SPEC traffics has for parameters: maximum packet size - M, peak rate - p, long term average rate - r, and burst size - b. The actual values used for all these parameters are shown in Table 1.

Table 1. Switch and Flow parameters

SW		WT1		TSPEC	WT2	WT3	WT4
R(bps)	$1.0X10^{9}$	L(bits)	576	b_i (bits)	4096	16384	48704
S(s)	$2.5X10^{-5}$	T(s)	10^{-4}	r_i (bps)	$4.0x10^{4}$	$1.0x10^{5}$	$1.0x10^{6}$
				M_i (bits)	2048	8192	12176
				p_i (bps)	$1.0x10^{8}$	$1.0x10^{8}$	$1.0x10^{8}$

For the case of this investigation, the 1P3Q policy is represented with the two step assumption model shown on figure 2. This assumption provides an easy way to calculate delays based on standard, simpler service curves.

Based on this two-step model, known network calculus can be used for calculating the delay in the switch with respect to the parameters from Table 1. The flow form WT1 is served with higher priority on the first step and it is serviced with the switch forwarding rate but it can wait for low priority packet already entered the queue (in the worst case the maximum size packet from all other flows) - equation (2) [1, 3]:

$$R_1 = R; S_1 = \frac{\max(M_i : j < i)}{R} \tag{2}$$

All other flows are served on the first step with the remaining rate and the latency for them depends on the burst size of WT1 - equation (3).

$$R_{2,3,4} = R^1 = R - r_1; S_{2,3,4} = S^1 = \frac{b_1}{R - r_1} \tag{3}$$

On the second step WT2, 3 and 4 are served in regular WRR policy but the maximum forwarding rate and slack term are substituted from equation (3) [1, 3]:

$$R_i = \left(R - \frac{L}{T}\right) \cdot \frac{\varphi_i - M_i}{\sum_j \varphi_j - M_i}; S_i = \frac{L + \sum_{j \neq i} \varphi_j}{R - \frac{L}{T}} \tag{4}$$

In equation (4) there is a limitation for the weights values. For each flow they must be bigger than the maximum packet size for that flow. Otherwise, the value for the rate can become negative or zero. In WRR theory the weights are cumulative. Nevertheless, in the Network Calculus the weights are static.

For the simulation model main parameters mapped on NS2: Switch to Transaction Server link set as CBQ/WRR, time allowed for other traffics to run without CBR 50.0 seconds. UTP Signal propagation speed taken as 59% of the 'Speed of Light' that is 177000000 m /s. Maximum distance between 2 nodes used in simulation is 100 m. Packet size for WT1 is kept as 72 bytes while for WT2 to WT4 is varying. Simulation carried out for each size of packet is 100 seconds.

4 Results and Discussions

The data is presented here based on the calculation of delay bounds in WRR. It is a question of assumption to obtain an exact value in Bits for Φ in Equation (2&3). The values for the packet sizes for WRR (WT2...WT4) are obtained by dividing the total bits (Φ) in the fixed ratio of 75:25:5.

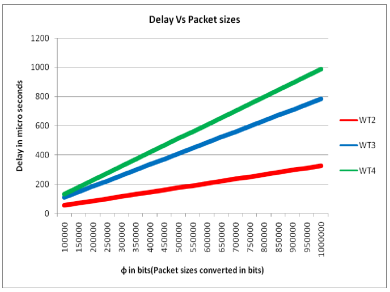
Φ [bits]	WT2 [bytes]	WT3 [bytes]	WT4 [bytes]
100000	8750	3125	625
150000	13125	4688	938
--	--	--	--
1000000	87500	31250	6250

The analytical model based on Network Calculus works with Bit flow, but the simulation model which is based on Network Switches works with packets in bytes. The Graph (1&2) placed under the Analytical and Simulation Results show the delay bounds obtained against the variable packet sizes for the different flows by plotting against the assumed values of Φ for the Analytical and Simulation model respectively. The packet sizes were distributed in this manner.

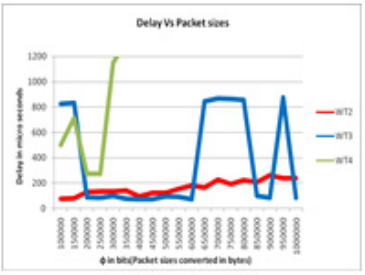
Analytical Results. The Analytical Results presented on Graph 1 show relation between packet sizes of different workload types and their delay in WRR queue. They are based on the two-step model (figure 2) and present the delay in the second step. The relation between the size of the packet and the delay, while keeping the ratio of 75/20/5, is almost linear for values of φ above 50000 and depends mostly on the slack term of the provided service curve. The forwarding rate provided to each flow is changing very little with the packet size. For values of φ below 50000 the relation between flows delay and packet size show instability and goes outside the possible values (e.g. negative delay, no delay). It is because of the limitation of Network Calculus equation explained in section 3.

Simulation Results. Simulation Results presented on Graph 2, obtained for the different flows show that WT2 which has a weight of 75% follows more or less same and uniform pattern that of the analytical model. While WT3, which has a weight of 25% and less number of packets has a non uniform pattern but average delay is similar to the analytical model. Delay pattern in case of WT4 (5%) is even more non uniform.

There are values for the weights that are not eligible for calculation of the delay. When the weight of the given flow is below or equal to the packet size of this flow, we cannot calculate the delay - it will become negative. Therefore, there is no data in some parts of the WT4 flow.



Graph 1. Analytical results



Graph 2. Simulation results

5 Conclusion

The analytical investigation show that packet size of the different traffic type influence the delay of flows traversing a WRR queue in linear manner in common case. It also shows that the actual values for the weights are the main parameter for the delay calculation and must be carefully chosen in modeling depending on the flow

parameters. Moreover, the main conclusion about the weights of the flows is: their values for the flow with smallest share must be bigger or equal to its bucket size and the weights for other flows can be calculated using the ratios. The simulation results show that there is some functional connection between packet sizes and delay for WRR, but the results depend very much on the implementation. In NS2, Round-Robin queue management is implemented in two ways: WRR and PRR. The later (PRR) works packet by packet and a packet gets into the forwarding path it cannot be preempted. But with WRR packets are transmitted bit-by-bit form each queue, depending on the weights. The actual work of WRR in the simulator environment is by estimating an average packet size for each flow and guaranteeing a share of the total bandwidth relative to the its weight. The long-term behavior is WRR but in small period it is unpredictable due to the random distribution of traffics.

The obtained results may further be validated through a test bed using actual switch CISCO 2950 and local controller network. Thus delays obtained by the corresponding packet size may be checked to find the best weights according to the packet sizes that still provide FAIR service to all flows.

References

1. Diouri, I., Georges, J.-P., Rondeau, E.: Accommodation of delays for networked control systems using classification of service. In: 2007 IEEE International Conference on Networking, Sensing and Control, pp. 410–415 (2007)
2. Le Boudec, J.-Y., Thiran, P.: A short tutorial on network calculus. I. Fundamental bounds in communication networks. In: Proceedings of the 2000 IEEE International Symposium on Circuits and Systems, ISCAS 2000, Geneva, vol. 4, pp. 93–96 (2000)
3. Kakanakov, N., Shopov, M., Spasov, G.: Time-Delay Simulation Analysis of Local Controller Networks. In: Proc. Intern. Conference CompSysTech, p. IIIA.12 (2008)
4. Shopov, M., Kakanakov, N., Spasov, G.: On the use of NS-2 is Simulations of Internet-based Distributed Embedded Systems. In: Proc. Industrial Electronics (INDEL 2008), Banja Luka, Bosnia, Herzegovina (2008)
5. Sarkar, S., Prasad, S.S.: Packet Delay analysis in operational network. In: Proc. ICCCNT 2010. IEEE (2010) 978-1-4244-6589-7/10
6. Azimi-Sadjadi, B.: Stability of networked control systems in the presence of packet losses. In: Proc. Conf. Decision Control, Maui, HI, pp. 676–681 (2003)
7. Zhang, H.: Providing end-to-end performance guarantees using non-work-conserving disciplines. Computer Communications, System Support for Multimedia Computing 18(10), 769–781 (1995)
8. Vatanski, N., Georges, J.-P., Aubrun, C., Rondeau, E., Jounela, S.-L.: Control compensation based on upper bound delay in networked control systems. In: 17th International Symposium on Mathematical Theory of Networks and Systems (MTNS), Kyoto, Japan (2006)
9. Zhang, Y., Harrison, P.G.: Performance of a Priority-Weighted Round Robin Mechanism for Differentiated Service Networks. In: Proc. of Computer Communications and Networks, ICCCN 2007, pp. 1198–1203 (2007)
10. Georges, J.-P., Divoux, T., Rondeau, E.: Strict Priority versus Weighted Fair Queuing in switched Ethernet networks for time-critical applications. In: Proc. Workshop on Parallel and Distributed Real-Time Systems (WPDRTS 2005 – IPDPS 2005), Denver, Colorado (2005)
11. Floyd, S., Jacobson, V.: NS-1 Simulator Tests for CBQ. IEEE/ACM Trans. on Networking 3(4), 366–386 (1995)

Component Based Software Development in Distributed Systems

Tiwari Umesh Kumar, Nautiyal Lata, Dimri Sushil Chandra, and Pal Ashish

Graphic Era University, Dehradun, India
{umeshtiwari22,lata.nautiyal1903,dimri.sushil2,
asheesh.pal}@gmail.com

Abstract. In the context of Software development, the idea of reusability has been used since the earliest days of computing. Component Based Software Engineering (CBSE) is an engineering methodology that aims to design and construct software systems using reusable software components. *Distributed component approach* is embraced in industry to reap the desired benefits, often looked for by a software development organization. Distributed applications require components of same or different applications on different machines to interact, especially when client and component reside on different machines. In this paper we propose that there is a need to develop a common Service Calculation module with the help of distributed component technology and component-based methodology. The high productivity is achieved by using standard components.

Keywords: Interface, Distributed Components, Distributed Components Architecture.

1 Introduction

Component-based software engineering (CBSE) intends to build large software systems by integrating pre-built, tested and functional software components. Clements describes CBSE as embodying *"the 'buy, don't build' philosophy". He also says that "in the same way that early subroutines liberated the programmer from thinking about details,[CBSE] shifts the emphasis from programming to composing software systems"*. The principles of CBSE can be described by the two guiding principles: reuse but do not reinvent (the wheel); and, assemble pre-built components rather than coding line by line.

1.1 Component

A component is any part of which something is made.

In our context this 'something' means system. In software engineering, this would allow a software system to have as components assembly language instructions, sub-routines, procedures, tasks, modules, objects, classes, software packages,

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 56–61, 2013.

processes, sub-systems, etc. The widely accepted goal of component-based development is to build and maintain software systems by using existing software components, e.g. [3, 8, 9, 4]. They must interact with each other in system architecture. Interaction between these independent components, to achieve their goal is based on some properties:

1. Specified *services*,
2. Fully explicit context *dependencies*,
3. Interactive *interfaces*,
4. Independent *deployment,*
5. Communication *protocols.*

1.2 Interface

It defines how a component can interact to other components to provide its services. It specifies what the parameters are and what results to expect. It is the way through which coupling is done. It may be Data, Content, Accidental, Common, External, Control, Stamp, and Message. Interface is an essential requirement to integrate (couple) two or more independent components. It works as a protocol among the components when they are assembled to provide their intended goal. Components in distributed, mobile or internet-based systems require their interfaces to include information about their locations or addresses.

Components should be developed and engineered in such a way that component could be used and implemented in several, possibly unpredictable, deterministic or in deterministic contexts. Therefore there must be some general information which is required by a generic user of a distributed component:

Component evaluation: for example, information on static and dynamic metrics computed on the components, such as cyclomatic complexity, in-out interactions, and coverage level achieved during testing.

Component deployment: for example, additional information on the interface of the component, such as pre-conditions, post-conditions, and invariants.

Component testing: for example, a finite state machine representation of the component, regression tests suites together with coverage data, and information about dependences between inputs and outputs.

Component analysis: for example, summary data-flow information, control-flow graph representations of part of the component, and control-dependence information.

2 Distributed Components

Component-based software development is based on the concept to develop complex software products by selecting appropriate off-the-shelf components and then to integrate them with a suitable, well-defined software architecture. In distributed

systems DCOM simply replaces the local inter-process communication with a network protocol. Neither the component nor its client is aware of the network protocols lying between them. Today component-based approach, introduces more benefits to this world in terms of reusability, flexibility, scalability and maintainability.

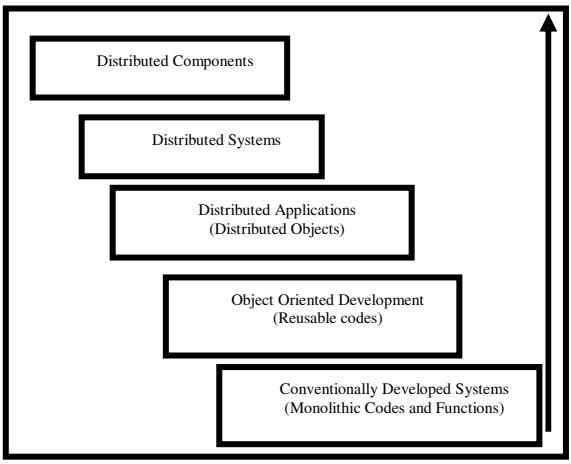


Fig. 1. Distributed Components

Large scale software systems need these features more than before. *Distributed component approach* is embraced in industry to reap the desired benefits, often looked for by a software development organization. This uses current build-time and run-time technologies to attempt to reduce cost and development time for distributed systems. It becomes apparent that something is needed that addresses both the challenge of distributed systems interoperability and the challenge of how to build individual systems that can be treated as atomic units and can easily be made to cooperate with each other [12].

3 Components in Distributed Systems

Component-based software development is understood to require reusable components that interact with each other and get into system architectures; there is so far no agreement on standard technologies for designing and creating components. Finding appropriate formal approaches for describing components, the architectures for composing them, and the methods for component-based software construction, is correspondingly challenging. Components can be checked out from a component repository, and assembled into a target distributed software system [10].

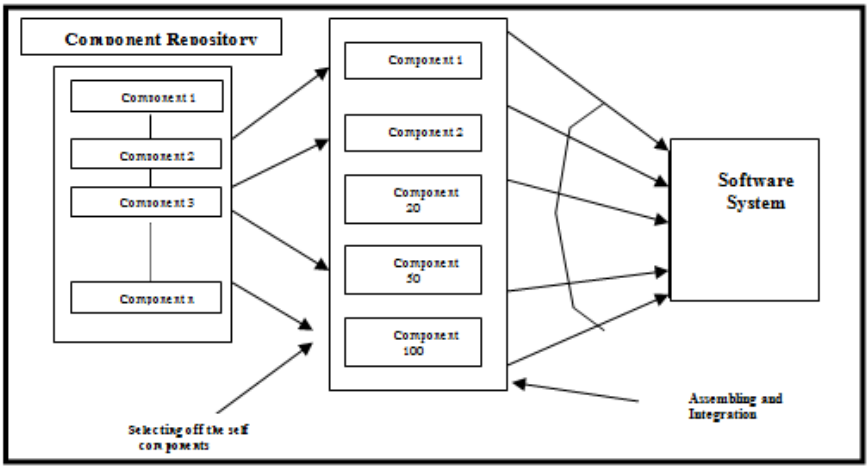


Fig. 2. Components in Distributed System

4 Distributed Component Architecture

To ensure that a component-based software system can run properly and effectively, the system architecture is the most important factor. The system architecture of distributed component-based software systems should be a layered and modular architecture.

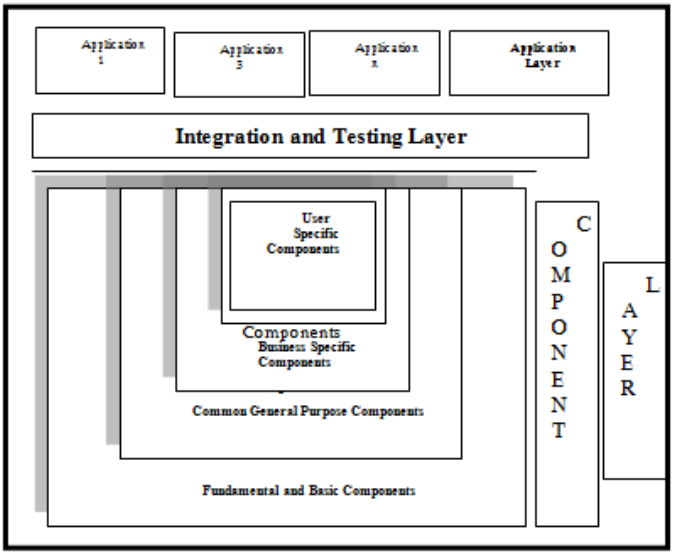


Fig. 3. Distributed Component Architecture

The questions of how to identify model and specify components, how to follow a component-based development process in a systematic and consistent manner, and how to assembly formally specified components into the component-based system architecture may be given as:

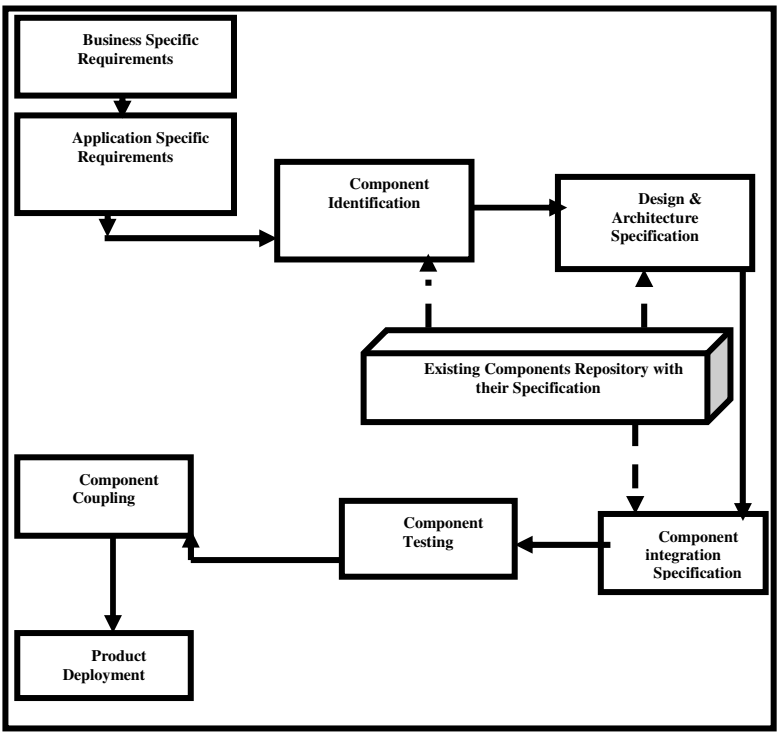


Fig. 4. Detailed Distributed Component Architecture

4.1 Component Certification

The objectives of component certification are to outsource, select and test the candidate components and check whether they satisfy the system requirement with high quality and reliability. The governing policies are: 1) Component outsourcing should be charged by a software contract manager; 2) All candidate components should be tested to be free from all known defects; and 3) Testing should be in the target environment or a simulated environment. The component certification process overview diagram is as shown in the following figure. The input to this phase should be component development document, and the output should be testing documentation for system maintenance.

5 Conclusion

There are two basic activities in component-based software development: First, develop components for reuse. The production process model for this activity involves component specification, design, coding, testing, and maintenance. Second, develop software with components. Components in a distributed application need to be notified if a client is not active anymore, even or especially in the case of a network or hardware failure. Component certification is the process that involves:

1) *Component outsourcing*: managing a component outsourcing contract and auditing the contractor performance;

2) *Component selection*: selecting the right components in accordance to the requirement for both functionality and reliability; and

3) *Component testing*: confirm the component satisfies the requirement with acceptable quality and reliability.

References

1. Vidger, M.: The Evolution, Maintenance and Management of Component Based Systems. Addison-Wesley, Boston (2001)
2. Narasimhan, V.L., Hendradjaya, B.: Some theoretical considerations for a suite of metrics for the integration of software components. Information Sciences: An International Journal 177(3), 844–864 (2007)
3. Szyperski, C.: Component Software: Beyond Object-Oriented Programming. Addison-Wesley (1997)
4. Hybertson, D.: A uniform component modeling space. Informatica 25, 475–482 (2001)
5. Pressman, R.S.: Software engineering: A practitioner's approach. McGraw Hill (2001)
6. Pfleeger, L.S., Fenton, N.E.: Software metrics: A rigorous and practical approach. Brooks/Cole Publications (1998)
7. Sommerville, I.: Software Engineering, 6th edn. Addison-Wesley (2001)
8. Schneider, J.-G., Nierstrasz, O.: Components, scripts and glue. In: Barroca, L., Hall, J., Hall, P. (eds.) Software Architectures Advances and Applications, pp. 13–25. Springer (1999)
9. Medvidovic, N., Taylor, R.N.: A classification and comparison framework for software architecture description languages. IEEE Transactions on Software Engineering 26(1), 70–93 (2000)
10. Cai, X., Lyu, M.R., Wong, K.-F., Ko, R.: Component-Based Software Engineering: Technologies, Development Frameworks, and Quality Assurance Schemes
11. Pooley, R., Steven, P.: Using UML: Software Engineering with Objects and Component. Addison-Wesley (1999)
12. ISO/IEC, Information Technology-Software Quality characteristics and metrics – Part 1: Quality characteristics and sub characteristics, p. 21 (1996); Weyuker: Testing Component Based Software: A cautionary Tale. IEEE Software 15(2) (2008)
13. IEEE, IEEE standard glossary of software engineering terminology, in Std 610.12-1990: IEEE, vol. 84 (1990)

Design of Open Equation Archive Server Resistant against Denial-of-Service Attacks

Shin-ya Nishizaki and Hiroshi Tamano

Department of Computer Science, Tokyo Institute of Technology
2-12-1-W8-69, Ookayama, Meguro-ku, Tokyo 152-8552, Japan
nisizaki@cs.titech.ac.jp

Abstract. We studied resistance against denial-of-service attacks in a web-based equation-archive server. Each set of equations to be archived should be verified not only by its contributor (on the client side) but also by the archive server (on the server side) to preserve the integrity of the archive. Checking an equation set entails a certain amount of computational cost, and if a malicious user submits an equation set whose verification is computationally expensive, excessive cost is placed upon the archive server, resulting in a denial-of-service attack. In this paper, we propose a new technique for improving a server's resistance against such attacks, known as the estimate-attaching method. This technique will also be applied to other kinds of web applications.

Keywords: Web application, theorem proving, denial-of-service attacks.

1 Introduction

We begin with an explanation of some of the background material used in this paper.

1.1 Equational Logic and Term-Rewriting Systems

Equational logic [1] is a formal system whose assertions are equations between first-order terms consisting of individual variables, constant symbols, and function symbols. A set of equations is called an *equational system*, and a basic query in equational logic is formulated as a *word problem* (i.e., a decision problem as to whether s = t is valid for two given terms s and t, assuming the equations in a given equational system *E*). For example, we can formulate a group as an equational system of the form

$$E_{Group} = \{(x * y) * z = x * (y * z), i(x) * x = 1, 1 * x = x\}$$

where the binary operator * denotes multiplication, the unary operator $i(-)$ denotes the inverse, and the constant 1 denotes identity.

A term-rewriting system (or TRS) [1] is a set of rewriting rules s → t such that if a part of a given term is matched with s, then the part is replaced by t. We may say that a rewriting rule is an oriented equation. If we know whether s = t is valid under an equational system E, we can know it by rewriting both s and t via the term-rewriting system corresponding to E. For example, if we represent the natural numbers 0,1,2,…

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 62–69, 2013.

as 0, S(0), S(S(0)), … in the style of Peano, we can formulate addition as an equational system

$$\{\mathrm{add}(\mathrm{x},0)=\mathrm{x},\mathrm{add}\big(\mathrm{x},\mathrm{S}(\mathrm{y})\big)=\mathrm{S}(\mathrm{add}(\mathrm{x},\mathrm{y}))\}$$

and a term-rewriting system can be obtained by imposing an appropriate direction on each equation, such as

$$\{\mathrm{add}(\mathrm{x},0)\rightarrow\mathrm{x},\mathrm{add}\big(\mathrm{x},\mathrm{S}(\mathrm{y})\big)\rightarrow\mathrm{S}(\mathrm{add}(\mathrm{x},\mathrm{y}))\}.$$

We can then determine an equation

$$\mathrm{add}\big(\mathrm{S}(0),\mathrm{S}(0)\big)=add(0,S\big(S(0)\big))$$

by rewriting both sides

$$add\big(S(0),S(0)\big)\rightarrow S(add(S(0),0)\rightarrow S\big(S(0)\big),$$
$$add\left(0,S\big(S(0)\big)\right)\rightarrow S\left(add\big(0,S(0)\big)\right)\rightarrow S\left(S\big(add(0,0)\big)\right)\rightarrow S\big(S(0)\big).$$

Consequently, we know the equation is valid. However, term rewriting is not generally a terminating procedure, and its results do not necessarily coincide with each other. When such coincidence occurs, it is called confluence or the Church–Rosser property. The combination of these two properties, confluence and termination, called completeness, is known as the criterion of usefulness for a TRS. The Knuth–Bendix completion algorithm [2] [1] constructs a term-rewriting system satisfying these two properties for a given equational system. Consider the term-rewriting system R_{Naive}, which is obtained by naively orientating E_{Group} :

$$\{(\mathrm{x}*\mathrm{y})*\mathrm{z}\rightarrow\mathrm{x}*(\mathrm{y}*\mathrm{z}),\mathrm{i}(\mathrm{x})*\mathrm{x}\rightarrow 1,1*\mathrm{x}\rightarrow\mathrm{x}\}.$$

Using R_{Naive}, i(y) * (x * (i(x) * y)) cannot be reduced any further. However, under the equational system E_{Group},

$$\mathrm{i}(\mathrm{y})*\big(\mathrm{x}*(\mathrm{i}(\mathrm{x})*\mathrm{y})\big)=\mathrm{i}(\mathrm{y})*\Big(\big(\mathrm{x}*\mathrm{i}(\mathrm{x})\big)*\mathrm{y}\Big)=\mathrm{i}(\mathrm{y})*(1*\mathrm{y})=\mathrm{i}(\mathrm{y})*\mathrm{y}=1.$$

Applying the Knuth–Bendix algorithm to E_{Group}, we obtain the complete term-rewriting system R_{Group}:

$$\{(\mathrm{x}*\mathrm{y})*\mathrm{z}\rightarrow\mathrm{x}*(\mathrm{y}*\mathrm{z}),1*\mathrm{x}\rightarrow\mathrm{x},\ \mathrm{x}*1\rightarrow\mathrm{x},\ \mathrm{i}(\mathrm{x})*\mathrm{x}\rightarrow 1,\mathrm{x}*\mathrm{i}(\mathrm{x})\rightarrow 1,$$
$$\mathrm{i}(\mathrm{i}(\mathrm{x}))\rightarrow\mathrm{x},\mathrm{i}(\mathrm{x}*\mathrm{y})\rightarrow\mathrm{i}(\mathrm{y})*\mathrm{i}(\mathrm{x}),\mathrm{x}*(\mathrm{i}(\mathrm{x})*\mathrm{y})\rightarrow\mathrm{y},\ \mathrm{i}(\mathrm{x})*(\mathrm{x}*\mathrm{y})\rightarrow\mathrm{y}\}$$

By using the term-rewriting system R_{Group}, $\mathrm{i}(\mathrm{y})*\big(\mathrm{x}*(\mathrm{i}(\mathrm{x})*\mathrm{y})\big)$is reduced to i(y)*y, and then to 1, which is coincident with the above result for E_{Group}.

We may say that an equational system defines equalities between expressions, and a term-rewriting system obtained via the Knuth–Bendix algorithm provides an algorithm for computing these equalities. Therefore, if we record an equational system in an archive, the term-rewriting system constructed for it via the Knuth–Bendix algorithm should be recorded simultaneously.

1.2 Important Features of an Equation Archive

In this paper, we study the design of *an equation archive*, a storage system for equational systems. The important features of an equation archive are *integrity* and *accessibility*.

- Integrity: every equational system accepted into the archive is guaranteed to be complete.
- Accessibility: it should be possible for equational systems to be submitted by as broad a range of users as possible without any authentication.

To ensure integrity, completeness of a submitted equational system should be verified on the server side. However, unassisted server-side verification is not realistic, because termination checking is an NP-complete problem, even when restricted to checking based on lexicographic path ordering [1].

To ensure accessibility, potential vulnerability to denial-of-service attacks must be carefully examined.

1.3 The Estimate-Attaching Method

As mentioned above, it is important to reduce vulnerability to denial-of-service attacks, and we may consider the following measures.

1. Restricting access to an equation-archive server by requiring authentication. Before initiating interaction with the equation archive, a user is asked to provide some authentication, and only registered users can access the server. In this way, the volume of requests can be controlled.
2. Imposing an upper bound on the computational cost of each request. If a fixed upper bound is imposed on the CPU time and amount of memory required for processing each request, the total load on the server will be controlled. However, if an attacker sends a large number of requests, each having a computational cost close to the upper bound, the server can still be overloaded.
3. Attaching the estimated computational cost of a request to the equation-archive server. Instead of imposing an upper bound, a client declares the computational cost of a request to the server. This method also controls the total load on the server and enables more appropriate resource allocation.

In our paper, we employ the third option, which we refer to as *the estimate-attaching method.*

2 A DoS Attack-Resistant Architecture for an Equation Archive

In this chapter, we describe the design of an equation archive using the estimate-attaching method.

A message submitted to the equation archive by a client includes the following two constructs:

- An equational system *E*,
- A term-rewriting system *R*.

The two sets *E* and *R* should satisfy the conditions of completeness (i.e., termination and confluence) and equivalence.

- *Termination*: there are no infinite rewriting sequences, such as $s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow \ldots$
- *Confluence*: if $s \rightarrow \ldots \rightarrow s_1$ and $s \rightarrow \ldots \rightarrow s_2$, then there is a term t such that $s_1 \rightarrow \ldots \rightarrow t$ and $s_2 \rightarrow \ldots \rightarrow t$.
- *Equivalence*: assuming E, s = t if and only if there exists a term v such that $s \rightarrow \ldots \rightarrow v$ and $t \rightarrow \ldots \rightarrow v$.

Verification of these three properties is undecidable. A Turing machine can be expressed as a term-rewriting system, and hence the termination decision problem of the system can be reduced to the termination decision problem of a Turing machine, which is undecidable. Therefore, these three properties should be verified with human assistance on the client side. Information collected during this verification process is sent to the equation-archive server as a cost estimate and its details, which are used as hints for server-side checking. The procedures to be carried out on the client side are:

- Input of the equational system E,
- Choosing the ordering used for termination checking,
- Constructing a term-rewriting system R via the Knuth–Bendix completion algorithm,
- Preparing a cost estimate and its details C,
- Sending E, R, and C to the equation-archive server.

In applying the Knuth–Bendix algorithm, the choice of ordering determines the class of termination. In the present study, we use lexicographic path ordering [1]. A user should set a priority among the function symbols appearing in E. The Knuth–Bendix completion algorithm is not guaranteed to succeed for every input equational system. If it fails, it may be made to succeed by changing the priority among the function symbols.

We next consider cost estimates and the detailed items associated with them.

2.1 Cost Estimate for Termination Checking

To check the termination of a term-rewriting system R, it is sufficient to show that each rewriting rule $s \rightarrow t$ in R satisfies $s > t$ with respect to lexicographic path ordering. Because checking that $s > t$ entails a time complexity of $O(|s| \cdot |t|)$, termination checking involves the same time complexity if a client provides lexicographic path ordering information as a "hint." Checking a proof document of termination entails a time complexity of $O((|s| + |t|)^2)$ because the number of lines of the proof document is proportional to $(|s| + |t|)$. Thus, if a client sends a skeleton of the proof document, the checking time complexity is reduced to $O(|s| + |t|)$. For example, the term-rewriting system R_{ack} in Fig. 1. Definition of the Ackermann function Fig. 1 describes the Ackermann function a(-,-).

$$\begin{cases} a(0,X) \to s(X) \\ a(s(X),0) \to a(X,s(0)) \\ a\big(s(X),s(Y)\big) \to a(X,a(s(X),Y)) \end{cases}$$

Fig. 1. Definition of the Ackermann function

A proof document (more precisely, a derivation tree) in Fig. 2 demonstrates termination of the first rewriting rule of R_{ack}.

$$\dfrac{\dfrac{s(X)|_1 = X}{\Gamma \vdash s(X) >_{lpo} X}\,(\mathsf{LPOa}) \quad \dfrac{\dfrac{a(s(X),0)|_2 = 0}{\Gamma \vdash a(s(X),0) >_{lpo} 0}\,(\mathsf{LPOa}) \quad a > s \in \Gamma}{\Gamma \vdash a(s(X),0) >_{lpo} s(0)}\,(\mathsf{LPOc})}{\Gamma \vdash a(s(X),0) >_{lpo} a(X,s(0))}\,(\mathsf{LPOd})$$

Fig. 2. Proof figure for termination of the first rewriting rule of R_{ack}

Due to lack of space, we do not explain the details of the derivation tree. Γ indicates an assumption of ordering among the function symbols; i.e., $a(\text{-},\text{-}) > s(\text{-})$.

To help the server check the derivation tree, the following skeleton is provided. The skeleton consists of information about rule names and locations in which the rules are applied. The skeleton is a "cost estimate" for termination checking because it enables the server to know the upper bound of the time required for termination checking.

$$\dfrac{(\mathsf{LPOa},1) \quad \dfrac{(\mathsf{LPOa},2)}{(\mathsf{LPOc},\_)}}{(\mathsf{LPOd},1)}$$

Fig. 3. Proof Skelton

2.2 Cost Estimation for Confluence

According to Newman's lemma [1], for a terminating term-rewriting system R, confluence is equivalent to local confluence, which is the property that if $s \to t'$ and s $\to$ t'', then there is a term u such that t' $\to \ldots \to$u and t'' $\to$ u. We can prove local confluence by showing that each *critical pair* (t', t'') is reduced to the same normal form. A critical pair (t', t'') consists of two distinct terms derived from another term by two distinct rewriting rules; i.e., s $\to$t' and s $\to$ t'', but t' $\neq$ t''. A "cost estimate of confluence" sent to the server consists of the critical pairs and the rewriting paths to the normal forms. Specifically, if we have two rewriting paths starting from a critical pair s and t,

$$s \rightarrow_{(R_1,p_1)} \rightarrow \cdots \rightarrow_{(R_l,p_l)} v \leftarrow_{(R_{l+1},p_{l+1})} \cdots \rightarrow_{(R_{l+m},p_{l+m})} t$$

we send the critical pair (s, t), the rewriting rules used in the above rewriting paths, and the locations in which the rewriting rules are applied, (R_i, p_i). The server check is conducted by using the information sent by the client, as follows.

- Obtaining the critical pairs from the given term-rewriting system R. This process entails a time complexity of $O(n \cdot |R|^2)$, where n is the number of rules in R, and $|R|$ the summation of sizes of the terms appearing in R.
- Matching the critical pairs obtained by the server with those sent by the client. The time complexity is O(*the number of critical pairs*).
- Verifying the convergence of the critical pairs. The server checks that each critical pair reaches the same normal form by tracing the paths (R_i, p_i) (i = 1, … , l + m) sent by the client. The time complexity of this process is $O(|(R_1, p_1)| + \cdots + |(R_n, p_n)|)$.

The total time complexity of these three procedures is proportional to the length of the message sent by the client, and thus the server knows the estimated cost of confluence from the length of the message.

2.3 Cost Estimate for Equivalence between Equational Systems and Term- Rewriting Systems

Let E be the equational system $\{l_1 = r_1, \cdots, l_n = r_n\}$, and let R be the term-rewriting system $\{l_1' \rightarrow r_1', \cdots, l_n' \rightarrow r_n'\}$ For equivalence between E and R, it is sufficient that (a) $l_i \leftrightarrow^* r_j (i = 1, \cdots, n)$ and (b) $l_i = r_j (j = 1, \cdots, m)$.

We can prove (a) by comparing the normal forms of l_i and r_i because the completeness of R is assured by the previous steps, and we can obtain their normal forms. Because the computation of normal forms can be weighty, we must send a cost estimate for this process. The cost estimate can be constructed as a reduction sequence in a manner similar to the construction of cost estimates for confluence.

We can prove (b) by using information obtained during the completion procedure. A rewriting rule in R is either an oriented equation that comes from E or a rewriting rule derived during completion. Given that both types of rewriting rule originate from equations in E, we can construct an equation sequence for $l_i' = r_j'$ from the rewriting sequence $l_i \leftrightarrow^* r_j$. Because the time complexity of this process is proportional to the size of the equation sequence, the client should send the equation sequence as a cost estimate.

2.4 Cost Estimate and Computational Cost on the Server Side

In this section, we explain how *the estimate-attaching method* defends against DoS attacks by malicious attackers in the equation-archive server.

As explained above, the time complexity for each processing step involved in checking a submitted equational system is proportional to the size of the message sent by the client. Hence, we may say that the server can know the rough computational cost of verification. If a malicious user tries to send an equational system that imposes

an enormous computational cost on the server, the accompanying cost estimate will have to be a very long message, and the server will be able to deduce that the submission is an attack.

3 Related Work

Necula [3] suggested the notion of "proof-carrying code," which guarantees the safety of code received from an untrusted site by attaching proof that the code satisfies the receiver's security policy. The idea of attaching proof to an imported code has something in common with our idea of attaching a cost estimate to an equational system.

Tsukada [4] proposed an improvement of proof-carrying code known as interactive proof-carrying code. In Necula's original scheme, an attached proof is sometimes larger than the code itself, which affects efficiency. Tsukada introduced the technique of *zero-knowledge proof* from cryptography to proof-carrying code. In place of the transmission of an entire proof, a small amount of interaction between server and client can ensure obedience to the security policy.

In our method, we improve resistance against denial-of-service attacks by revealing the clients' computational costs. However, there is an alternate approach to improving resistance to denial-of-service attacks, known as *client puzzle.* The idea is that by sending the client a problem that must be solved, the client's computational cost increases, and thus it becomes more difficult to launch a denial-of-service attack.

4 Concluding Remarks

In this paper, we introduced new software architecture to make an equation-archive server resistant to denial-of-service attacks. The main idea is that a client must attach a cost estimate to a submitted equational system. Since the size of the cost estimate is proportional to computational cost of verifying the submitted item, the server can assess whether the submission is an attack based on the size of the cost estimate description in the message. Although the proposed architecture is intended for an equation archive, we believe that it can be extended to more general web applications in which servers are open to unspecified clients and are subject to substantial client costs.

References

1. Baader, F., Nipkow, T.: Term Rewriting and All That. Cambridge University Press (1999)
2. Knuth, D., Bendix, P.: Simple word problems in universal algebras. In: Computational Problems in Abstract Algebra, pp. 264–296 (1970)
3. Necula, G.: Proof-Carrying Code. In: Proceedings of the 24th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, pp. 106–109 (1997)
4. Tsukada, Y.: Interactive and provablistic proof of mobile code safety. Automated Software Engineering 12(2), 237–257 (2005)

5. Meadows, C.: A formal framework and evaluation method for network denial of service. In: Proceedings of the 12th IEEE Computer Security Foundations Workshop, pp. 4–13 (1999)
6. Tomioka, D., Nishizaki, S., Ikeda, R.: A Cost Estimation Calculus for Analyzing the Resistance to Denial-of-Service Attack. In: Futatsugi, K., Mizoguchi, F., Yonezaki, N. (eds.) ISSS 2003. LNCS, vol. 3233, pp. 25–44. Springer, Heidelberg (2004)
7. Aura, T., Nikander, P., Leiwo, J.: DOS-Resistant Authentication with Client Puzzles. In: Christianson, B., Crispo, B., Malcolm, J.A., Roe, M. (eds.) Security Protocols 2000. LNCS, vol. 2133, pp. 170–177. Springer, Heidelberg (2001)
8. Ikeda, R., Narita, K., Nishizaki, S.: Cooperative model checking and network simulation for cost analysis of distributed systems. International Journal of Computers and Applications 33(4), 323–329 (2011)
9. Ikeda, R., Nishizaki, S., Ohata, T.: Formalization of broadcast communication in process calculus and its model checking. In: Proceedings of the 4th International Conference on Software and Data Technologies, ICSOFT 2009, vol. 1, pp. 348–352 (2009)
10. McDowell, M.: Understanding Denial-of-Service Attacks. Cyber Security Tip ST04-015, US-CERT (2004)
11. Denial of service attacks, CERT (1997), `http://www.cert.org/tech_tips/denial_of_service.html`

Real-Time Model Checking for Regulatory Compliance

Shin-ya Nishizaki and Takuya Ohata

Department of Computer Science, Tokyo Institute of Technology,
2-12-1-W8-69, O-okayama, Meguro-ku, Tokyo, 152-8552, Japan
nisizaki@cs.titech.ac.jp

Abstract. Nowadays, regulatory compliance is one of the most important issues in Japan. Due to the increasing number of regulations, it will not be easy to ensure that all governance requirements are fulfilled by the business processes of an information system. In this paper, we propose a new method of strengthening the compliance controls in information systems using model checking. We formulate an information system as a timed automaton and compliance requirements as CTL formulas. We employ the model checker UPPAAL to check whether the automaton satisfies the requirements. We apply our method to an example taken from Japanese banking regulations.

Keywords: regulatory compliance, model checking, timed automaton.

1 Introduction

1.1 Regulatory Compliance

Nowadays, regulatory compliance is one of the most important issues in many countries. For example, the Gramm–Leach–Bliley Act [1], the Sarbanes–Oxley Act (SOX) [2], and the USA Patriot Act [3] are in effect in the United States. In Japan, most of the internal control portions of the Financial Instruments and Exchange Act (FIEA) [4] were enacted in 2008. Documentation of company business processes is generally required by these acts, and formalization and improvement of these processes is also promoted. Today's business processes and their documentation are supported by existing IT systems, which therefore play an important role in ensuring compliance.

1.2 Model Checking

In software engineering, model checking is regarded as a genuine breakthrough, especially in regard to the improvement of software design and coding. Model checking is a technique for verifying whether a model satisfies a given specification. Models are extracted from descriptions presented as state-transition diagrams or in concurrent programming languages. The specifications are often represented by temporal logic formulae. A number of model checkers have been developed, including SPIN [5] and UPPAAL [6]. In UPPAAL, models are described in terms of timed automata using a GUI and specifications expressed by temporal logic CTL (Computational Tree Logic). The most distinctive feature of UPPAAL in comparison with other model checkers is its use of real-time clocks. We can write conditions for the branches of an

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 70–77, 2013.

automaton and verification queries using real-time clock variables to represent real-time constraints. UPPAAL can thus be applied to the verification of real-time systems [7,8].

Model checking is utilized to verify software correctness. Liu et al. [9] extended the uses of model checking to include checking compliance of business processes. We focus on real-time constraints in compliance controls and introduce a technique for checking the compliance controls in information systems to make them more rigorous. We have chosen the UPPAAL model checker, as it can be used to describe models with real-time constraints in terms of timed automata.

Purpose of Our Paper

In this paper, we propose a new method for checking the specifications of enterprise IT systems using a model checker. We present a case study of a bank accounting system in an imaginary bank called TinyBank. This case study is utilized to investigate the feasibility of our technique and to clarify some of the issues involved.

2 Modeling of an IT System and Its Compliance

2.1 Modeling with Timed Automata

A timed automaton is a variant of a finite automaton, extended via clocks that progress continuously and synchronously with absolute time. Each transition between nodes is labeled by a constraint (known as a guard) expressed as an arithmetic condition with integer variables and clock variables, and by an update statement expressed as an assignment statement. The formal definition of a timed automaton is presented in [10].

A model of the main processes of TinyBank as a timed automaton is illustrated in Fig 1 and Fig 2 represents the account-opening procedure, and Fig. 2 represents the payment procedure. To simplify the model for checking, we focus on a single account and omit procedures other than account opening and payment. The automaton comprising these two procedures is summarized as follows.

When the branch shown in Fig. 1 is selected, the automaton transits the states CustomerAccountReq, ObtainCustomerInfo, and VerifyCustomerInfo in turn. If the result of the customer information inquiry is unsatisfactory, the automaton transits to the state NotVerify. Otherwise, it transits to the states RetainCustomerDoc, OpenAccount, and End, in that order.

The branch shown in Fig. 2 is selected when the customer initiates a payment transaction. The automaton transits the state CustomerTransactionReq and then sends the SelectTransaction message to the automaton depicted in Fig. 3, which describes the current status of the transaction.

If the message is accepted, the automaton in Fig. 3 immediately sends the InputAmount message, which requests a payment amount for the transaction. The message is received by the automaton shown in Fig. 4, which models the customer's choice of payment amount. The automaton then transits to CheckAccount, which determines whether the customer has an account. If the customer has an account, the automaton transits to CheckAmount; otherwise, it jumps to ImplementTransaction, and the information-recording process for a doubtful customer follows. In the former case, the legitimacy of the payment amount is checked. If it is legitimate, the automaton

transits to CheckTransaction, ObtainCustomerInfo, VerifyCustomer, RetainCustomerInfo, and ImplementTransaction, in sequence. The path from ObtainCustomerInfo to RetainCustomerInfo is shared with the account-opening process shown in Fig. 1. These two types of path are distinguished by a flag variable x. The process is finished after a record of the transaction has been created.

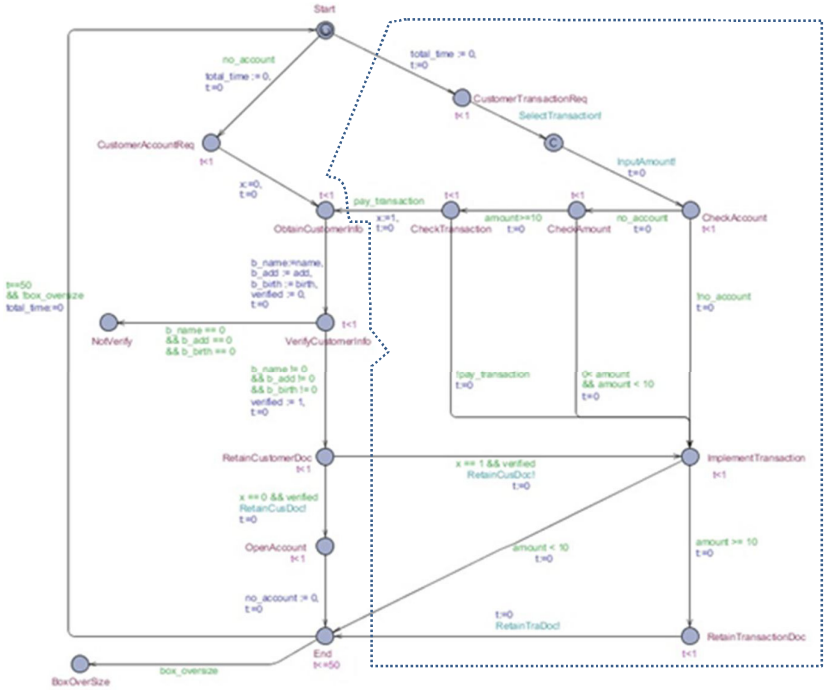


Fig. 1. Account-opening process

Fig. 2. Transaction process

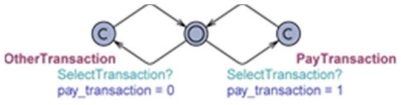


Fig. 3.

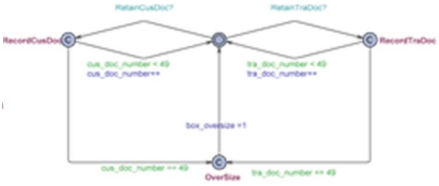


Fig. 4.

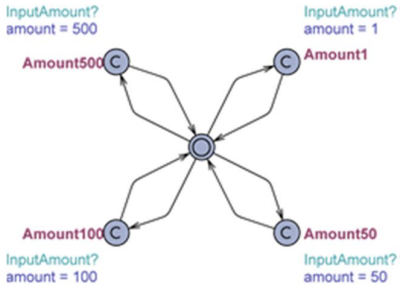


Fig. 5.

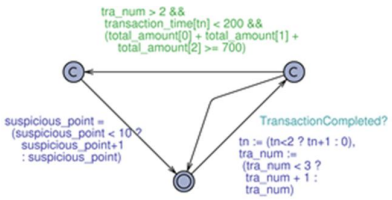


Fig. 6.

2.2 Compliance in the Model

Monitoring regulatory compliance is one of the important concerns in enterprise IT systems. In the early days, software verification was focused on correctness. Over the past 20 years, however, its domain has been extended as far as software security and is currently being further extended to include regulatory compliance.

As a case study of regulatory compliance verification, we incorporate compliance enforcement into our timed automata. In the model, surveillance for suspicious transactions is conducted during payment transactions. Specifically, we consider the following cases.

Personal Information Verification. Personal information is verified at the state VerifyCustomInfo. This is based on the Act on Prevention of Transfer of Criminal Proceeds [11]:

Article 5 (Immunity of Specified Business Operators from Obligations). A specified business operator may, when a customer, etc. or representative person, etc. does not comply with a request for customer identification upon conducting a specified transaction, refuse to perform its obligations pertaining to the said specified transaction until the customer, etc. or representative person, etc. complies with the request.

Personal Information Recording. After personal information has been verified, it is recorded at the state RetainCustomerDoc according to the Act [11]:

Article 6 (Obligation *to* Prepare Customer Identification Records, etc.).(1)A specified business operator shall, having conducted customer identification, prepare immediately, by a method specified by an ordinance of the competent ministries, records on customer identification data, on measures that have been undertaken for conducting customer identification, and on other matters specified by the ordinance of the competent ministries (hereinafter referred to as "customer identification records").

Reporting Suspicious Transactions. Suspicious transactions should be reported:
Article 9 (Reporting of Suspicious Transactions, etc.).

(1) A specified business operator shall, when property accepted through specified business affairs is suspected to have been criminal proceeds, or a customer, etc. is suspected to have been conducting acts constituting crimes set forth in Article 10 of

the Act on the Punishment of Organized Crime, or crimes set forth in Article 6 of the Anti-Drug Special Provisions Law with regard to specified business affairs, promptly report the matters specified by a Cabinet Order to a competent administrative agency, pursuant to the provisions of the Cabinet Order.

1. The automaton checks the payment amount in the state CheckAmount according to this article.
2. If frequent and consecutive monetary transactions are made during a short period, such transactions should be considered suspicious. The required surveillance is conducted by the timed automaton depicted in Fig.6. Specifically, if the automaton detects that the total amount of the last three payment transactions exceeded 700 in 200 units of time, a suspiciousness index is incremented.

3 Compliance Surveillance by Model Checking

3.1 Temporal Logic TCTL

The queries of the model checker UPPAAL are written in Timed Computational Tree Logic (TCTL). The formulas should have one of the following forms:

- A [] φ Invariantly φ,
- E <> φ Possibly φ,
- A◊ φ Always Eventually φ,
- E□ φ Potentially Always φ,
- φ →ψ φ always leads to ψ, which is an abbreviation for the implication formula φ ⇒A◊ ψ,

where φ and ψ are Boolean expressions over predicates on locations and integer variables, with clock constraints that can be checked locally at a state.

3.2 Describing Requirements via Temporal Logic Formulas

In this section, we enumerate six regulations for the system and express them in TCTL. These regulations are examples of regulatory compliance requirements for TinyBank.

Regulation 1. “If an account is open, its customer identification data should be verified.”

This regulation is required by Article 4 (Obligation to Conduct Customer Identification, etc.) of the Act on Prevention of Transfer of Criminal Proceeds [11], and is formalized by the TCTL formula

Customer1.CustomerAccoutReq → Customer1.VerifyCustomerInfo

The identity Customer1 denotes an instance of the timed automaton illustrated in **Fig.** and Fig. . We consider only the instance Customer1. This regulation says, “if the automaton arrives at the state CustomerAccountReq, then it eventually reaches the state VerifyCustomerInfo.”

Regulation 2. “After the personal identification data are verified, the data must be recorded.”

This regulation is required by Article 6 (Obligation to Prepare Customer Identification Records, etc.) of the Act [11] and is formalized by the TCTL formula
Customer1.VerifyCustomerInfo → Customer1.RetainCustomerDoc.

Regulation 3. "In handling a transaction of less than 100,000 JPY, the bank is not required to record personal identification data and the content of the transaction."

This is required by Article 7 (Obligation to Prepare Transaction Records, etc.) of the Act [10] and is formalized by the formula
InputAmountForm1.Amount1 →
!Customer1.VerifyCustomerInfo & !Customer1.RetainTransactionDoc,
where InputAmountForm1 is an instance of the timed automaton shown in Fig. 4. The unary operator ! and the binary operator && denote negation and conjunction, respectively. The states Amount1, Amount50, Amount100, and Amount500 in the instance InputAmountForm1 of the timed automaton shown in Fig. 4 indicate respective amounts of 10,000 JPY, 500,000 JPY, 1,000,000 JPY, and 5,000,000 JPY.

Regulation 4. "In handling a transaction of more than 100,000 JPY, the bank must record the content of the transaction." This is required by the same article and is formalized by the formula

(InputAmountForm1.Amount50
|| InputAmountForm1.Amount100
|| InputAmountForm1.Amount500)→
Customer1.RetainTransactionDoc.

Regulation 5. "If a customer with no account in the bank makes a payment transaction of more than 100,000 JPY, his/her identification confirmation is required." This and the next regulation are required by Article 6 as well as by Regulation 2. It is expressed by the formula

(InputTransaction1.PayTransaction &&
(InputAmountForm1.Amount50 || InputAmountForm1.Amount100 ||
InputAmountForm1.Amount500) &&
Customer1.no_account) → Customer1.VerifyCustomerInfo.

Regulation 6. "If a customer with no account in the bank makes a transaction of more than 2,000,000 JPY, his/her identification confirmation is required."
Customer1.no_account && InputAmountForm1.Amount500 →
Customer1.VerifyCustomerInfo.

3.3 Verifying the Formalized Requirements via the Model Checker UPPAAL

We verify the six queries of Section 3.2 in the model given in Section 2.2. The verification is carried out automatically by the model checker UPPAAL, and all execution paths of the model are exhaustively searched. If the number of possible states is small enough, the checker can verify the model efficiently.

Regulations 1, 2, 3, 4, and 5 were successfully verified. However, Regulation 6 failed. The model checker UPPAAL provided a counterexample against this regulation as a transition trace of the timed automaton. Fig. 7. shows a part of the transition trace.

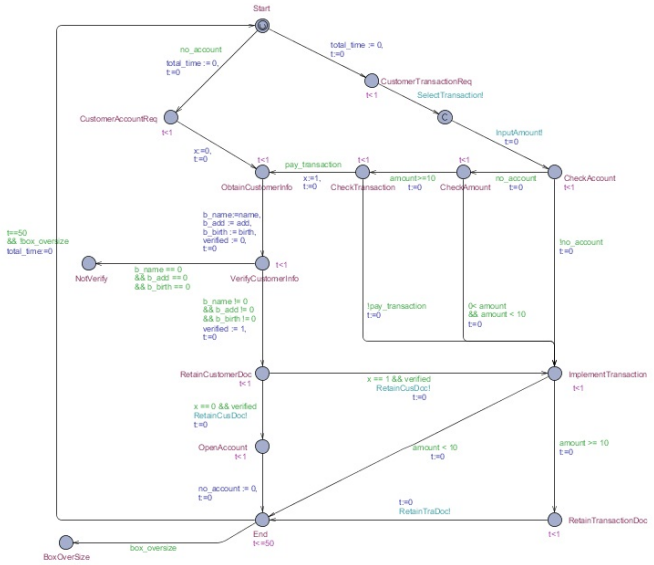


Fig. 7.

We know from the trace that the bank did not confirm the personal identification, even though the amount of the transaction was more than 5,000,000 JPY. This fault resulted from an error in the definition of the model of TinyBank and can easily be corrected by rewriting part of Fig. 1 and 2.

4 Concluding Remarks

In this paper, we proposed a method for strengthening compliance controls in information systems using model checking. We applied the technique to a case study of a tiny Japanese bank accounting system, formulating the bank's information system as a timed automaton and expressing Japanese banking regulations as CTL formulas.

4.1 Discussion

We recognize that the case study presented in the previous section is very elementary in comparison to the information systems involved in real bank accounting. However, it demonstrates the feasibility of software design driven by model checking, especially in regard to regulatory compliance. Our work differs from the pioneering work [9] of Liu et al. in terms of the logical framework employed. In [9], models are described in pi-calculus, and queries in BPSL (Business Property Specification Language), and these are translated into finite-state automata and LTL, respectively. Our work is based on timed automata and timed CTL, enabling us to describe real-time properties. The approach of Liu et al. seems to be more practical, and we therefore regard the introduction of their approach to our work as an interesting research direction.

4.2 Future Work

Software Design and Regulatory Compliance. The relevant laws are passed by a national assembly, and their interpretation is comparatively unambiguous. However, when we impose these laws on a software system for regulatory compliance, ambiguities sometimes occur. The boundary between software designers and compliance professionals is unclear and complicated. We should study methodologies for cooperative development of software systems.

Execution of Time Compliance Checking. Our method is statically applied. In other words, the compliance requirements are checked prior to their execution. Although such static time checking is valuable for developing reliable software, dynamic time checking is also helpful for improving the quality of software's compliance. One of the most successful instances of dynamic time checking is seen in the security automata [11] proposed by Schneider, which check security properties in execution time. This type of checking can also be regarded as promising in the present area. As the applicability of model checking is limited by the exhaustive search routines employed, it is essential to simplify the models to be checked. Hence, a combination of static time checking (such as model checking) and execution time checking (such as that conducted by security automata) seems to be a practical idea.

References

1. Gramm-Leach-Bliley Act of 1999 (GLBA), Public Law 106-102 (113 Statute 1338). United States Senate Committee on Banking Housing and Urban Affairs (1999)
2. Sarbanes-Oxley Act of 2002, Public Law 107-204 (116 Statute 745). United States Senate and House of Representatives in Congress (2002)
3. USA Patriot Act of 2001, Public Law 107-56, HR 3162 RDS. United States Senate and House of Representatives in Congress (2001)
4. The Financial Instruments and Exchange Law. Financial Services Agency, The Japanese Government (2006)
5. Holzmann, G.: The SPIN Model Checker: Primer and Reference Manual. Addison-Wesley Professional (2003)
6. Pattersson, P., Larsen, K.: UPPAAL 2k. Bulletin of the European Association for Theoretical Computer Science 70, 40–44 (2000)
7. Ravn, A.P., Srba, J., Vighio, S.: Modelling and Verification of Web Services Business Activity Protocol. In: Abdulla, P.A., Leino, K.R.M. (eds.) TACAS 2011. LNCS, vol. 6605, pp. 357–371. Springer, Heidelberg (2011)
8. Iversen, T., Kristoffersen, K., Larsen, K., Laursen, M., Madsen, G., Mortensen, S., Pettersson, P., Thomasen, C.: Model-checking Real-time Control Prgorams. In: Proceedings of the 12th Euromicro Conference on Real-time Systems (ECRTS 2000), pp. 145–155 (2000)
9. Liu, Y., Müller, S., Xu, K.: A static compliance-checking framework for business process models. IBM System Journal 46(2) (2007)
10. Act on Prevention of Transfer of Criminal Proceeds, Act No. 22 of 2007. The Japanese Government (2007)
11. Bengtsson, J., Yi, W.: Timed Automata: Semantics, Algorithms and Tools. In: Desel, J., Reisig, W., Rozenberg, G. (eds.) Lectures on Concurrency and Petri Nets. LNCS, vol. 3098, pp. 87–124. Springer, Heidelberg (2004)
12. Schneider, F.: Enforceable security policies. ACM Transactions on Information and System Security 3(1), 30–50 (2000)

A Biometric Based Design Pattern for Implementation of a Security Conscious E-Voting System Using Cryptographic Protocols

Nirmalya Kar, Sharmistha Roy, Ashim Saha, Kunal Chakma, and Anupam Jamatia

Computer Science & Engineering Department,
National Institute of Technology, Agartala, 799055, Tripura, India
sharmistharoy11@gmail.com,
{nirmalya,ashim.cse,kchakma,anupamjamatia}@nita.ac.in

Abstract. With the recent explosion of new technologies in the field of Cryptography, there comes the need for increasing the Security of the present voting system of India. The use of cryptographic protocols as well as biometrics provides a technical response to the security flaws of the current voting system in India. Voting is an instrument of democracy that provides an official mechanism for people to express their views to the government. And e-Voting is considered as one of the most intensely debated subjects in Information Technology. Thus, the voting system should be much more secure and it should also have the stronger authentication mechanism so as to remove fraud and cheats completely from the voting process. This paper proposes and discusses the design of a secure e-voting system based on cryptographic protocols and biometrics that can replace the traditional voting system of India.

Keywords: E-Voting, Biometrics, Ballot, Cryptography, Security, Encryption, Public key.

1 Introduction

Election is a fundamental tool that allows people to choose who would govern them. Voting is one of the most critical features in our democratic process. E-voting is a term used to describe any of several means of determining people's collective intent electronically. In a way, the need for a secure electronic voting system is an obvious demand [1]. The construction of electronic system is one of the most challenging security-critical tasks, because of the need for finding a trade-off between many seemingly contradictory security requirements like privacy vs. audit ability. The aim of this paper is to study the flaws in the existing voting system in INDIA and to propose new design pattern and implementation of a security conscious e-voting system.

2 Security Requirements

This section elaborates the security requirements that an e-voting system should meet to make it secure. We claim that our system satisfies these requirements.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 78–85, 2013.

Table 1. Security Requirements for the e-voting system

Security Requirements	Description
Confidentiality	The voter's ballot should be kept confidential. In addition, a voting protocol must not allow any voter to prove that he or she voted in a particular way.
Authentication	The system must ensure that the person is an eligible voter.
Integrity	The protocol must ensure that only valid votes are counted in the final tally and no one can change anyone else's
Uniqueness	Each eligible voter has voted only once. Avoid double voting.
Verifiability	Voter can validate that his vote is recorded correctly.
Receipt-Freeness	Voter is not identifiable from the receipt. Vote is not revealed from the receipt. Voter cannot prove his vote.
Anonymity	Voter identity is not revealed to anyone.
Fairness	Result is not published till the end of the election. Counting comes after the voting stage. No one can guess the content of any cast vote.

3 Security Flaws in the Existing Voting System of India and Its Related Works

India is the world's largest democracy. In paper based voting scheme, the risk of an information leak is several degrees higher than in an electronic environment where frauds on a similar scale can be executed in an automated manner by just a few people. From a security perspective, the use of electronic voting machines in elections around the world continues to be concerning. The present voting system in India includes EVM (Electronic Voting Machine) which consist of many flaws and lacks many security related issues. Registration process cannot guarantees the uniqueness of the system. Among the other security related flaws, Verifiability is one of them. In voting process, once vote has been casted EVM cannot convince the voter in future that his vote has been received properly and counted properly. The second and the important feature is Security. Because capturing of the EVM machine can cause the temperament of the whole voting process. The third feature that EVM lacks is Authenticity, i.e. EVM cannot authenticate that only eligible voter can cast vote and each voter has cast only once. There is no such mechanism that guarantees that only the eligible person has cast his own vote. Another important flaws lie in the security of the EVM machine itself. Though EVM manufacturers and election officials have attempted to keep the design of the EVMs secret, this presents only a minor obstacle for would-be attackers. There are nearly 1.4 million EVMs in use throughout the country and criminals would only need access to one of them to develop working attacks [3].

In the last few decades, many e-voting protocols have been proposed but none of them could satisfy the voting requirements of India. The e-voting system proposed by

Rachid Anane [9] cannot meet the requirements because it was very difficult to implement an e-voting system in India where computer literacy rate is very low. Moreover, one more e-voting system is proposed by Mohammed Khasawneh [4] which lacks security as well as confidentiality. In principle, many security issues can be allayed with cryptography. Several cryptographic solutions have been proposed to meet these requirements, but their effectiveness is predicated on several assumptions [7]. All voting schemes mentioned above suffer from the "abstaining voters" problem: the authority can cast bogus votes on behalf of voters who register but then decide to abstain from subsequent steps [2]. Obviously, both the first invulnerability and the third accuracy requirements are violated with this attack.

4 The Proposed E-Voting Architecture

The proposed system is subdivided into three modules: Registration phase, Voting phase and Counting phase. There are mainly three different servers: Local Server, Global Server, and Mirrored Server. Local Server performed three roles for three different processes. The three roles include Administrator, Validator, and Counter. The Global Server received information from different Local Servers and kept the record for future execution. The Mirrored Server is mainly used in case of any loss of data or further verification purpose.

Client Machine: The Client program performed the functions on behalf of the voter. These functions include exchanging messages with the servers, processing user input and performing necessary cryptographic transformations.

Server: The cryptographic operation as well as the biometric authentication was done by the server. Both the Registration process and Voting process were controlled by Administrator Server and Counter Server.

Administrator: During Registration process the Administrator maintained the records of all the voters who were eligible for voting. It also generates public-private key-pair corresponding to each voter. A database for generating the unique Voter ID number for each candidate was also maintained by the Administrator. During Voting, it checks whether the person had already registered or not. It also checked whether the person was the eligible voter or not by using the biometric of the person. It maintained the voting records of each voter who cast the vote successfully. It also maintained the Candidate list for each region.

Validator: The main task of Validator was to keep voting record of those entire candidates who casted the vote successfully.

Counter: The Counter server was responsible for collecting the valid votes and counting the votes corresponding to each candidate.

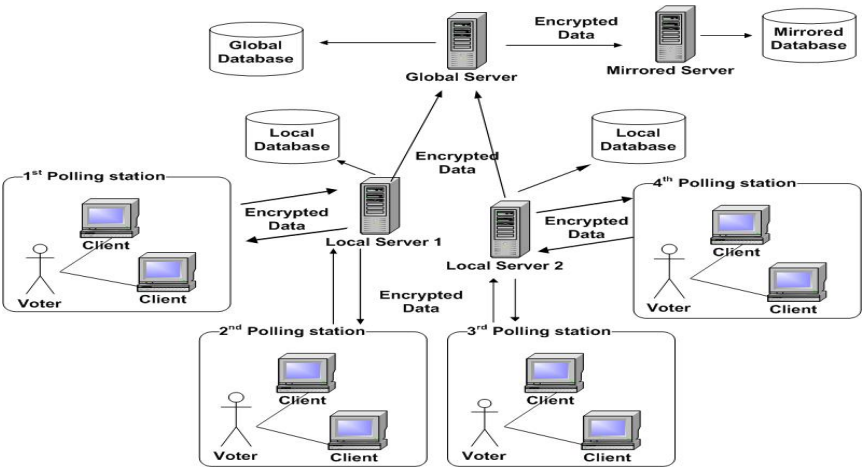


Fig. 1. General Schematic diagram of the proposed system

The proposed architecture was designed in such a way that both the Registration & Voting Procedure had been conducted in various Polling booths. Each Polling station maintained several Client machines. Local server performed operation on behalf of two-three Polling stations. The Global server maintained the entire voting records from all the Local Servers and update the Global database throughout the whole voting process.

4.1 Cryptographic Protocols

Public key cryptography, also called asymmetric key cryptography, is a family of cryptographic algorithms which use two keys. One key is the private key that must be kept secret, while the other key is the public key which is advertised [6]. In our proposed scheme we used RSA as one of the strongest asymmetric key algorithm.

4.1.1 Key Generation

Select two prime numbers p & q.
ϕ (n) = (p-1) (q-1)
Public Key (e) = {e | e $\perp$ (p-1) (q-1) and e< ϕ (n): $\perp$ denotes relatively prime}
Private Key (d) = {de ≡ 1 (mod ϕ (n)) and d< ϕ (n)}

4.1.2 Encryption Procedure

Input comprises of a Plain Text (PT) and Public key (e).
Cipher Text (CT) = {PT^e (mod n)}

4.1.3 Decryption Procedure

Input is Cipher Text (CT) and Private key (d).
Plain Text (PT) = {CT^d (mod n)}

4.2 Authenticity in the E-Voting System

In our proposed scheme, the authentication of the voter was done using biometric [5]. We employed fingerprint as our biometric [8]. During Registration process, the user gave samples of fingerprint which send through the channel in an encrypted format. In the server side, these information's were being processed and the statistical result was stored. During Voting process fingerprint of the person was send through the channel in an encrypted format which was then compared with the previously stored template. Whenever, the matching was found then the person was allowed for casting vote. The matching was done using a standard biometric matching system. It thus helped the Administrator to remove the fraud and made the authentication much more stronger and challenging.

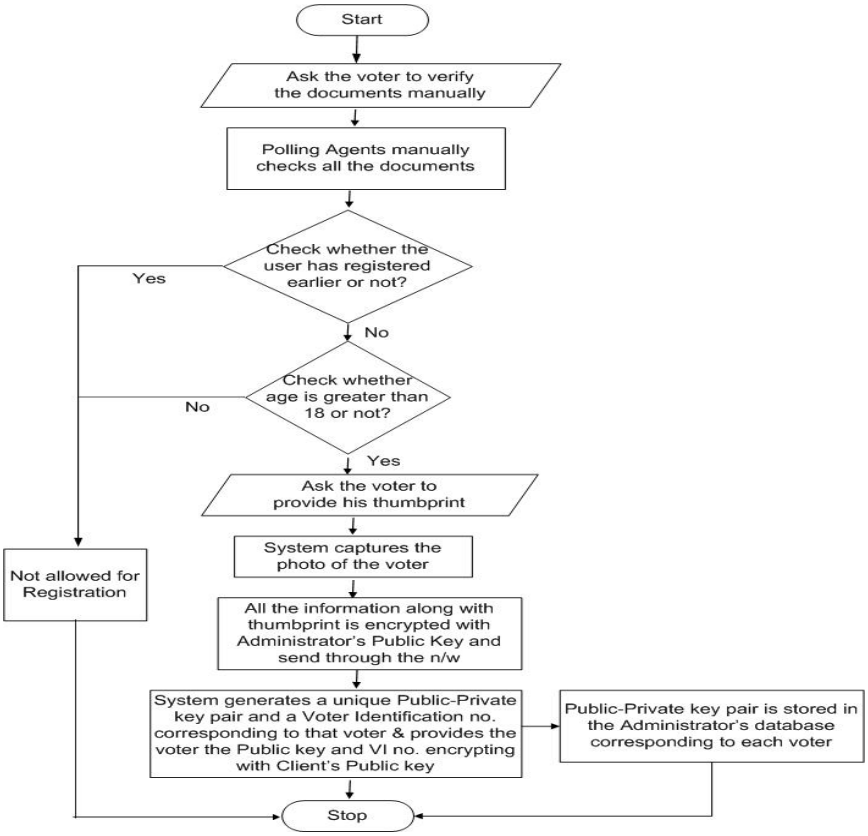


Fig. 2. Flowchart representation of Registration Process

4.3 Registration Procedure

In any election, every individual must register to be an eligible voter. The voter's registration is a phase which enables a voter to make an entry in to the voting process

and generates an authentication id for casting vote later on. The steps necessary for registration procedure is described below:

i) First of all, the user has to bring all his documents which are manually verified by the polling agents.
ii) If the documents are verified, then the system checks the registration status of the particular user and gives permission for further registration.
iii) In the next step, all his details will be entered into the system and checks the age status.
iv) If eligible, the user is prompt to provide his fingerprint into the system [5].
v) System also captures the user's photo.
vi) Now, all the information's are send in an encrypted form to the server which keeps a record for every individual and generates a unique Public-Private key pair which is also maintained into the database.
vii) Finally a unique Voter ID no. is generated corresponding to every individual.

The user is provided with a card which stores the necessary information necessary for voting process along with the Public key embedded into it. Finally, the registration procedure is over and the Administrator server maintains a record of all the individuals who have registered successfully for further process.

4.4 Voting Procedure

After the registration process is over, next comes the main task i.e. the Voting procedure which involves the participation of the voter for giving their choices for construction of the government. The steps are described below:

i) First of all, the user shows the ID card given during registration process. The system captures the Voter Identification no. and checks whether this no. is already registered or not.
ii) If registered, then the system checks the voting status of that individual. If the voting status is false, then the user is asked to provide his fingerprint.
iii) If the fingerprint is matched then the system recognizes that the individual is an authentic voter and he is then provided with a unique vote ballot. The ballot is depended on the location or area of the voter.
iv) Next the voter cast the vote according to his choice only once corresponding to his favorite candidate.
v) The vote along with Voter Identification number and location is then send in an encrypted form to the Administrator server who maintains a separate record of all the votes along with voter authentication number and location. The information is encrypted with voter's Public key which is decrypted only by his Private key and the information of key is known by the Administrator only.
vi) The voting status along with VI no. is also been updated in Validator's database for future verification.
vii) The vote received by Administrator after decryption consists only a serial no. which is then send to the Counter server by encrypting with Counter's Public key.

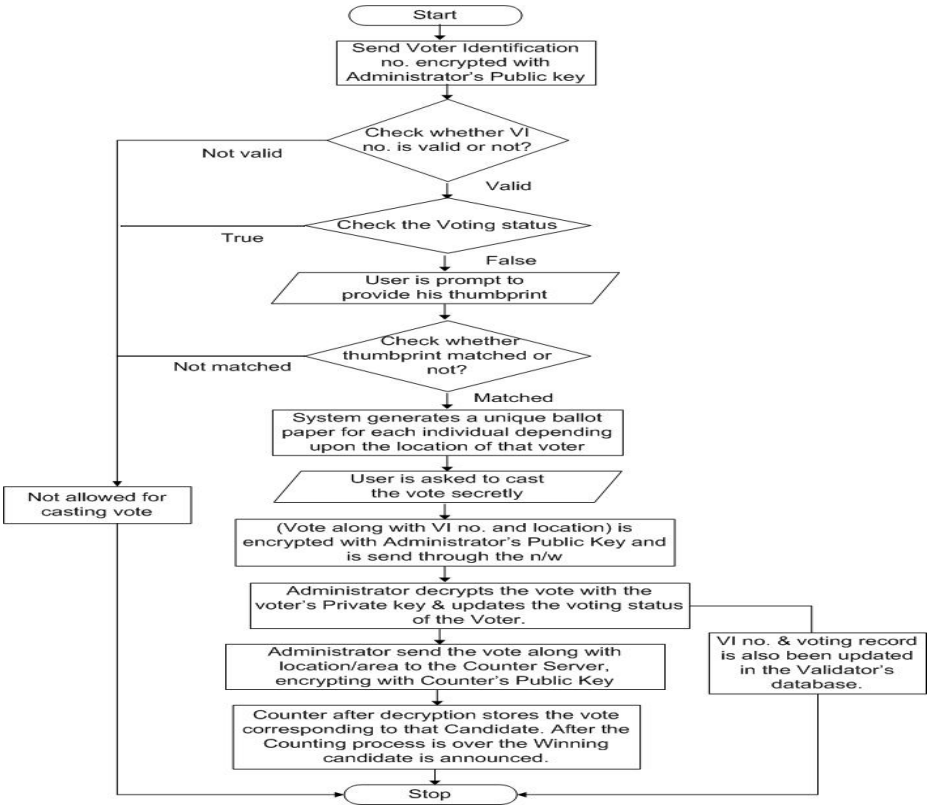


Fig. 3. Flowchart representation of Voting & Counting Process

4.5 Counting Procedure

The Counting procedure of the present e-Voting system includes the following steps:

i) Counter after receiving the vote, decrypts it with its Private Key and receives the Serial number.
ii) The serial no. corresponds to a particular Candidate of that location, which is maintained by Counter server.
iii) It then updates the voting record of that particular Candidate for whom the vote has been cast.
iv) Then among all the votes received Counter declared the Winning Candidate of that particular region. This process could not co-relate the Voter with his Vote.

5 Conclusion

With the use of a biometric-secure e-voting system, as the one proposed in this paper, many of the issues, that have challenged traditional voting systems in the past, are

bound to be resolved providing peace of mind to both voters and election candidates. At every step the system maintains confidentiality, security, privacy & authentication. The new system preserves the integrity of the voting process from the minute a voter steps in to cast his/her vote until the cast vote is registered in favor of the chosen candidate at a globally allocated DB repository. The proposed system is capable of denying access to any illegal voter/s, preventing multiple votes by the same voter, and blocking any introduced forms of malice that would adversely affect the voting process altogether. Thus, the proposed system is capable to increase the citizens trust and confidence and increases the voting population altogether.

References

1. Omidi, A., Azgomi, M.A.: An Architecture for E-Voting Systems Based on Dependable Web Services, Islamic Azad University, Makoo Branch, Makoo, Iran School of Computer Engineering, Iran University of Science and Technology, Tehran, Iran. IEEE (2009)
2. Ondrisek, B.: E-Voting System Security Optimization. Vienna University of Technology IEEE (2009)
3. Election Commission of India. Information under RTI on EVMs. No. RTI/2009-EMS/39 (July 2009)
4. Khasawneh, M., Malkawi, M.: A Biometric-Secure e-Voting System for Election Process, IEEE Senior Member, College of Engineering, Vice President, AIM Wireless, USA (May 2008)
5. Nanavati, S., Thieme, M., Nanavati, R.: Biometrics: Identity Verification in a Networked World. John Wiley and Sons, Inc. (2002)
6. Ibrahim, S., Kamat, M., Salleh, M., Aziz, S.R.A.: Secure E-Voting With Blind Signature, Faculty of Computer Science & Information Technology, University Technology of Malaysia
7. Gritzalis, D. (ed.): Secure Electronic Voting. Springer, Berlin (2003)
8. Smith, J., Schuckers, S.: Improving Usability and Testing Resilience to Spoofing of Liveness Testing Software for Fingerprint Authentication (2005)
9. Anane, R., Freeland, R., Theodoropoulos, G.: e-Voting Requirements and Implementation, Computer and Network Systems, Coventry University, UK School of Computer Science, University of Birmingham, UK. (CEC-EEE 2007) (2007)

Fingerprint Matching Based on Texture Feature

Ravinder Kumar[1], Pravin Chandra[2], and M. Hanmandlu[3]

[1] Deptt. of Computer Science and IT, Ansal Institute of Technology
Gurgaon - 122003, India
[2] University School of Information Technology, GGSIP University
Delhi - 110075, India
[3] Department of Electrical Engineering, Indian Institute of Technology,
Hauz Khas, New Delhi -110016, India.
ravinder.kumar@aitgurgaon.org, chandra.pravin@gmail.com,
mhmandlu@ee.iitd.ac

Abstract. This paper presents a texture feature based algorithm for fingerprint matching. Our proposed fingerprint-matching algorithm employs texture features like Correlation, Inverse Difference Moment, and Entropy measure of fingerprint images. The proposed textural features of two fingerprints have been compared to compute the similarities at a given threshold. The algorithm has been tested on the FVC2002 DB2_B database. The proposed algorithm is evaluated using GAR and FAR. GAR of 97.5 % is observed with 8.53% of FAR at a threshold value of 0.9. The proposed textural Feature based matching will enhance GAR at the cost of slightly higher value of FAR and hence gives the best GAR at reasonable value of FAR.

Keywords: Fingerprint Matching, Textural Feature, Genuine Acceptance Rate, False Acceptance Rate, Co-occurrence Matrix.

1 Introduction

In recent years biometrics physiological and/or behavioral traits have been widely accepted and implemented for individual identification [1]. Biometrics system can operate in two modes depending on the types of application [2] i.e. v*erification mode, and identification mode.* From all the biometric traits, fingerprints have been popularly used and accepted biometrics, because of its highest levels of reliability [3] and have been extensively used by forensic experts in criminal investigations [4]. The methods of fingerprint recognition have been mainly classified as minutiae based, ridge feature based (like orientation), correlation-based methods. Minutiae – based matching is a popular and widely used technique for fingerprint matching. In this technique minutiae points are extracted from two fingerprint images and stored as a set of points in two – dimensional plane. This technique essentially consists of alignment of template and the input minutiae set to determine the total number of matched minutiae [5]. In correlation based matching technique two fingerprint images are superimposed, and then pixel intensity level matching will be done at different alignments [6, 7]. These techniques are suffering from non-linear distortion, different finger pressure and alignment, and skin conditions.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 86–91, 2013.

2 Related Work

The pattern of ridges and valleys in a fingerprint image can be observed as an oriented texture pattern [8]. Jain et al. [8] proposed a hybrid algorithm that uses both minutiae and texture information for fingerprint matching. Aggarwal et al. [9] proposed a gradient-based approach to extract texture features around each minutiae location. Recently, co-occurrence matrices are employed to extract texture features of a fingerprint image. A fingerprint verification method proposed by Arivazhagan et al. [10], uses both Gabor wavelets and co-occurrence matrices to extract textural features. Yazdi et al. [11] also used co-occurrence matrices for fingerprint classification. We have proposed a fingerprint-matching algorithm using the three textural feature of fingerprint image characterized by the co-occurrence matrix. This method assumes that tone and texture are always present in an image. In order to extract these features a co-occurrence matrix is to be formed from the images. Haralick et al. [12] have proposed the 14 textural features. Out of these 14 textural features, three features seem to be relevant in context of fingerprint images. The main step of the proposed method includes: fingerprint image enhancement, core point detection, and ROI extraction. In this method, we have cropped the image around its core point of size (101× 101) [13]. All the rotations are performed by considering the core point as the central point [14]. Finally the textural features (Correlation, Inverse Difference Moment, and Entropy) are extracted by statistical analysis of the co-occurrence matrix formed from the cropped image at various angles (0^o, 45^o, 90^o, 135^o). This paper is organized in the following manner: section 3 describes the proposed algorithm. Section 4 discusses co-occurrence matrix and proposed feature extraction. Section 5 highlights the experimental results. And finally, the section 6 concludes the paper.

3 Proposed Algorithm

The flowchart of the textural feature based fingerprint-matching algorithm is shown in Figure 1.

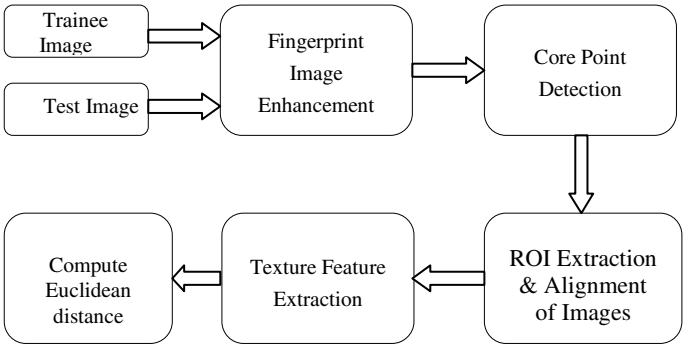


Fig. 1. Flowchart of Proposed Algorithm

The steps of proposed algorithm are explained in following sub-sections.

3.1 Fingerprint Image Enhancement

Automatic fingerprint matching is dependent upon the comparison of local ridge characteristics and their relationships to make a personal identification [15]. The performance of a feature extraction algorithm heavily depends on the quality of the input fingerprint images. To deal with poor quality image, fingerprint enhancement is used to improve the performance of the matching algorithm. In this paper we have used the algorithm proposed by Hong, Wan and Jain [16] with different filter to increase the contrast between ridge and valley. The enhanced fingerprint image is shown in figure 2.

Fig. 2. (a) Typical Fingerprint image and (b) Enhanced image

3.2 Core Point Detection

Many approaches for detecting singular points were proposed in the literature. Poincare index method for detecting singular points in fingerprint images has been widely used [16]. It has been shown that Poincare index method usually detects almost all true singular points when the index is computed along small region boundaries. In this work we have used Poincare index method to detect core point [15].

3.3 ROI Extraction

To reduce the computational complexity, we have cropped the image of size 101×101 (m×n) around core point and considered as the region of interest (ROI) for computing texture feature [13].

3.4 Alignment of Fingerprint Images

After extraction of region of interest fingerprint images are aligned using rotation about core point by an angle of rotation θ_r given in equation (1) [13].

$$\theta_r = \frac{\left(\theta'(1,1) - \theta(1,1)\right) + \left(\theta'(m,n) - \theta(m,n)\right)}{2} \tag{1}$$

4 Co-occurrence Matrix

First order statistics based features only provide the gray value distribution but lack in providing topographical distribution of the pixels. Second order statistics considers

pixels in pair and can be useful in extracting this type of information. Gray scale co-occurrence matrix also known as gray tone spatial dependence matrix, is used to extract second order texture information. It has been shown that texture information is specified by the matrix of relative frequencies $[P(i,j,d,\phi)]$ in which two pixels separated by distance d occur on the image, one with gray tone i and the other with gray tone j in the direction specified by an angle ϕ [13]. Co-occurrence matrices for angles quantized to 45^o interval are defined in [13].

4.1 Texture Feature Computation

It is assumed that all the texture information is contained in Gray scale co-occurrence matrices. Haralick et al. [12] defines the measures of 14 different textural features, which can be extracted from the co-occurrence matrices. In this paper we have employed 3 of the 14 textural features defined as given below.

Correlation	Inverse Difference Moment	Entropy
$f_1 = \frac{\sum_i \sum_j (i,j) p_{ij} - \mu_x \mu_y}{\sigma_x \sigma_y}$	$f_2 = \sum_i \sum_j \frac{1}{1+(i-j)^2} p_{ij}$	$f_3 = -\sum_i \sum_j p_{ij} \log_2 p_{ij}$

Above defined features are extracted for each gray level co-occurrence matrices in four direction specified by an angle 0^o, 45^o, 90^o, 135^o respectively. Generally the relative distance between the ridges in fingerprint images are at 4 pixels. Hence 16 gray level co-occurrence matrices are obtained with distance d = 4 pixels for each angle as shown in figure 4.

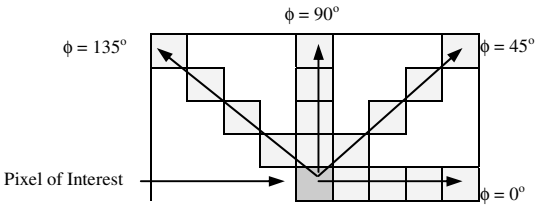


Fig. 3. Various directions and the pixel of interest

Here core point is assumed to be the pixel of interest. Each texture feature has been computed by taking the mean of 16 features computed for each gray level co-occurrence matrix obtained by varying the values of d (1 to 4) and ϕ (0^o, 45^o, 90^o, 135^o). The matching score has been computed by taking the mean of three features (Correlation, Inverse Difference Moment, and Entropy). The computed matching score is further compared with the threshold to take the matching decision.

5 Experimental Results and Discussion

Any biometrics traits have been evaluated using the following four parameters: Genuine Acceptance Rate (GAR) Genuine Rejection Rate (GRR); False Rejection Rate

(FRR); False Acceptance Rate (FAR). The FVC2002 DB2_B database [17] has been taken to test the proposed algorithm. Matching performance of the proposed algorithm has been measured in terms of genuine accept rate (GAR) and False acceptance Rate (FAR) for predefined threshold. Total fingerprint images in DB2_B database are T = 8 × N, where N is the number of subject (images of different person) in the database. A single image from each subject has been considered as trainee image and remaining T −1images are taken as test images for experimentation. Total trials carried out for finding genuine claims and imposter claims are N × (T−1), in which genuine claims are N × 7 and imposter claims are (total trials − genuine claims). GAR and FAR are measured in percentage and given in equation (2).

$$\begin{aligned} &\text{FAR (in \%age)} = (\text{Imposter claims accepted / Total imposter claims}) \times 100, \text{ and} \\ &\text{GAR (in \%age)} = (\text{Genuine claims accepted / Total genuine claims}) \times 100, \text{ and} \end{aligned} \tag{2}$$

In FVC2002 DB2_B database there are 10 subjects with 8 fingerprints in each subject. The subject 104_1.tif has been rejected for computing GAR; due to their core point being very close to the boundary region but these subjects have been considered for computing FAR. The test results of the experiment have been shown in table 1. The matching has been performed with the threshold value of 0.9. Weighted GAR and FAR has been computed for each subject and also shown in the table.

Table 1. Test results of FVC2002 DB2_B databaseat a threshold of 0.9

Database	Subjects	GAR (in %age)	FAR (in %age)
FVC2002 DB2_B	101	100	8.3
	102	100	9.7
	103	87.5	4.1
	105	87.5	4.1
	106	100	5.5
	107	87.5	6.9
	108	100	6.9
	109	100	5.5
	110	87.5	6.9
Weighted Mean		97.5	5.83

The Euclidean distance is employed for computing the similarity, between the trainee and test image.

6 Conclusion

In this proposed method fingerprint images are being matched using second order statistical features. The experimental results are very encouraging. To reduce the computational efforts, we have extracted the region of interest (ROI) around core point. The result shows that the performance of this method is comparable with minutiae based method. The proposed algorithm can only be applied to the images having core point. Currently we are investigating the hybrid method, where textural features can be combined with minutiae features for fingerprint matching.

References

1. Jain, A.K., Ross, A., Prabhakar, S.: An introduction to biometric recognition. IEEE Transactions on Circuits and Systems for Video Technology 14(3), 4–20 (2004)
2. Wayman, J.L.: Fundamentals of biometric authentication technologies. Int. J. Image Graphics 1(1), 93–113 (2001)
3. Berry, J., Stoney, D.A.: The history and development of fingerprinting. In: Lee, H.C., Gaensslen, R.E. (eds.) Advances in Fingerprint Technology, 2nd edn., pp. 1–40. CRC Press, Florida (2001)
4. Federal Bureau of Investigation, The Science of Fingerprints:Classification and Uses, Washington, D.C, U.S. Government Printing Office (1984)
5. Jain, A.K., Prabhakar, S., Pankanti, S.: Filterbank – based fingerprint matching. IEEE Transactions on Image Processing 9, 846–859 (2000)
6. Ravichandran, G., Casasent, D.: Advanced in-plane rotation-invariant correlation filters. IEEE Transactions on Pattern Analysis and Machine Intelligence 16(4), 415–420 (1994)
7. Bazen, A.M., Verwaaijen, G.T.B., Gerez, S.H., Veelenturf, L.P.J., van der Zwaag, B.J.: A Correlation-Based Fingerprint Verication System. In: Proc. Workshop on Circuits Systems and Signal Processing, pp. 205–213 (2000)
8. Jain, A.K., Ross, A.: Fingerprint Matching Using Minutiae and Texture Features. In: Proceeding of International Conference on Image Processing (ICIP), pp. 282–285 (2001)
9. Aggarwal, G., Ratha, N.K., Jea, T.-Y., Bolle, R.M.: Gradient based textural characterization of fingerprints. In: Proceedings of IEEE International Conference on Biometrics: Theory, Applications and Systems (September-October 2008)
10. Arivazhagan, S., AruIFiora, T.G., Ganesan, L.: Fingerprint Verification using Gabor Co-occurrence Features. In: International Conference on Computational Intelligence and Multimedia Applications, pp. 281–285 (2007)
11. Yazdi, M., Gheysari, K.: A New Approach for the Fingerprint Classification Based on Gray-Level Co-Occurrence Matrix. International Journal of Computer and Information Science and Engineering (2008)
12. Haralick, R.M., Shanmugam, K., Dinstein, I.H.: Textural Features for Image Classification. IEEE Transactions on Systems, Man and Cybernetics 3, 610–621 (1973)
13. Kumar, R., Chandra, P., Hanmandlu, M.: Fingerprint Matching Based on Orientation Feature. Advanced Materials Research Journal 403-408, 888–894 (2011), doi:10.4028/www.scientific.net/AMR.403-408.888
14. Kumar, R., Chandra, P., Hanmandlu, M.: Fingerprint Singular Point Detection Using Orientation Field Reliability. Advanced Materials Research Journal 403-408, 4499–4506 (2011), doi:10.4028/www.scientific.net/AMR.403-408.4499
15. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of Fingerprint Recognition. Springer (June 2009)
16. Hong, L., Wan, Y., Jain, A.: Fingerprint Image Enhancement: Algorithm and Performance Evaluation. IEEE Transactions on Pattern Analysis and Machine Intelligence 20(8) (August 1998)
17. Fingerprint Verification Competition (FVC), `http://bias.csr.unibo.it/fvc2002/`

A Comparative Study and Design of Full Adder Cells Using Carbon Nano Tube FET

V. Malleswararao, P.H.S. TejoMurthy, and V. Nooka Raju

Department of EIE, GIT, GITAM University, Visakhapatnam
teja_usha@rediffmail.com

Abstract. Adders are the core elements of the complex arithmetic operations. These are the essential component which limits the speed of the operation of the overall systems like DSP Processors, Microprocessors etc. In this paper, we are proposing a new full adder based on FTL Logic using Carbon Nano Tube FET's which are suitable for biomedical sensor applications. The proposed method significantly reduces the problem of Threshold voltage loss. The results are compared with 8-T, 10-T and 16-T full adders using CMOS and Carbon Nano tube FET's

Keywords: feed through logic, Carbon naotube FETS, Adders, Bio medical sensors, Full adders.

1 Introduction

The basic building blocks of full adders are XOR, XNOR and AND gates. In the literature, various XOR, XNOR have been proposed which are motivated by three design goals viz. Minimizing the Transistor Count, Minimizing the Power Consumption and Increasing the Speed of the system. Therefore the low power design is of the major consideration. The conventional adder requires 28-Transistors are implemented in CMOS technology and it has full rail to rail voltage. But the problem with 28-Transistor full adder is the more dynamic power consumption and large area requirement. These problems are avoided by various full adders proposed in the literature. These low power full adders have the problem of threshold voltage loss thus produce weak 0 and weak 1 However; these are very useful in many applications like Multipliers etc. In the following section we shall discuss elaborately about the advantages and disadvantages of these gates.

2 4-Transistor XOR and XNOR Gates

The 4-Transistor XOR and XNOR Gates are shown in the figure (1). The XNOR gate uses pass transistor logic. The pull down network is built in such a way that, the input of one transistor is used as source to another transistor. A is always compared with B and vice versa. If both are same then only the transistors will be ON and creates the discharging path.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 92–96, 2013.

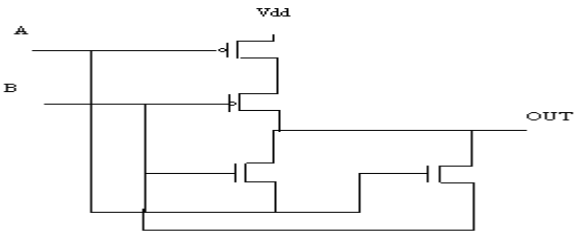


Fig. 1. XNO

Fig. 2. XNOR Using CNFET

Fig. 3. XNOR Using 50nm technology

As shown in the above figure it is shown that there is significant threshold loss. It is implemented both with CMOS 50nm technology and Carbon Nano Tube FET's. The advantage of using CNFET here is it increases the speed of operation with out threshold loss for logic 1.The following table explains the comparison between the two technologies.

Table 1. Comparison of XNOR between 50nm and CNFET

A	B	OUT 50nm	CNFET
0	0	Weak 1	Strong 1
0	1	Strong 0	Strong 0
1	0	Weak 0	Weak 0
1	1	Weak 1	Strong 1

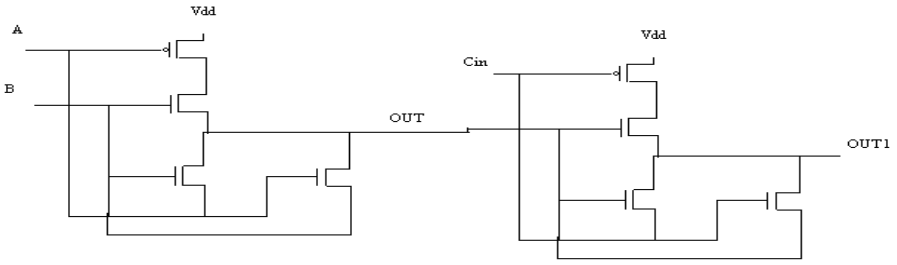


Fig. 4. Full Adder Sum Implementation Using 8-Transistors

3 Proposed Circuit

P The circuit shown in figure () is based on feed through evaluation concept. It consists of NMOS logic network and a pull down transistor (Td) for discharging the output node to low logic level and a pull up PMOS load transistor (Tp). Td and Tp are operated by a clock signal. The advantages of FTL logic over Dynamic logic are the elimination of cascading problem and charge sharing problem. FTL always resets the out put node during the resetting phase and during the evaluation phase it MP1 is ON and MP2 will be OFF. Irrespective of the input values cascaded gats first rise to Vdd/2 and depending on the input values the output may rise to Vdd or remains same. So there is only a partial transition which increases the speed of operation. It has the advantage of domino type cascaded stages. So FTL speed is high because it reduces both low to high and high to low propagation delays. The below circuit is implemented in such a way that A is given as source to B transistor or vice versa.

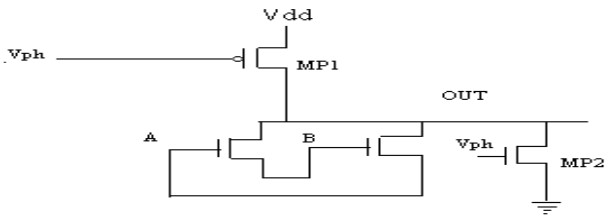


Fig. 5. FTL Based XNOR Gate

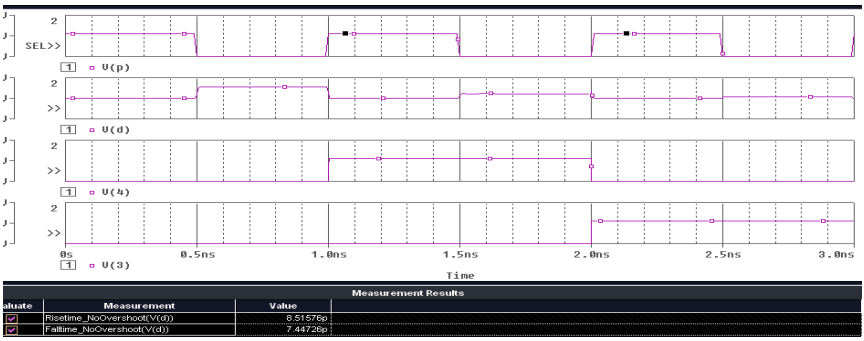


Fig. 6. Waveforms of FTL Based XNOR Gate

As shown in the above figure the circuit is operated with V p as clock during the low clock cycle (evaluation period) circuit may produce the logical levels Vdd and Vdd/2. Here Vdd/2 is treated as low logic level and Vdd is treated as high logic level. This logic is suitable for biomedical applications where compact circuit with high speed is required.

The circuit is tested for both CNFET and 50nm Technology. The main advantage of FTL based XNOR gate is it always produce Strong logic 1. The circuit is tested for several combinations and functionalities are verified. The following comparison table shows the significant improvement of FTL based circuits over 4-Transistor XOR.

Table 2. Comparison of XNOR gate using 50nm and CNFET Technology

TECHNOLOGY	RISE TIME(ps)	FALL TIME
FTL (50nm)	8.515	7.447
4-T XNOR(50nm)	12.35	11.6
FTL CNFET	6.21	5.34
4-T XNOR CNFET	7.1	6.3

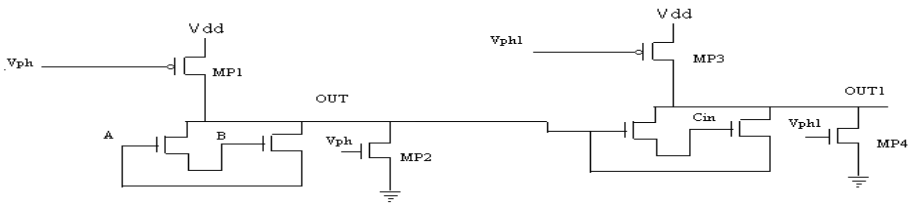


Fig. 7. FTL based FULL ADDER

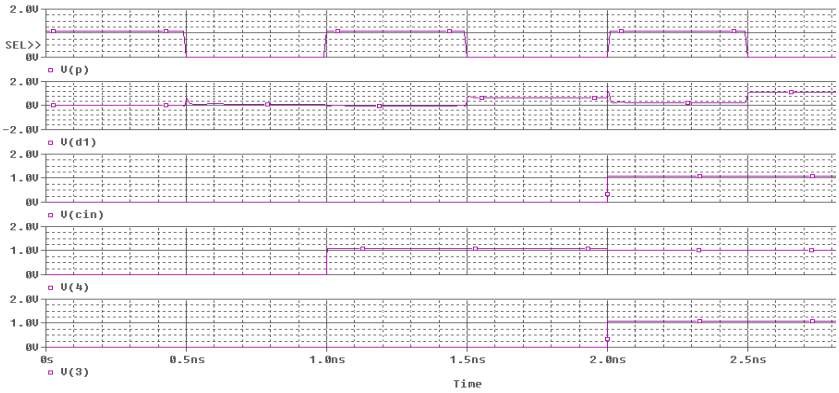


Fig. 8. Waveforms of FTL based FULL ADDER

4 Conclusion

In this paper, we have compared the performance of FTL based XNOR gates and 4-Transistor XNOR gate in 50nm and CNFET technologies. The results are compared first with 2V operation and then with 1V operation. FTL based circuit's shows significant performance improvement over the 4T Xnor gates. The results are verified for full adder also. The maximum propagation delay observed for FTL based circuis observed is 6.31ps.

References

1. Bozorgzadeh, B., Zhian-Tabasy, E., Afzali-Kusha, A.: Low-Power High-Performance Logic Style for Low-Voltage CMOS Technologies. Nanoelectronics Center of Excellence, School of Electrical and Computer Engineering, University of Tehran,Tehran, Iran
2. Wei, Y., Shen, J.-Z.: Design of a novel low power 8-transistor 1-bit full adder cell (Department of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China)
3. Lundstrom, M.: Device Physics of Carbon Nanotube FETs. Network for Computational Nanotechnology and Purdue University 465 Northwestern Avenue West Lafayette, IN 47907
4. Bagherizadeh, M., Eshghi, M.: Two novel low-power and high-speed dynamic carbon nanotube full-adder cells

Reduction of Data Size in Intrusion Domain Using Modified Simulated Annealing Fuzzy Clustering Algorithm

Nandita Sengupta[1], Amit Srivastava[2], and Jaya Sil[3]

[1] University College of Bahrain, Manama, Bahrain
[2] IBM India Pvt Ltd, Bangalore, India
[3] Bengal Engineering and Science University Shibpur, Howrah, WB, India
ngupta@ucb.edu.bh, amit.aot02@yahoo.com, js@cs.becs.ac.in

Abstract. Network security is becoming an important issue as the size and application of the network is exponentially increasing worldwide. Performance of Intrusion Detection System (IDS) is greatly depends on the size of data and a systematic approach to handling such data. In the paper, modified simulated annealing fuzzy clustering (SAFC) algorithm has been proposed using the concept of Rough set theory that removes randomness of the SAFC algorithm and applied on intrusion domain for data size reduction. The reduced data set increases classification accuracy in detecting network data set as 'anomaly' or 'normal' compared to the original data set. Davies-Bouldin (DB) validity Index is evaluated to measure the performance of the proposed IDS.

Keywords: Intrusion Detection System, Simulated Annealing Fuzzy Clustering algorithm, Rough set theory.

1 Introduction

The NSL KDD Dataset consists of 42 attributes and 11,850 instances. To work with such large Knowledge database, system complexity as well as computational complexity increases steadily. Hence data reduction is the obvious choice that eliminates unimportant and redundant instances of network data set for detecting the data as 'normal' or 'anomaly'. In the paper modified simulated annealing fuzzy clustering (SAFC) algorithm [1] has been proposed using the concept of rough set theory that reduces data size by removing randomness of the SAFC algorithm.

The limitations of fuzzy *c*-means clustering algorithm (FCM) [2] has been tackled in the SAFC algorithm using simulated annealing (SA) method [3]. In the SAFC algorithm, a configuration or state of the energy level is encoded as the centres of a variable number of fuzzy clusters. The Xie and Beni (XB) index [4] of the corresponding partition is used to measure the energy level of different states. Since the number of clusters is variable, the searching technique applied on different dimensions creating different number of clusters until the algorithm converges. However, the algorithm perturbs the current centres (current configuration) by calling three functions namely, perturb_centre, split_center and delete_centre, which are

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 97–102, 2013.

randomly chosen at each perturbation resulting different partitions of data depending on the random number. The SAFC algorithm was extended by Xiao Ying Wang et. al. [5] for applying in gene expression data set. In the paper modified SAFC algorithm has been proposed where the instances are clustered with respect to each attribute. Among the clusters corresponding to a particular attribute, most significant cluster is selected using attribute dependency concept of RST. The selected cluster contains features essential to classify the instances based on that attribute. For further reducing the data size, redundant or overlapped instances are removed from the most significant clusters. The remaining instances show comparable classification accuracy with that of the complete data set using DB index value [6].

The paper is divided into five sections. Section 2 describes rough set preliminaries required in the paper and section 3 explains the Modified SAFC algorithm. Results are listed in section 4 while conclusions are arrived at section 5.

2 Rough Set Preliminaries

In rough set theory (RST) an information system is represented as a pair $I = (U, A)$ where U is the universe of objects, A is a nonempty finite set of attributes such that $a:U \rightarrow V$ for every $a \in A$ and V_a is the set of values of attribute a. With any $P \subseteq A$. An associated equivalence relation, called indiscernibility relation: $\mathrm{IND}(P)=\{(x,y)\in U^2 \mid a\in P,\ a(x)=a(y)\}$ for which two objects (x and y) are equivalent. Due to lack of information, objects cannot be treated individually but as a granule consisting of equivalent objects. The family of all equivalence classes of $I(P)$, i.e., the partition is determined by P and denoted by $U/I(P)$. The block of the partition $U/I(P)$, containing x is denoted by the notation: $[x]_P$. RS is an approximation of crisp set where lower and upper approximations are two crisp sets used to deal with the vagueness of information table. The lower approximation of the set of interest X with respect to the attribute subset say, E is defined as $\underline{E}X=\{x\in U \mid [x]_E \subseteq X\}$ representing the objects of X which are certainly belong to the subset E. Upper approximation is a description of the objects that possibly belong to the subset E and defined as $\bar{E}X=\{x\in U \mid [x]_E \cap X\neq\phi\}$. The boundary region is the difference between the upper and lower approximation. The set X is called *rough* if the boundary region is nonempty, otherwise crisp. The information table partitions A into two classes C and Q where $Q \subseteq A$ of attributes, called condition and decision (action) attributes, respectively. The tuple $I = (A,C,Q)$ is called a decision system. For data analysis another important concept of RST is to find out dependencies between the attributes. If a set of attributes Q depends totally on a set of attributes P ($P \subseteq A$), dependency $\gamma_P(Q)$ is defined with a degree $k = \gamma_P(Q) = \sum_i \bar{P}Q_i / |U|$ where $i = 1 \ldots N$, number of values of attribute Q.

3 Modified SAFC Algorithm

In the algorithm, T represents current temperature while *max_T* and *min_T* are maximum and minimum temperatures respectively. At each temperature cooling rate

r is used to decrease the temperature. The parameters of the algorithm are set as: max_T=100, min_T=10.5, k = 40 and r = 0.9, determined through experimentations.

For each attribute the instances are clustered using FCM algorithm with variable c. As a next step, the attribute dependency concept of rough set theory [7] is applied on the clusters to select the important and unique instances. The most significant cluster (MSC) with respect to each attribute having highest attribute dependency value is chosen and corresponding objects are considered for detecting inclusion in the reduced network data set.

Algorithm: *MODIFIED_SAFC*

1. Randomly initialise the number of clusters to c where 2 <=c<=n and n is the number of data points.
2. Initialise data points with fuzzy membership μ_{ij}, where i=1to n and j= 1 to c
3. Calculate cluster centres $v_j = \frac{\sum_{i=1}^{n}(u_{ij})^m * x_i}{\sum_{i=1}^{n}(u_{ij})^m}$
4. Calculate current XB value (S) : $S = \frac{\sum_i \sum_j u_{ij} \| x_i - v_j \|^2}{n * \min_{ij} \| x_i - v_j \|^2}$ (1)
5. Calculate attribute dependency of decision attribute with respect to each attribute considering complete dataset using c-means algorithm and stored in an array.
6. Calculate mean (m) and standard deviation (std) of the array elements and set $std_{min} = m - std$ and $std_{avg} = m + std$
7. T=max_T, setting maximum temperature as starting temperature.
8. If T > min_T , for i = 1 to k
9. Cluster the objects with respect to each attribute considering initial number of clusters, supplied a priori
10. Calculate attribute dependency of decision attribute considering objects in each cluster.
11. Calculate mean (m) and standard deviation (std) of the attribute dependency value of each cluster
12. If std<std_{min} call *preturb_centre*, else if std<std_{avg} call *split_centre*, Else call *delete_centre*.
13. Calculate new XB (XB_{new}) value using (1) based on the perturbation
14. If new XB (XB_{new}) < current XB ($XB_{current}$), accept new centre. Set XB_{new} to $XB_{current}$, save best XB and best centres position.
15. If $XB_{new} > XB_{current}$, accept new centre with probability: $exp(-(XB_{new}-XB_{current})/T)$.
16. $T = rT$, $0 < r < 1$
17. $i = i + 1$. If $i > k$, stop, Else go to step 8.
18. Return the best XB and best centre positions.

The instances or objects appearing in different MSC with a finite number of times is determined and based on a threshold value objects are selected for forming the network intrusion data set. The algorithm *reduced_dataset* is described below.

Algorithm: *reduced_dataset*
Input: Clusters with respect to each attribute
Output: Reduced number of objects
Begin
for (*j*=1 to *no_of*_objects)
Initialize counter[*j*] ;
for (*i*=1 to *no_of_attributes*)
{
1. Modified_SAFC();
2. Calculate attribute dependency for each cluster and select the cluster with highest attribute dependency value, called most significant cluster (MSC)
3. For each objects '*j*' in MSC, increment the respective counter[*j*].

/**********Setting up the threshold (*Th*) **********/
sum =0 ;
for (*p*=1 to *no*_of *_objects*)
{
sum = *sum* + counter[*p*]
}
Th=*sum*/ *no*_of *_objects*

/**selection of objects based on the respective content of counter**/
for (*i*=1 to *no*_of *_objects*)
{
If (counter[*i*]>*Th*)
select corresponding object
}

4 Experimental Results

This algorithm was tested on network data NSL KDD data set [8]. The initial and final number of clusters is shown in table 1. The DB index [6] for initial and correspondingly final clusters obtained are shown below in the table1.

Table 1. Comparison of MODIFIED_SAFC algorithm and SAFC algorithm

Initial Number of Cluster	Initial DB Index	Final Number of Clusters	Final DB Index	Final No. of clusters in SAFC Algorithm	DB Index in SAFC Algorithm
7	000.397	2	0.015	3	0.022
13	002.963	3	0.195	3	0.195
25	008.695	3	0.198	3	0.182
45	304.086	3	0.195	2	0.195
50	876.928	3	3.042	3	3.269

The network data set consists of 41 conditional attributes and 1 decision attribute. There are 11850 objects. The modified SAFC algorithm is run on the reduct. The number of objects in Most Significant Cluster (MSC) to be included in reduced dataset for each attribute is given in table 2. The classification accuracy of complete dataset is 64.64%. The reduced dataset consists of only 7182 objects and its classification accuracy is 79.88% as shown in figure 1.

Table 2. Number of Significant Objects for each Attributes

Attributes	Final Number of Clusters	Reduced Instance in MSC
1	4	476
2	4	4597
3	6	89
4	3	632
5	6	494
6	2	1974
7	2	5175
8	2	5901
9	2	2705

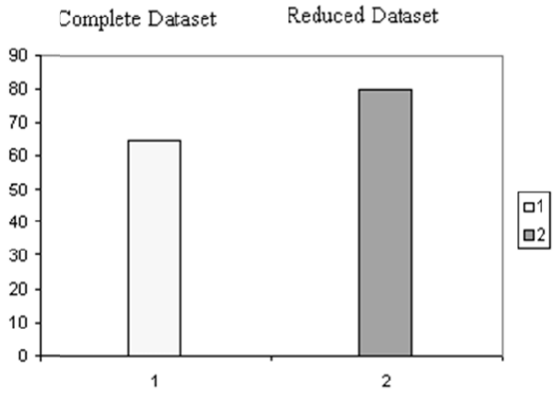


Fig. 1. Comparison of Classification Accuracy of Complete Dataset and Reduced Dataset

5 Conclusions

The concept of rough set theory and simulated annealing are used in the paper to develop the modified SAFC algorithm, which do not require any additional information other than the network data. Here, clusters are generated with respect to each attribute that preserves profiles of traffic data and at the same time reduces size of the data. The limitations of FCM clustering algorithm, generation of sub-optimal solution and finally drawback of randomness in preparing clusters are removed in the proposed method. It is clear from fig.1 that accuracy increases when data set is reduced, and consequently computation time becomes less.

References

1. Bandyopadhyay, S.: Simulated Annealing Using a Reversible Jump Markov Chain Monte Carlo Algorithm for Fuzzy Clustering. IEEE Transactions on Knowledge and Data Engineering 17(4) (April 2005)
2. Zhu, L., Chung, F.L., Wang, S.: Generalized Fuzzy C-Means Clustering Algorithm With Improved Fuzzy Partitions. IEEE Transactions on Systems, Man and Cybernetics 39(3), 578–591 (2009)
3. Kirkpatrik, S., Gelatt, C., Vecchi, M.: Optimization by Simulated Annealing. Science 220, 671–680 (1983)
4. Xie, X.L., Beni, G.: A validity measure for fuzzy clustering. IEEE Trans. Pattern Analysis and Machine Intelligence 13, 841–847 (1991)
5. Wang, X.Y., Whitwell, G., Garibaldi, J.M.: The Application Of A Simulated Annealing Fuzzy Clustering Algorithm For Cancer Diagnosis. In: IEEE 4th International Conference on Intelligent Systems Design and Application, Budapest, Hungary, August 26-28, pp. 467–472 (2004)
6. Davies, D.L., Bouldin, D.W.: A cluster separation measure. IEEE Transactions on Pattern Recognition and Machine Intelligence 1(2), 224–227 (1979)
7. Guo, P., Tanaka, H.: Upper and lower possibility distributions with rough set concepts. In: Inuiguchi, M., Hirano, S., Tsumoto, S. (eds.) Rough Set Theory and Granular Computing. LNCS, pp. 243–250. Springer (2002)
8. The NSL-KDD Data Set, `http://iscx.ca/NSL-KDD/`

Using Limited Flooding in On-Demand Distance Vector Junior for Reduction Power Consumption in ZigBee Networks

Arman Zare, Hasan Taheri, and Meisam Nesary Moghaddam

Electrical Engineering Faculty, Amirkabir University of Technology, Tehran, Iran
{arman.zare,htaheri,nesari}@aut.ac.ir

Abstract. ZigBee network uses the simplified version of on-demand distance vector (AODV) routing which is named AODV junior (AODVjr). There are three main differences between it and AODV. It does not have sequence number of destination and eliminates the sending error packet when a node crashes. The third difference is preventing form sending hello messages periodically for detecting the active nodes and use connect message that it sends by destination node. The other features are the same in both of them. One of them is the broadcasting of Route Request (RREQ) in the network for finding the shortest path. It causes that high energy is consumed in the network. As a result, ZigBee nodes that are supplied by batteries will die early. Therefore, an AODVjr based on Limited Flooding (FLAODVjr) in the network is suggested in this paper. This approach uses the information of tree routing algorithm for limiting the broadcasting of RREQ packets. The results of simulations prove the improvement of using this method.

Keywords: ZigBee, Limited Flooding, AODVjr, Power Consumption.

1 Introduction

ZigBee network is a wireless network technology which provides low data rate, low power consumption, low cost and reliable communications. It is used in a variety of applications due to these features. It is one of the best choices for many networks that they do not need to high data and have restrictions in the accessing in power supplies such as automation and monitoring applications [1].

The Physical (PHY) and Medium Access Control (MAC) layers of ZigBee are based on IEEE 802.15.4 [2]. It is developed by ZigBee Alliance [3]. This technology operates in unlicensed frequency range of 868MHz, 915MHz and 2.4GHz with data rate 20-250kbps and also has a coverage range of 10-100m [2].

The nodes are divided into two types that include Full Function Device (FFD) and Reduced Function Device (RFD). A FFD node can act as a coordinator, router and end-device. In fact, it can form a network and specify address to joint nodes, if it acts as a coordinator. Also, it enables to route the packets and plays the role of end-device, namely, send its information to parent node. Another type of node, RFD, can just be

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 103–108, 2013.

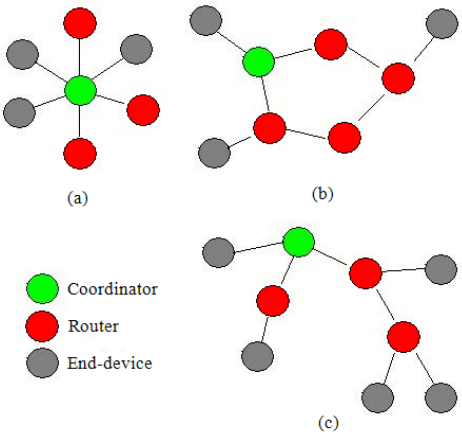


Fig. 1. Topologies of ZigBee Technology: (a). Star, (b). Mesh and (C). Tree.

end-device in the network. From topological point of view, ZigBee has three topologies including star, tree and mesh that they are shown in figure 1. In the star topology, there are a coordinator and several other nodes which are connected to it by a hop. Tree is constituted based on parents-child relationships. There are a coordinator as a root of tree and other nodes that join as child to it. If a children node is a router, it can accept some child under certain circumstances. The last topology is mesh. This topology has a coordinator and some routers and end-devices. The FFD nodes can communicate with others, if they are in the coverage range of together. Thereby, there are maybe some paths between source and destination [3].

For finding shortest path, ZigBee uses routing algorithm based on AODV [4]. It can find the appropriate route by the use of broadcasting Route Request (RREQ). The destination node may receive RREQ packet from several paths. Then, it sends Route Reply (RREP) through shortest path. Also, this algorithm sends hello message periodically for detecting active nodes. This feature consumes energy that is not suitable for ZigBee technology. Therefore, the use of AODVjr [5] is suggested in [3]. The main difference between them is that sending hello message and precursor list are eliminated in AODVjr and just destination node transmits connect message. It is for understanding the situation of a route. If source node receives this packet, it can send their packet from that route. This difference reduces the number of sent packets and power consumption but AODVjr is similar to AODV in broadcasting RREQ packet that it consumes a high level of energy. Thus, finding a suitable method for reduction in sending routing packets is very important that is suggested in this paper. It use information of tree routing in AODVjr.

The paper is organized as follows: section 2 introduces the overview of ZigBee routing protocol, section 3 is about the suggested routing method, section 4 presents the simulations of the discussed issues and the results of them; and the last section concludes this paper.

2 Overview of ZigBee Routing Protocols

ZigBee technology has two main routing protocols. One of them is tree routing and another is AODV that AODVjr is introduced in this paper rather than it due to the above discussed reason. These approaches are explained as follow.

2.1 Tree Routing

Tree routing algorithm is based on addressing scheme that is named "Cskip". It assigns an address to each joint node to the network. It operates based on three parameters including the maximum number of router children of a parent R_m, the maximum depth of network L_m and the maximum number of children of a parent C_m. The Addresses are calculated by Cskip(d) and assigned to nodes by their parents. It is calculated from below formula:

$$Cskip(d) = \begin{cases} 1 + C_m.(L_m - d - 1) & R_m = 1 \\ \dfrac{1 + C_m - R_m - C_m.R_m^{\,L_m - d - 1}}{1 - R_m} & R_m \neq 1 \end{cases} \quad (1)$$

In this formula, d is the depth of a router node in the network. If Cskip(d) is 0, the router node cannot accept other nodes as child of it. A router node can distribute addresses to its child, if its Cskip(d) is greater and equal than 1. The address of the first node that joins to a node is one more than its parent node and the next router children node will have an address that has Cskip(d) difference with previous router children node. If the address of parent node is supposed Ap, the address of its n^{th} router child Arp will be:

$$A_n^{\,r} = A_p + Cskip(d) \times (n-1) + 1$$
$$1 \le n \le R_m \quad (2)$$

and the address of its m^{th} end-devices will be:

$$A_n^{\,e} = A_p + Cskip(d) \times R_m + m$$
$$1 \le m \le C_m - R_m \quad (3)$$

After the assigning address to all the nodes, tree routing algorithm uses it and parent-children relationships between nodes for sending packets. When a node receives a packet, it must send it based on destination address to its child or parent. If the address of destination is between the address of its child, it will send that packet to it else it will send it to its parent.

2.2 AODVjr

This algorithm uses broadcasting routing packets for finding a path between source and destination as it was said before. When destination node receives RREQ, it sends RREP to source node for creating a suitable node between them that is shown in figure 2. For understanding that the path is active, destination transmits connect message to source periodically.

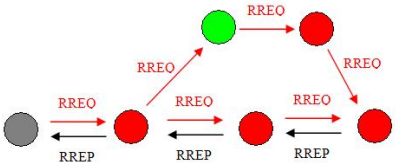


Fig. 2. Route Discovery in AODVjr

3 Using Limited Flooding in AODVjr

Broadcasting the RREQ packet in the network consumes a high level of power and in the ZigBee technology which uses batteries as a power supply generally; it reduces the lifetime of the network.

For decreasing the number of sent routing packets in the network, a method that has suggested in [6] is applied to AODVjr. In this method, the information of tree routing is used for limiting the flooding of RREQ packet. This packet will be sent by nodes, if the amount of hop count is not 0. When each node transmits it, their amounts of hop count decrement one unit. Therefore, we can use it for limitation of broadcasting RREQ packet. This method is named Flooding Limitation AODVjr (FLAODVjr).

Each node has a unique address based on Cskip addressing scheme. When a node wants to send its information to a destination, it set the amount of hop count based on information of tree routing. In the tree routing, when node A wants to send packet to node B, the packet send to the first common ancestor of node A and B. Then, it sends the packet to node B. Hence, the number of hops between them H_t is equal with the sum of the number of hops from A until the first ancestor and between that ancestor and B. H_t can be calculated by finding the difference of depth A and B with the first common ancestor.

For finding the depth of nodes, if the address of sender is 0, it is coordinator with depth 0. If it is greater than 0, will start searching in the depth 0. If the parent of node is in that depth, its depth will be 1 else, will continue searching in next depth until L_m. when a node can find the depth of its parent d_p, its depth is d_p+1. For finding the depth of first common ancestor, it starts searching from d=0 until L_m-1 (because, nodes in the last depth is leaves of tree and cannot be parents of other nodes). At first, searching starts in the child of depth 0, namely, in the router child of the coordinator. It supposes the address of the first common ancestor is 1. If the address of A and B

are in the range of 1 and 1+Cskip(0), it will start searching in the router child of it and this trend will use for them. If they do not have common ancestor in the child of the first found ancestor, the depth of it is the depth of the first common ancestor. In the same depth, searching happens in router child and is started with one of them that it has smallest address.

After calculating the H_t and setting the hop count equal to it, destination node receives RREQ packet. Sometimes the optimal hop count is not H_t. Because, in tree routing, it is possible that source and destination are in the coverage range of together or destination can receive packets, if it is in the coverage range of a router node that is not its parent but it can ensure that packets receive to their destinations. It can reduce the number of sent RREQ packet. As a consequence, the amount of consumed energy decreases.

4 Simulations and Results

For simulation, OPNET [7] simulator is used. In this case, area is considered 100m×100m and operational frequency is 2.4GHz with 250Kbps. Also, the number of nodes increases up to 200 and the all amounts of Rm, Cm and Lm are the same, namely, 7. Also, achieving to comparison of the amounts of sent overhead is the goal of experiments in the AODVjr and FLAODVjr methods when the destination node is random. Reduction in mentioned parameter causes reduction in power consumption.

Figure 3 demonstrates the amount of communication overhead of both approaches. In this case, FLAODVjr sends communication overhead less than another. This difference rises with increasing the number of node in the network. The high number of nodes causes the rising in the depth of the network. It increases the number of hops between source and destination. Thereby, sent routing packets or communication overhead will climb.

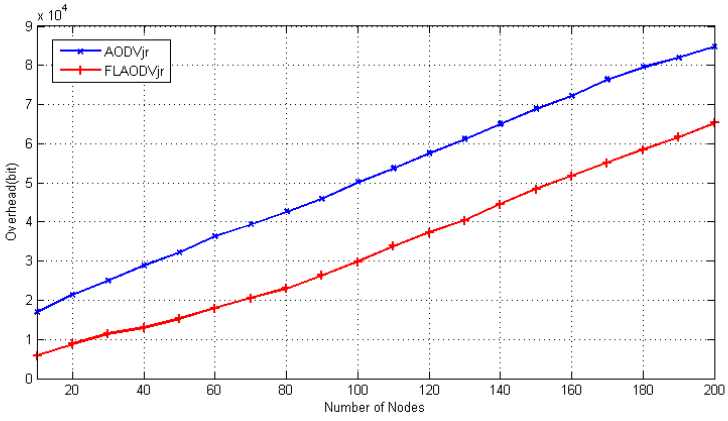


Fig. 3. Comparison of Overhead in AODVjr and FLAODVjr

5 Conclusion

In this paper, a method is suggested for reducing in power consumption of ZigBee Technology. It limits the broadcasting routing packets that use in AODV routing algorithm by the use of tree routing's algorithm. The amount of hop count is set based on counterpart parameter in tree algorithm. Doing this work decreases the number of sent packets for finding the path between source and destination. Thereby, the consumed energy for sending mentioned packets is reduced and it helps to increase the lifetime of the network. Also, the result of simulation proves the above discussed issues.

References

1. Lin, S.: ZigBee Based Wireless Sensor Networks and Its Applications in Industrial. In: IEEE International Conference on Automation and Logistics, pp. 1979–1983. IEEE Xplore, Jinan (2007)
2. IEEE Standard 802.15.: Standard for Information Technology-telecommunications and information exchange Systems between systems - Local and metropolitan area networks (2006), http://www.standards.ieee.org/
3. Li, J., Zhu, X., Tang, N., Sui, J.: Study on ZigBee Network Architecture and Routing Algorithm. In: International Conference on Signal Processing Systems, pp. 389–393. IEEE Xplore, Dalian (2010)
4. Perkins, C.E., Royer, E.M.: Ad Hoc On-Demand Distance Vector Routing. In: Second IEEE Workshop Mobile Computing Systems and Applications, pp. 90–100. IEEE Xplore, New Orleans (1999)
5. Chakeres, I.D., Klein-Berndt, L.: AODVjr, AODV Simplified. ACM Mobile Computing and Communications Review, 100–101 (2002)
6. Lin, Z., Meng, Q.H., Liang, H.: A Route Discovery Method Based on Limited Flooding in ZigBee Network. In: IEEE International Conference on Automation and Logistics, Qingdao, pp. 3039–3044 (2008)
7. OPNET Network Simulator Software, http://www.opnet.com

Optimization for Agent Path Finding in Soccer 2D Simulation

Amir Tavafi[1], Narges Majidi[1], Michael Shaghelani[2], and Amir Seyed Danesh[3]

[1] University of Guilan, Rasht, Iran
amir23t@gmail.com,
azin_majidi@yahoo.com
[2] Payame Noor University of Rasht, Iran
michael.sh90@yahoo.com
[3] Department of Software Engineering
Faculty of Computer Science and Information Technology
University of Malaya, 50303, Kuala Lumpur, Malaysia
amir_s_d@siswa.um.edu.my

Abstract. This paper describes an optimized algorithm for a simulated agent in the RoboCup Soccer 2D Simulation environment to find a better way dribbling with ball and simultaneously avoiding opponents who are blocking our agent's way to the goal. At first an optimized algorithm for finding the best way for each cycle is introduced, which gives us an optimum direction toward the goal, avoiding the opponent agents. Then this algorithm is improved using a Reinforcement Learning (RL) method. The experimental results using Soccer 2D Simulator shows the improvement of the behavior of the agent in this multi-agent system.

Keywords: Soccer 2D Simulation, Reinforcement Learning, Multi-Agent System, Path Finding, Dribbling.

1 Introduction

In RoboCup Soccer 2D Simulation which is a multi-agent system, the goal of each team is to play a realistic soccer game and win the game against the opponent team. Each team has 11 agents and each agent has to play its role in the field and try to lead the ball to the offense part of the team and score goals [3]. In order to achieve this goal, each agent have different situations e.g. with ball situation or without ball situation, and for each situation it can make different decisions e.g. passing the ball or dribbling with ball. In Fig.1 a simple decision tree for one agent in a simulated soccer field is shown.

When the game is in play on mode, most of the times the ball owner decides to dribble with ball unless an opponent is so near that it can't continue dribbling or a high recommended pass e.g. a through pass is available [1]. This shows that dribbling is so effective in the gameplay of a team. In this skill agent should find a target to start or continue its movement with ball to create a good attacking situation [6].

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 109–114, 2013.

This paper introduces an optimized algorithm to find the optimum target for the current state of the environment which is a soccer field including 11 teammate agents and 11 opponents as in a realistic soccer game, and then this algorithm is improved with RL algorithm and the experimental result is shown. It is organized as follows: section 2 briefly describes the RL algorithm. Section 3 describes the optimized static algorithm used for finding an optimum target regarding the current state without any further knowledge from last actions and decisions. Section 4 describes adding the RL algorithm to this method in order to improve the results for dribbling skill and shows the results after using this method in Soccer 2D Simulation environment. Finally, Section 5 summarizes some important points learned from this research and outlines future work.

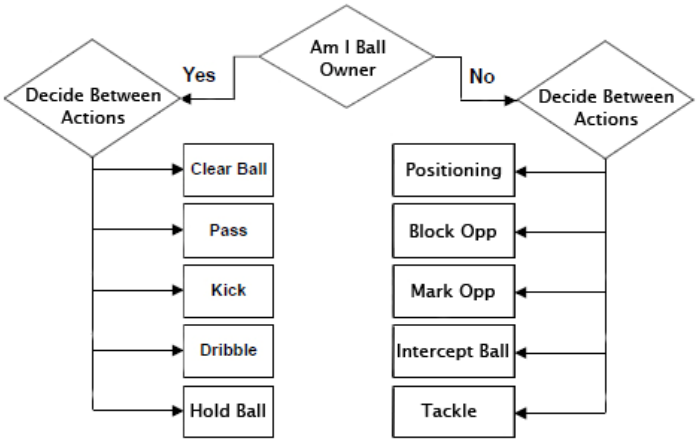


Fig. 1. A simple decision tree for one agent

2 Reinforcement Learning Algorithm

Reinforcement Learning is one of the studies in Machine Learning field that is concerned with an autonomous agent interacting with its environment via perception and action [4]. The basic Reinforcement Learning model consists of:

1. A set of environment states;
2. A set of actions;
3. Rules of transitioning between states;
4. Rules that determine the scalar immediate reward of a transition;
5. Rules that describe what the agent observes.

The agent senses the current state of the environment in each step of the interaction, and decides an action to execute. The action alters the state of the environment, and the agent receives a scalar reinforcement signal to indicate the admissibility of the next state. In this way, "The RL problem is meant to be a straightforward framing of the problem of learning from interaction to achieve a goal" [2].

The task of each RL agent is to find out a static plan of actions that maps the current state of the environment into an optimal action(s) to be performed in the current state, maximizing the expected long term sum of values of the reinforcement signal, from any starting state [5]. The RL learner agent must expressly explore its environment to find the best decision in each state to perform the best possible action in its environment and remember its previous experiences to improve its future decision making.

3 The Optimized Static Path Finding Algorithm for Ball Owner

This algorithm is planned for the ball owner to find an optimum path to dribble with ball and create a better situation to pass the ball or create a dangerous situation in front of opponent's goal to score a goal. In this algorithm, the agent is making a decision based on the current situation of the environment without any further knowledge about its last experiences.

In every different part of the field the agent must have a default dribble target which is not affected by opponent agents yet and then the algorithm will change the path in order to achieve the best effective movement. A simple default target for each part of the field is shown in Fig. 2.

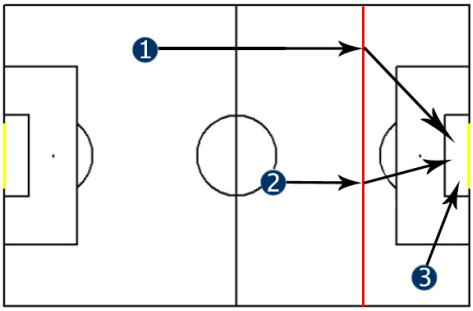


Fig. 2. Default dribble target for an agent in the soccer field

If ball owner is before the red line, dribble target will be set to a point in front of the agent toward the X dimension only, but after the red line where the agent is near the opponent's goal, dribble target will be set to the center of the goal line in order to carry the ball to a position that a shoot may be available. In dribble skill, the ball owner agent at last executes a dribble to a particular direction which is calculated by the algorithm.

After setting the default target, a comparative factor d is defined. This factor is the main part of the algorithm which affects the final direction of dribbling and is associated with the distance of the opponent to the ball owner agent. It is calculated with a comparative ratio. This d should be converted from distance to an angle with a

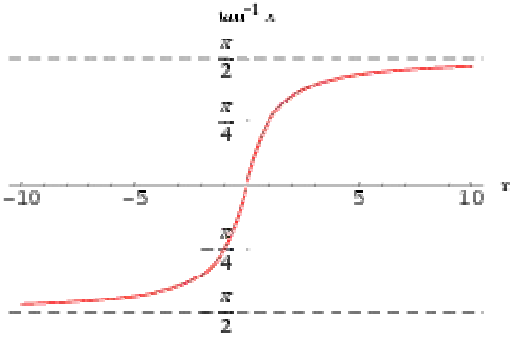


Fig. 3. The graph of inverse tangent function

high sensitivity for low distances and less sensitivity for far distances so the agent can have a good reflex while the opponent moves near him and to have less consideration while the opponent moves in far distances e.g. when opponent's distance from agent changes from 4 meters to 2 meters, it should affect more on the dribble direction than a change from 15 meters to 13 meters. The inverse tangent function is appropriate for this situation. As shown is Fig. 3, the inverse tangent function has more sensitivity for low values and low sensitivity for high values in both minus and positive values.

After *d* is converted to the desired angle with the suitable function, the range of *d* should be changed to match with all possible movements for agent.

The agent should do this process for every single opponent to find one effective direction for each opponent and at last add the average of these calculated directions to the default dribble direction which was created from the defined dribble targets. The algorithm for this process is shown in Table 1.

Table 1. The optimized path finding algorithm

Initialize the environment's current situation *s*.
Define a *modifier* with the value 0.
Repeat for every opponent:
Declare *difference* with value of the difference of angle between these two vectors:
- current opponent's position to ball
- agent's dribble target to ball
Define *d*, the distance between the opponent and the ball owner.
Normalize the value of *d*.
Convert the type of *d* with inverse tangent function.
Normalize the range of *d*.
Update the value of modifier according to:
modifier ← *modifier* + *difference*d.*
Until next seen opponent is not checked.
Update the final dribble direction according to:
Final direction ← *default direction* + *(modifier / number of checked opponents)*

4 Adding Reinforcement Learning to Improve the Results

The algorithm discussed in Section 3 can be improved by using the experiences of agent from its last actions and percepts. In order to optimize the dribble routes, the RL algorithm can be used. There are some comparative factors in the algorithm used to normalize the range of the final direction which are some integers that had needed to be tuned for an acceptable result. Without a learning algorithm it is difficult to find the best ratio. Best rates in the current algorithm are not the same against different team with different strategies and block algorithms.

The agent should derive a better ratio after each cycle of this algorithm's execution by investigating the start and finish position of the dribble action. Every time agent starts to dribble, it saves the information of the beginning point and continues its action until it decides not to execute the dribble skill according to its decision tree. At last the dribble route is compared to other possible routes that might have been chosen if the ratio was different from the current value. If the agent thinks that a better route is available for the next dribble action, it would change the ratio to improve the result of dribbling. To compare different situations and actions, the success percent *SP* is used which is calculated from current dribble action's start position distance to goal *(s1)* and dribble's finish position distance to goal(*s2*):

$$SP = (s1 - s2) / s1. \quad (1)$$

After using this algorithm and improving it with the RL algorithm, the experimental results in the simulated 2D soccer field showed improvement of the RL learner's behavior and its effect on team's performance. A comparative graph of the performance for 3 different tests is shown in Fig. 4. A simple agent is an agent which dribbles with ball toward the opponent team's goal without considering opponent agents. Success percent of a simple agent is also shown in the following figure.

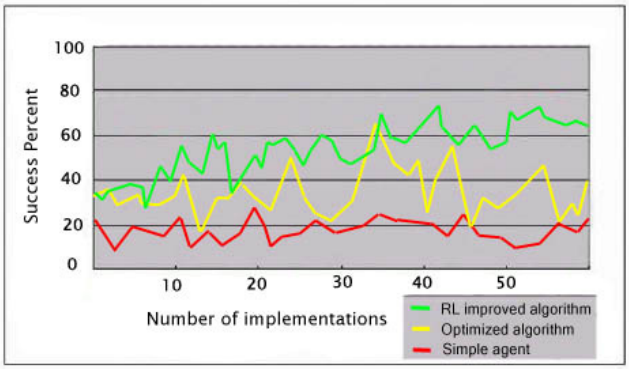


Fig. 4. Results of different algorithms against static opponent agents in Soccer 2D Simulation domain

5 Conclusions and Future Work

In this paper an optimized algorithm for an agent in a Soccer 2D Simulator environment was presented. The presented algorithm had better results and improved the performance of a simple agent dribbling toward the opponent team's goal without considering opponent agents. Although there are different algorithms which consider opponent agents and affect them on the dribbling path for agent, the presented algorithm has a higher performance than other implemented algorithms and gives a satisfying result with high flexibility which makes it easier to improve this existing algorithm with learning methods such as Reinforcement Learning. Results of adding the RL algorithm to the optimized algorithm showed improvement in agent's behavior.

Future works include adding other well-known learning methods for the simulated agent and also extending this algorithm to other complex domains, such as Soccer 3D Simulation and Small Size League robots.

References

1. Tavafi, A., Nozari, N., Vatani, R., Rad Yousefi, M., Rahmatinia, S., Piredeyr, P.: MarliK 2011 Team Description Paper. In: RoboCup 2011 Symposium and Competitions, Turkey (2011)
2. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction. MIT Press, Cambridge (1998)
3. de Boer, R., Kok, J.: The Incremental Development of a Synthetic Multi-Agent System: The UvA Trilearn 2001 Robotic Soccer Simulation Team. Master's thesis, University of Amsterdam, The Netherlands (2002)
4. Mitchell, T.: Machine Learning. McGraw Hill, New York (1997)
5. Bianchi, R.A.C., Ribeiro, C.H.C., Costa, A.H.R.: Heuristically Accelerated Q–Learning: A New Approach to Speed Up Reinforcement Learning. In: Bazzan, A.L.C., Labidi, S. (eds.) SBIA 2004. LNCS (LNAI), vol. 3171, pp. 245–254. Springer, Heidelberg (2004)
6. Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach. Prentice-Hall, Upper Saddle River (1995)
7. Akiyama, H., Noda, I.: Multi-agent Positioning Mechanism in the Dynamic Environment. In: Visser, U., Ribeiro, F., Ohashi, T., Dellaert, F. (eds.) RoboCup 2007. LNCS (LNAI), vol. 5001, pp. 377–384. Springer, Heidelberg (2008)
8. Ko, J., Klein, D.J., Fox, D., Hähnel, D.: Gaussian processes and reinforcement learning for identification and control of an autonomous blimp. In: Proceedings of the IEEE International Conference on Robotics & Automation, ICRA (2007b)
9. Crites, R., Barto, A.: Improving elevator performance using reinforcement learning. In: Advances in Neural Information Processing Systems 8 (NIPS 1995), Denver, USA, pp. 1017–1023. MIT Press, Cambridge (1995)
10. Kitano, H., Minoro, A., Kuniyoshi, Y., Noda, I., Osawa, E.: Robocup: A challenge problem for ai. AI Magazine 18(1), 73–85 (1997)

Compressive Sensing for Pulse Position Modulated UWB Signal

D. Abhilash, Akshaya Sankar, S. Janakiraman, G. Ravi Kiran, Roshni Kaur Sudan, and V. Mekaladevi

Amrita School of Engineering, Coimbatore, India
{desh.abhi07,johny.manasee,ricky.rocks}@gmail.com

Abstract. A band-limited signal must be uniformly sampled at lea st twice its bandwidth to guarantee reconstruction at the receiver's end according to Shannon/Nyquist theorem. Compressive sensing suggests that sparsity helps representing the signal in much lesser dimensions. Analog-to-information convertor utilizes this sparsity in order to sample signals at information rate that enhances the efficiency of ADC which otherwise has to sample at very high rates for wideband signals. In this paper, we explore the combination of Compressive Sensing with M-ary PPM Ultra Wideband modulation scheme. The signals are detected without having to estimate the channel. Orthogonal matching pursuit is used as the reconstruction algorithm for the recovery of sparse signal from its compressed measurements.

Keywords: Analog to information converter, Compressive sensing, Orthogonal matching pursuit, Pulse-position modulation.

1 Introduction

Sampling rate is one of the key features of any communications system and its reduction increases the overall performance of the communications system. At the receiver end, a big hurdle is the efficiency of the Analog-to-Digital Converter (ADC). The ADC process is based on the nyquist sampling theorem, which guarantees the reconstruction of a band-limited signal .Growth in applications like ultra-wideband communication are pushing the performance of ADC's. Due to this, Pulse Position Modulation scheme has gained quite an importance in the realization of such systems.

Compressive Sensing (CS) with pulse position modulation yields one step forward in this direction as Pulse position modulated (PPM) signal is sparse in time domain. Also it is used as ultra wideband modulation scheme as PPM modulated signal is more immune to noise. Furthermore, it has been found that most of the signals with large bandwidth have a small rate of information. This property of wideband signals makes them sparse in information and has led to methods of sampling based on the amount of information. Compressive sensing offers more flexible options to deal with sparse signals in terms of the location of the information and the non-uniformity of measurements. Using this CS theory, an Analog-to-information convertor can be employed that uses random demodulation to compress the signal and using a reconstruction algorithm like the orthogonal matching pursuit and basis pursuit.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 115–121, 2013.

2 Proposed Block Diagram and Brief Discussion

Ultra Wide Band (UWB) operates at base bands and transmits digital data at very high rates with very low power densities. The pulses transmitted are of very short duration (in the order of nanosecond or less) and low duty cycle making this appropriate for high speed short range indoor wireless communications.

Pulse Position Modulation (PPM) is an M-ary orthogonal modulation technique where the data modulates the position of the transmitted pulse within an assigned time period. Advantages of PPM include its simplicity, immunity to noise, non-coherent detection, the ease of controlling delays [5] and the property of sparsity in time domain useful for compressed sensing.

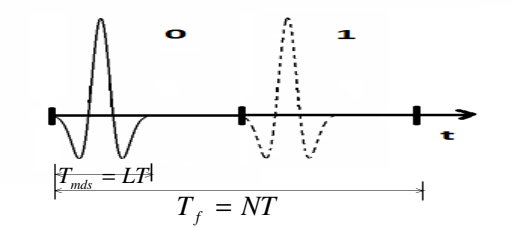


Fig. 1. PPM symbol in UWB

A 0 message bit generates a solid line and a 1 message bit generates a dotted line.

Let x (t) represent the chain of transmitted PPM symbols. The PPM signal is then passed through a pulse shaping filter for UWB communication. The maximum channel delay spread possible without any conflict is $T_{mds} < Tf/2$, so each pulse does not spill over to other pulse positions.

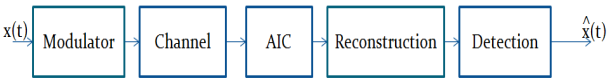


Fig. 2. Generation, reconstruction and detection scheme

This block diagram shows the generation, sub-nyquist sampling and reconstruction and detection process for PPM. Additive white Gaussian noise is added after the channel. It is seen that y measurements are obtained from the AIC block; and from $\hat{s}$ is the reconstructed vector. For reducing the overall system complexity, non-coherent receiver is used. Since the transmitted power is spread over a large bandwidth; this detection causes very low interference to narrowband signals. The detection process is more like a generalized maximum likelihood detector. Non-coherent receivers are more robust to synchronization errors as a slight shift in integration region will not result in significant performance degradation as the receiver is less sensitive to clock jitter.

In the case of 2-PPM, if N is the total number of Nyquist rate samples of the symbol, then the decision rule is based on the following relation,

$$\left\| \hat{S}_{1:\frac{N}{2}} \right\|_2^2 \underset{<}{\geq} \left\| \hat{S}_{N/2+1:N} \right\|_2^2 . \tag{2.1}$$

where $\hat{S}_{i:j}$ denotes the sub-vector of $\hat{s}$ from position i to j with i, j ∈ Z^+.Since the channel effect is limited to the duration of the pulse, multiple paths in fact contribute to a correct decision. Sparsity expresses the idea that the information rate of a continuous time or discrete signal may be much smaller than suggested by its bandwidth [1].

Consider a real-valued, discrete time signal, $x \in R^N$ which can be expressed in orthonormal basis $\psi = [\psi_1, \psi_2, ..., \psi_M]$. Where the vector $s = [s_1, s_2, ..., s_M]$ is a sparse vector having few non-zero components. Using matrix notation it may be represented as $x = \psi s$ .This theory extends to signals that are well approximated with a signal that is K-sparse in ψ . Considering the classical linear measurement model for the above signal,

$$y = \phi x = \phi \psi s = As \tag{2.2}$$

Matrix A is called measurement matrix with rank M, lesser than the rank N of the signal x. It satisfies incoherence and Restricted Isometric Property [1]. To identify the ideal signal s, determine which columns of A participate in the measurement vector y. After each iteration, the column of A that is most strongly correlated with the remaining part of y is chosen. Then subtract its contribution from y and iterate on the residual and after m iterations, the algorithm will identify the correct set of columns.

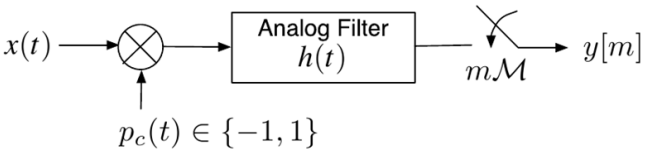


Fig. 3. Block diagram of Random Demodulator

The random demodulator [3] consists of three main components: demodulation, filtering, and uniform sampling. The signal x (t) is multiplied by a chipping sequence, which alternates between -1 and 1 at the Nyquist rate or higher. The chipping sequence is used to create a continuous-time demodulation signal via the formula

$$t \in [\frac{N}{W}, \frac{n+1}{W}) and \ \ n = 0,1,.....,W-1 \, , \, p_c(t) = (\in_1, \in_2, ...\in_n)$$

The altered signal is then sent through an anti-aliasing low- pass filter that acts as an integrator that sums the demodulated signal y(t) for 1/R seconds, which prepares the signal for the ADC. The ADC samples to obtain y_m .This system can be modeled as a discrete vector x operated on by a matrix Φ containing N/M pseudo-random $\pm 1's$ per row. For example, with N = 6 and M = 3 and a sequence p = [-1, 1, -1, -1, 1, -1], over a period of 1 second such a Φ might look like,

$$\phi = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

Such that y= Φ x. The sampling rate M is much lower than the Nyquist rate. M depends primarily on the number K of significant frequencies that participate in the signal in.

3 Simulation Results

The PPM signal generated should be sampled at a Nyquist rate of 1 GHz to retrieve the signal back in an UWB application [5]. Hence the ADC should sample at a frequency of 1GHz and according to the quantization levels the rate of the bits coming will be much faster and that such should be processed by the system at that rate to detect the message bit sent.

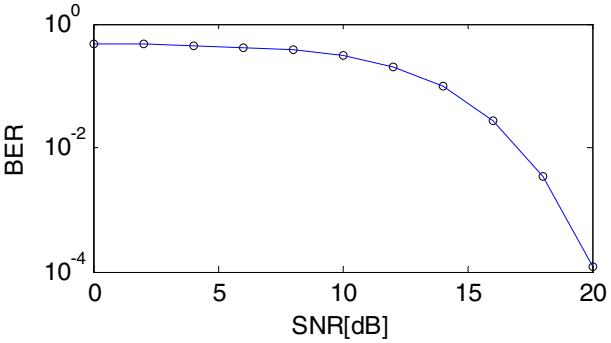


Fig. 4. 2-PPM BER using Random Sampling and reconstructed using OMP algorithm

If the sampling rate is reduced it reduces a lot of stress on ADCs which can be compensated to provide better quantization levels and reducing the stress on the digital systems involved in later stage. The PPM signal is sampled at sub Nyquist rate using random sampling and random demodulation. The above step produces the compressed measurement matrix. The signal is then reconstructed using basis pursuit or orthogonal matching pursuit algorithm.

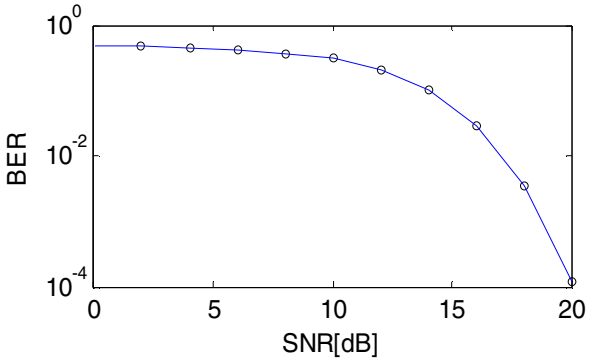


Fig. 5. 2-PPM BER using Random Demodulation and reconstructed using OMP algorithm

The same signal is compressed with a Random demodulator and then reconstructed with OMP algorithm. The results show that in both Random demodulation and random sampling the graph converges in same rate. But the hardware involved in making a Random sampling is very complicated. The hardware of a Random sampling is huge as it involves a parallel bank of ADC's and for each ADC exclusive chipping sequence should be generated and then reconstructed with OMP algorithm [4].On the other hand, a random demodulator just employs only one ADC, one chipping sequence generator and provides the same result as that of a random sampler. Here the hardware is much simpler. By using a random demodulator resource needed, energy used and complications can be reduced to a great extent.

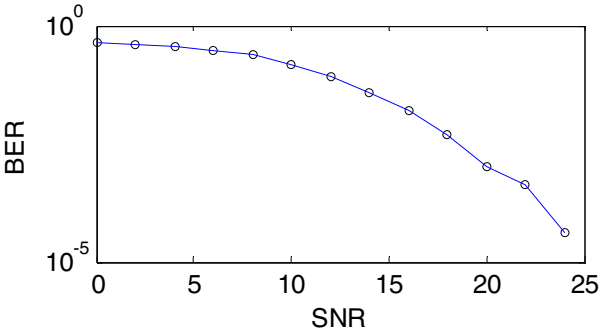


Fig. 6. 2-PPM BER using Random Demodulation and reconstructed using Basis pursuit algorithm

From the Fig.6 it is observed that in both basis pursuit and OMP algorithm the signal is reconstructed using almost the same convergence. But the OMP reconstruction is very faster than Basis pursuit reconstruction technique. Taking the computational time taken by the basis pursuit in the reconstruction part, OMP performs better for a real time system. With a compression ratio of 25 % that is used here signal can now be sampled at just .25 GHz. Hence the ADC can now increase the precision of the quantization level which gives better signal detection.

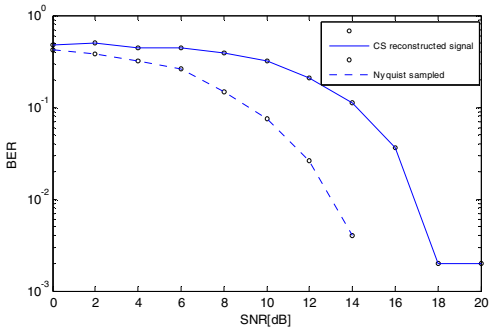


Fig. 7. Signal reconstruction by Compressive sensing and Nyquist sampling

The Fig.7 shows the plot of the signal sampled at Nyquist rate and another signal sampled by compressed sensing technique. From the graph it can be observed that signal reconstructed by CS is only a little SNR below the signal reconstructed by Nyquist rate. But the stress on the digital systems is just $1/4^{th}$ of the original stress.

The Table1 gives the comparison of BER for different compression ratio

Table 1. Comparison of BER for different compression ratios

Compression ratio	BER for SNR 20	BER for SNR 25	BER for SNR 15
m/N=.50	0.001	0	0.0180
m/N=0.25	0.0040	0	0.07
m/N=0.125	0.0330	0.005	0.1240
m/N=0.1	0.0920	0.0190	0.2060
m/N=0.05	0.1310	0.0580	0.2770

4 Conclusion

In this paper, the effect of CS for PPM is explored. The model can be used for recovering a PPM signal with the help of analog to information converter implementations namely, random demodulation and random sampling. CS has been utilized to reduce sampling rates for these signals. The AIC samples at a fraction of the nyquist rate exploiting the property of sparsity of the PPM signal and hence the computational stress over the ADC at the receiver is reduced. In this paper it is shown that the signal detection and reconstruction can be done without having to estimate the channel. In situations where the channel estimate is crucial, CS can still be applied to estimate the channel as well. Using a random matrix as the measurement matrix, the message signal can be reconstructed exactly using OMP and BP algorithms amongst which OMP gave better results. Research can be carried out to ascertain and enhance the performance of compressed sensing in situations with practical abnormalities e.g., timing errors, quantization errors, integrator abnormalities etc.

References

1. Donoho, D.L.: Compressed sensing. IEEE Transactions on Information Theory 52(4) (2006)
2. Laska, J., Kirolos, S., Massoud, Y., Baraniuk, R., Gilbert, A., Iwen, M., Strauss, M.: Random sampling for analog-to-information conversion of wideband signals. In: IEEE Dallas/CAS Workshop, pp. 119–122 (2006)
3. Boggess, B.S.: Compressive Sensing Using Random Demodulation. Master's Thesis, University of Tennessee (2009)
4. Tropp, J., Gilbert, A.: Signal recovery from random measurements via orthogonal matching pursuit. IEEE Transactions on Information Theory 53(12), 4655–4666 (2007)
5. Wentzloff, D.D.: Pulse-Based Ultra- Wideband Transmitters for digital communication. Ph.D Thesis, Massachusetts Institute of Technology (2007)

NN Based Ontology Mapping

K. Manjula Shenoy[1], K.C. Shet[2], and U. Dinesh Acharya[1]

[1] Department of Computer Science and Engineering, Manipal University, MIT, Manipal, 576104
[2] Department of Computer Science, National Institute of Technology, Suratkal
kcshet@rediffmail.com,
{manju.shenoy,dinesh.acharya}@manipal.edu

Abstract. The Semantic Web presents new opportunities for enabling modelling, sharing and reasoning with knowledge available on the web. These are made possible through the formal representation of the knowledge domain with ontologies. Ontology is also seen as a key factor for enabling interoperability across heterogeneous systems. Ontology mapping is required for combining distributed and heterogeneous ontologies. This paper introduces you to the problem of heterogeneity, and need for ontologies and mapping. Also gives an overview of some such ontology mapping systems together with a new system using neural network. The system designed takes OWL files as input and determines all kinds of matching such as 1:1,1:n,n:1,n:n between the entities.

Keywords: Semantic web, OWL, Ontology, Ontology Mapping, Neural Network.

1 Introduction

The Semantic Web[1], is a vision where data has structure and semantics. Ontologies describe the semantics of the data.When data is marked up using ontologies, software agents can better understand the semantics and therefore more intelligently locate and integrate data for a wide variety of tasks. Although Semantic Web offers a compelling vision, but it also raises many difficult challenges. Most important of it is finding semantic mapping of ontologies. Software agents will find great difficult in making an alignment in two different ontologies from two different domains while surfing the semantic web and trying to comprehend the semantic contents over it. Our software agent surfing the semantic web should understand that the various terminologies used in these pair of ontologies are same though they look and read different. This lack of standardization, which hampers this communication and collaboration between agents on semantic web, creates interoperability problem and requires some ontology mapping mechanism.

2 Ontology Mapping Systems

Mapping systems[4][5][6][7] are classified into manual, semi-automatic and automatic systems based on human intervention during mapping. They may also be categorized

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 122–127, 2013.

into metadata based and instance based methods. Metadata based methods rely on the name, attributes ,structure of ontology concepts to generate similarity and instance based methods match the instances under the concepts to generate similarities.

2.1 Instance Based Mapping Methods

GLUE [2]uses a machine learning technique to find semantic mapping between ontologies.

FCA-merge [9]uses formal concept analysis techniques to merge two ontologies sharing the same set of instances. The overall 3 steps in this method are instance extraction, concept lattice computation and interactive generation of the final merged ontology.

2.2 Metadata Based Mapping Methods

Cupid[10] implements an algorithm comprising linguistic and structural schema matching techniques and computing similarity coefficients with the assistance of domain specific thesauri.

COMA,COMA++[11][12] consists of parallel composition of matchers. It uses string based techniques such as n-gram,editdistance ,thesauri look up like Cupid etc. It is more flexible than Cupid. COMA++ is the advanced implementation of COMA with graphical interface.

2.3 Combined Methods

RiMOM[13] discovers similarities within entity descriptions, analyses instances, entity names, entity descriptions, taxonomy structure and constraints prior to using Bayesian decision theory in order to generate an alignment between ontologies and additionally accepts user input to improve mappings. Both instances and metadata play important role in mapping ontology concepts. Some ontologies metadata may be erroneous, there instances may play important role in identifying correspondences; in some cases instances may not be available there structure of ontology i.e metadata may give good response. If both are well defined then accuracy of mapping will be improved by considering both instances and metadata. Also majority of the methods designed earlier do not take OWL as input and find only concept mappings of the form 1:1. Our aim is to find all kinds of mapping between concepts ,between properties such as 1:1, 1:n,n:1,n:n. By keeping this in mind we designed a system which computes various types of similarities based on metadata and instances and combines them as a weighted sum to get resulting similarity using neural network learning.

3 Overview of Our Approach

For our study, an ontology consists of a hierarchical information about various concepts and the various attributes and instances associated with each concepts. An

instance can be associated with multiple concepts, e.g. when the ontology contains concepts of orthogonal aspects. Moreover, an instance may be assigned not only to leaf level concepts but also to inner concepts of the ontology. The key idea of our approach is to derive the similarity between concepts from the similarity of the associated instances. Determining such instance matches is easy in some domains, e.g. by using the non-ambiguous EAN in e-commerce scenarios. Moreover, instance matches may be provided by hyperlinks between different data sources and, thus, can easily be extracted. In the absence of unique identifiers, instance matching can be performed by general object matching (duplicate identification) approaches, e.g. by comparing attribute values. An important advantage for instance-based ontology matching is that the number of instances is typically higher than the number of concepts. This way, we can determine the degree of concept similarity based on the number of matching instances. Furthermore, the match accuracy of the approach can become rather robust against some instance mismatches. We use name, attribute and structure based similarity under metadata section.

3.1 Similarity Computation

First we find dice similarity between pairs of concepts of ontology O1 and O2 as per following formula.

$$Sim_{DICE}(c_1, c_2) = \frac{2 \cdot |I_{c_1} \cap I_{c_2}|}{|I_{c_1}| + |I_{c_2}|} \in [0...1], \forall c_1 \in C_{O1}, c_2 \in C_{O2} \quad (1)$$

In the formula (1), |Ic1| (|Ic2|) denotes the number of instances that are associated to the concepts c1 (c2). |Ic1∩Ic2| is the number of matched instances that are associated to both concepts, c1 and c2. In other words: the similarity between concepts is the relative overlap of the associated instances.

The dice similarity values do not take into account the relative concept cardinalities of the two ontologies but determine the overlap with respect to the combined cardinalities. In the case of larger cardinality differences the resulting similarity values thus can become quite small, even if all instances of the smaller concept match to another concept. We therefore additionally utilize the minimal similarity metric which determines the instance overlap with respect to the smaller-sized concept:

$$Sim_{MIN}(c_1, c_2) = \frac{|I_{c_1} \cap I_{c_2}|}{\min(|I_{c_1}|, |I_{c_2}|)} \in [0...1], \forall c_1 \in C_{O1}, c_2 \in C_{O2} \quad (2)$$

For comparison purposes we also consider a *base similarity* which matches two concepts already if they share at least one instance.

$$Sim_{Base}(c_1, c_2) = \begin{cases} 1 & , \ if \ |I_{c_1} \cap I_{c_2}| > 0 \\ 0 & , \ if \ |I_{c_1} \cap I_{c_2}| = 0 \end{cases} \in [0...1], \forall c_1 \in C_{O1}, c_2 \in C_{O2} \quad (3)$$

Next we consider the fourth and fifth metrix as

$$Sim(c1, c2) = \frac{|I_{c1} \bigcap I_{c2}|}{|I_{c2}|} \tag{4}$$

$$Sim(c1, c2) = \frac{|I_{c1} \bigcap I_{c2}|}{|I_{c1}|} \tag{5}$$

Two attribute based matrices are determined based on ratio of number of common attribute types to maximum number of attributes and to the minimum number of attributes.

$$Sim(c1, c2) = \frac{|A_{c1} \bigcap A_{c2}|}{max(|A_{c1}|, |A_{c2}|)} \tag{6}$$

$$Sim(c1, c2) = \frac{|A_{c1} \bigcap A_{c2}|}{min(|A_{c1}|, |A_{c2}|)} \tag{7}$$

Path difference consists of comparing not only labels of objects but the sequence of labels of entities to which those bearing the label are related. Suppose we want to compare two labels we construct two strings s and t consisting of names of their super classes on the path from common root and use them to find similarity. Path based similarity is determined as per following equation. Here δ' is the similarity computed based upon semantic relation between leaves of the tree. Here λ is between 0 to 1 value. This measure is dependent on the similarity between the last element on each path: this similarity is affected by λ penalty but every subsequent step is affected by a $\lambda(1-\lambda)^n$ penalty.

$$\begin{aligned} &\delta\left(\langle s_i\rangle_{i=1}^{n}, \langle t_j\rangle_{j=1}^{m}\right) = \lambda \times \delta'(s_n, t_m) + (1-\lambda) \times \delta\left(\langle s_i\rangle_{i=1}^{n-1}, \langle t_j\rangle_{j=1}^{m-1}\right) \\ &\delta\left(\langle\rangle, \langle t_j\rangle_{j=1}^{k}\right) = \delta\left(\langle s_i\rangle_{i=1}^{k}, \langle\rangle\right) = k \end{aligned} \tag{8}$$

3.2 Concept Similarity Matrix

If $s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8$ are the similarities computed between two concepts C_1 and C_2 of O_1 and O_2 respectively then the similarity value s between C_1 and C_2 is obtained as weighted sum of $s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8$ as weights are randomly initialized and will be adjusted through the learning process. For two ontologies being mapped O_1 and O_2 we calculate the similarity values for pairwise concepts (one from O_1 and the other from O_2, considering all combinations). Then we build $n_1 X n_2$ matrix M to record all values calculated , where n_i is the number of concepts in O_i. The cell M_{ij} stores similarity between i^{th} concept in O_1 with j^{th} concept in O_2.

3.3 Learning through NN

The main purpose of NN is to learn the different weights for instance based and metadata measures during matching process.

We adopt a three layer feed forward neural network.The network has 8 units in the input layer, 12 units in the hidden layer and one output unit. The input to the network is vector s representing different type of similarity. The output is s the similarity value between concept C_1 and C_2 of O_1 and O_2.

Initially we obtain 8 concept similarity matrices M_i as explained in section A. Then we pick up randomly a set of concepts and an expected similarity value as per manual matching and give this to train the network with initially set random weights and biases.

3.4 Other Kinds of Mappings

Once concepts are mapped the system proceeds to find mappings between other entities and other type of mapping . For properties three kinds of similarity are found. First one is lexical similarity between labels such as Levenshteins distance, lexical similarity for comments using TF-IDF[14] score and the similarity given by the ratio of number of token operations needed to maximum of number of tokens in comments of both. Another matrix used is similarities between domain and range. In place of attributes their values in instances are considered for properties. Again NN is used to give different weightage to different measures calculated for properties.

3.5 Experimental Results

The test was conducted on benchmark data sets 101,103,104,201…210. The accuracy is measured using precision,recall and f-measure. Precision is the ratio of correctly found correspondences over the total number of returned correspondences. Recall is the ratio of correctly found correspondences over expected correspondences. F-measure is computed using precision and recall. The plot of precision, recall and

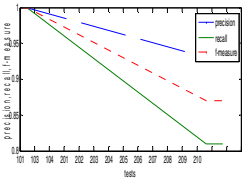


Fig. 1. Results of proposed method

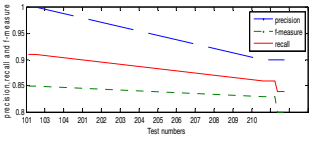


Fig. 2. Result of RiMOM

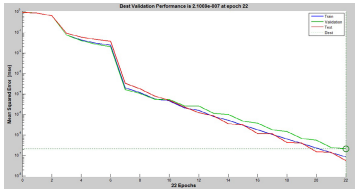


Fig. 3. NNperformance

f-measures are as per Figure 1 and 2 for proposed method and RiMOM respectively. Figure 3 shows performance of neural network. It is found that precision, recall and f-measures are good for initial benchmark sets and as the complexity of the data set increases the precision, recall and f-measures decrease slightly. Also the accuracy is good when compared to RiMOM[16].

4 Conclusion

This Paper has presented an approach for ontology mapping using neural network. The accuracy of the system proposed is found to be remarkable with respect to another similar system RiMOM.

References

1. Berners_Lee, T., Hendler, J., et al.: The Semantic Web. Scientific American (2001)
2. Doan, A., Madhavan, J., Domingos, P., Halevy, A.Y.: Learning to Map between Ontologies on the Semantic web. In: Proceedings of the WWW Conference (2002)
3. Ehrig, M., Staab, S.: QOM – Quick Ontology Mapping. In: McIlraith, S.A., Plexousakis, D., van Harmelen, F. (eds.) ISWC 2004. LNCS, vol. 3298, pp. 683–697. Springer, Heidelberg (2004)
4. Kalfoglou, Y., Schorlemmer, M.: Ontology mapping: The State of the Art. The Knowledge Engineering Review (2003)
5. Taye, M.M.: State of the Art: Ontology matching techniques and Ontology Mapping Systems. International Journal of ACM (2010)
6. Choi, N., Song, I-Y., Han, H.: A Survey on ontology mapping. SIGMOD Record (2006)
7. Noy, N.: Semantic Integration: A Survey of ontology based approach. SIGMOD Record (2004)
8. Noy, N., Musen, M.: Anchor-PROMPT: Using non local context for Semantic matching. In: Proceedings of the IJCAI (2001)
9. Stumme, G., Maedche, A.: FCA-MERGE: bottom up merging of ontologies. In: Proceedings of the IJCAI (2001)
10. Madhavan, J., Bernstein, P., Rahm, E.: Generic Schema matching with cupid. VLDB Journal (2001)
11. Do, H.H., Rahm, E.: COMA-A system for flexible combination of schema matching approaches. VLDB Journal (2002)
12. Aumueller, D., Do, H.H., Massmann, S., Rahm, E.: Schema matching and ontology matching with COMA++. In: SIGMOD Conference (2005)
13. Li, J., Tang, J., Li, Y., Luo, Q.: RiMOM: a dynamic multistrategy ontology alignment framework. IEEE Transactions on Knowledge and Data Engineering (2009)
14. Jean-Mary, Y.R., Patrick Shironoshita, E., Kabuka, M.R.: Ontology matching with semantic verification. Journal of Web Semantics: Science, Services and Agents on the World Wide Web (2009)
15. Ramos, J.: Using TF-IDF to determine word relevance in document queries. In: Proceedings IIRC 2003 (2003)
16. Wang, Z., Huo, L., Zhao, Y., Li, J., Qi, Y., Tang, J.: RiMOM Results for OAEI2010. In: Proceedings of Ontology Matching Workshop (2010)

A Load Reduction and Balancing Scheme for MAG Operating in PMIPv6 Domain

Juneja Dimple and Chander Kailash

M.M. Institute of Computer Technology & Business Management,
Maharishi Markendeshwar University, Mullana, Haryana, India
{Dimplejunejagupta,kailashchd}@gmail.com

Abstract. Researchers have been emphasizing on employing mobile agents to assist vertical handover decisions in 4G mainly because it reduce consumption of network bandwidth, delay in network latency and the time taken to complete a particular task in PMIPv6 environment is minimized as well. However, when more than the desired number of mobile nodes is attached to the PMIPv6 domain including a particular MAG, the MAG gets overload. In fact, with this increasing load, the end-to-end transmission delay and the number of packet loss also increases. Since the number of mobile users and the associated applications are increasing exponentially, the problem stated above is obvious. Yet, the current specification of PMIPv6 does not provide any solution for this problem. Thus, this work extends the previous works wherein the employment of mobile agents in PMIPv6 has been justified but no efforts have been put towards reducing and hence balancing the load of overloaded MAG. The paper proposes a skilful load balancing scheme for the PMIPv6 network to balance the load over MAGs in the domain.

Keywords: Load Balancing Schemes, Mobile Agents, PMIPv6, Vertical Handover.

1 Introduction

The communication technology has observed exponential progression from fixed line analogue connection to high speed connections including wireless interfaces at a touch of a button. In fact, these advancements have propagated to mobile devices which are capable of handling multimedia and other real time applications. In addition to providing new technological opportunities, new challenges such as increased usage of services and maintaining QoS, utilization of available resources and most importantly seamless handover have posed into. In order to address the issues stated above authors in their previous works had introduced agent-based PMIPv6[16] with the aim for efficient service dissemination and seamless handover. The idea of injecting mobile agents was implemented after scrutinizing the literature [12] thoroughly and it was discovered that mobile agents can greatly reduce the battery-consumption, communication cost, especially over low bandwidth links by

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 128–136, 2013.

moving the processing functioning to the agents rather than bringing the data to central processor. However, the handover latency and packet loss could improve in some conditions only.

Injecting mobile agents in PMIPv6 architecture lead to new computing paradigm which is a noticeable distinction to the traditional client/server-based and hardwired computing. The mobile agent is a special kind of software that travels the network either periodically or on demand. It performs data processing autonomously while migrating from node to node. Literature [10] reflects that mobile agents offer advantages such as scalability, extensibility, energy awareness and reliability making them more suitable for heterogeneous wired/wireless networks than their counterparts. Mobile agents found their utility for data fusion tasks in distributed networks and this characteristic have been exploited in this paper to reduce the load of overloaded MAG in PMIPv6 domain.

Authors in their previous works proposed agent-based PMIPv6 and extensive simulations reflected that in certain conditions a particular MAG gets overloaded. The load of a MAG has to be less than or equal to the specified threshold. The present load (PL) of a MAG can be measured on packet arrival rate [7] at every certain time interval and the arrival rate measured at j^{th} time interval is denoted by $\lambda_{i,j}$ ($1<=i<=N_m$). At every N_m measurements, the average arrival rate $\lambda`_i$ is estimated. Supposing that μ_i is the average service rate of MAG_i, the average packet arrival rate at MAG_i is calculated using weighted moving average method [7] given as follows:

$$\lambda`_i = \frac{\sum_{j=1}^{N_m}(N_m - j+1)\lambda_{N_m - j+1}}{\sum_{j=1}^{N_m} J} \quad (1)$$

Then, the load $pi = \lambda`_i / \mu_i$ is calculated. If the value of pi exceeds the MAG threshold (θ) it gets overloaded and need for either reducing or balancing the load of MAG becomes apparent. Therefore, this work aims to propose a load reduction and balancing mechanism.

The structure of the paper comprises of four sections. Section 2 presents the related work. Section 3 presents the proposed solution for load reduction and balance. Section 4 shows the evaluation of the proposed work and finally conclusion of the work is presented in Section 5.

2 Related Work

Recently, mobile agents have been proposed for efficient migration of data in telecommunications [8]. Migration in distributed computing can be performed for enhancement in performance. Load balancing with resource migration [14] partitions the distributed system into regions to get the optimal solutions in many instances. Novel load balancing architecture [3] has employed mobility agent who, acts as a front-end for the load balancing devices hiding mobility related issues from the load balancer and from the servers. Another efficient load balancing AP selection algorithm [5] considers load at each of the access points (APs) and its suitability with respect to the direction of advancement of the MN. Talent-based and Centrality-based [4] task allocation models are also used for load balancing. In Talent-based model the

tasks are assigned as per the talent of the agent. However centrality-based method is implemented according to the agent's centralities in network. A distributed agent strategy [13] allow the system to scale well rather than using a centralized control, as well as achieving a reasonable good resource utilization and meeting application execution deadlines. An analytic model [1] of average queuing delay per packet was proposed. The model also includes the Media Independent Handover (MIH) protocol to prevent the overload from being relocated to the overloaded point of attachments (PoA). Another approach for PMIPv6 based on measurement of packet arrival rate at every certain time interval by MAG. MAG sends its current load information to the related LMA periodically by utilizing the Heartbeat message [17]. Extension in heartbeat message [6] by adding 'Load Information' field after 'Heartbeat Interval' was proposed. Another approach for load balancing [7] proposed by setting a threshold value of MAG. LMA computes the fairness index (FI) value of all the connected MAGs. If the computed FI value is less than threshold value it sends a load balancing message to most overloaded MAG.

Generating candidate MAG list [9] is another approach proposed for balancing the load. MAG list can be built by exchanging the MAG load information among MAGs in the PMIPv6 domain. The order preference of the candidate MAG list is based on the technique for order preference by similarity to ideal solution (TOPSIS) algorithm [18]. Since heartbeat messages are exchanged between each network entity, extending the heartbeat messages enables each MAG gather the load information of other MAGs to build up a candidate MAG list. This candidate MAG list will serve as a reference for load balancing mechanism that will be triggered when the load of an MAG exceeds its load threshold.

Now, focusing on the existing solutions that are basically hardwired and are not intelligent enough to provide independence to service providers as well as to users, we had proposed an agent based smart solution [16] mainly comprising of agents injected at all the core elements namely LMA, MAG and MN of PMIPv6 domain. The new agent-based solution is powerful enough to reduce the battery-consumption and communication cost however, handover latency and packet loss during handover could be reduced conditionally. The agents had been named as LMA_{agent}, MAG_{agent}, MN_{agent}, compute agent and information agent [16]. As shown in figure 1, the smart solution operates in three modules which can be executed autonomously and iteratively as well.

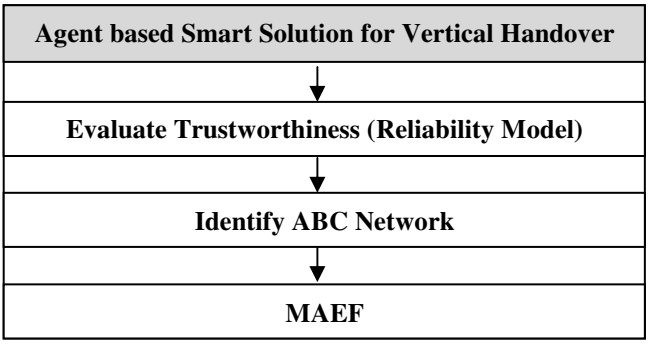


Fig. 1. Smart Solution for Vertical Handover

When two agents (at any level) intend to communicate, the first module is activated to check the trustworthiness of agents [11] by evaluating the trust certificate for mobile agents. In fact, trust is one of the basic parameter of evaluating a mobile agent-based system and usually computed by analyzing various parameters such as direct experience, third party references and confidentiality, persistence etc. A trusted agent migrates through the neighboring MAGs populating weight for every MAG which, in turn helps the mobile users to choose always best network [15]. When the best connected network is identified, the emigration framework (MAEF) [2] finally transfers the MN to the target MAG. Apart from these basic functionalities these agents manages network policy, managing MN logs, buffering data packets, performing AAA for MN, computing Remaining Transaction Time (RTT) and managing user preferences etc. All these tasks are performed by agents deployed at different level of PMIPv6 environment. For instance, MN_{agent} is responsible for managing the user's preferences, which is computed by MAG_{agent} by calculating the network weight based on weight rules [15]. MAG_{agent} collects this information by interacting with other neighboring MAG_{agents}. It also computes Remaining Transaction Time (RTT) for the present transactions for shifting to other MAGs (in case of lesser battery life as compared to the required) without losing the data. In parallel MAG_{agent} buffered the data packets to reduce the packet lost in case of handover situation. It also manages the MN log for authentication and authorization on the behalf of MN to overcome the latency period. Figure 2 depicts the high level of various responsibilities of MAG_{agent} and hence it is evident that the MAG_{agent} is usually overloaded.

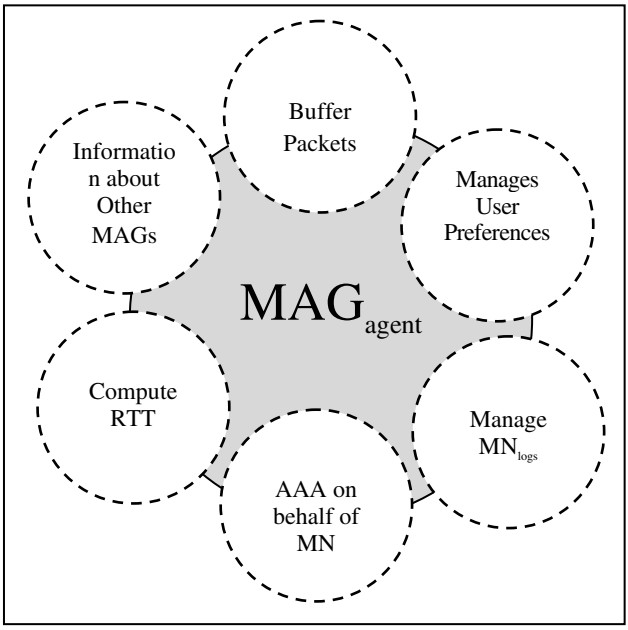


Fig. 2. Responsibilities of MAG_{agent}

On the other side LMA_{agent} is responsible for managing the network policies and in case of any change in network policy the same should be intimated to MAG_{agent} to check that MN connected to the MAGs are able to fulfill the criteria or handover is needed for those MN. So the responsibility of MAG in network based model is high as compare to MN and LMA. Now since, MAG is a device which works on behalf of LMA and manages number of MNs, therefore when the number of MNs connected to a particular MAG, becomes large, the MAG and its associated agent gets overloaded and hence, the MNs will suffer from queuing delay at the MAG, resulting in increased end-to-end delay. To solve this issue an intelligent load reduction and balancing mechanism is required in PMIPv6 environment.

Next section aims to propose criteria for reducing load of MAG with the help of MAG_{agent} and hence proposes to balance the load at MAG in PMIPv6.

3 Proposed Solution

This section initially proposes a strategy to reduce the load of MAG and further balances the overloaded MAG.

The proposed strategy employs the mobile agent's, specifically MAG_{agent}'s ability to reduce the communication and cost in an application specific zone. MAG_{agent} shall adjust their behaviors depending on the QoS needs such as handover latency. The upcoming section proposes to reduce the MAG load by filtering the irrelevant data , reducing the information redundancy and communication overhead.

3.1 The Load Reduction Strategy

With the introduction of mobile agents in PMIPv6 domain, the trend has changed to "One Deployment and Multiple Applications" and such a trend requires mobile agents to have various capabilities to handle multiple applications at the same time. In general, one cannot expect to store every possible application to the cache associated with either mobile agent or even MAG itself. The MAG can assign the local code based on the requirement of a specific application to the associated MAG_{agent} which is then responsible for locally processing the information. In addition MAG_{agents} can be replicated by MAG, if desired. This capability results into the transmission of only relevant data to MAG for further processing. Let α be the data reduction ratio contributed by the MAG_{agent} and β_i be the size of actual data at the mobile node i (MN_i), then the size of reduced data (γ_i) is given by

$$\gamma_i = \alpha . \beta_i \tag{2}$$

Where, R [0,1] defines the range of α.

Now, the MAG_{agent} initially moves to one mobile node to gather the data and then traverses the complete list of all mobile nodes attached to the respective MAG at that

particular instant of time. If for the first node the size of reduced data is given by γ_1, then the total Load Reduced (ω) at MAG is given by

$$\omega = \gamma_1 + \sum_{i=2}^{k} \pi.\gamma_i \tag{3}$$

Where k is the number of mobile nodes connected with MAG and π ($0< \pi<1$) is the aggregation factor multiplied with the sequence of reduced data.

Further, the communication overhead can also be reduced if the shorter packets can be combined to form bigger packet before transferring on to the network. However, this part is little difficult to implement at his level and is out of the scope of this work. Now, ω shall be considered as the current load of MAG and is compared with θ i.e. the threshold load set for the MAG. If $\omega > \theta$then the load LMA_{agent} is activated and executes the load balancing strategy as proposed in next section else the MAG continues to handle the tasks allocated to it.

3.2 The Load Balancing Strategy

The proposed strategy executes if the reduced load of MAG is still more than the setlimits of MAG. LMA_{agent} initially computes the overload as per the given equation:

$$\epsilon = \omega - \theta \tag{4}$$

Now based on the ϵ LMA_{agent} select the mobile nodes on the basis of following criteria:

Criteria 1: MNs initiating maximum number of packets to be transferred (high priority)
Criteria 2: Mobile nodes causing queuing delay
Criteria 3: Mobile node running real time applications (Lowest priority)

The mobile nodes satisfying the criteria 1 shall be the nodes which must be handover on top priority. However, if the same node satisfies the criteria 3 also, then the same nodes shall be ignored and next node shall be checked. The process of selecting nodes will continue until it satisfies MAG's threshold. The MAG_{agent} also sends a De-registration request to LMA_{agent} for the selected node. These nodes should be shifted to best available network as per the algorithm given by Chander[15]. Further, the new updated load is computed after one mobile node has been handed over to the neighboring MAG. The load of MAG would now be given as:

$$\omega = \omega - \beta_i \tag{5}$$

The figure 3 shows the complete working of the proposed solution

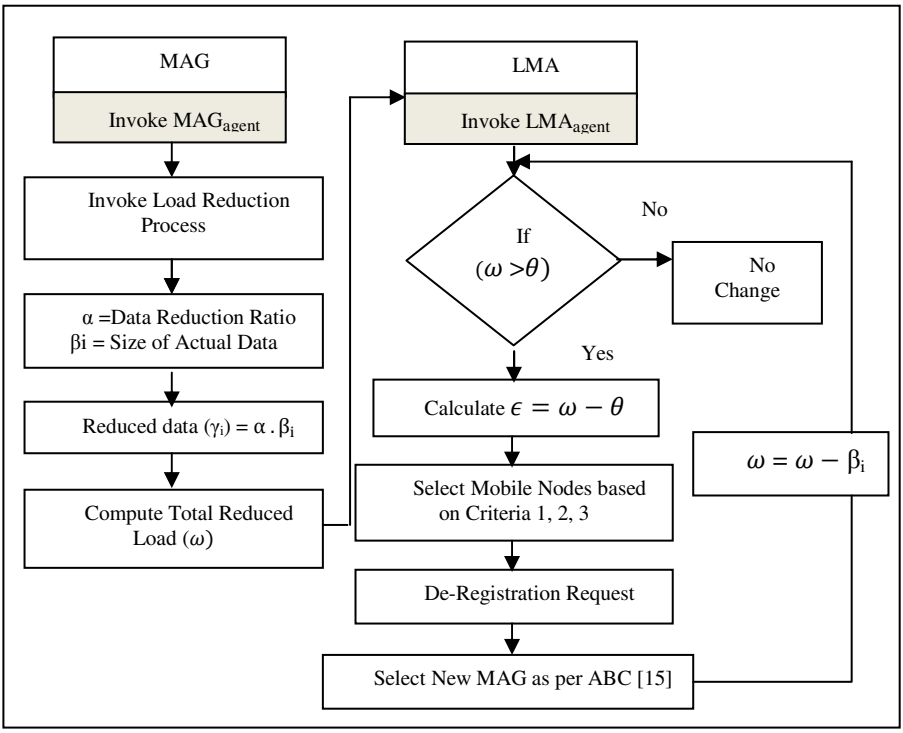


Fig. 3. Flow Diagram showing working of MAG and LMA Agent

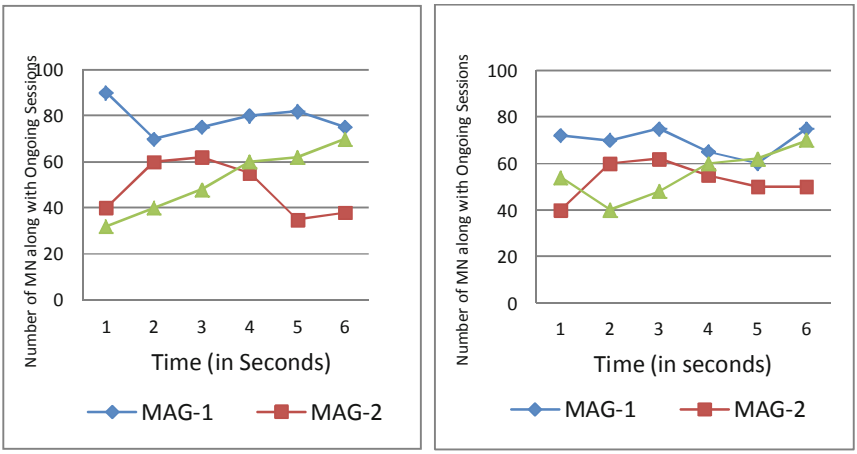
4 Performance Evaluation

We proposed proficient model to reduce average queuing delay at MAG level in PMIPv6 environment using Intelligent Agents. For evaluating the proposed model a simulator was developed in JAVA. The parameters used for simulation are shown in Table 1.

Table 1. Parameters used for simulation

Parameters	Values
MAG Capacity	100
MAG Load	0-100
MAG Threshold	75
Simulation Time	In Seconds

In our simulation study we took load of three MAGs assuming that each MAG is having capacity of 100 MN's sessions. The threshold value of the MAG is set to 75. The MAG initiates load reduction and load balancing whenever overload is considered overloaded MAG is shown in the figure 4 (a).



(a) Before Load Balance (b) After Load Balance

Fig. 4. Ongoing session Load over MAGs

As shown in Figure 4(b) all the three MAGs are significantly balanced within their threshold value and also reduce the queuing delay at the MAG compared to no-load balancing case.

5 Conclusion

This work uniquely contributed a load reduction and balancing strategy for reducing the load of MAG. The previous works had been indicative that MAG gets overloaded often loading to queuing delay. The work proposes two algorithms initially for reducing the loading & later for balancing the load, if still required. The work has been simulated in java and the evaluation reflects that results are in comparison with other existing strategies.

References

1. Kim, M.-S., Lee, S.: Load Balancing based on Layer 3 and IEEE 802.21 Frameworks in PMIPv6 Networks. IEEE, 788–792 (2009)
2. Chander, K., Juneja, D.: Mobile Agent based Emigration Framework for 4G: MAEF. Communicated to International Journal of Information and Computing Science, ISSN: 1746–7659
3. Cabellos-Aparicio, A., Pascual, J.D.: Load Balancing in Mobile IPv6's Correspondent Networks with Mobility Agents. In: IEEE Communications Society Subject Matter Experts for Publication in the ICC 2007 Proceedings, pp. 1827–1832 (June 2007)
4. Jiang, Y., Zhaofeng: Locality-sensitive task allocation and load balancing in networked multiagent systems: Talent versus centrality. Journal of Parallel Distributed Computing 71, 822–836 (2011)

5. Misra, I.S., Banerjee, A.: A Novel Load Sensitive Algorithm for AP selection in 4G Networks
6. Kim, M., Lee, S.: Load Balancing for Proxy Mobile IPv6 Networks. In: 6th IEEE Consumer Communications and Networking Conference, CCNC 2009, pp. 1–2 (2009)
7. Kim, M.-S., Lee, S.: A novel load balancing scheme for PMIPv6-based wireless networks. Journal of Electronics and Communications (AEU) 64, 579–583 (2010)
8. Sharma, A.K., Dimple, J.: User Authentication Through Intelligent Agents. In: Proceedings of National Conference MOBICOMNET 2004, Mobile Computing and Networking (MobiComNet 2004) Vellore Institute of Technology, Chennai, India, September 30- October 2 (2004)
9. Premec, D.: NetExt Working Group, Internet Draft Inter-technology handover in PMIPv6 domain at `http://www.ietf.org` (March 9, 2009)
10. Sodnik, J., Stular, M., Milutinovic, V., Tomazic, S.: Encyclopedia of Wireless and Mobile Communications (2008)
11. Chander, K., Juneja, D.: A New Reliability Model for Evaluation Trustworthiness of Intelligent Agents in Vertical Handover. International Journal of Advanced Computer Science and Applications (IJACSA) 2(9) (2011)
12. Chander, K., Juneja, D.: Vertical Handover Decision Management in Heterogeneous Network: Developments and Challenges. In: International Conference on "Innovation Practices in Management and Information Technology for Excellence" (2010)
13. Cao, J., et al.: Grid load balancing using intelligent agents. Elsevier-Future Generation Computer Systems 21, 135–149 (2005)
14. Varadarajan, R.: An efficient approximation algorithm for load balancing with resource migration in distributed system. Computer and Information Sciences Department, University of Florida, Gainsesville
15. Chander, K., Juneja, D.: A Novel Approach for Always Best Connected in Future Wireless Networks. Global Journal of Computer Science and Technology 11(15), Version 1.0 , 49–53 (2011)
16. Chander, K., Juneja, D., et al.: An Agent Based Smart Solution for Vertical Handover in 4G. International Journal of Engineering Science and Technology 2(8), 3381–3390 (2010)
17. Devarapalli, V., Lim, H., Kant, N., Krishnan, S.: Heartbeat Mechanism for Proxy Mobile IPv6, draft-devarapalli-netlmm-pmipv6-heartbeat-03.txt (July 2008)
18. Farahani, R., Asgari, N.: Combination of MCMD and covering techniques in a hierarchical model for facility location: a case study. European Journal of Operational Research 176, 1839–1858 (2005)

Document Image Segmentation for Kannada Script Using Zone Based Projection Profiles

Siddhaling Urolagin[1], K.V. Prema[2], and N.V. Subba Reddy[2]

[1] Department of Computer Science & Engineering, MIT, Manipal, India
[2] Mody Institute of Technology and Science, Rajastan, India
[3] Dept. of E&C, MIT, Manipal, India
siddesh_u@yahoo.com, drprema.mits@gmail.com,
dr_nvsreddy@rediffmail.com

Abstract. In Kannada document image analysis, the document segmentation into individual lines, words and characters poses many challenges. The *aksharas* of Kannada language are formed by combination of glyphs of consonant, consonant conjuncts and vowel modifiers; therefore a lot variation in widths and heights of *aksharas* can be observed. Moreover an *akshara* is composed of two or more disconnected components. The consonant conjuncts pose more problems as they occasionally overlap with next line or next character. In this research we have proposed a Kannada document segmentation method using zone based projection profiles. The experiments are conducted on several scanned Kannada document with different fonts and character segmentation rate of 99.2% and consonant conjunct segmentation rate of 92.73% on an average is observed.

Keywords: Document image analysis, document segmentation, Kannada script.

1 Introduction

During the past few decades, substantial research efforts have been devoted to document image analysis using Optical Character Recognition (OCR) [l]. Recently there are several OCR systems developed for Indian languages such as Devanagari characters [2], Bengali characters [3], and Tamil characters [4] etc. The segmentation is one of the crucial steps in OCR. In the document image analysis, the segmentation process involves three steps viz: Line, Word and Character segmentation. For the languages like Kannada, segmentation is more challenging due to the large collection of *aksharas* and presence of conjunct consonant (subscript or Vatthu) [5]. The Kannada character segmentation is also difficult because Kannada characters have higher similarity in shapes and higher variability across fonts in character belonging to same class [6]. As in any Indian language, the Kannada characters are too inflectional and agglutinative in nature. The Kannada script has 19090 characters which are formed by grammatical combination of basic symbols. In [7] authors have proposed a two stage character segmentation technique for Kannada. At first stage consonant conjuncts are separated using connected component algorithm and in the next stage characters are segmented using projection profile. The work reported in

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 137–142, 2013.

this paper is motivated by the fact that high efficient segmentation can be carried out by projection profile method. The rest of the paper is organized as follows. In section 2 we have discussed the challenges involved in the Kannada document segmentation. Our proposed segmentation method is discussed in section 3. The experimental results are given in section 4 and finally the conclusion is covered in section 5.

2 Challenges Involved in Kannada Document Segmentation

The segmentation of Kannada document poses many challenges. In a language like English, usually the space between two lines, words and characters is sufficient to segment them individually. But Kannada document this is more difficulty as the width of the *aksharas* formed by combination of a consonant and vowels vary widely also a single *akshara* may be composed of two or more disconnected components. Moreover the conjunct occurring at the bottom of an *akshara* may sometimes intercept with boundary of next line or with next character, causing difficulty while selecting the segmentation position. To handle such cases in this paper a segmentation method is developed using projection profiles of the document image.

3 Proposed Segmentation

In this section we discuss the Kannada document image segmentation, which involves in identifying boundary position of individual lines, words and characters and separate them from the document.

3.1 Lines and Words Segmentation

The digitized image of the document is segmented into lines and words using Horizontal Projection Profile (HPP) and Vertical Projection Profile (VPP).

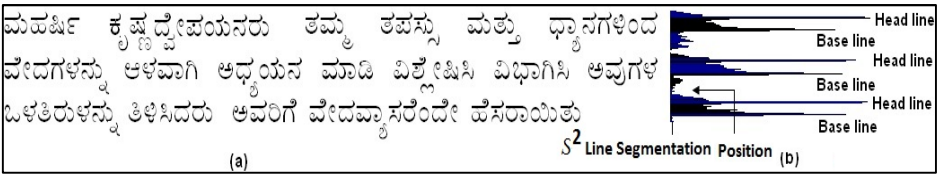


Fig. 1. (a) A portion of Kannada document. (b) HPP of document.

A HPP is a one-dimensional array where each element denotes the sum of pixels along a row in the image and VPP gives the column sums. By looking at the HPP of a document it is possible to separate lines of it. However, for Kannada document consonant conjuncts (*vatthus*), which occur below the base consonant, may overlap with vowel modifier of next line as depicted in Fig. 1(a). Therefore distinct zero valleys may not occur in HPP. The approach proposed in this paper is to identify Head line (*Hl*) and Base line (*Bl*) position from HPP as shown in Fig. 1(b), where

more strength in projection profile is observed. Then the segmentation position S^l for l^{th} line is taken as position where a minimum profile in HPP between the present *Bl* to next *Hl* position, i.e.,

$$S^l = \min\{\mathrm{HPP}(k) \mid Bl_l \le k \le Hl_{l+1}\} \tag{1}$$

For last line $Hl_{l+1} = n$ where n is number of profiles in HPP. A VPP is applied to a line for word segmentation as shown in Fig. 2. The space between each word, which is sufficiently large compared to inter character space, is used to determine the word boundaries.

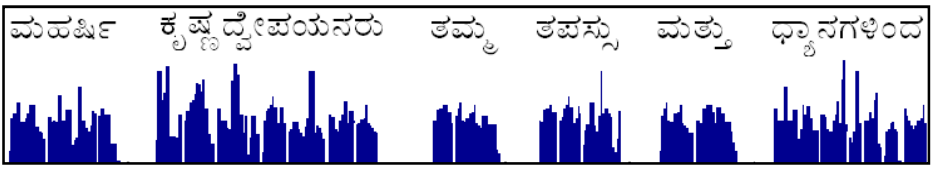


Fig. 2. Word segmentation

3.2 Characters Segmentation

In Kannada, *aksharas* are formed by combination of glyphs corresponding to vowels, consonants, conjuncts and vowel modifiers. The conjuncts appear at the bottom of a character while vowel modifiers appear either at right side or at bottom. Occasionally, the consonant conjuncts, which appear at below the base consonant, may overlap with the base consonant of next character in the VPP as shown in Fig. 3(a) or overlap in the middle zone in HPP as shown in the Fig. 3(b). Due to these peculiarities of conjuncts, a single VPP of a word does not yield proper segmentation of characters.

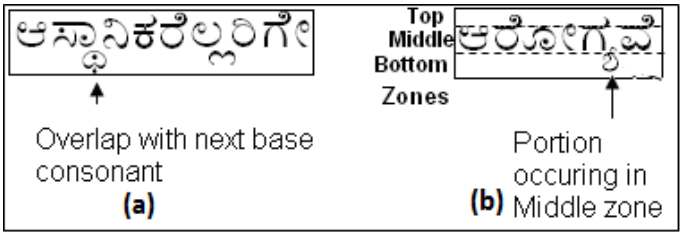


Fig. 3. Examples for overlap of consonant conjunct

To overcome these problems, we have proposed to use two VPPs of a word for the segmentation. From the top and middle zones of the word a VPP (VPP_{TM}) is obtained. By analysing VPP_{TM} base consonants are located. For a word of Fig. 4(a) three zones are depicted in Fig. 4(b). In the Fig. 4(c) top and middle zones along VPP_{TM} of the word are shown. The space in VPP_{TM} is used to separate base consonants. To separate consonant conjuncts, we take VPP of bottom zone (VPP_B) of the word. The consonant

conjuncts can readily be located in VPP_B as shown in Fig. 4(d). Furthermore, it is important to know for which base consonant a particular conjunct is associated. We achieve this by analysing the corresponding positions of base consonant in VPP_{TM} and position of conjunct in VPP_B. Suppose a consonant conjunct is present between the locations xc_1 and xc_2 in VPP_B. Then this conjunct is associated with i^{th} base consonant if

$$xb_1^{\,i} \le xc_1 \le xb_2^{\,i} \tag{2}$$

where i^{th} base consonant is present between the locations $xb_1^{\,i}$ and $xb_2^{\,i}$ in VPP_{TM}. In Fig. 4(c) the locations of consonant ನ are shown. The conjunct ್ನ is present between locations xc_1 and xc_2 as shown in Fig. 4(d). Here $xb_1^{\,2} \le xc_1 \le xb_2^{\,2}$ therefore conjunct ್ನ is associated with ನ. We also utilize width of the vowel modifier to make it associated with base consonant. In Fig. 4(e) the segmented base consonant, vowel modifier and consonant conjuncts are shown.

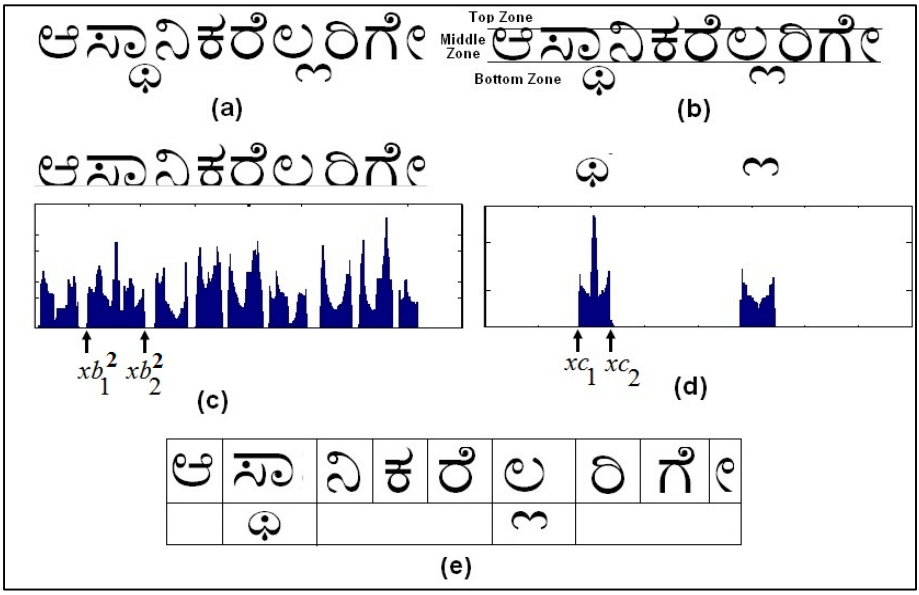


Fig. 4. (a) A word. (b) Word with three zones. (c) VPP for top and middle zones. (d) VPP for bottom zone. (e)Segmented *aksharas.*

4 Experimental Results

To conduct the experiments, we have prepared the text pages containing different fonts using multilingual Baraha editor. The prepared pages are scanned using HP

Laserjet 3055 scanner. An example for Kannada document is shown in Fig. 5. The document collection has pages of four major font types and for the segmentation experiments two pages of each font are considered. The total 8 pages with the text taken from the various weekly, monthly magazines, newspapers and several text books are collected. Each page has around 30 to 35 lines and around 220 words in it.

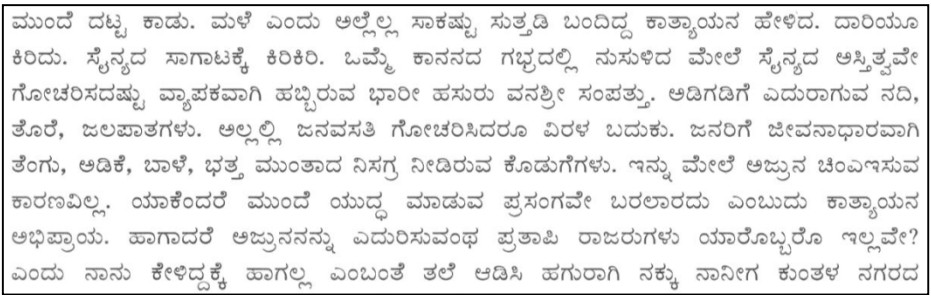

ಮುಂದೆ ದಟ್ಟ ಕಾಡು. ಮಳೆ ಎಂದು ಅಲ್ಲೆಲ್ಲ ಸಾಕಷ್ಟು ಸುತ್ತಾಡಿ ಬಂದಿದ್ದ ಕಾತ್ಯಾಯನ ಹೇಳಿದ. ದಾರಿಯೂ ಕಿರಿದು. ಸೈನ್ಯದ ಸಾಗಾಟಕ್ಕೆ ಕಿರಿಕಿರಿ. ಒಮ್ಮೆ ಕಾನನದ ಗಭ್ರದಲ್ಲಿ ನುಸುಳಿದ ಮೇಲೆ ಸೈನ್ಯದ ಅಸ್ತಿತ್ವವೇ ಗೋಚರಿಸದಷ್ಟು ವ್ಯಾಪಕವಾಗಿ ಹಬ್ಬಿರುವ ಭಾರೀ ಹಸುರು ವನಶ್ರೀ ಸಂಪತ್ತು. ಅಡಿಗಡಿಗೆ ಎದುರಾಗುವ ನದಿ, ತೊರೆ, ಜಲಪಾತಗಳು. ಅಲ್ಲಲ್ಲಿ ಜನವಸತಿ ಗೋಚರಿಸಿದರೂ ವಿರಳ ಬದುಕು. ಜನರಿಗೆ ಜೀವನಾಧಾರವಾಗಿ ತೆಂಗು, ಅಡಿಕೆ, ಬಾಳೆ, ಭತ್ತ ಮುಂತಾದ ನಿಸರ್ಗ ನೀಡಿರುವ ಕೊಡುಗೆಗಳು. ಇನ್ನು ಮೇಲೆ ಅಜ್ರುನ ಚಿಂಎಇಸುವ ಕಾರಣವಿಲ್ಲ. ಯಾಕೆಂದರೆ ಮುಂದೆ ಯುದ್ಧ ಮಾಡುವ ಪ್ರಸಂಗವೇ ಬರಲಾರದು ಎಂಬುದು ಕಾತ್ಯಾಯನ ಅಭಿಪ್ರಾಯ. ಹಾಗಾದರೆ ಅಜ್ರುನನನ್ನು ಎದುರಿಸುವಂಥ ಪ್ರತಾಪಿ ರಾಜರುಗಳು ಯಾರೊಬ್ಬರೊ ಇಲ್ಲವೇ? ಎಂದು ನಾನು ಕೇಳಿದ್ದಕ್ಕೆ ಹಾಗಲ್ಲ ಎಂಬಂತೆ ತಲೆ ಆಡಿಸಿ ಹಗುರಾಗಿ ನಕ್ಕು ನಾನೀಗ ಕುಂತಳ ನಗರದ

Fig. 5. A scanned Kannada document

To evaluate the proposed method each scanned Kannada document is subject to segmentation. At the first level the line segmentation is carried out using HPP, the line segmentation is very accurate and our segmentation method is successful in segmenting the lines in all the documents. Then word segmentation is carried out after line segmentation using VPP. Standard word space is used to segment a word. VPP method is successful in word segmentation with 100 % accuracy.

Table 1. Character segmentation results

Page No.	Font Type BRH	Character Segmentation			Vathhu segmentation		
		Characters Present	Segmented Properly	Segmentation Rate in %	Vatthu Present	Segmented Properly	Segmentation Rate in %
1	Amerika kannada	898	896	99.78	92	84	91.30
2	Amerika kannada	977	974	99.69	109	101	92.66
3	kasturi	919	896	97.50	77	73	94.81
4	kasturi	922	902	97.83	79	75	94.94
5	Kannada	874	871	99.66	132	121	91.67
6	Kannada	931	926	99.46	136	124	91.18
7	Vijay	976	975	99.90	157	145	92.36
8	Vijay	845	843	99.76	127	118	92.91

Thereafter the character segmentation is carried out. For each scanned document, the segmented characters are manually inspected and the accuracy of the character segmentation is found out, the results are summarized in Table 1. Minimum character segmentation rates of 97.5% and 97.83% for BRH Kasturi are observed. For the other

documents character segmentation rate of more than 99.0% is observed. On comparison with the segmentation technique of [7] in which the character segmentation rate reported as 98.2% and our segmentation method produced character segmentation rate of 99.2% on the average. For vatthu segmentation rate of 92.73% on an average is obtained.

5 Conclusion

In document analysis, a document is segmented into individual lines, words and characters. The Kannada document segmentation poses many challenges due to peculiarities of the script. In Kannada, *aksharas* are quite complicated with considerable variation in widths and heights. Also they are composed of two or more disconnected components. Consonant conjunct which occurs at the bottom of the consonant may sometimes overlap with boundary of next line or next character; causing difficulty to select segmentation position. To overcome all these challenges, we have proposed an efficient Kannada document segmentation method using projection profiles. Experiments have been conducted on scanned Kannada document with different fonts and character segmentation rate of 99.2% and vatthus segmentation rate of 92.73% on an average are observed.

References

1. Nagy, G.: Twenty years of Document Image Analysis. PAMI, IEEE Transactions on Pattern Analysis and Machine Intelligence 22(1), 1, 38–63 (2000)
2. Bansal, V., Sinha, R.M.K.: A survey of OCR in Indian Languages and a Devanagari OCR scheme. In: Proceedings of the STRANS 2001, IIT, Kanpur (2001)
3. Chaudhuri, B.B., Pal, U.: A complete printed Bangla OCR system. Pattern Recognition 3(1), 531–549 (1998)
4. Anna Durai, S.: Tamil character recognition using multilayer neural network. In: Indian Conference on Pattern Recognition, in Image Processing and Computer Vision (ICPIC), pp. 155–160 (1995)
5. Zand, M., Nilchi, A.N., Amirhassan Monadjemi, S.: Recognition-based Segmentation in Persian Character Recognition. WASET 38, 183–188 (2008)
6. Kumar, B.V., Ramakrishnan, A.G.: Radial basis function and subspace approach for printed Kannada text recognition. In: Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 321–324 (2004)
7. Sanjeev Kunte, R., Sudhaker Samuel, R.D.: A two-stage Character Segmentation Technique for Kannada Text. Journal of Graphics, Vision and Image Processing 6, 1–8 (2006)

An Approach for Improving Signal to Interference Plus Noise Ratio in MC DS-CDMA Downlink System

A.K. Gnanasekar, D. Agilandeswari, and V. Nagarajan

Department of Electronics and Communication Engineering,
Adhiparasakthi Engineering College, Melmaruvathur-603 319

Abstract. Transmitter based linear precoding scheme that outperforms conventional precoding by making use of a portion of the interference between the users in a multicarrier direct sequence code-division multiple-access (MC DS-CDMA) system downlink is proposed. A part of interference utilization is achieved by selectively orthogonalizing the desired symbols to destructive interference by means of precoding while allowing interference that constructively contributes to the useful signal's energy. Existence and exploitation of constructive interference effectively spreads the signal constellation and enhances the signal-to-interference-plus-noise ratio (SINR) at the receiver. The scheme introduced in this project applies to the downlink of cellular phase-shift keying (PSK)-based MC DS-CDMA systems.

Keywords: Multi carrier code-division multiple access (MC DS-CDMA), downlink interference utilization, precoding.

1 Introduction

The performance of a multi carrier code-division multiple-access (MC DS-CDMA) system is afflicted with various forms of interference. To reduce the complexity of the mobile units (MUs) of a MC DS-CDMA communication system, the current trend is mitigating these system impediments at the base station (BS) prior to downlink transmission by use of precoding techniques. Various methods have been proposed by a number of researchers toward this end, following an initial idea proposed by Esmailzadeh*et al.* [1] introduced a single-user precoding technique by transferring the Rake processing to the BS. This technique's main advantage is the simplification of the MU's architecture. This is done in the time-division duplex CDMA mode [2], where uplink and downlink have highly correlated. A multiuser transmitter precoding (TP) scheme is presented in [3], similar to the conventional receiver-based decorrelator–detector. A joint pre- and postdecorrelating scheme is introduced in [4] by Vojcic and Jang [3], which improves performance, compared with [3], but limits the complexity reduction at the MU. A further linear-precoding technique based on numerical optimization is introduced in [6], where the restrictions on the received symbols relevant to the decision thresholds of the modulation's constellation are

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 143–148, 2013.

relaxed to attain a more-flexible optimization of the precoding matrix to reduce the transmitted power. Complementary to the preceding methods, further precoding schemes have been investigated in [5].Christos Masouros and EmaudAlusa proposed a system in where precoding is done in the account of constructive interference in [7]. A further achievement is obtained in this system.

2 Selective Precoding

Selective precoding is applied onthe downlink transmission of a multicarrier direct sequence CDMA system. The modification is usually made in the transmitter of the downlink to achieve good performance in the transmission. The Fig.1 represents the block diagram of the proposed system in which binary phase shift keying is used for modulation, the code Walsh Hadamard code for coding, Rayleigh scattering effect and additive White Gaussian noise affects the signal when it passes through the channel.

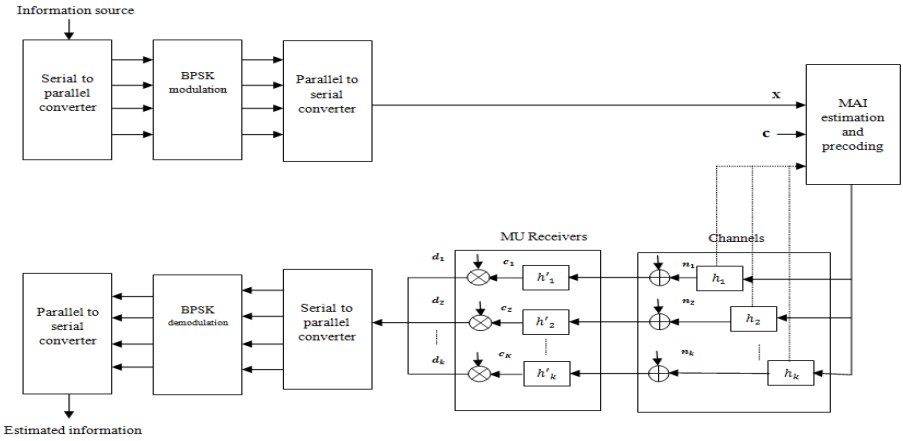


Fig. 1. Block diagram of the proposed system

The channel impulse response is estimated in the downlink since the downlink transmission takes after the uplink transmission which transmits the signal through the same channel. In such case a preferred system implementation is to take N user bits (possibly but not necessarily for different destinations), to transform these using a Walsh Hadamard Transform, followed by an inverse fast Fourier transform.

2.1 Downlink Signal Model

Considering K numberusers, the channels' path delays are assumed to be multiples of the chip intervals with rectangular pulse shapes. All codes are assumed to have a fixed processing gain of L and a normalized energy of one. we use the following definitions: $X^{(i)} = [x_1^{(i)} x_2^{(i)} \cdots x_k^{(i)}]$ is the $1 \times K$ data vector, with the kth element $x_k^{(i)}$ being the modulated symbol of the kth user for the ith symbol period;

$\mathbf{A} = \text{diag}([a_1\ a_2 \cdots a_K\ ])$ is the K × K diagonal matrix of amplitudes, with scalar $a_k$being the amplitude of the kth user. $\mathbf{C} = [C^{(1)}C^{(2)}\ \cdots C^{(K)}]^T$is the K × L matrix containing the users' codes, with $C^{(k)} = [c_1^{(k)}c_2^{(k)} \cdots c_L^{(k)}\ ]$ being the 1 × L code vector of the kth user, where the lth element $c_l^{(k)} \in \{\ -1/\sqrt{L}, +1/\sqrt{L}\ \}$ is the lth chip of the kth user's code sequence, and $H^{(K)}$ is the kth user's channel matrix, with $h_p^{(k)}$being the pth tap of the kth MU's (user's) downlink channel. Using the preceding equation, we can further define $S^{(u)} = [S^{(1u)}S^{(2u)} \cdots S^{(Ku)}]^T$ as the K × (L + P − 1) matrix of the codes of all users corrupted by the multipath channel of the uth MU, with $S^{(ku)} = C^{(k)} \cdot \mathbf{H}^{(u)}$being the code of user k corrupted by the multipath of the uth MU. In addition, let us model the additive white Gaussian nose at the uth MU at the ith symbol period by the 1 × (L + P − 1) vector $\mathbf{n}^{(iu)} = [n_1^{(iu)}n_2^{(iu)} \cdots\ n_{L+P-1}^{(iu)}]$, with the lth element $n_l^{(iu)}$being the noise component at the lth received chip of the uth MU's received signal for the ith symbol period. The received signal at the uth MU at the ith symbol period can then be modeled by the 1 × (L + P − 1) vector

$$r^{(iu)} = X^{(i)}.A.C.H^{(u)} + n^{(iu)} =.A.S^{(u)} + n^{(iu)} \tag{1}$$

Employing Rake processing at the uth receiver, the decision symbol at the ith symbol period can be written as

$$d_u^{(i)} = r^{(iu)}.S^{(uu)H}$$

$$= X^{(i)}.A.C.H^{(u)}.H^{(u)H}.C^{(u)H} + n^{(iu)}.H^{(u)H}.C^{(u)H} \tag{2}$$

Which is effectively the received signal multiplied by the code of user u corrupted by his downlink channel. In (2), the operator $(.)^H$represents the Hermitian transpose of a matrix. The output of the Rake receivers at all MUs can be combined in the 1 × K vector

$$d^{(i)} = X^{(i)}.A.R + \eta^{(i)} \tag{3}$$

Where **R** is the K × K cross correlation matrix of the multipath corrupted non modulated signature waveforms, and $\eta^{(i)} = [\eta_1^{(i)}\eta_2^{(i)} \cdots \eta_K^{(i)}\ ]$, with the uth element $\eta_u^{(i)} = \mathbf{n}^{(iu)}.\mathbf{H}^{(u)H}.C^{(u)H}$being the noise component at the uth Rake output. It is assumed that **R** is positive definite in order for the inverse to exist. The element on the kth row and uth column of the matrix **R** is given by

$$d_u^{(i)} = a_u x_u^{(i)} \rho^{(uu)} + MAI_u^{(i)} + \eta_u^{(i)} \tag{5}$$

with the first term of the right side of the equation being the desired signal and the second term being the MAI at the uth MU at the ith symbol period given as

$$MAI_u^{(i)} = \sum_{k=1,k\neq u}^{K} a_k x_k^{(i)} \rho_{ku} \tag{6}$$

For PSK modulation, the MAI is constructive when it adds to the desired user's signal energy, thus improving the effective signal-to-interference-plus-noise ratio (SINR). Hence, the instantaneous effective SINR can be written as

$$SINR = \frac{s+Jc}{J_d+N} \tag{7}$$

Where $J_c = E\{\sum_{k=1}^{U_c} a_k x_k^{(i)} \rho_{ku}|^2\}$ and $J_d = E\{\sum_{k=1}^{U_d} a_k x_k^{(i)} \rho_{ku}|^2\}$ are the average power of the instantaneous constructive and destructive MAI, respectively, and $U_c$and $U_d$denote the number of constructive and destructive interferers, respectively. For binary phase shift keying (BPSK) modulation, the desired user's symbol $x_u^{(i)} \in \{-1, +1\}$, and therefore, interference from a specific (kth) useris constructive when it has the same sign as the desired data $x_u^{(i)}$.For the user-to-user interference, this translates to

$$\mathrm{Re}\left(a_u . x_u^{(i)}\right).\mathrm{Re}\left(a_k . x_k^{(i)} \rho_{ku}\right) > 0 \tag{8}$$

Therefore, to characterize the MAI, the left part of (8) should be evaluated for all interfering users.

$$M^{(i)} = Re(\mathrm{diag}(X^{(i)}.A).R.\mathrm{Re}\left(\mathrm{diag}(X^{(i)}.A)\right) \tag{9}$$

The equation (9) defines the cross-correlation matrix of the multipath affected signature waveforms. The preceding criterion translates differently to various higher order PSK constellations. Here the orthogonalizing is applied for every user but only users that impose destructive interference to the useful signal at each symbol period. This would completely remove all destructive interference while allowing all constructive interference.

Loop for u=1 to K

Loop for k=1 to K, $k \neq u$

If $[M^{(i)}]_{k,u} < 0$ then $R_{k,u}^c = 0$

Else

$$R_{k,u}^c = \rho_{ku} (10)$$

$$XR^c = [x_1 + \rho_{14}x_4x_2 + \rho_{23}x_3x_3 + \rho_{23}x_2 + \rho_{34}x_4x_4 + \rho_{14}x_1 + \rho_{34}x_3] \tag{11}$$

Obviously, this method aims at enhanced performance by retaining purely constructive MAI, but it requires higher amount of decorrelation. The trade off to useful MAI energy, however, is better than other techniques, which yields better performance. However, since decorrelation is fuller, the complexity is increased.

3 Results and Discussion

The simulation of bit error rate for the DS CDMA system that uses precoding technique is shown in the Fig.2.

The simulation shows the comparison of different PSK techniques and it is efficient in the case of BPSK where the BER is significantly reduced. In the case of QPSK and 8PSK, the reduction of BER is observed but not as much as of BPSK modulation.

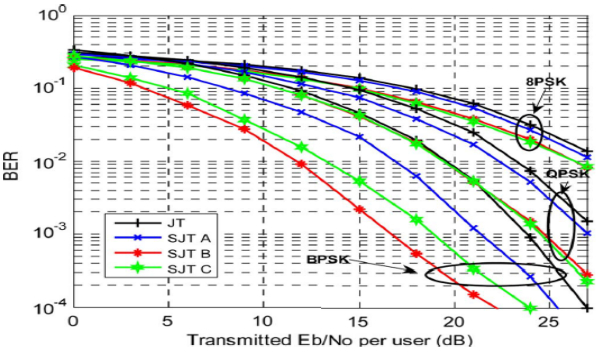


Fig. 2. BER for DS CDMA system

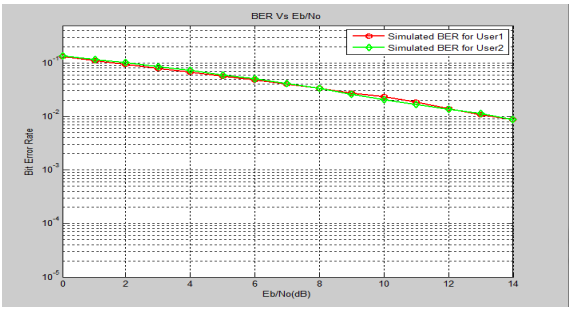


Fig. 3. Non selective precoding technique

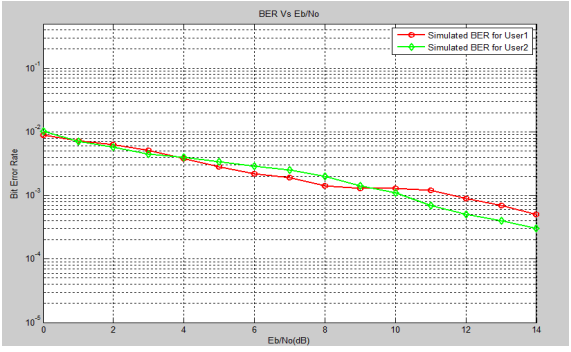


Fig. 4. Selective coding technique

Analytically additive white Gaussian noise (AWGN) provides a qualitative reference for the proposed system performance. The restriction of normalized equal power users is added to the previously described system to further simplify the analysis. The simulation of MC-DS CDMA system for different users is shown in the Fig.3. The BER for the system is 10^{-1} when the signal is transmitted without any

special type of technique with PSK modulation. The channel state information is not considered similar to the transmission in the uplink system. Users signal experience the average bit error rate.

The simulation of MC-DS CDMA system with the proposed precoding technique is shown in the Fig.4. At the transmitter the selective precoding technique is applied and the signals that pass through the channel and reaches the receiver. In Fig 3, the simulation output of MC-DS CDMA system which applies spreading to the entire user's modulated signal. The simulation is done with two users whose signal has to be transmitted to the receivers. The bit error rate of the MC-DS CDMA between users yields10^{-1}.In the Fig.4, the simulation for the selective precoding technique is done where the users data gives constructive interference are passed through the channel without coding and the interference users data is coded . Here the bit error rate is reduced from $10^{-1}$to nearly10^{-2}. Thus the signal to interference plus ratio is improved with Reduced Bit Error Rate.

4 Conclusions and Future Work

The proposed technique offers a significant improvement in yielding an acceptable performance and limiting the need for power control, which is present, even in the majority of receiver-based decorrelation techniques in MC DS-CDMA system. The selective precoding technique used in multi carrier direct sequence code division multiple access system is to reduces the burden in mobile station. it has been shown that a part of the MAI in Multi carrier direct sequence code division multiple access communication systems can be exploited toward BER performance improvement. The bit error rate can be still improved by using efficient data correction technique and increasing the rake fingers at the receiver side by which the signal reach receiver at various directions makes the deduction of transmitted signal efficient.

References

1. Esmailzadeh, R., Sourour, E., Nakagawa, M.: Pre-rake diversity combining in time division duplex CDMA mobile communications. IEEE Trans. Veh. Technol. 48(3), 795–801 (1999)
2. Esmailzadeh, R., Nakagawa, M., Sourour, E.A.: Time-division duplex CDMA. IEEE Pers. Commun. 4(2), 51–56 (1997)
3. Vojcic, B.R., Jang, W.M.: Transmitter precoding in synchronous multiuser communications. IEEE Trans. Commun. 46(10), 1346–1355 (1998)
4. Reynolds, D., Madsen, A.H., Wang, X.: Adaptive transmitter precoding for time division duplex CDMA in fading multipath channels: Strategy and analysis. EURASIP J. Appl. Signal Process. 2002(1), 1325–1334 (2002)
5. Tang, Z., Cheng, S.: Interference cancellation for DS-CDMA systems over flat fading channels through pre-decorrelating. In: Proc. PIMRC, pp. 435–438 (June 1994)
6. Keskin, F., Hahn, J., Baier, P.W.: Minimum energy soft precoding. Electron. Lett. 43(6), 55–56 (2007)
7. Masouros, C., Alsusa, E.: Soft Linear Precoding for the Downlink of DS/CDMA Communication Systems. IEEE Transactions on Vehicular Technology 59(1) (January 2010)

Prediction of Course Selection in E-Learning System Using Combined Approach of Unsupervised Learning Algorithm and Association Rule

Sunita B. Aher[1] and L.M.R.J. Lobo[2]

[1] ME (CSE)-II, Walchand Institute of Technology, Solapur University, India
Sunita_aher@yahoo.com
[2] Department of IT, Walchand Institute of Technology, Solapur University, India
headitwit@gmail.com

Abstract. Clustering also called unsupervised learning algorithm & Association rule algorithm are the data mining techniques which can be used to discover rules & patterns from data. Course recommender system in E-Learning is used to predict the best combination of courses based student's choice. Here in this paper we present how the combination of clustering algorithm- EM clustering Algorithm & association rule algorithm- Apriori Association Rule is useful in Course Recommender system. So we present this new approach & show how its result differs from the result of using only the association rule algorithm.

Keywords: Simple K-means Clustering Algorithm, Apriori association Rule, Weka, Moodle.

1 Introduction

The course recommendation system in e-learning is a system that suggests the best combination of courses in which the students are interested [9] e.g. If student is interested in studying the data structure then he/she is interested in studying the C-programming language also. The architecture of this Course Recommender System & preprocessing technique is already explained in [8].

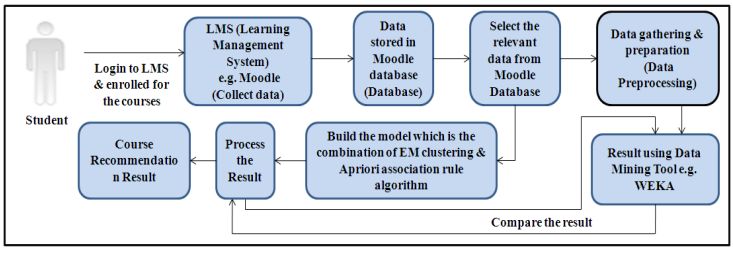


Fig. 1. Course Recommender System in E-Learning

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 149–154, 2013.

In the combined approach of clustering & association rule algorithm, after collecting the data from Moodle database, we clustered the data using clustering algorithm e.g. EM algorithm. After clustering data, we use the Apriori Association Rule algorithm to find the best combination of courses. To find the result using only the Apriori association rule algorithm, we need to preprocess the data from Moodle database but if we consider the combination of clustering & association rule algorithm then there is no need to preprocess the data.

2 Literature Review

In paper [2], they proposed a method for grouping and summarizing large sets of association rules according to the items contained in each rule. They used hierarchical clustering to partition the initial rule set into thematically coherent subsets. This enabled the summarization of the rule set by adequately choosing a representative rule for each subset, and helped in the interactive exploration of the rule model by the user. Rule clusters can also be used to infer novel interest measures for the rules.

In paper [3], they explained Aspect oriented programming which offers a unique module, an aspect, to encapsulate scattered and tangled code, which made it hopeful to solve the problem of crosscutting concerns. A novel aspect mining method which combined clustering and association rule technology is provided in this article.

In the paper [4], they proposed a system that integrated Web page clustering into log file association mining and used the cluster labels as Web page content indicators. They experimented with several approaches to content clustering, relying on keyword and character n-gram based clustering with different distance measures and parameter settings.

The research [5] proposed a bracing approach for increasing web server performance by analyzing user behavior, in this pre-fetching and prediction was done by pre-processing the user access log and integrating the three techniques i.e. Clustering, Markov model and association rules which achieved better web page access prediction accuracy.

In paper [9], they presented a new approach to cluster web images. A tree-distance-based evaluation measure was used to evaluate the quality of image clustering with respect to manually generated ground truth.

3 Expectation Maximization (EM) Clustering Algorithm

An expectation–maximization (EM) algorithm is an iterative method for finding maximum likelihood or maximum a posteriori (MAP) estimates of parameters in statistical models, where the model depends on unobserved latent variables. The EM iteration alternates between performing an expectation (E) step, which computes the expectation of the log-likelihood evaluated using the current estimate for the parameters, and maximization (M) step, which computes parameters maximizing the expected log-likelihood found on the E step. These parameter-estimates are then used to determine the distribution of the latent variables in the next E step [10]. EM assigns a probability distribution to each instance which indicates the probability of it belonging to each of the clusters.

4 Apriori Association Rule Algorithm

Apriori Association rule is used to mine the frequent patterns in database. Support & confidence are the normal method used to measure the quality of association rule. Support for the association rule X->Y is the percentage of transaction in the database that contains XUY. Confidence for the association rule is X->Y is the ratio of the number of transaction that contains XUY to the number of transaction that contain X.

The Apriori association rule algorithm is given as [6]:

Input : Database of Transactions D={t_1,t_2, …, t_n}
Set if Items I= {I_1 ,I_2 ,….I_k}
Frequent (Large) Itemset L
Support,
Confidence.

Output : Association Rule satisfying Support & Confidence

Method :

C_1 = Itemsets of size one in I;
Determine all large itemsets of size 1, L_1;
i = 1;
Repeat
i = i + 1;
C_i = Apriori-Gen(L_{i-1});
Apriori-Gen(L_{i-1})

1. Generate candidates of size i+1 from large itemsets of size i.
2. Join large itemsets of size i if they agree on i-1.
3. Prune candidates who have subsets that are not large.

Count C_i to determine L_i;
Until no more large itemsets found;

5 Result and Implementation

Here we are considering the sample data extracted from Moodle database as shown in figure 2. Here we consider 45 student & 15 courses. We consider the courses like C-programming (C), Visual Basic (VB), Active Server Pages (ASP), Computer Network (CN), Network Engineering (NE), Microprocessor (MP), Computer Organization (CO), Database Engineering (DBE), Advanced Database System (ADS), Operating System (OS), Distributed System (DS), Finite Automata System (FSA), Data Structure (DS-I), Software Engineering (SE), and Software Testing & Quality assurance (STQA). In this table yes represent that the student is interested in that course & no represent that student do not like that course . In preprocessing step, we delete those rows & columns from sample data shown in figure 2, having very less student count & less course count. After preprocessing of data we got 8 courses & 38 rows i.e. 38 students. These 8 courses are C-programming (C), Visual Basic (VB), Active Server Pages (ASP), Computer Network (CN), Network Engineering (NE), Operating System (OS), Distributed System (DS), Data Structure (DS-I).

The result of applying Apriori association rule before & after preprocessing of data is in table 1. Before preprocessing of data, we got the association rule containing “no”

only. As we are recommending the course, we preprocess the data. The result after preprocessing of data is also shown in table 1. Now the association rule contains only "yes". The meaning of the association rule "DS=yes -> OS=yes" is that we can recommend to new student who has recently enrolled for DS course, the operating system as a course to be opted. If we consider combination of clustering & association rule then there is no need preprocess the data. First we apply EM clustering algorithm to data selected from Moodle course. EM algorithm give four clusters out of which only cluster 2 gives the correct association rule containing "yes" only. After clustering of data, we got the data which is shown in figure 3. The result of applying the Apriori association rule on clustered data is shown in table 1.

Courses→ Roll_No↓	C	VB	ASP	CN	NE	MP	CO	DBE	ADS	OS	DS	FSA	DS-I	SE	STQA
1	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	yes	no	no
2	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
3	yes	yes	yes	yes	yes	no	no	no	no	yes	yes	yes	yes	yes	yes
4	no	no	no	yes	yes	no	yes	no	no	no	no	no	no	no	no
5	yes	yes	yes	yes	yes	no	no	yes	no	yes	yes	no	yes	no	no
6	yes	yes	yes	no	no	no	no	no	no	yes	no	no	yes	no	no
7	no	no	no	yes	yes	yes	yes	no	no	no	no	no	no	yes	no
8	no	no	no	no	no	no	no	yes	yes	yes	yes	no	yes	no	no
9	no	no	no	yes	yes	yes	yes	no	no	no	no	yes	no	no	no
10	yes	no	no	no	no	no	no	no	no	no	no	no	no	no	no
11	yes	yes	yes	no	no	no	no	no	no	yes	yes	no	yes	no	no
12	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no
13	no	no	no	no	no	no	no	yes	yes	yes	yes	no	yes	yes	yes
14	yes	yes	yes	yes	yes	no	no	no	no	yes	yes	no	no	no	no
15	yes	yes	yes	no	no	no	no	no	no	no	no	no	yes	no	no
16	no	no	no	yes	yes	no	no	yes	yes	yes	yes	no	yes	no	no
17	yes	yes	yes	no	no	no	no	no	no	yes	yes	no	yes	yes	yes
18	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	no	no
19	no	no	no	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no
20	yes	no	no	no	no	no	no	no	no	yes	yes	no	yes	yes	yes
21	yes	no	yes	no	no	yes	yes	no	no	yes	yes	yes	no	no	no
22	no	no	no	no	no	no	no	yes	yes	yes	yes	no	yes	no	no
23	yes	yes	yes	yes	yes	yes	yes	no	no	yes	yes	no	yes	no	no
24	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
25	no	yes	yes	no	no	yes	yes	yes	yes	yes	yes	no	no	no	no
26	yes	yes	yes	no	no	no	no	no	no	yes	yes	no	yes	no	no
27	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no
28	no	no	no	yes	yes	no	no	no	no	yes	yes	no	yes	no	no
29	no	no	no	no	no	yes	yes	yes	yes	no	no	no	no	no	no
30	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	yes	yes
31	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
32	yes	yes	yes	no	no	no	no	yes	yes	yes	yes	no	yes	no	no
33	no	no	no	yes	yes	no	no	no	no	yes	yes	no	yes	no	no
34	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	no	no
35	no	no	no	no	no	no	no	no	no	yes	yes	no	no	no	no
36	no	no	no	yes	yes	no	no	no	no	no	no	no	yes	no	no
37	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no
38	no	no	no	no	no	no	no	no	no	yes	yes	yes	yes	yes	yes
39	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
40	no	no	no	no	no	no	no	no	no	no	no	no	no	yes	yes
41	yes	yes	yes	no	no	no	no	no	no	yes	yes	no	yes	no	no
42	no	no	no	yes	yes	no	no	no	no	no	no	no	no	no	no
43	no	no	no	no	no	no	no	no	no	yes	yes	no	yes	no	no
44	no	no	no	no	no	no	no	no	no	no	no	no	no	no	yes
45	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no

Fig. 2. Sample data extracted from Moodle Database

Courses→ Roll_No↓	C	VB	ASP	CN	NE	MP	CO	DBE	ADS	OS	DS	FSA	DS-I	SE	STQA
3	yes	yes	yes	yes	yes	no	no	no	no	yes	yes	yes	yes	yes	yes
5	yes	yes	yes	yes	yes	no	no	yes	no	yes	yes	no	yes	no	no
11	yes	yes	yes	no	no	no	no	no	no	yes	yes	no	yes	no	no
17	yes	yes	yes	no	no	no	no	no	no	yes	yes	no	yes	yes	yes
21	yes	no	yes	no	no	yes	yes	no	no	yes	yes	yes	no	no	no
23	yes	yes	yes	yes	yes	yes	yes	no	no	yes	yes	no	yes	no	no
24	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
25	no	yes	yes	no	no	yes	yes	yes	yes	yes	yes	no	no	no	no
26	yes	yes	yes	no	no	no	no	no	no	yes	yes	no	yes	no	no
32	yes	yes	yes	no	no	no	no	yes	yes	yes	yes	no	yes	no	no
39	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
41	yes	yes	yes	no	no	no	no	no	no	yes	yes	no	yes	no	no

Fig. 3. Sample Table after application of clustering algorithm- EM algorithm

Table 1. Result after application of data mining algorithms

Course	Courses considered in table 1	C, VB, ASP, CN, NE, OS, DS, DS-I	
Result	Result of Apriori Association Rule before preprocessing & application of combination of Clustering & Association Rule	Result of Apriori Association Rule after preprocessing & before application of combination of Clustering & Association Rule	
	Minimum support: 0.7 confidence: 0.9 Best rules found: 1. CO=no → MP=no 2.DBE=no → ADS=no 3.CO=no FSA=no → MP=no 4.MP=no → CO=no 5.STQA=no → SE=no 6.SE=no → STQA=no 7.ADS=no → DBE=no 8.MP=no FSA=no →CO=no 9.FSA=no STQA=no → SE=no 10. FSA=no SE=no → STQA=no	Minimum support: 0.5 confidence: 0.9 Best rules found: 1. DS=yes →OS=yes 2. VB=yes →ASP=yes 3. NE=yes → CN=yes 4. CN=yes → NE=yes 5. C=yes VB=yes → ASP=yes 6. DS=yes DS-I=yes → OS=yes 7. OS=yes →DS=yes 8. ASP=yes →C=yes 9. C=yes →ASP=yes 10.ASP=yes →VB=yes	Minimum support: 0.6 confidence: 0.9 Best rules found: 1. DS=yes →OS=yes 2.OS=yes →DS=yes
Course	Courses considered in table 1		
Result	Result After Application of Clustering algorithm-Simple K-means	Result After Application of Clustering algorithm-Simple K-means & Association Rule-Apriori Association Rule(Cluster 2 correct result)	
	Number of Cluster:-1 Seed: 100 Clustered Instances 0 10 (22%) 1 13 (29%) 2 12 (27%) 3 10 (22%)	Minimum support: 0.95 confidence: 0.9 Best rules found 1. OS=yes → ASP=yes conf:(1) 2. ASP=yes → OS=yes conf:(1) 3. DS=yes → ASP=yes conf:(1) 4. ASP=yes → DS=yes conf:(1) 5. DS=yes → OS=yes conf:(1) 6. OS=yes → DS=yes conf:(1) 7. OS=yes DS=yes→ ASP=yes conf:(1) 8. ASP=yes DS=yes → OS=yes conf:(1) 9. ASP=yes OS=yes → DS=yes conf:(1) 10. DS=yes→ ASP=yes OS=yes conf:(1)	

6 Conclusion and Future Work

In this paper, we try to present the combination of EM clustering & Apriori association rule algorithms in Course Recommender System. If we consider only Apriori association rule mining then we need to preprocess the data from Moodle database. But if we use this combine approach of clustering & association rule algorithm then there is no need to preprocess the data. With only the association rule algorithm, as we increase the support, we get the less number of results. But with this combined approach, we find that the combination of clustering & association rule algorithm gives more number of association rules for increased support as compare to the number of association rules got using only the association rule. Future work includes the atomization of this combination algorithm & testing the result on data.

References

1. Castro, F., Vellido, A., Nebot, A., Mugica, F.: Applying data mining techniques to e-learning problems: A survey and state of the art. In: Jain, L.C., Tedman, R., Tedman, D. (eds.) Evolution of Teaching and Learning Paradigms in Intelligent Environment. SCI, vol. 62. Springer, Heidelberg (in press)
2. Jorge, A.: Hierarchical Clustering for thematic browsing and summarization of large sets of Association Rules. Supported by the POSI/SRI/39630/2001/Class Project
3. He, L., Bai, H.: Aspect Mining Using Clustering and Association Rule Method. IJCSNS International Journal of Computer Science and Network Security 6(2A)
4. Guo, J., Kešelj, V., Gao, Q.: Integrating Web Content Clustering into Web Log Association Rule Mining? Supported by NSERC
5. Makker, S., Rathy, R.K.: Web Server Performance Optimization using Prediction Prefetching Engine. International Journal of Computer Applications (0975 – 8887) 23(9) (June 2011)
6. Dunham, M.H.: Data Mining Introductory and Advanced Topics
7. Aher, S.B., Lobo, L.M.R.J.: Data Mining in Educational System using WEKA. In: IJCA Proceedings on International Conference on Emerging Technology Trends (ICETT), vol. (3), pp. 20–25. Published by Foundation of Computer Science, New York (2011)
8. Aher, S.B., Lobo, L.M.R.J.: Preprocessing Technique for Association Rule Based Course Recommendation System in E-learning. Selected in ICECT 2012, Proceeding published by IEEE (2012)
9. Malik, H.H., Kender, J.R.: Clustering Web Images using Association Rules, Interestingness Measures, and Hypergraph Partitions. In: ICWE 2006, Palo Alto, California, USA, July 11-14. ACM 1-59593-352-2/06/0007 (2006)
10. `http://en.wikipedia.org/wiki/Expectation%E2%80%93maximization_algorithm` (accessed on January 23, 2012)

Context-Aware Information Processing in Visual Sensor Network

H.H. Kenchannavar[1], Shridhar G. Domanal[2], and Umakant P. Kulkarni[3]

Visvesvaraya Technological University, Belgaum
[1] Department of Computer Science and Engineering
Gogte Institute of Technology, Belgaum
[2] Maratha Mandal Engineering College, Belgaum
[3] SDMCET, Dharwad

Abstract. A visual sensor network is one of the streams of sensor network in which an image from the sensor node is transmitted to the server to be processed further without human intervention. Object Recognition in specific applications such as agriculture, defense etc, requires only the object of importance to be captured and transmitted to the processing center. This can be done by training the visual sensor node using the multilayer feed-forward technique. In the proposed work, a fruit object recognition system has been developed using the multilayer feed-forward technique of the neural network by extracting features from the sample fruit images. The experimental results reveal average recognition rate of 94.23%.

Keywords: Wireless sensor network, Neural network, pattern recognition.

1 Introduction

Wireless sensor networks being the current research topics find wider applications in surveillance systems, weather monitoring, target detection and classification etc. Visual sensors have the advantage of incorporating the image processing and machine learning techniques. This feature helps in improving fruit detection and classification depending on their quality. Limitations of visual sensors are that the resources of energy, memory, computational speed and bandwidth are severely constrained [1]. Hence wireless visual sensor networks should be designed such that, the lifetime of sensor nodes is maximized. Fruit detection system is primarily developed for robotic fruit harvesting. However this technology can easily be tailored for other applications such as on tree yield monitoring, crop health status monitoring, disease detection, and other operations which require a visual sensor. Object Recognition has become the real challenge for the detection of objects. Many recognition techniques used for object detection are being explored to achieve the human recognition level for tens of thousands of categories under a wide variety of conditions.

Currently, sorting of fruits mostly depends on manual method which results in low efficiency, and on annual basis requires lots of labor and financial assistance. During 1970s, machine vision was widely used in the field of varietal recognition because of

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 155–162, 2013.

the advancements in the field of image processing, lower cost and higher speed of computer and its many features [2]. Machine vision system was mainly used for obtaining information of the working environment and performed detection and orientation of the operating target.

Various methods for image recognition have been presented by many researchers. Image matching is one of the most widely used techniques. To build an energy model that accurately defines the energy consumption of sensor node is extremely difficult while designing a protocol [5]. In this work we propose a fruit recognition approach based on images, which combines three different categories like shape, texture and color features. An attempt has been made to develop an efficient recognition system for wireless visual sensor network combining the aforementioned features. A theoretical approach gives an idea as to how the algorithm helps in increasing the WSN lifetime.

2 Related Work

The process of using the high level information of an object is known as Pattern classification, which includes assigning an object to a category of a class using the extracted features. In computer science and engineering, Pattern recognition is an important area which aids in classifying/recognizing the different patterns, and also includes sound and vision patterns. Energy consumption is the core issue in the WSN. Combined effect of object recognition and training of sensors will lead to the enhancement in the WSN lifetime.

Woo Chaw Seng et al. [2] described new fruit recognition, which combines three features analysis methods: color-based, shape-based and size-based in order to increase accuracy of recognition. The proposed method classifies and recognizes fruit images based on the obtained feature values by using nearest neighbors classification. The proposed fruit recognition system analysis classifies and identifies fruits successfully up to 90% accuracy. Shah Rizam M. S. B. et al. [3] describes the watermelon ripeness determination using image processing techniques and artificial neural networks. The significance of this study is to determine the ripening stages of the watermelon with the mean value that is obtained from the image by using color space of the watermelon. The proposed work is successful up to 86.51% in determining the ripeness of watermelon. Hetal N. Patel et al. [4] proposed a fruit detection using improved multiple features based algorithm. Different features such as intensity, color, orientation and edges are extracted using the image processing toolbox. The Detection Efficiency is achieved up to 90% for different fruit image on tree, captured from different positions. The input images are the sections of tree images captured. This proposed approach can be applied for targeting fruits for robotic fruit harvesting. Hai-Ying Zhou et al. [5] describes the current simulation models and platforms such as OPENET, NS2, SHAWN, SensorSim, EmStar, OMNet, GloMoSim etc does not give appropriate results of energy consumption in sensor nodes and its applications. This paper addresses on WSN energy model and aiming to provide accurate energy models for simulations to be carried out.

3 Proposed Method to Detect Fruit

The fruit detection technique consists of mainly 3 stages – preprocessing, feature extraction and training using the multilayer feed-forward neural network. The algorithm for the proposed work is

Input data:
Training parameter set: **x1, x2,……x11** features from **a1,a2,** and **a3** categories
Multilayer feedforward back propagation network
Transfer functions: *tansig, trainlm, purelin and learngdm*

Output data:
Z1 and **Z2** are 2 bit values to recognize a particular fruit

Procedure: Training
Extract features **x1** to **x11** from fruit and store it in training file
Training the neural network with: **x1, x2,……x11** extracted features
Repeat **N** times

Testing
Extract features **x1** to **x11** and store it in testing file
Z1 and **Z2** values gives required output

Assumptions:
Fruit images are taken in a closed room. (No lighting effects)
Fruits are placed on white surface
Images are taken from fixed height i.e 15 inches with the help of stand. (Figure 7)
All images are of same size

In this work, before feature extraction, an object has to be preprocessed and it includes cropping of an image, RGB (figure 1) to B/W (figure 2) conversion, Edge detection (figure 3) and lastly Filling (figure 4).

The proposed work is carried out using MATLAB software. Converting an image in to gray scale has several advantages in finding out the different features accurately. To find out a particular shape of an object, detecting an edge is necessary, here the edge is detected by using *canny* method. The area covered by an image is easily determined by filling.

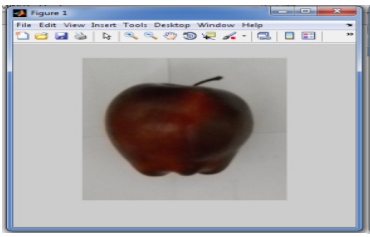


Fig. 1. (RGB image)

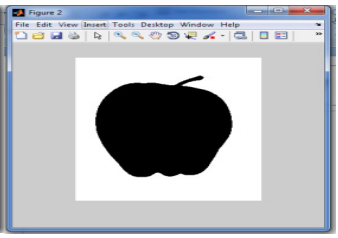


Fig. 2. (RGB to B/W)

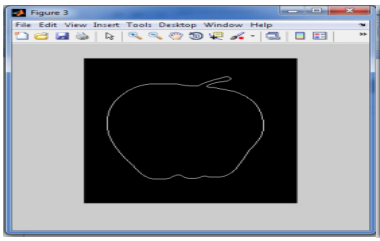


Fig. 3. (Edge detection)

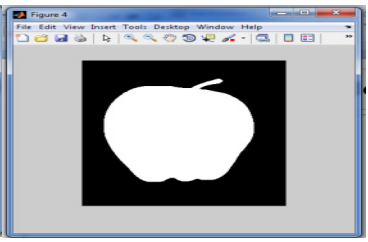


Fig. 4. (Filling)

3.1 Feature Extraction

In the proposed work, features are taken from mainly three categories like Shape, Texture and Color respectively. In the shape category we have extracted features like Area, Majoraxislength, Minoraxislength and Perimeter. These will act as the external features of an object. These features are helpful for the detection at the initial level, as they hold only for physical appearance. The Table 1 shows the sample shape values of fruits.

Table 1. Shape values

Fruits	Area	Majoraxis length	Minoraxis length	Perimeter
Apple	28046	207.29	173.55	667.41
Orange	22573	172.76	166.44	562.35
Sapota	19923	161.54	157.38	523.87

The second category is the Texture. From this category we have extracted features such as Contrast, Energy, Homogeneity and Entropy. Since many years Texture plays a lead role in pattern analysis. It differentiates the pattern by visual properties of objects and background, difference in the intensity of pixels. The following Table 2 shows the texture values of three different fruits.

Table 2. Texture values

Fruits	Contrast	Energy	Homogeneity	Entropy
Apple	0.0655	0.2559	0.9672	6.5958
Orange	0.0476	0.3712	0.9762	6.2086
Sapota	0.0554	0.2917	0.9723	6.9332

From the color category we have extracted the features like Hue, Saturation (from HSV color space) and Blue (from RGB color space). From Hue and Saturation we can get the dominance of a color that we readily experience when we look at the color of an object. For example Blue signifies the blue color content from the color of an object. The Table 3 shows the color values for three different fruits.

Table 3. Color values

Fruits	Hue	Saturation	Blue
Apple	0.6479	0.2828	121.59
Orange	0.0954	0.3912	98.21
Sapota	0.6234	0.3337	118.27

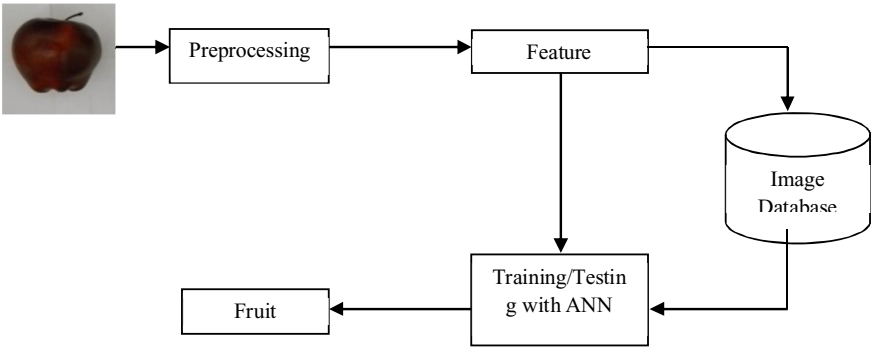


Fig. 5. Fruit detection system

3.2 Training with Artificial Neural Network

A neural network is a massively parallel distributed processor made up of simple processing units, which has the natural propensity for storing the experimental knowledge and making it available for use [09]. The network model used for training of fruits is as shown in figure 6 [11].

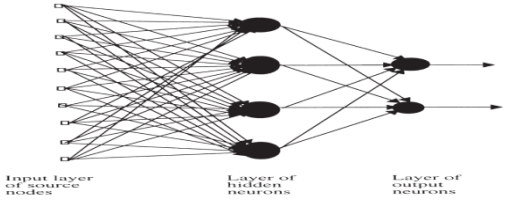


Fig. 6. Multi layer feed-forward network

The input layer signifies the no. of inputs given for training, in the present work eleven inputs will act as source nodes. Four hidden neurons are present in the middle layer and in the output layer we have used two bits.

There are four types of inputs (fruits) given to the learning machine: they are Apple, Orange, Sapota and other things which are not fruits. Eleven different features are extracted from each of the type and are used for training. Testing is done by the query image and it will be identified by the ANN.

The functions that determine the I/P and O/P behaviors are called transfer functions. The transfer functions used in the proposed work are *tansig, trainlm, purelin and learngdm*. And for creating our network we have used *newff* training function. Levenberg-Marquardt backpropagation algorithm is used for training and it is considered to be fastest, supervised backpropagation algorithm.

4 Experimental Results

To carry out the experiment we have used the hardware device for taking the fruit image is as shown in figure 8. The digital camera used for taking image is *Nikon coolpix* L20 with 10 megapixels resolution and 3.6 zoom and has 3inch LCD. To carry out the simulation we have used MATLAB software. In this work, for three different types of fruits a separate database has been created, and they are used for further processing.

Fig. 7. Camera stand

Table 4. Results of proposed system

Sl.No	Fruits	Training fruit images	Testing fruit images	Recognition rate
1	Apple	50	71	94.36%
2	Orange	40	60	93.33%
3	Sapota	40	60	95%

The table 4 contains the fruit image database and recognition rate of the proposed system. The proposed system achieves an average recognition rate of 94.23%.

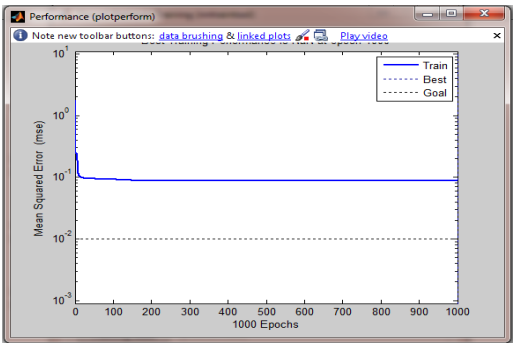


Fig. 8. Simulation result of training the ANN

The simulation results for multilayer feed-forward neural network used to train the data set. More specifically Figure 8 depicts the performance of the ANN for a particular training data. The tern "Train" specifies the network's performance according to the mean of squared errors, which resembles the more favorable nature of training indicated by the term "Best" in the above graph.

5 Energy Factor

In this work we have considered three types of energy consumption. Namely Sensor sensing, Sensor logging and Micro controller processing [11]. The amounts of energy consumed by the three energy sources are given by the following equations.

$$E_{sens_N}(b) = bV_{sup}I_{sens}T_{sens} \quad (1)$$

$$E_{loggN}(b) = E_{write} + E_{read} = \frac{bV_{sup}}{8}(I_{write}T_{write} + I_{read}T_{read}) \quad (2)$$

$$E_{proN}(b, N_{cyc}) = b_1 N_{cyc} C_{avg} V_{sup}^2 + b_1 V_{sup}\left(I_0 e^{\frac{V_{sup}}{n_p V_t}}\right)\left(\frac{N_{cyc}}{f}\right) \quad (3)$$

The energy consumed by different energy sources in single iteration is as follows.

Sensor sensing: 0.675 W

Sensor logging: 8.74 μW

Microcontroller processing: 51.94 W

In a single iteration energy consumed by a single sensor node is 52.615 W. In a particular environment 10 sensor nodes are deployed, the energy consumed by these ten sensor nodes will be 526.15 W. The lifetime of the sensor network can be increased by training the sensor nodes, to increase the sensor network lifetime, if we train the four sensor nodes then the amount of energy consumed by them will be 210.46 W. By this 315.69 W energy can be saved because the other six nodes will be in sleep/inactive mode. So with the proposed method the lifetime of the sensor network can be increased considerably.

6 Conclusion

The proposed experimental system for visual sensor node is trained to detect fruits. An experiment has been carried out by extracting features from different categories. An analysis has been done by taking each category feature individually as well as by combining all the features. More accurate results have been drawn considering all the features rather than taking them individually. An efficiency of 94.23% has been achieved with the proposed work. The proposed method can be applied to sensor nodes so that it increases the lifetime of the sensor network considerably.

References

1. Chen, F.: Simulation of wireless sensor nodes using SMAC, Master's thesis, Department of Computer Science, University of Erlangen-Neuremberg (September 2005), http://dcg.ethz.ch/theses/ss05/mics-embedding-report.pdf
2. Seng, W.C., Mirisaee, S.H.: A New Method for Fruits Recognition System. In: 2009 International Conference on Electrical Engineering and Informatics, Selangor, Malaysia, August 5-7 (2009)
3. Shah Rizam, M.S.B., Farah Yasmin, A.R., Ahmad Ihsan, M.Y., Shazana, K.: Non-destructive Watermelon Ripeness Determination Using Image Processing and Artificial Neural Network (ANN). International Journal of Electrical and Computer Engineering 4(6) (2009)
4. Patel, H.N., Jain, R.K., Joshi, M.V.: Fruit Detection using Improved Multiple Features based Algorithm. International Journal of Computer Applications (0975 – 8887) 13(2) (January 2011)
5. Zhou, H.-Y., Luo, D.-Y., Gao, Y., Zuo, D.-C.: Modeling of Node Energy Consumption for Wireless Sensor Networks. Journal of Scientific Research, Wireless Sensor Network 3, 18–23 (2011)
6. Song, W.-G., Guo, H.-X., Wang, Y.: A Method of Fruits Recognition Based on SIFT Characteristics Matching. In: 2009 International Conference on Artificial Intelligence and Computational Intelligence (2009)
7. Jiménez, A.R., Jain, A.K., Ceres, R., Pons, J.L.: Automatic fruit recognition: A survey and new results using Range/Attenuation images. Pattern Recognition 32(10), 1719–1736 (1999)
8. Yang, L., Dickinson, J., Wu, Q.M.J., Lang, S.: A Fruit Recognition Method for Automatic Harvesting. IEEE (2007)
9. Basu, J.K., Bhattacharyya, D., Kim, T.-H.: Use of Artificial Neural Network in Pattern Recognition. International Journal of Software Engineering and Its Applications 4(2) (April 2010)
10. Zilan, R., Barceló-Ordinas, J.M., Tavli, B.: Image Recognition Traffic Patterns for Wireless Multimedia Sensor Networks. EuroNGI Network of Excellence and CICYT TEC2004-06437-C05-05
11. Halgamuge, M.N., Zukerman, M., Ramamohanarao, K., Vu, H.L.: An estimation of sensor energy consumption. Progress In Electromagnetics Research B 12, 259–295 (2009)

Software Program Development Life Cycle Standards for Banking and Financial Services IT Industry

Poyyamozhi Kuttalam[1], N. Keerthika[2], K. Alagarsamy[3], and K. Iyakutti[4]

[1] Project Manager, Larsen and Toubro Infotech
poyyamozhi@gmail.com
[2] Department of Information Technology, Sethu Institute of Technology, India
ns.keerthika@gmail.com
[3] Computer Center, Madurai Kamaraj University, India
alagarsamymku@gmail.com
[4] School of Physics, Madurai Kamaraj University, India
iyakutti@gmail.com

Abstract. In general we have Software Program Development Life Cycle standards for the IT projects in common, considering all the industries or domains or businesses. This Program Lifecycle outlines the structure by which a program manager can effectively plan, execute, monitor and control a program and the projects of which it is comprised. It defines high level activities that can be applied to all organizations within a Banking and Financial Services Company with the ability to add line of business level activities and deliverables to meet the needs of that particular business.

Keywords: Software program development, bfs software program structure, Program development standards.

1 Introduction

This Program Management is the process of managing multiple ongoing projects which may be inter-dependent. This management includes the coordination and prioritization of cross-functional resources to obtain the desired benefits of the program. [1]. This defines a standard set of delivery lifecycles across all Banking and Financial Services Industry's Line of Business as well as terminology and high level processes. In this effort, programs are defined with the following criteria:

- Scale – Large / Extra Large (Size categories defined at LOB level)
- Breadth – Multiple processes, multi-disciplinary, multiple sponsors, multi-project and/or multiple LOB impacts
- Complexity – High (Significant Architectural changes, independent testing team)
- Technology – Re-engineering, re-architecting, multiple conversions
- Strategic Importance – Capital expansion, repositioning, executive sponsorship
- Governance and Oversight – Steering Committee and/or Guidance Management Team

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 163–180, 2013.

- Release Structure – Quarterly
- Examples – Acquisitions or mergers, new business ventures, or new technology platform
- Each LOB will utilize the Program Criteria as defined and add additional criteria specific to their LOB as necessary for proper categorization.

1.1 Audience

The intended audience for the Program Lifecycle is program directors, managers, and Project Management Office.

1.2 Scope

The scope of this document will be the program lifecycle structure and controlling processes. These include:

- Program Phases
- ETVX Model
- Program Phase Gates
- Activities and Deliverables

1.3 Out of Scope

The following items are out of scope of this document:

- LOB level implementation processes
- LOB specific activities and deliverables
- Roles and responsibilities

1.4 Assumptions

The following assumptions apply:

- Prioritization model ensures project alignment with business objectives
- Project Sponsors engaged to provide direction, address escalated issues and mitigate risks
- Quality Assurance practices will support the proposed model

2 Related Supporting Documents

2.1 BFS IT MP

The Banking and Financial Services Information Technology Management Policy (BFS IT MP) is the BFS standard for managing Information Technology within the

organization; the document includes Architecture, Change Management, Contracting and Outsourcing, Information Security, Problem Management, Project Management, Resource Management, and Software Management. Many of these topics touch upon activities executed during the program lifecycle. [4] [9]

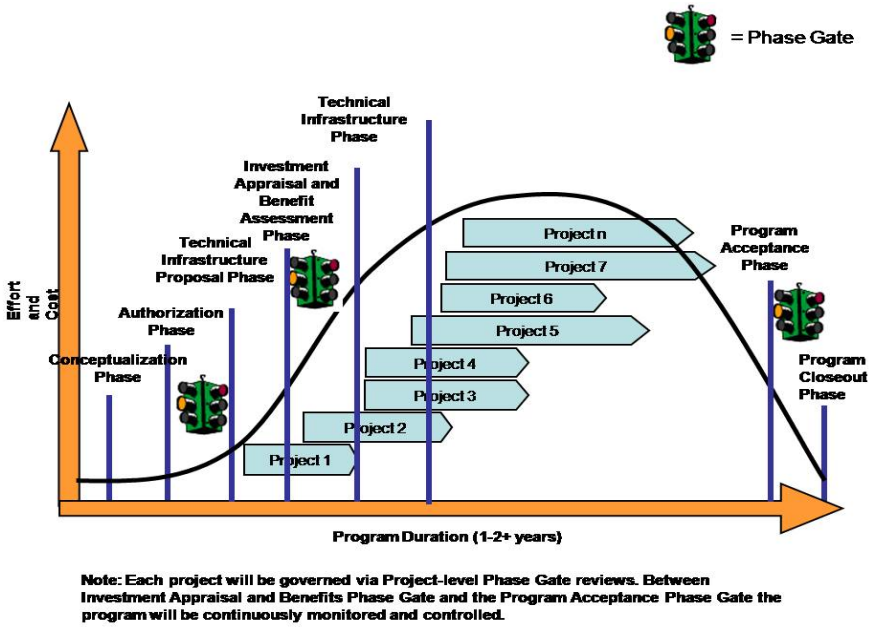
2.2 Program Management Methodology Standard (PMMS)

The Program Management Methodology Standard document is a partner to this effort. The management methodologies defined in that document can be applied to manage the activities and deliverables in this effort. [1] [2] [3]

2.3 Program Management Handbook

The Program Management Handbook outlines a standard framework for managing and delivering enterprise-wide flagship initiatives and programs. The content of the document is based on a combination of internal and external leading practices. These leading practices are tightly aligned with industry standard methodologies published by the Project Management Institute (PMI). [2]

3 Overview



Note: Each project will be governed via Project-level Phase Gate reviews. Between Investment Appraisal and Benefits Phase Gate and the Program Acceptance Phase Gate the program will be continuously monitored and controlled.

3.1 Program Life-Cycle Structure

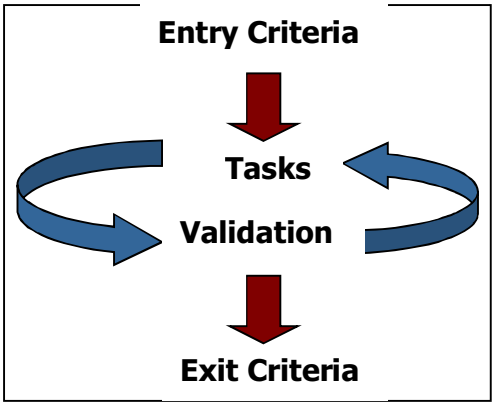
The ETVX (Entry, Tasks, Verification and Validation, and EXit) model outlines a structured approach for accomplishing activities and ensuring that tasks meet defined criteria. By utilizing this model, at each phase in the Program Lifecycle it creates a "soft stop" allowing activities of following phases to be started without the a final signoff of all required deliverables of the past phase. [5] [6] [7] [8]

Entry Criteria: A checklist of conditions that must be satisfied before the beginning of phase activities.

Tasks: A set of tasks that need to be carried out.

Verification and Validation: A list of validation tasks to verify the work items produced by the activity.

Exit Criteria: A checklist of conditions that must be satisfied before the activity is completed.



3.2 Program Life-Cycle Structure

Program gates are placed at critical locations in the program lifecycle. They represent "hard stops" where no work in later phases can be started without the completion of all deliverables of phases up to that Phase Gate. It provides the program manager and release managers with an opportunity to closely review the progress and quality of the program up to that point to ensure that it is meeting its objectives and quality standards. [5] [6] [7] [8]

3.3 Monitoring and Controlling

Monitoring and controlling is done over all major program variables; cost, time, scope and quality of deliverable throughout the program lifecycle. This is an extension of the monitoring and controlling that is done at the project level. The health of each of

the projects is reported to the Program Manager to be used for determining the health of the program. [5] [6] [7] [8]

4 Time Tracking

Time tracking at the Program level will be done via Time Tracking system used in the organization where program time will be logged against a project specific to the time and deliverables of the program. [10]

5 Conceptualization Phase

5.1 Purpose

The purpose of this phase is for the program to undergo analysis of feasibility and program design which includes the program charter, initial cost benefit analysis and risk assessments. Upon completion of this phase the program should be at a state where it is ready to be pitched for sponsorship and business approval. [10] [11] [12]

5.2 ETVX

5.2.1 Entry Criteria

The entry criteria's are

- Business Need - A business problem or initiative that requires the management methodologies of a program. This is used to create the "Solution Concept".
- Solution Concept (The Program) - The end result of the solution concept is what the program will be created to deliver.

5.2.2 Tasks

a. Activity: Create Program Charter

Purpose: Define the reason for the program, objectives and scope of the program.
Items to be included in the charter:

- Business Case
- Benefits
- Problem Statement
- Goal Statement
- High-level Program Statement of Scope
- Program Assumptions and Limitations
- High-level Program Timeline

All are required deliverables

b. Activity: Perform Feasibility Study

Purpose: Review the program technology plans to ensure that it fits with the current technology roadmap and is feasible to implement within the environment. This is a required deliverable.

c. Activity: Perform Initial Cost Benefit Analysis

Purpose: Initial review to determine if the benefits to be obtained from the program will outweigh the costs incurred from it or by not doing it. This is a required deliverable.

d. Activity: Perform Initial Risk Assessment

Purpose: Initial review to determine if the risks that will be taken while executing the program are within acceptable tolerance levels and can be appropriately mitigated to ensure program success. This is a required deliverable.

5.2.3 Validation and Verification

Action - Definition of what is to be accomplished has been documented

5.2.4 Exit Criteria

The deliverables are:

- Program Charter - Required
- Feasibility Study - Required
- Initial Cost / Benefit Analysis – Required
- Initial Risk Assessment – Required

6 Authorization Phase

6.1 Purpose

The purpose of this phase is aimed defining the necessary components to obtain business approval. Such things include proof of concept, level 0 cost and effort estimates, governance model, and sponsorship. Upon completion of this phase the program should be ready for a technical infrastructure proposal. [10] [11] [12]

6.2 ETVX

6.2.1 Entry Criteria

The entry criteria's are,

- Completed Program Charter
- Completed Feasibility Study
- Completed Initial Cost Benefit Analysis
- Completed Initial Risk Assessment

6.2.2 Tasks

a. Activity: Define SC and obtain sponsorship

Purpose: The Steering Committee (SC) is comprised of senior level people within the IT organization, this may include the CIO and directors as well as senior level people from Operations and Front-End. They are responsible for providing the sponsorship of the program and the high level of management oversight. This is a required deliverable.

b. Activity: Define Program Governance Model

Purpose: The Program Governance Model (PGM) provides the governance structure of the program including roles and responsibilities. This is a required deliverable.

c. Activity: Obtain ITLC notification and authorization

Purpose: Information Technology Leadership Council notification and authorization is required for all projects/programs that involves excess budget. This deliverable is complying as necessary.

d. Activity: Define program setup and organization

Purpose: The program setup and organization is used to define the structure of the overall program and those involved. The Program Management Handbook should be utilized for this activity. This is a required deliverable.

e. Activity: Determine Level 0 cost and effort estimate

Purpose: Level 0 estimates are used to provide an initial view of the cost and effort of the program based on the high level information available. This is a required deliverable.

f. Activity: Perform resource skill assessment

Purpose: An assessment of current workforce's skills and those required to complete to program must be done. If there are gaps in the skills available and those required investigation should be done to determine if external contracts should be issued to fill the gaps or provide training for the existing resources. A plan of how to obtain the cross-functional resources should be created. This is a required deliverable.

g. Activity: Obtain pre-funding approval

Purpose: Pre-funding approval is the "go-ahead" for doing in-depth research, program structuring, vendor selection, etc. and the use of resources to perform such activities. This is a required deliverable.

h. Activity: Perform proof of concept

Purpose: This in-depth research is done to determine the best method(s) by which to obtain the benefits of the program. This may include Vendor-Software selection and technology mapping to determine integration capabilities. The need for this is to be determined by the Steering Committee. This is a required deliverable.

i. Activity: Obtain architecture review and approval

Purpose: The Chief Technology Office must sign off on the proof of concept and technologies being used to ensure they align with the technology roadmap. This deliverable is complying as necessary.

j. Activity: Obtain vendor approvals

Purpose: Vendor approvals are to show that vendor selection activities have been completed and approved before working with a new vendor. This deliverable is complying as necessary.

6.2.3 Validation and Verification

The actions are,

- Review of the Program Governance Model and program setup should be reviewed by the PMO to ensure compliance with the Program Lifecycle and Management Methodology Standard.
- Resource skill assessments should have a plan of how to obtain the required skilled resources if they are not present or available in the current environment.
- The proof of concept should have details regarding the technology being leveraged to obtain the benefits of the program in the most efficient manner as well as the details of the technologies that were not selected to minimize future proof of concept activities for related projects/programs.

6.2.4 Exit Criteria

The Deliverables are:

- SC and Sponsorship - Required
- Program Governance Model - Required
- ITLC Notification / Authorization - Comply As Necessary
- Program Setup and Organization - Required
- Cost and Effort Estimate (level 0) - Required
- Resource Skill Assessment - Required
- Pre-Funding Approval - Required
- Proof of Concept - Comply As Necessary
- Architecture Review and Approval - Comply As Necessary
- Vendor Management – Vendor Approvals - Comply As Necessary

7 Technical Infrastructure Proposal Phase

7.1 Purpose

This is meant to develop the complete plan for the technical infrastructure. This includes defining plans for architecture, development environment, facilities, obtaining vendor contracts, and offshore plans. [10] [11] [12]

7.2 ETVX

7.2.1 Entry Criteria

The entry criteria's are,

- SC and Sponsorship is available
- Program Governance Model is complete
- ITLC Notification / Authorization is complete if required
- Program Setup and Organization is complete

- Cost and Effort Estimate (level 0) is complete
- Resource Skill Assessment is complete
- Pre-Funding Approval is complete
- Proof of Concept is complete if required
- Architecture Review and Approval is complete if required
- Vendor Management – Vendor Approvals is complete if required

7.2.2 Tasks

a. Activity: Define plans for architecture, development environment and facilities

Purpose: Document plans for the technical architecture of the program, with collaboration from the CTO architects to ensure compliance with the technology roadmap, the development environment and the facilities required such as office space or data center space for completion of the program. This is a required deliverable.

b. Activity: Obtain vendor contracts

Purpose: Obtain signed agreements with the selected vendors to provide specific services to products to the program. This is complying as necessary.

c. Activity: Define offshore plan, security and organization

Purpose: Document plans for the use of offshore resources i.e. development teams, required security, and organization. This is a required deliverable.

7.2.3 Validation and Verification

The actions is,

All plans and agreements have been documented and stored in the proper knowledge base.

7.2.4 Exit Criteria

The Deliverables are:

- Plans for architecture, development environment and facilities - Required
- Vendor Management – Vendor contract - Comply as Necessary
- Offshore plan, security and organization – Required

8 Investment Appraisal and Benefit Assessment Phase

8.1 Purpose

The purpose of this phase is to perform Investment and benefits assessments which are required to determine if the structure of the program will provide the benefits that it is designed to and to ensure that the investment required obtaining these benefits falls within a tolerable range. [10] [11] [12]

8.2 ETVX

8.2.1 Entry Criteria

The entry criteria's are,

- Plans for architecture, development environment and facilities are complete
- Vendor Management – Vendor contract are complete
- Offshore plan, security and organization are complete

8.2.2 Tasks

a. Activity: Obtain projected ROI

Purpose: Return On investment is required to determine if the program will provide benefits that compensate for the investment required by the program. This is a required deliverable.

b. Activity: Define incremental benefits plan

Purpose: The incremental benefits plan is to state the expected benefits that are to be obtained over the duration of the program as each of the phases within the program is implemented. This is a required deliverable.

c. Activity: Define fund allocation plan

Purpose: The fund allocation plan defines how the funds from the CEP are to be allocated over the program and how the spending is tracked. This is a required deliverable.

d. Activity: Obtain approved CEP

Purpose: The CEP represents a commitment to the program from the business and provides the capital funding required continuing the program through completion. This is a required deliverable.

8.2.3 Validation and Verification

The action is,

Ensure that the Incremental Benefits Plan, Fund Allocation Plan, ROI and CEP are stored in the program knowledge base.

8.2.4 Exit Criteria

The Deliverables are:

- ROI Validated - Required
- CEP Approved - Required
- Incremental benefits plan - Required
- Fund allocation plan – Required

9 Technical Infrastructure Phase

9.1 Purpose

This phase's purpose is to create the technical infrastructure required for the program and define the COB plan. [10] [11] [12]

9.2 ETVX

9.2.1 Entry Criteria

The entry criteria's are,

- ROI projection is complete
- CEP Approved
- Incremental benefits plan is complete
- Fund allocation plan is complete

9.2.2 Tasks

a. Activity: Create architecture and development environment

Purpose: Install the architecture and development environment required for program development and testing. It is a required deliverable.

b. Activity: Define technology processes

Purpose: Define the processes that are to be used to construct and support the new technologies being implemented. It is a required document.

c. Activity: Create program COB plan

Purpose: Define the Continuity of Business plan for this program ensuring it meets the COB requirements. It is a required deliverable.

d. Activity: Create program facilities

Purpose: If facilities are required for the program, acquire and setup the environment for the needs of the program. It is comply as necessary.

9.2.3 Validation and Verification

The actions are,

- Validate the architecture and development environment meet the requirements of the program.
- Validate that the technology processes and COB plans are complete and stored in the specified program knowledge base.
- If program facilities are required, validate that they meet the requirements of the program.

9.2.4 Exit Criteria

The Deliverables are:

- Architecture and development environment - Required
- Technology processes - Required
- Program COB plan - Required

Program facilities - Comply as necessary

10 Organizational Readiness and Quality Assurance Plan Phase

10.1 Purpose

This phase's purpose is to provide the business preparation, communication and oversight to the impacted components of the business for the impending cultural changes that will be introduced by the program. [10] [11] [12]

10.2 ETVX

10.2.1 Entry Criteria

The entry criteria's are,

- Program Charter is completed
- Risk Assessment is completed and available

10.2.2 Tasks

a. Activity: Discovery sessions and concept days

Purpose: Executive level discussions where they decide to move forward with the program. R

b. Activity: Quality Assurance Plan

Purpose: Reviews of the scripts that are to be used to validate the components of the program to ensure quality. It is a required deliverable.

c. Activity: Blueprinting/Gap analysis

Purpose: To determine and document the gaps between the current system and the proposed system as well as business operational processes and determine a resolution and ensure system gaps are incorporated into requirements. It is a required deliverable.

d. Activity: Training Plan

Purpose: Defining and documenting the training plan and policy as it pertains to the program. It is a required deliverable.

e. Activity: Event Management Model

Purpose: The planning and documenting of the event, command centers, issue reporting, etc. It is a required deliverable.

f. Activity: Functional Model

Purpose: Document the equipment and operational and business functions that you have and the changes that are being made to them to determine what functions are new to the systems. Document the operational and business functions and the changes that are being. It is a required deliverable.

g. Activity: Communications Plan

Purpose: Defining and documenting a communications plan for informing whoever is being impacted by the program. It is a required deliverable.

h. Activity: Transitional Support

Purpose: Ensuring that those who are familiar with the program are available to the business to facilitate knowledge transfer. It is a required deliverable.

i. Activity: Balancing and Reconciliation Process

Purpose: Defined and documented plan to ensure business and operations have accounted for inventory and/or dollars as a result of the change resulting from the program. It is comply as necessary.

j. Activity: Interface Plan

Purpose: To determine what interfaces need to be changed based on the program i.e. letters, forms, system entitlement. It is a required deliverable.

k. Activity: Pipeline Freeze Plan

Purpose: Cutover Plan. It is a required deliverable.

10.2.3 Validation and Verification

The action is to verify that the documents from the Blueprinting/Gap analysis, Training Plan, Functional Model, Communication Plan, Event Management Model, Production validation and script reviews, floor support and process balancing are all available in the specified program knowledge base.

10.2.4 Exit Criteria

The Deliverables are:

- Discovery sessions and concept days - Required
- Blueprinting/Gap analysis - Required
- Training Plan - Required
- Functional Model - Required
- Communications Plan - Required
- Event Management Model - Required
- Production validation and script reviews - Required
- Transitional support - Required
- Balancing and Reconciliation Process - Comply as Necessary
- Interface Plan - Required

Pipeline Freeze Plan – Required

11 Program Acceptance Phase

11.1 Purpose

The purpose of this phase is to transition from the program into the operations environment and to validate the new functionality. [10] [11] [12]

11.2 ETVX

11.2.1 Entry Criteria

The entry criteria's are,

- Program projects development and testing are complete.
- Implementation plan has been documented and is available.

11.2.2 Tasks

a. Activity: Non-Public Pilot Test

Purpose: A pilot test of the components of the program and/or the program as a whole to ensure quality before making it publicly available. It is comply as necessary.

b. Activity: Transition and Acceptance Plan

Purpose: Transition of the knowledge of the new systems to support teams for operations. It is a required deliverable.

c. Activity: Program Evaluation

Purpose: Evaluate if the business case, goals and objectives were met. It is a required deliverable.

d. Activity: Program Debrief or Lessons Learned

Purpose: To gather high level metrics of the program to aid in determining room for improvement in the program process. Lessons learned are to be gathered from discussions with the teams and management involved in the program.

Lessons learned may be gathered after each project and provided to the program management team to be rolled into other projects that are either in progress or haven't started yet. This enables a composite strategic view to be compiled at the end of the program. It is a required deliverable.

11.2.3 Validation and Verification

The actions are,

- Verify that documentation is available regarding the tests on new and impacted functionality and the documentation is stored in the specified program knowledge base.
- Verify that all system knowledge has been transferred to support teams to minimize the need for development teams to be brought back into the support role after the program has been put into the production environment.

11.2.4 Exit Criteria

The Deliverables are:

- Non-Public Pilot Test - Comply as Necessary
- Transition and Acceptance Plan - Required
- Program Evaluation - Required
- Program Debrief or Lessons Learned – Required

12 Program Closeout Phase

12.1 Purpose

The purpose of this phase is to dismantle the program. This includes tasks such as closing reports, finalizing any remaining documentation, reallocation of unneeded infrastructure, etc. [12]

12.2 ETVX

12.2.1 Entry Criteria
The entry criteria is, all deliverables from prior phases have been completed

12.2.2 Tasks
a. Activity: Closeout reports

Purpose: Ensure that all reports and program information is complete. It is a required deliverable.

b. Activity: Perform Investment and Benefit Realization / Validation

Purpose: Perform an assessment on the benefits of the program to determine if they meet the expectations that were defined at the beginning of the program. It is a required deliverable.

12.2.3 Validation and Verification
The action is to verify that all reports and program information is complete.

12.2.4 Exit Criteria
The Deliverables are:

- Closeout reports - Required
- Investment and benefit realization / validation – Required

13 Program Acceptance Phase

13.1 Authorization Phase Gate

13.1.1 Purpose
To ensure the Program Charter is defined, sponsorship is received and the program organization and governance is established. [10] [11] [12]

13.1.2 Deliverables
The deliverables are,

- Program Charter - Required
- Risk Assessment - Required
- SC & Sponsorship - Required
- Program Setup & Organization - Required
- Program Governance - Required
- Architecture Approval – Required

13.2 Investment/Benefit Analysis Phase Gate

13.2.1 Purpose
Purpose is to ensure the proposed technical infrastructure and architecture is approved, investment and benefit assessments are completed and CEP is approved.

13.2.2 Deliverables

The deliverables are,

- Architecture / Infrastructure Plans - Required
- ROI Validated with Benefits Plan - Required
- CEP Approved – Required

13.3 Program Acceptance Phase Gate

13.3.1 Purpose

Purpose is to ensure the proposed technical infrastructure and architecture is approved, investment and benefit assessments are completed and CEP is approved.

13.3.2 Deliverables

The deliverables are,

- Program Evaluation - Required
- Program Debrief or Lessons Learned – Required

14 Glossary

Architecture Documentation: It is a special breed of design document. In a way, architecture documents are third derivative from the code (design document being second derivative, and code documents being first).

Complexity: The level of intricacy of the proposed task relating to design and testing requirements.

Pipeline Freeze Plan: The transitional timeline from point A to point B in a project.

Scale: Consists of effort and cost and is implemented at LOB level due to vast differences in capacity but must have each of the following size categories defined; small, medium, large, and extra-large.

Technology: A holistic view of the technologies to be implemented or modified and the associated impacts and risks they give to the environment.

Phase Gate: A “hard stop” where the project cannot enter the next phase of the life cycle without having completed all activities/deliverables of the phases prior to that Phase Gate.

Banking and Financial Services (BFS): Banking and Financial Services (also known as BFS) is an industry name. This term is commonly used by IT/ITES/BPO companies to refer to the services they offer to companies in these domains. Banking may include core banking, retail, private, corporate, investment, cards and the like. Financial Services may include stock-broking, payment gateways, mutual funds etc. A lot of data processing, application testing and software development activities are outsourced to companies that specialize in this domain.

Change Management: It is a structured approach to shifting/transitioning individuals, teams, and organizations from a current state to a desired future state. It is an organizational process aimed at helping employees to accept and embrace changes in their current business environment. In project management, change management refers to a project management process where changes to a project are formally introduced and approved

Line of Business (LOB): It is a general term which often refers to a set of one or more highly related products which service a particular customer transaction or business need.

Business Requirement Document: In IT projects business requirements is often a precursor to designing and building a new business application/system, or changing an existing one, and often sets the context for business process modeling, requirements analysis. The business requirements may encompass pre-requisites for both functional requirements and non-functional requirements that are desirable in the new or to be changed system.

Contract and Outsourcing: Contract is a legally enforceable agreement between two or more parties with mutual obligations, which may or may not have elements in writing. Outsourcing or sub-servicing often refers to the process of contracting to a third-party.

The University of California **Information Technology Leadership Council (ITLC)** is the system-wide IT governing body that works in partnership with UC leadership to articulate goals, strategies, priorities and resource allocations for investment in, and deployment of, technology-based solutions that support the University's missions of teaching, research, public service, and patient care.

Phase Review: A “soft stop” where a project can continue into the next phase’s activities without fully completing the requirements of the last phase. However, the last phase will remain incomplete until all activities/deliverables are completed.

Project Management: It is the discipline of planning, organizing, securing, and managing resources to achieve specific goals

Steering Committee: A steering committee is a committee that provides guidance, direction and control to a project within an organization. The term is derived from the steering mechanism that changes the desired course of a vehicle. Project Steering Committees are frequently used for guiding and monitoring the information technology projects in large organizations, as part of project governance. The functions of the committee might include building a business case for the project, planning providing assistance and guidance, monitoring the progress, controlling the project scope and resolving conflicts.

15 Acronyms

CEP Capital Expenditure Policy

BFS IT MP Banking & Financial Services Information Technology Management Policy

ETVX Entry Tasks Validation and Verification eXit

LOB Line of Business

BFS IT Banking & Financial Services Information Technology

PCM Process Control Manual
PM Project Manager
SQM Software Quality Manager

Acknowledgements. We are grateful to Larsen and Toubro Infotech –India, Madurai Kamaraj University-India, Citigroup-USA and Nordea AB-Sweden for this effort to get realized.

References

1. Kuttalam, P., Iyakutti, K., Alagarsamy, K.: Software Development Life Cycle Standards for Banking and Financial Services IT Industry. International Journal of Wisdom Based Computing, ISSN: 2231-4857
2. http://www.softwarecertifications.org/cboks/cmsq/skillcategories.html Also available for download for QAI members registered for Certified Manager for Software Quality QAI Global Institute
3. PMP Exam Prep, 6th edn. Rita Mulcahy PMP (2009)
4. Pressman, R.S.: Software engineering: a practitioner's approach, 5th edn. McGraw-Hill series in computer science (2001)
5. A Guide to the Project Management - Body of Knowledge (PMBOK Guide) 4th edn. An American National Standard ANSI/PMI 99-001-2008 Also available for download by PMI members at http://www.pmi.org
6. Adelson, B., Soloway, E.: The role of domain experience in software design. IEEE Trans. Softw. Eng. 11(11), 1351–1360 (1985)
7. Curtis, B., et al.: A Field Study of the Software Design Process for Large Systems. IEEE Trans. Software Engineering SE-31(11) (November 1988)
8. Software Maintenance - As Part of the Software Life Cycle, Comp180: Software Engineering Prof. Stafford, Department of Computer Science, Tufts University, http://hepguru.com/maintenance/Final_121603_v6.pdf
9. Software Development Program Characteristics Pekka Forselius, http://www.compaid.com/caiinternet/ezine/forselius-characteristics.pdf
10. The global and independent source of data and analysis for the IT industry. Also available for download by members at http://www.isbsg.org/, http://www.isbsg.org/isbsgnew.nsf/webpages/~GBL~Development%20&%20Enhancement
11. Finnish Software Measurement Association, http://www.fisma.fi/in-english/methods/
12. Guidance on Choosing a Sampling Design for Environmental Data Collection United States Environmental Protection Agency Office of Environmental Information, Washington, DC 20460 EPA/240/R-02/005 (December 2002), http://www.epa.gov/quality/qs-docs/g5s-final.pdf
13. Software Assurance Standard Nasa-Std-8739.8 w/Change 1 (July 28, 2004), http://www.hq.nasa.gov/office/codeq/doctree/87398.pdf
14. Galin, D.: Software Quality Assurance: From Theory to Implementation. Pearson Education Limited (2004)
15. Awad, E.M.: Systems Analysis and Design, 2nd edn.
16. Ho-Stuart, C., Thomas, R.: Laboratory practice with software quality assurance. In: International Conference on Software Engineering: Education and Practice (SE:EP 1996), Dunedin, New Zealand (1996), http://doi.ieeecomputersociety.org/10.1109/SEEP.1996.534003

Non-linear Mathematical Equations for PAPR Reduction in OFDM

R. Sundararajan, M. Dilip Reddy, D. Vinurudh, Aparna Anil,
Devika Mohan, and N.K. Neelima

Communications Engg. Research Group
Dept. of Electronics and Communication Engineering
Amrita School of Engg., Ettimadai
Amrita Vishwa Vidyapeetham
Coimbatore, India - 641112
r_sundar@cb.amrita.edu,
{dilip426426,Vinurudh,aparnaanil24,devikam2007}@gmail.com,
neelima_nk@hotmail.com6

Abstract. High Peak to Average Power Ratio (PAPR) is the major disadvantage in Orthogonal Frequency Division Multiplexing (OFDM). OFDM is extensively used in wireless communication systems due to its excellent performance. To reduce high PAPR, various techniques like clipping, companding, coding, etc. have been proposed. Companding technique effectively reduces the PAPR. In this paper, we propose new non-linear companding transforms which reduce PAPR efficaciously and are much more flexible when compared to other existing transforms. The theoretical analysis and simulation results of the proposed scheme are presented.

Keywords: OFDM, PAPR, and Companding.

1 Introduction

OFDM, a special case of Multi-carrier modulation is a popular modulation technique used in many broadband wireless communication systems. The main advantage of OFDM is its high spectral efficiency, high data rates, its robustness to multipath fading and its easy implementation. One of its main disadvantages is its high Peak to Average Power Ratio (PAPR). When the OFDM signal with high PAPR passes through a non-linear device, the signals may suffer significant non-linear distortion which in turn degrades the performance of the system. In order to reduce PAPR, several PAPR reduction techniques have been proposed, out of which non-linear companding transforms are more effective. Non-linear companding effectively reduces the PAPR by transforming the power or amplitude of the original signals into uniform distribution.

Advantages of nonlinear companding transform include good system performance, simplicity of implementation, no restriction on the number of subcarriers, type of

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 181–186, 2013.

constellation and it can be used in both large and small signals without any ramifications and bandwidth expansion.

The rest of the paper is organized as follows: Section 2 describes the OFDM system and PAPR. Section 3 describes reduction of PAPR in OFDM using existing non-linear transforms. Section 4 describes newly proposed transforms. Section 5 gives performance analysis of the companding transforms.

2 System Model

2.1 OFDM

Orthogonal frequency division multiplexing (OFDM) is a special case of multi-carrier modulation (MCM). The basic idea of multicarrier modulation is to divide the bit-stream into many different sub-streams and send these over many different sub-channels. In OFDM, these sub-channels overlap. But these sub channels are orthogonal to each other. Main advantage of OFDM is its very high data rate of up to 100Mbps and very less Inter-symbol Interference (ISI).

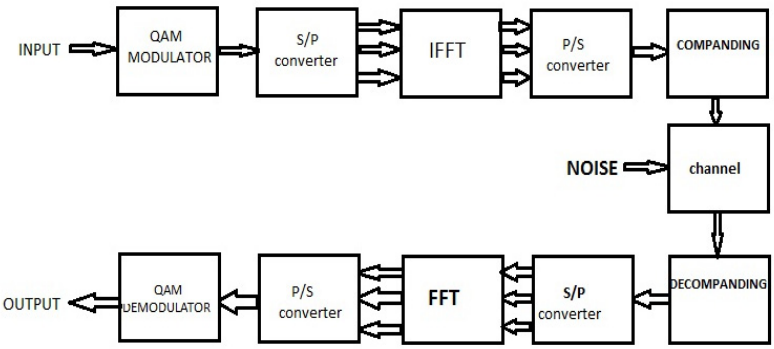


Fig. 1. OFDM system with companding

2.2 PAPR

The PAPR for the discrete-time signal $x[n]$ is defined as the ratio of its maximum instantaneous power to its average power and can be expressed as

$$PAPR(x[n]) = \max \frac{|x[n]|^2}{E[|x[n]|^2]}, 0 \leq n \leq N-1 \quad (1)$$

Where [ ] denotes the expectation operator. The PAPR reduction capability is measured by the empirical complementary cumulative distributive function (CCDF), which indicates the probability that the PAPR is above a certain threshold.

3 Companding Transforms

Firstly, companding equations must adjust the power levels of low amplitude signals and reduce the high peaks to minimize the overall PAPR. It should amplify the low amplitudes and compress the high peaks. So it should be a non-linear equation which is monotonically increasing with decreasing slope.

Specifically, the companding transform should satisfy the following two conditions:

$$(1)\ E(|s(n)|^2) \approx E(|x(n)|^2) \tag{2}$$

$$(2)\ |s(n)| \geq |x(n)| \text{ when } |x(n) \leq v|; \tag{3}$$

$$|s(n)| \leq |x(n)| \text{ when } |x(n) \geq v| \tag{4}$$

Where v is the average amplitude of the signal, s(n) is the companded signal and x(n) is the input signal.

4 Proposed Companding Transforms

Newly proposed equations are more efficient than the existing transforms. First one is the n^{th} root equation and second equation is a non-linear fraction equation. They are designed using basic mathematical knowledge. They have been designed for satisfying basic requirements of non-linear companding transforms.

4.1 n^{th} Root Transform

$$y = k1.\sqrt[n]{\frac{x}{k2}} \tag{5}$$

It amplifies the amplitudes below the amplitude value of Δ and compresses the amplitudes beyond that value. Δ is given by

$$\Delta = \left(\frac{k1}{k2^{\frac{1}{n}}}\right)^{\frac{n}{n-1}} \tag{6}$$

Conditions:
Conditions for the equation to perform companding are:

- *k1*, *k2* and *n* should be positive
- $n \geq 1$

Advantages:
Advantages of the equation are:

- More flexible equation where the slope can be determined by the values of *n*
- *k2* determines the point of transition from expansion to compression.
- Power of companded signal is determined by *k1*.

4.2 First Order Fraction Transform

$$y = \frac{k1.x}{1+(k2.x)} \tag{7}$$

The amplification and compression values are determined by the value of (*k1/k2*) because for larger values of *x*, equation tends to have a slope of zero at a max value of y = (*k1/k2*).

Conditions:
Conditions for the equation to perform companding are:

- *k1* and *k2* are positive
- $k2 \geq 1$

Advantage of the equation is it is more flexible equation where the slope of the curve and its amplitude can be varied by the values of *k1* and *k2*

These equations (5) and (7) are monotonically increasing functions with decreasing slope in the interval [0,-∞) and they have following common advantages:

- Level of companding can be easily modified by appropriate selection of parameters
- They have very simple inverse functions

5 Performance Evaluation

In this section, new companding transforms are simulated and their performance is evaluated by comparing them with few existing companding transforms. These simulations were carried out on an OFDM system with 16-QAM modulation and IFFT size of 64 in an Additive White Gaussian Noise (AWGN) channel.

Fig. 2 is the simulation of the nth root equation (7) for different values of *n* with *k1*=1 and *k2* = 1. It shows that since Δ=1, amplitude levels below the value of 1 will be amplified and beyond the value of 1 will be compressed and it also shows that the amplification and compression depends on the value of *n*.

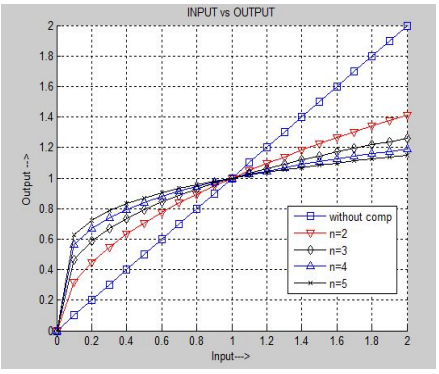


Fig. 2. Transform curves of nth root equation for different values of *n*

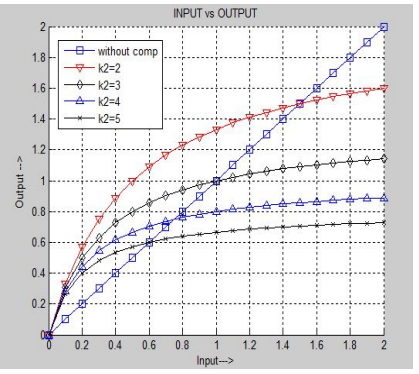


Fig. 3. Transform curves first order fraction equation for different values of *k2*

Fig. 3 is the simulation of the first order fraction equation (9) for different values of *k2* and *k1*=4. It shows that the equation is tending to be line of slope zero y = (*k1/k2*) and this determines the compression of the peaks.

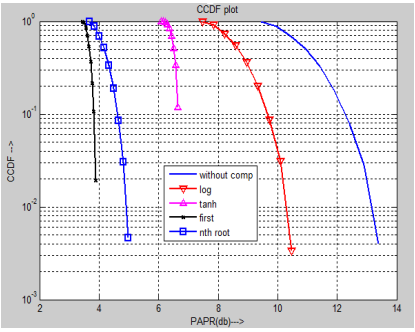


Fig. 4(a)

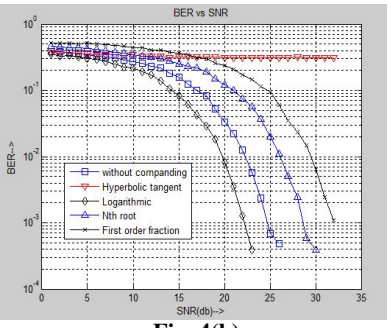


Fig. 4(b)

Fig. 4. Comparison of performance of companding transforms. (a) Comparison of PAPR (b) Comparison of BER.

Fig. 4 compares the performance of various companding transforms on an OFDM system. Proposed techniques have a relatively better performance than the existing techniques. Firstly, nth root transform has a 13db less PAPR and first order fraction has a 6db less PAPR than the logarithmic transform.

Performance of nth root transform and first order fraction equation transform is analyzed in Fig. 5 and Fig. 6. As *n* and *k2* increases PAPR decreases, but BER increases. There is tradeoff between PAPR and BER. The degradation in BER is because the equations (5) and (7) tend to have a slope of zero with increase in *n* and *k2* values and the peaks get compressed to a constant value irrespective of their amplitude. So, the information in these peaks cannot be retrieved easily by decompanding.

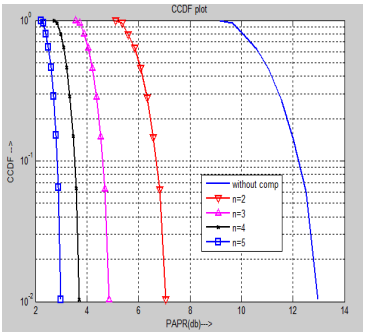


Fig. 5(a)

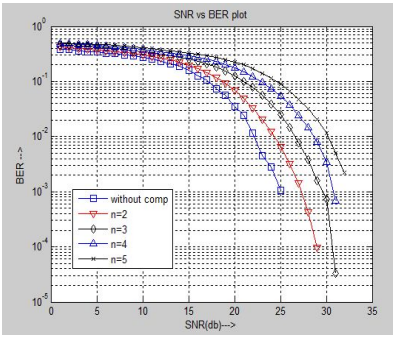


Fig. 5(b)

Fig. 5. Performance analyses of nth root equation. (a) Comparison of PAPR (b) Comparison of BER.

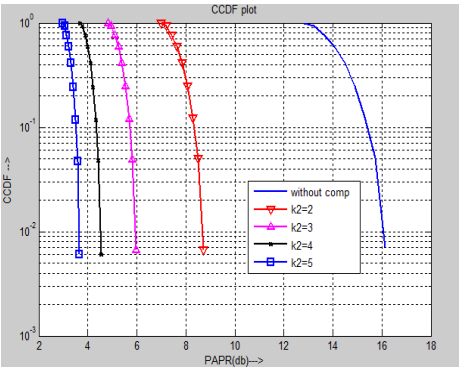


Fig. 6(a)

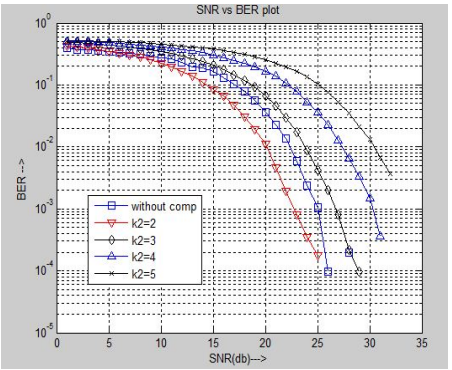


Fig. 6(b)

Fig. 7. Performance analyses of first order fraction equation. (a) Comparison of PAPR (b) Comparison of BER.

6 Conclusion

This paper is focused on PAPR reduction techniques which use non-linear companding to reduce the high PAPR of OFDM systems. This technique is very effective in reducing the PAPR. Simulations show that the proposed transforms have better PAPR reduction capability than the existing non-linear companding transforms and these transforms can be tuned according to their application by selecting appropriate parameters. But these transforms have relatively more BER.

So, this work has paved a new path for further work in reducing the BER in the proposed scheme.

References

1. Jiang, T., Wu, Y.: An Overview: Peak-to-Average Power Ratio Reduction Techniques for OFDM Signals. IEEE Trans. Broadcasting 54(2), 257–268 (2008)
2. Manjith, R., Ramesh, S.C., Mohamed Ismail Majeed, M.: PAPR reduction in OFDM & MC-CDMA system using nonlinear companding techniques. In: IEEE Region 8 SIBIRCON 2010, Irkutsk Listvyanka, Russia, July 11-15 (2010)
3. Jeng, S.-S., Chen, J.-M.: Efficient PAPR Reduction in OFDM Systems Based on a Companding Technique with Trapezium Distribution. IEEE Transactions on Broadcasting 57(2), 291–298 (2011)
4. Jiang, T., Yao, W., Guo, P., Song, Y., Qu, D.: Two novel nonlinear companding schemes with iterative receiver to reduce PAPR in multicarrier modulation systems. IEEE Trans. Broadcasting 52(2), 268–273 (2006)
5. Jiang, T., Xiang, W.D., Richardson, P.C., Qu, D.M., Zhu, G.X.: On the nonlinear companding transform for reduction in PAPR of MCM signals. IEEE Trans. Wireless Communications 6(6), 2017–2021 (2007)
6. Jiang, Y.: New Companding Transform for PAPR Reduction in OFDM. IEEE Communications Letters 14(4), 282–284 (2010)

Effect of Temperature on Gate Leakage Current in P4 and P3 SRAM Cells at Deep Sub-micron CMOS Technology

Rajesh Singh[1], Neeraj Kr. Shukla[3], S. Birla[2], Naveen Yadav[3], Ritu[3], and Ankit Goel[3]

[1] ST Microelectronics, Greater Noida (Uttar Pradesh) India
[2] Sir Padampat Singhania University, Bhatewar, Udaipur, (Rajasthan), India
[3] ITM University, Department of EECE, Gurgaon, (Haryana), India

Abstract. The explosive growth of battery operated semiconductor devices has made low-power design a priority in recent years. The effect of temperature on the leakage currents is coming as a future challenge in the high density CMOS based portable mobile multimedia rich applications. In this paper the standby leakage power analysis at different operating temperature for the Conventional 6T SRAM, P3 SRAM and P4 SRAM cell has been carried out. It has been observed that for 6T SRAM there is an increase in standby leakage power with increase in temperature and the gate leakage current is found to be reduced with temperature variations, but this effect of temperature variation is found to be significantly lowered in P3 and P4 SRAM cells due to p-MOS stacking. The design simulation has been performed on CMOS deep sub-micron technology node, 45nm, at V_{DD}=0.7V and 0.8V, for n-MOS and p-MOS threshold voltages as 0.24V and 0.224V, respectively on temperature range from -25^{0}C to +125^{0}C.

Keywords: Conventional SRAM Bit-Cell, P4 Cell, P3 Cell, Leakage Current, Gate Oxide Tunneling, Temperature.

1 Introduction

The demand of multimedia rich applications in portable and mobile devices is increasing at a starling rate, and the most basic need of such applications are Performance and Power. These two parameters performance and power are primary design issues for systems ranging from server computers to handheld devices. Performance is affected by both temperature and supply voltage because the circuit delay is dependent on temperature and voltage.

By consequently scaling technology down further, we now reached a state where parasitic leakage effects do not only dominate the power budget but also become a concern as important as the performance which used to be the most important scaling concern since the advent of MOSFETs. According to the ITRS roadmap [1], the top challenges towards the 32nm technology node are controlling leakage while increasing performance. Furthermore, as semiconductor technology scales down, leakage power's exponential dependence on temperature and supply voltage becomes

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 187–192, 2013.

significant. Therefore, future design studies call for temperature and voltage aware performance and power modeling.

In this paper, we investigate the effects of the temperature on the characteristics of the static random access memory (SRAM) bit cells. The rest of the paper is organized as follows. In section 2, basic leakage current mechanisms and detailed description of gate leakage currents is presented. The section 3 reviews the P4 and the P3 SRAM cell designs. The simulation results and conclusions are discussed in section 4 and 5 respectively.

2 Leakage Current Mechanisms

High leakage current in deep-submicron regimes is the major contributor of power dissipation of CMOS circuits as the device is being scaled [2]. Various leakage current mechanisms are shown in Figure 1.

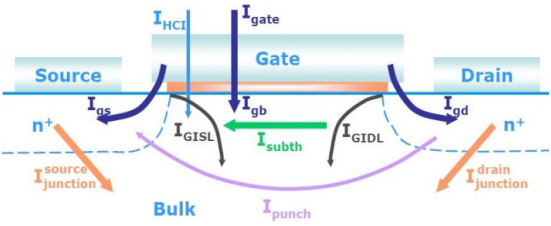


Fig. 1. Leakage current mechanisms of deep-submicron transistors

2.1 Gate Leakage (I_G)

A vast fraction of the total leakage of sub-100nm systems is due to gate current. The gate leakage can be divided into the major effect of gate tunneling (I_{GD}) and hot-carrier injection I_{HCI}.

A. Gate Direct Tunneling Leakage (I_{GD})

The dominant source of gate currents is gate tunneling, which can again be subdivided by current path, origin of carriers or the shape of potential to be passed.

• From an electrical point of view, the possible current paths are directly through the oxide to source I_{gs} or drain I_{gd}, or into the channel I_{gc}. From the channel, the current can then flow to source I_{gcs}, to drain I_{gcd} or to bulk I_{gb}. Figure 2 presents these five possible gate-tunneling current paths.
In a simplified form, it results as

$$I_{gx} = K_x W L_{eff,x} T_{OX}^{-2} V_{gx} V_Y . e^{-\beta_x . T_{ox} (\alpha_{0,x} + \alpha_{1,x} . V_{ox} + \alpha_{2,x} . V_{ox}^2)}$$

Where x is the respective terminal (source, drain or bulk), k_x, $\alpha_{i,x}$, and β_x are constants, depending on the terminal x. Gate-source and gate-drain leakage are described identically. As tunneling probability gives the charge transfer per electron

trying to pass the barrier, it has to be multiplied with an area to compute current from current density. Gate tunneling always occurs over the entire width W of a transistor, but for gate-drain and gate-source leakage only over a small fraction of the channel length, denoted $L_{eff,x}$. Another asymmetry between gate-source/drain and gate-bulk leakage is V_Y, which is V_{gx} for x = source/drain but V_T (thermal voltage) for x = gate. Usually $V_T << V_{gx}$ for x = source/drain, thus tunneling to bulk is much smaller and temperature dependent.

Effect of Temperature: The behavior as described by above equation can be observed, when plotting the gate current I_{gx} versus the respective voltage V_{gx} for different oxide thicknesses and temperatures as presented in Figure 3: Gate leakage is exponentially dependent on oxide thickness as well as potential difference. Gate to bulk tunneling is almost two orders of magnitude smaller than gate to source or gate to drain leakage and only gate to bulk leakage is showing a significant thermal dependency.

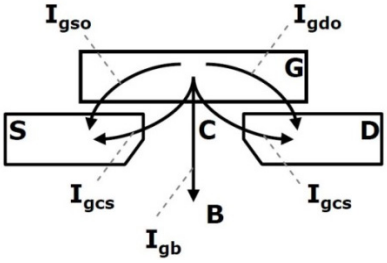


Fig. 2. Current paths in MOSFET transistors

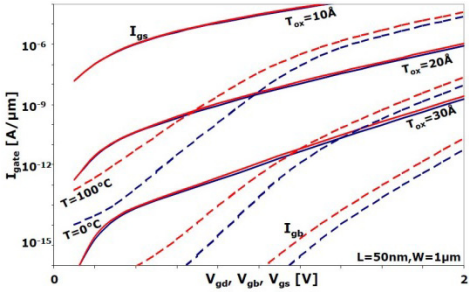


Fig. 3. Gate leakage of the 45nm predictive technology model NMOS device [3]

B. Hot Carrier Effects

In comparison to other isolators, silicon oxide, the recent MOSFET gate dielectric, has a comparatively small band-gap of 3.3eV @ 300K. As soon as a carrier inside the channel manages to obtain this energy, it can enter the oxide's conduction band and carry a current towards the gate without the exponential decrease per oxide thickness. Especially in short channel devices, the electric fields may become large enough to accelerate the carriers so that they become hot, what means that their kinetic energy exceeds the oxide band-gap. Holes from the valence band additionally need to overwhelm the silicon band-gap (1.12eV @ 300K) and have a larger effective mass.

• Drain avalanche hot carrier injection (DAHC): When switching, for $V_{ds} > V_{gs} > 0$ (this condition is typically reached when switching), a high field occurs at the drain. The accelerated carriers collide with the silicon lattice, generating electron-hole pairs (impact ionization). The resulting electron-hole pairs may have sufficient energy to becoming hot, thus reaching the oxide conduction band.

• Channel hot electron injection (CHE): If drain and channel potential both are significantly higher than the source, carriers traversing the channel are attracted by the gate potential. Some of them may reach channel-oxide junction before reaching drain.

• Substrate hot electron injection (SHE): In case of extreme biasing of the substrate, carriers from the substrate are accelerated towards the gate if V_{gb} is large. As they move upwards, they gain more energy in the higher field close to the surface.

• Secondary generated hot electrons injection (SGHE): The SGHE equals the DAHC effect. But additionally, while the majority carriers move towards the gate oxide, the minority carriers move towards the substrate creating a field which is driving the majority carriers towards the gate.

3 A Review of Related Work

With new concept of "PMOS stacking" introduced in P4-SRAM Bit-Cell [4]. The main ideas behind the bit-cell were:

- The barrier height of the holes is more than that of the electrons, i.e, 4.5eV versus 3.1eV, respectively. So, the gate leakage current will be lower in the off state of the p-MOS transistors whereas the full-supply substrate body-bias will involve in reduction of the sub-threshold leakage during standby mode of proposed design.
- Concept of the transistor stacking for the leakage power reduction. These, extra p-MOS transistors of the Gated-GND, produces the stacking effect in conjunction with the SRAM Bit-Cell transistors when the Gated-GND transistors are turned-off by opposing flow of leakage current through them.

Thus hot carrier injection is much stronger in NMOS devices.

According to [5], hot carrier effects can be sub-divided into four effects:

In [6], a P3 SRAM bit-cell structure at 45nm technology has been proposed for high activity factor based applications. The cell has been proposed for reduction of leakage power through gate leakage current and sub-threshold leakage current reduction in both active and standby mode of memory operation. The p-MOS stacking transistor, connected in series (in line), is kept off in standby mode and kept on in active (read/write) mode. The p-MOS transistors are used to lower the gate leakage current [7]. In [8], PP SRAM-cell has been proposed which uses p-MOS transistors as pass transistors to reduce gate leakage on the expense of some performance degradation. This performance degradation occurs due to different mobility coefficients for n-MOS and p-MOS.

4 Results and Discussion

4.1 Static Power Consumption

To analyze the effect of temperature on standby power and Gate Leakage currents in 6T, P4, and P3 SRAM Cells, the simulation work is being performed at 45nm CMOS technology with gate oxide thickness of 2.4nm at different temperatures (-25°C to 125°C) to ensure different environmental temperatures at supply voltage of V_{DD}=0.8V

and 0.7V. It has been observed that due to change in operating temperature there is an increase in standby leakage power in 6T SRAM cell, but in P3 and P4 SRAM cell, due to p-MOS stacking, effect of temperature has been lowered as compare to 6T SRAM cell by 71.02%, 85.03%, 94.12% and 94% in P3 SRAM cell at -25°C, 25°C, 75°C and 125°C respectively, and 72.69% , 87.45%, 96.07% and 96.27% in P4 SRAM cell at -25°C, 25°C, 75°C and 125°C, at V_{DD}=0.8V respectively.

At V_{DD}=0.7V, the standby leakage power is found to be reduced by 71.77%, 85.93%, 94.03% and 93.14% in case of P3 SRAM cell (w.r.t 6T SRAM cell) at -25°C, 25°C, 75°C and 125°C and in case of P4 SRAM cell, decrease in standby power w.r.t conventional 6T SRAM cell is 82.88%, 92.58%, 97.34% and 96.19% at -25°C, 25°C, 75°C and 125°C.

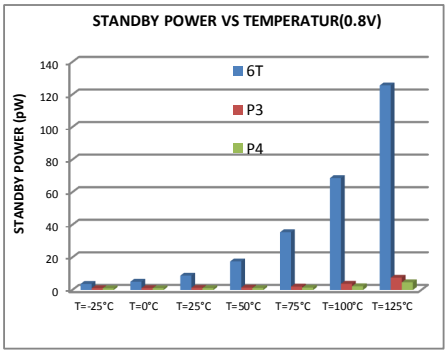


Fig. 4. Standby Power at V_{DD}= 0.8V

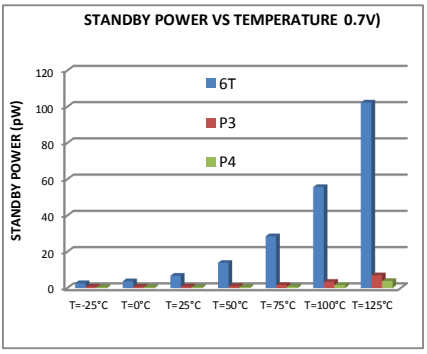


Fig. 5. Standby Power at V_{DD}= 0.7V

Figure 4 and 5 shows the comparison of standby power at different temperatures for 6T, P3 and P4 SRAM Bit cells.

4.2 Gate Leakage Current (I_G)

Fig. 6 and 7 shows the comparison of Gate leakage current at different temperatures in standby mode for 6T, P3 and P4 SRAM Bit cells. Because of the p-MOS stacking effect in P3 and P4 SRAM cells, gate leakage current is found to be significantly reduced at different operating temperatures.

At V_{DD}=0.8V, in P3 SRAM cell, the gate leakage current reduction is 88.50%, 69.07%, 67.61% and 50.8% at -25°C, 25°C, 75°C and 125°C respectively and in P4 SRAM cell the percentage value of gate leakage current reduction is 88.57%, 79.77%, 73.62% and 66.3% at -25°C, 25°C, 75°C and 125°C respectively. And at V_{DD}=0.7V, in P3 SRAM cell, the gate leakage current reduction is 59.45%, 83.6%, 70.83% and 37.74% at -25°C, 25°C, 75°C and 125°C respectively and in P4 SRAM cell the percentage value of gate leakage current reduction is 59.95%, 81.34%, 72% and 68.87% at -25°C, 25°C, 75°C and 125°C respectively.

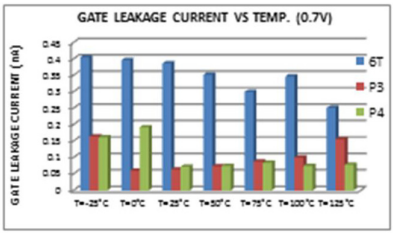


Fig. 6. Gate Leakage Current at V_{DD}=0.7V

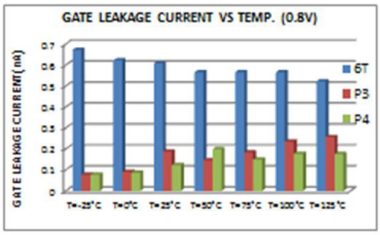


Fig. 7. Gate Leakage Current at V_{DD}=0.8V

5 Conclusion

In this paper the standby leakage power analysis at different operating temperature for the Conventional 6T SRAM, P3 SRAM and P4 SRAM cell has been carried out. It has been observed that for 6T SRAM there is an increase in standby leakage power with increase in temperature and the gate leakage current is found to be reduced with temperature variations, but this effect of temperature variation is found to be significantly lowered in P3 and P4 SRAM cells due to pMOS stacking.

Acknowledgements. The authors are grateful to their respective organizations for their encouragement and support.

References

1. International Technology Roadmap for Semiconductors (2003), http://www.publicitrs.net
2. Roy, K., Mukhopadhyay, S., Mahmoodi-Meimand, H.: Leakage Current Mechanisms and Leakage Reduction Techniques in Deep-Submicrometer CMOS Circuits. Proceedings of the IEEE 91(2), 305–327 (2003)
3. Nanoscale Integration and Modeling Group. 45nm BSIM4 model card for bulk CMOS v1.0 (February 2006), http://www.eas.asu.edu/~ptm/
4. Elakkumanan, P., Thondapu, C., Sridhar, R.: A gate leakage reduction strategy for sub-70 nm memory circuit. In: Proc. IEEE Dallas/CAS Workshop, pp. 145–148 (2004)
5. Silicon Far East. Hot carrier effects. Technical report, siliconfareast forum (2004), http://www.siliconfareast.com/hotcarriers.html
6. Shukla, N.K., Singh, R.K., Pattanaik, M.: A Novel Approach to Reduce the Gate and Sub-threshold Leakage in a conventional SRAM Bit-cell Structure at Deep-Sub Micron CMOS Technology. International Journal of Computer Applications (IJCA) 23(7), 23–28 (2011)
7. (Steve) Kang, S.-M., Leblebici, Y.: CMOS Digital Integrated Circuits-Analysis and Design, 3rd edn. Tata McGraw-Hill Edition, New Delhi
8. Razavipour, G., Afzali-Kusha, A., Pedram, M.: Design and Analysis of Two Low-Power SRAM Cell Structures. IEEE Transaction on VLSI Systems 17(10), 1551–1555 (2009)

Ensemble Approach for IP Auto-configuration in Ad Hoc MANETs

S. Zahoor Ul Huq[1], S. Shabana Begum[2], N. Geethanjali[3], and K.E. Srinivasa Murthy[4]

[1] Dept. of CSE, G. Pulla Reddy Engineering College, Kurnool,
[2] Dept of CSE, Sri Kottam Tulasi Reddy Memorial Engineering College, Kondair
[3] Computer Science &Technology, Sri Krishnadevaraya University, Anantapur
[4] Sri Kottam Tulasi Reddy Memorial Engineering College, Kondair,
Andhra Pradesh, India
{s_zahoor_2000,zahoor2saba}@yahoo.com,
{geetanjali.sku,kesmurthy}@gmail.com

Abstract. The main task of an address allocation protocol is to manage the address allocation to the nodes in the ad hoc MANETs. All routing protocols assume nodes to be configured a priori with a unique IP address. Allocating addresses to mobile nodes is a fundamental and difficult problem. A mobile device cannot participate in unicast communications until it is assigned a conflict-free IP address. So addressing in MANETs is of significant importance, and the address configuration process should be fast, as the algorithm must be able to select, allocate and assign a unique network address to the unconfigured node before with a unique IP address. As the network is an ad hoc network no centralized control of the network is possible. Since every address change may break transport layer connections, unnecessary address changes should be avoided. Here we present an Ensemble approach in which a group of nodes can be assigned IP Addresses when the network is initiated.

Keywords: Address Allocation, Mobile Ad-Hoc Network, Node Joining, Address Character, Temporary Address, Permanent Address.

1 Introduction

In the real work in order to communicate with other communication entities, each of the communication entity should be equipped with a unique address using which it can be reached. Addresses are important in computer communications, as both identifiers and locators of nodes requires addresses for communication. One of the important criterions when allocating the addresses is that each host must be assigned with a unique address. Typically, there are three ways that a host might obtain a unique address. The first is to ask a central server, such as a dynamic host configuration (DHCP) server [2], where the server is responsible for ensuring that all allocated addresses are unique. The second one is, a host may use a link-local (LL) protocol [2] to generate itself an address and carry out duplicate address detection (DAD) to ensure that the chosen address is unique. However, LL protocols assume all nodes are of link-local scope and are thereby available to defend their addresses at

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 193–199, 2013.

all times. Finally, a node may rely on the user to ensure that the assigned addresses are unique.

A Mobile Ad Hoc Network (MANET) [5] is an independent self organizing network in which each node functions as both an end host and a router. This form of wireless network is created by mobile nodes without any existing or fixed infrastructure. The topology of a mobile Ad hoc network is typically highly dynamic because its nodes are free to move independently and randomly. The size of a MANET is constantly changing as nodes come in and out of the network range. A node is not part of a MANET until it is within the transmission range of an already configured node in the MANET. During the time a node is present in the MANET; it may or may not participate in communication or packet forwarding.

Nodes in the MANET need some form of identity before participating in communication. Using traditional methods, such as DHCP [3], is not possible because nodes in the MANET are highly mobile and a central authority is not always reachable. Mobile IP [4] is also not a solution because MANETs are wired.

2 Literature Survey

The address allocation protocols are divided into three categories (1)Stateful protocols (2) Stateless protocols (3) Hybrid protocols. Stateless [15] approaches use conflict detection mechanism for allocating addresses to new nodes. A new node randomly chooses an address from a prescribed address block and then verifies its uniqueness by either querying that address to receive a reply or by performing duplicate address detection. No address allocation tables are maintained. An unconfigured node self-assigns an address randomly and verifies its uniqueness with duplicate address detection (DAD) [12].

According to [6], recent address allocation protocols can be divided into stateless and stateful approaches. In a stateless approach [7], a node chooses an IP address and performs Duplicate Address Detection (DAD) [8][12] is the stateless proposal for distributed protocol in which every node that wants to join the network, floods the network with Address Request message (AREQ) to find the existence of the IP in the network.

Stateful [13] approaches can be further divided into leader based-approaches, distributed common address space approaches and distributed disjoint address space approaches. In leader- based approaches [9] the workload on a leader is too heavy and a node may not be always reachable in the MANET. In MANETconf [10], address assignment is based on a distributed mutual exclusion algorithm that treats IP addresses as a shared resource. As in MANET-conf, the dynamic address configuration protocol [7] also requires nodes to perform DAD during address assignment, which increases the latency of nodes joining the MANET. Complete synchronization is required between all nodes to avoid duplicate addresses. Dynamic Host Configuration Protocol (DHCP) [3] is not feasible because a server may not be permanently reachable by all nodes. In DHCP, there is also a single point of failure.

All the stateless approaches and common distributed allocation table approaches can be categorized as conflict avoidance approaches. Addresses are assigned from a pre-authorized or calculated address pool. Then permission is granted by all the nodes for the address allocation, or all the nodes in the MANET are informed of the allocated address to purge it from the pending address table entry.

3 IP Address Assignment

In order to establish the Communication between the nodes it is mandatory that the nodes to be in the network. To uniquely identify the nodes present in the network, these nodes must be assigned with unique IP. If these IP addresses are not unique misrouting takes place and the information cannot be delivered or sent to the right receiver. So the nodes must be configured with the non duplicate IP address. This can be implemented by Ensemble Approach. The name ensemble is used here as we are assigning the group of nodes IP addresses together at the same time. This method makes use of the MAC-IP table, which contains the information regarding the MAC addresses of the nodes and the random number generated for that node, and the new IP address that has been generated is maintained.

Every Node in the MANET maintains the MAC-IP table. Whenever the network is instantiated, hello messages are exchanged by the nodes to tell their existence and also to know about the existence of the neighbors. Embed in the hello messages is the MAC-IP table. The MAC-IP table has the entries of MAC Address of the nodes, Nonce and the IP Addresses that will be generated. The format of MAC-IP table is given in [11].

Nonce are the Random numbers they are used because the MAC Addresses are not unique as some manufacturers sell network adapters with non-registered MAC addresses or the MAC address may be corrupt while manufacturing. Most of the network adapters allow users to change the MAC address to arbitrary values and some of the adapters may not be having IEEE MAC addresses at all. Because of the above reasons the random number is used to recognize the uniqueness.

3.1 Duplicate Address Detection Scheme

The new node after it enters into the network gets the information regarding the IP address to use. It assigns the IP address and creates the packet as Fig. 1 and broadcasts the packet to the subnet.

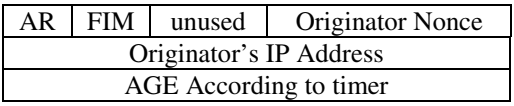


Fig. 1. Address Request/Reply Packet Format

In the Fig. 1 the first bit tell whether it is Address Request or Address Reply packet. If the bit is set to 0 (zero) it does mean that it is the Address Request, if it is set to 1 (one) it means that it is Address Reply. The next bit FIM represents Found In the MAC-IP table. By default this bit is set to 0. If this bit is set to 1 it means that the requested IP Address is present in the MAC-IP table and there is an address conflict. It maintains the age of the IP address that has been assigned by the originator. The next 14-bit field is unused one and may be used in the future for more further enhancement. The 16 bits are for the Nonce generated by the Originator, this is maintained because, it may so happen that more than one node can assign the same address, and the Address Reply packet may be routed back wrongly. So by means of the nonce the packets can determine whether it has

reached the correct node or not. The requested IP Address is present in the Originator's IP Address field. The next 32 bits represents AGE field and is capable of counting up to 1,193,046 hours ≈ 49,710 days.

3.2 Conflict Detection Allocation

There do already exist some nodes for which IP addresses are assigned, and there are these two new nodes which entered into the network for which IP Addresses are not assigned. As said already both of them have temporarily been assigned the same IP address and they need to check for the address conflict. The two new nodes in the network one of the node we call them as the new node and the other node we call it as current node.

The new node broadcasts the packet by assigning the following things AR field is set to 1, FIM to 0, AGE field is incremental timer which has started when the node has entered into the network, the Originators IP Address which it wants to assign and the Nonce generated by the node. When the current node receives this AREQ packet, it first checks whether it has processed this packet before or not, this is done by maintaining the AREQ (Address Request) seen table which will be associated with time-to-live field, if it has processed before it simply discards the packet. If it has not done before it identifies this as the new packet and makes an entry in the AREQ seen table. Now the current node checks with its own MAC-IP table whether the assigned IP address is already assigned by some other node and this information has not reached to the new node. If an entry is found in the MAC-IP table then it sets the AR field to 1 and sets the (Found in MAC-IP table) FIM bit to 1 and sends it back to the new node, now the new node changes its IP address and repeats the process again, that is by selecting the new IP from the list of the free IP Addresses available and broadcasting the AREQ/AREP packet again. The nodes upon receiving the AREQ packet checks with the entry in the MAC-IP table, if no match is found, then the new node check with its own IP addresses whether it matches or not. If the new nodes IP Addresses does not match then it simply forwards the packet to its neighbors, else

- If the current node's AGE timer is greater than that of the AREQ packets AGE timer generated by the new node, then the current node's IP address remains unchanged and the current node changes the AGE field with that of its own and changes the AR bit to 1 and sends back to the new node. The current node changes the AGE field to that of its own because, to let know the new node that the IP address conflict has happened and the current node is elder than itself and it has to change its IP address. The new node can now change its IP address and initiate the AREQ packet and the process is repeated again.
- If the current node AGE timer is less than that of the AREQ packet AGE timer, generated by the previous node, then the new node drops its own IP Address and initiates the AREQ for a new IP Address. Now the current node will include its current IP Address Timer into the AGE field of the AREQ packet.

Fig 2 represents the flowchart for Duplicate Address Detection Scheme and Conflict Detection Allocation.

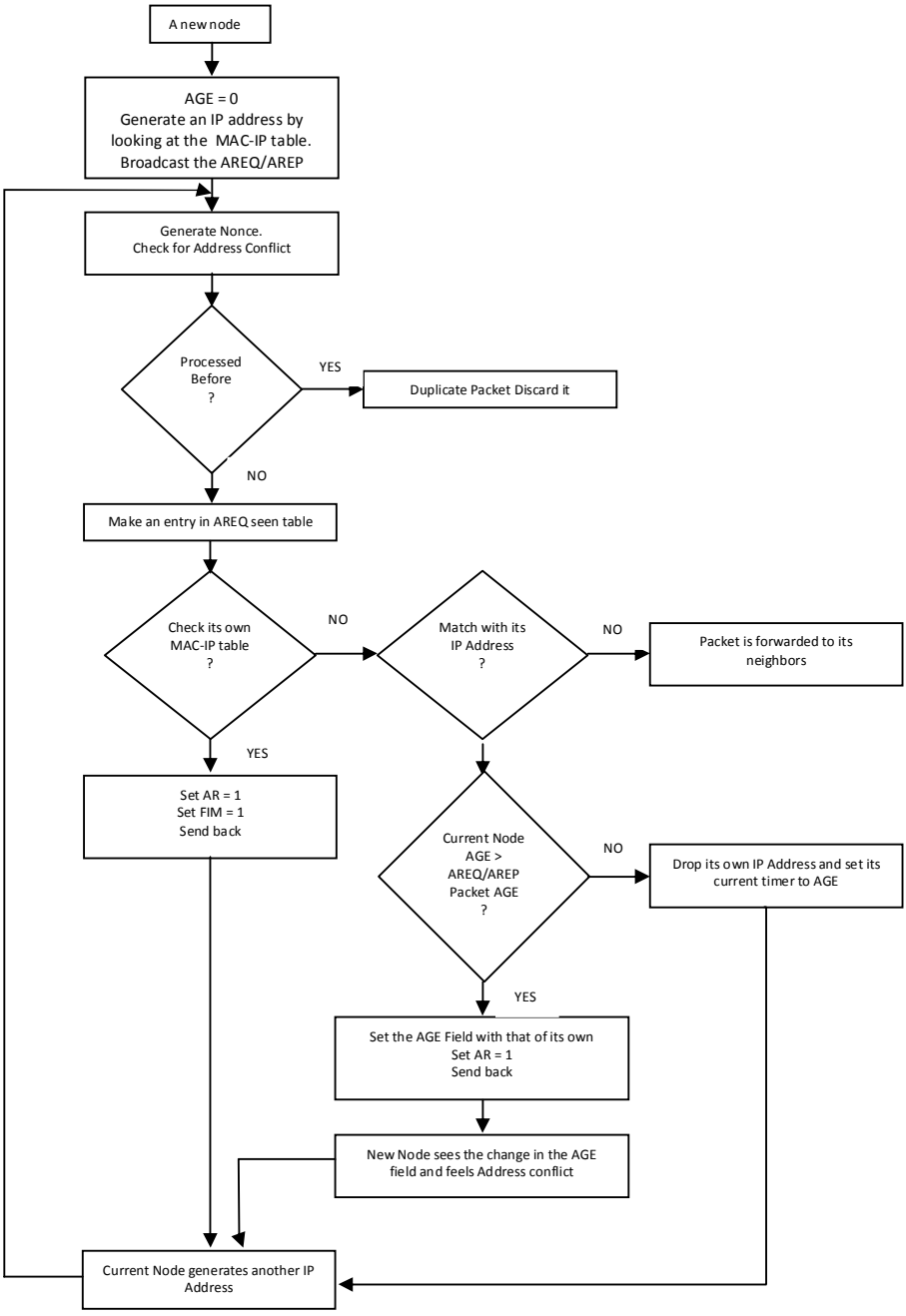


Fig. 2. Flow Chart for the Duplicate Address Detection Scheme & Conflict Detection Allocation

4 Experimental Results

The entire simulation was run on ns-2[1] version 2.30 simulator installed in Intel Core 2 CPU of 2.00 GHz Processor and 256 MB of RAM. Fig. 3 and Fig. 4 represent the graphs drawn against Number of nodes and Time required in seconds. Fig. 5 and Fig. 6 represents the graphs drawn against Number of nodes and number of Broadcasts required. All the Figures from Fig. 3 through 5 are shown in the terrain range of 500 x 500 and 750 x 750. Whatever may be the number of nodes and terrain area it takes less than 1 second to build the MAC-IP table.

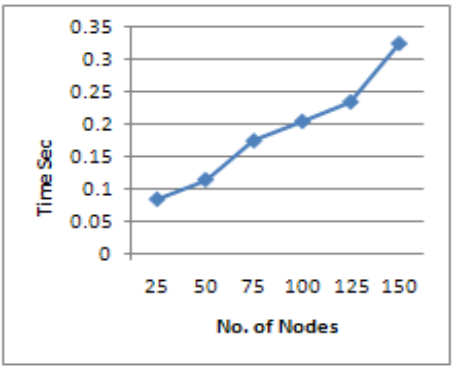


Fig. 3. Time required when Terrain Range is 500 x 500

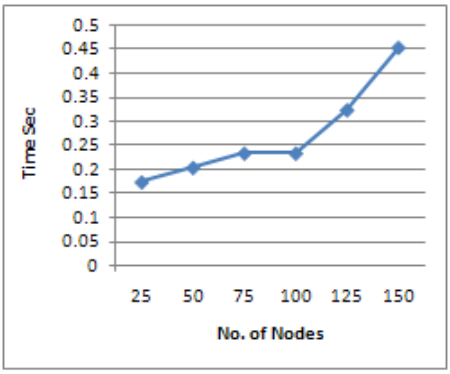


Fig. 4. Time required when Terrain

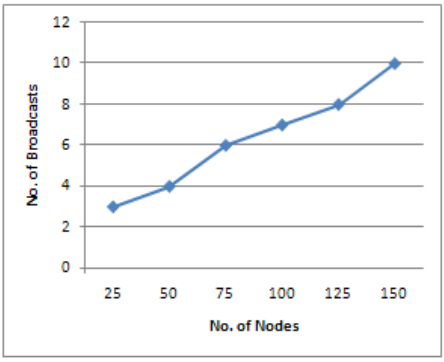


Fig. 5. Broadcasts required when Terrain

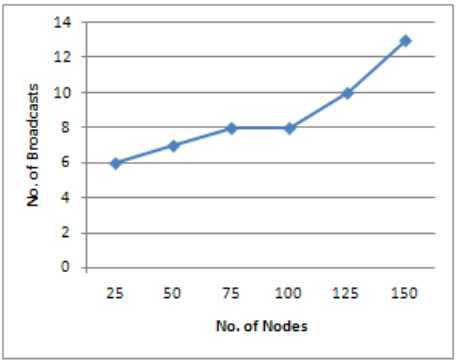


Fig. 6. Broadcasts required when Terrain Range is 750 x 750

5 Summary

Thus we are successfully able to allocate the IP addresses to a group of nodes when the network is initiated or to the individual nodes when the nodes add up after the initial configuration, Reasons why Nonce is used, Duplicate Address Detection Scheme, Conflict detection allocation is clearly shown. Thus we have effectively build a MAC-IP table which is used to auto-configure the IP addresses of the nodes in the network, which takes less than 1 second of the time.

References

1. Wireless and mobility extensions to ns-2. Software Documentation (1999), http://monarch.cs.cmu.edu
2. Cheshire, Aboba, B., Guttman, E.: Dynamic configuration of link-local IPv4 addresses, Internet Draft: draft-ietf-zeroconf-ipv4-linklocal-13.txt (February 2004)
3. Droms, R.: Dynamic host configuration protocol, Internet Engineering task Force (IETF), RFC 2131 (March 1997)
4. Perkins, C.E.: IP mobility support. In: Internet Engineering Task Force (IETF), RFC 2002 (October 1996)
5. Mobile Ad Hoc Networks (MANETs), Advanced Network Technologies Division, Wireless Ad hoc Networks (2008), http://w3.antd.nist.gov/wahn_mahn.shtml
6. Weniger, K., Zitterbart, M.: Address autoconfiguration in mobile ad hoc networks: current approaches and future directions. IEEE Network Magazine 18 (August 2004)
7. Perkins, C.E., et al.: IP Address autoconfiguration for ad hoc networks. Internet Engineering Task Force (IETF), Internet Draft (November 2001), http://people.nokia.net/charliep/txt/aodvid/autoconf.txt
8. Perkins, C.E., Royers, E.M., Das, S.R.: IP address autoconfiguration for ad hoc networks, Internet Draft (July 2000)
9. Guines, M., Reibel, J.: An IP address configuration algorithm for zeroconf mobile multihop ad hoc networks. In: Int'l Workshop, Broadband Wireless Ad Hoc Networks and Services, Sophia Antipolis, France (August 2004)
10. Prakash, R., Nesargi, S.: Manetconf: Configuration of hosts in a mobile ad hoc network. In: Proc. of IEEE INFOCOM 2002, New York, NY (June 2002)
11. Zahoor Ul Huq, S., Kavitha, D., Sreenivas Murthy, K.E., Satyanarayana, B.: Dynamic IP Address Auto-configuration in MANETs. In: Das, V.V., Stephen, J., Chaba, Y. (eds.) CNC 2011, Part III. CCIS, vol. 142, pp. 452–457. Springer, Heidelberg (2011), http://www.springerlink.com/content/u094q70u704r715v/
12. Zahoor Ul Huq, S., Sreenivasa Murthy, K.E., Satyanarayana, B., Kavitha, D.: Dynamic IP Address Auto-configuration in MANETs. Global Journal of Computer Science and Technology 10(1), 23–36 (2010), http://globaljournals.org/GJCST_Volume10/gjcst_vol10_issue_1_paper12.pdf
13. Zahoor Ul Huq, S., Sreenivasa Murthy, K.E., Satyanarayana, B., Kavitha, D.: Analysis of Efficient Address Allocation Schemes in Mobile Ad Hoc networks. International Journal of Engineering Science & Technology (IJEST) 2(3), 227–231 (2010), http://www.ijest.info/docs/IJEST10-02-03-16.pdf

HMM Based Enhanced Dynamic Time Warping Model for Efficient Hindi Language Speech Recognition System

Sharma Krishna Kumar[1,*], Lavania Krishan Kant[2], and Sharma Shachi[2]

[1] Department of Computer Science, Central University of Rajasthan, Kishangarh, Ajmer, India
[2] Department of CS, AIET, Rajasthan Technical University, Jaipur, India
{krisshna.sharma,shachi.sharma14}@gmail.com, k@lavania.in

Abstract. Dynamic Time Warping (DTW) is template based cost minimization technique. We propose Hidden Markov Model (HMM) based enhanced DTW technique to efficiently recognize various speaking rate signals and for recognizing closely similar utterances. We extend the derivation of Viterbi and forward algorithms for finding optimized path alignment in new propose technique and extend the Baum-Welch algorithm to optimize the model parameters. The proposed technique is compared with conventional DTW technique, and from comparative results analysis we find that it improves the results from 84% to 94 % using DTW technique for Hindi spoken words for various speech utterances in different environmental conditions or for varying speakers.

Keywords: Dynamic Time Warping, Hidden Markov Model.

1 Introduction

In informal way, different people speak Hindi in different way like with fast, slow, and normal speaking rate. Then this variation degrades the recognition rate [5]. There are three techniques prevalent for automatic speech recognition (ASR): The template based technique, the statistical based technique, and artificial neural network (ANN) based technique. In template based technique, each word of dictionary is denoted by one or more templates. And we find out the most time aligned path between reference template and testing word, using DTW technique. Statistical based technique is sounder in mathematical formulation, this is based on HMM technique.

Many authors worked for Hindi language speech recognition system. K. Kumar et al [1] have presented an approach for designing an isolated-word, Hindi speech recognition system. The features are derived by employing linear prediction coding coefficients and the speech recognition is performed using continuous HMM. The system provides good performance results for speaker independent environments and

* Corresponding author.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 200–206, 2013.

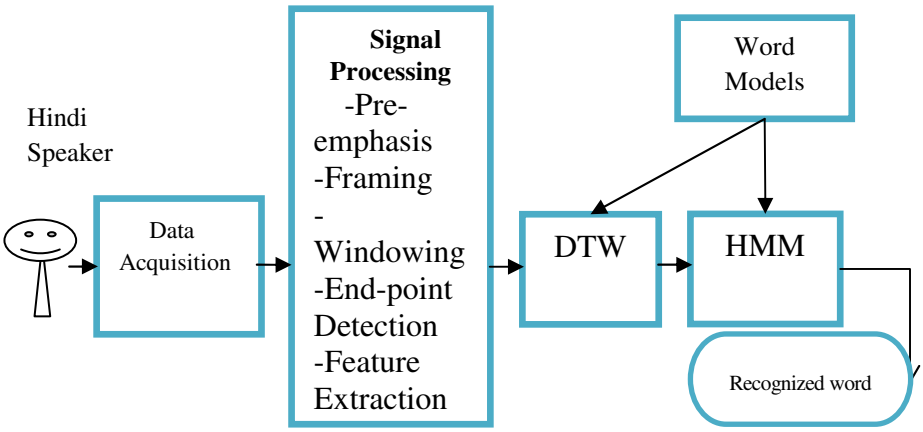


Fig. 1. Block Diagram of Hindi Speech Recognition (HSR)

is highly efficient for large vocabulary size. R K Agarwal et al [2] also presented a HMM with Gaussian mixture based automatic ASR using MFCC features, but this technique was also evaluated for simple utterances. Shivesh Ranjan [3] has also applied LPC over Discrete Wavelet transform coefficients for designing a speech recognition system for Hindi words, but it focuses over computing LPC coefficients separately for the wavelets coefficients. These authors focused on features extraction and matching techniques.

Above authors presented various speech features and HMM based speech recognition system and evaluate them on various situations. If we evaluate them for varying speaking rate then their recognition result degrade. In order to enhance results for varying speaking rate condition, we are proposing model to compress and stretch the testing sample according the reference template and HMM will calculate sum of densities along all possible alignment paths [4].

This paper presents unified approach of vector quantization, DTW, and HMM as shown in Fig. 1. Second section describes signal pre-processing, feature extraction techniques and vector quantization. Third section presents about DTW, enhanced DTW, and HMM based enhanced DTW. In the fourth section we discussed experiments and result analysis. And in the fifth section, conclusion is discussed.

2 Speech Feature Extraction

Speech is a quasi-static signal. Initially Speech is stored in amplitude and time domain at 16KHz and with mono channel in .wav (file extension) form. In order to resemble human-auditory system in our mechanics, we have to extract features in the similar manner. Human auditory system behaves logarithmically for high-frequency signal and linearly for low-frequency signal. So, we used the MFCC (Mel Frequency Cepstral Coefficients) coefficients [6], and their first and second derivatives [7] including energy of the signal. Initially a Hindi dictionary of 13 words is created. Each word is spoken by 40 different

users, 10 times each word. Database of spoken words is created by different-different speakers of varying speaking rate like slowly, moderate, and fast in order to make it efficient. Through random selection, we used 32 speakers' data to extract features and rest data are used for testing purpose. Features are extracted as shown in the block diagram. Linde–Buzo–Gray *algorithm* [10] is used to quantize features. This is an iterative technique of vector quantization. If each word contains n utterances, then n utterances are represented in the following manner:

$$V^1, V^2, \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, V^{(n-1)}, V^n$$

And each utterance contain $l \times m$ feature vector:

$$V_{11}^1, V_{12}^1, \ldots\ldots, V_{1(m-1)}^1, V_{1m}^1 \ldots\ldots\ldots V_{l(m-1)}^1, V_{lm}^1$$

After LBG vector quantization, k utterances are generated:

$$V^1, V^2, \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, V^{(k-1)}, V^k$$

An average template is created from the all k feature vectors to lessen the varying factors of the speech signals. If a word is spoken different times by a single speaker, different-different signal will generate every time as given in the fig. 2. So, this difference will reduce by averaging the signal [4].

$$\bar{V}_{ij} = \frac{\sum_{l=1}^{l=k} {V_{ij}}^l}{K} \tag{1}$$

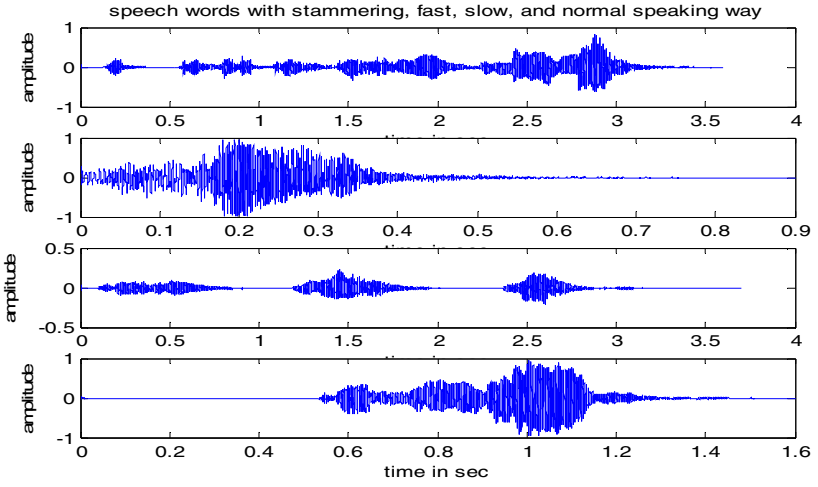


Fig. 2. Single word 'chunana (in English 'Select')' with different speaking rates

3 DTW and Enhanced DTW

The DTW algorithm provides a procedure to align optimally in time the test and the references patterns and to give the average distance associated with the optimal warping path as shown in Fig. 3. Dynamic Time Warping is a pattern matching

algorithm based on Bellman's principle of optimality with a non-linear time normalization effect [11].

The DTW technique creates alignment between the two sequences of feature vectors. $T_1, T_2, \ldots\ldots\ldots, T_n$ and $R_1, R_2, \ldots\ldots\ldots, R_n$:

A distance between two feature vectors T_i and R_i can be represented by a function $f(i,j)$, this is a local distance. For computing global distance, we compute recursively local distances and add them for best predecessor. Best global distance is:

$$D(i,j) = \min_{m \le i, k \le j}[D(m,K)] + f(i,j) \tag{2}$$

DTW is used to establish time scale alignment between two feature vectors. DTW system can capture long-range dependencies in acoustic data, and can potentially adapt to differences in speaker, and accent. From the above analysis, DTW is effective in wide scale observation and HMM is used for solving the analysis of the detail. Hence we can combine these two methods for designing efficient system as shown in Fig. 3. Enhanced DTW [12] can be designed with following elements:

A DTW technique can be applied on given input feature vector sequence $T_1, T_2, \ldots\ldots\ldots, T_n$ and DTW template which is a sequence of states $R_1, R_2, \ldots\ldots\ldots, R_n$. In enhanced DTW each state is characterized by the set of parameters.

1. π is the set of model parameters.
2. Quality of the model is calculated by the following form:

$$f(y|\pi) = \sum_{q \epsilon Q} f(y, q|\pi) \tag{3}$$

   Here, Q is the set of all valid paths.
3. Path quality is measured by

$$f(y, q|\pi) = f(q|\pi) f(y|q,\ \pi). \tag{4}$$

   $f(q|\pi)$ is a prior measure for traversing through path q, and $f(y|q,\ \pi)$ is a posterior quality measure.

Prior path traversal measure is given as:

$$f(q|\pi) = \left[\sum_{i=1}^{L(q)} a_{q_i}^{W(q_i)}\right]^{\rho(q)}, \tag{5}$$

$L(q)$ is the sequence of the connected segments $q_1, q_2, \ldots\ldots\ldots, q_n$ of valid path.

$W(q_i)$ is the weight of the i^{th} segment q.

$W(q) = \sum_{j=1}^{r_q} w_q . j$ (6), r_q is the number of grid points.

$\rho(q)$ is the path independent so it can be denoted by ρ.

The posterior path quality calculation is the summation of quality of segment that make path.

$$f(y|q,\pi) = \prod_{i=1}^{L(q)} b\left(v_{q_i}|\,\pi\right), \tag{7}$$

$b\left(v_{q_i}|\,\pi\right)$ is the quality measure of the segment q and its associated grid points.

$$b(v_{q_i}|\,\pi) = \prod_{j=1}^{L(q)} b_q\left(v_q(j)\pi\right) \tag{8}$$

Where$b_q(v|\pi) = [p(v|\pi)]^{\rho(q)w_q(v)}$, $b_q(v|\pi)$ is the weighted and normalized local similarity calculation for grid point. $p(v|\pi)$ is the Gaussian probability density function.

Our technique is not statistical, we measure quality of the path that may depend upon the path and weighting and normalization violate the condition of statistical behavior. Problems associated with enhanced DTW model are optimal path resolution, and parameter estimation for the model [12].

Path quality can be calculated using the forward and backward algorithms [11]. We define forward variable $\alpha(u)$ as the sum of the path quality measure over all valid path from starting to end point.

$$\alpha(u) \triangleq \sum_{q\epsilon Q_{start,u}} \left\{\prod_{i=1}^{L(q)} a_{q_i}^{\rho W(q_i)}.\, b(V_{qi}|\pi)\right\}. \tag{9}$$

And for the backward algorithm, we start from the last grid point and reached at the start. That can be represented by the $\beta(u)$.

Optimal path can be measured by maximizing the all valid path originating from the starting point.

$$\delta(u) \triangleq \max_{q\epsilon Q,start,u} \left\{\prod_{i=1}^{L(q)} a_{q_i}^{\rho W(q_i)}.\, b(V_{qi}|\pi)\right\}. \tag{10}$$

We calculate the parameter set π and given training patterns are $V = V^1, V^2, \ldots\ldots\ldots\ldots\ldots\ldots\ldots, V^{(n-1)}, V^n$. Our objective is to maximize$f(V|\pi)$ [12]. i.e.

$$\hat{\pi} = \arg\max_{\pi} f(V|\pi) \tag{11}$$

4 Experiments and Result Analysis

We developed a Hindi speech recognition system in order to increase user-machine interaction for non-technical and for rural users. They can use this interaction system in knowing anaj mandi (a place of selling and buying of agriculture goods) rate of cereals, vegetables, and other agriculture products and in other technical gadgets. So, we stored speech utterances of various rural users from different places. We choose eight places and five users from each place. Each user speak ten times single word so, we get 400 utterances for each word. We have 13 words dictionary so, total 400*13=5200 spoken words.

DTW technique helps in normalizing the time of utterances and this improves the ability of Baum Welsh algorithm and in learning the data because sequence of feature vectors aligned in time and it favours in discriminating the speech feature vector from other templates than similar templates. For the phonetically close utterances, after the alignment they have their own size and it increases the discrimination of the models at the transition matrix level.

Eighty spoken words for each thirteen Hindi words from the training data are used for the testing purpose of the recognition system in the three different environments. We

compared the results as shown in the fig. 4, from both enhanced and conventional DTW technique and find out that results improved from 84% to 94% recognition rate in closed class room environment. In the conference room results improved from the 81.5% to 92.5% and in the open air environment improved from 80% to 91.5%.

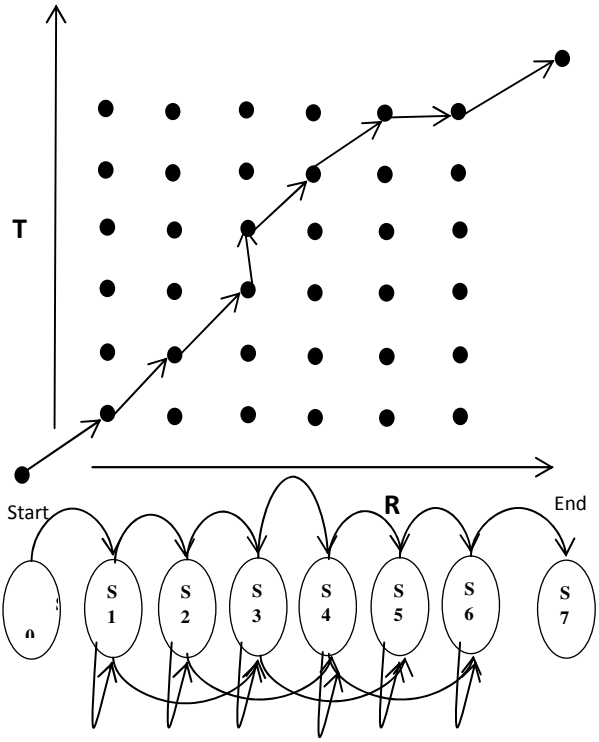


Fig. 3. Unified model of DTW and HMM

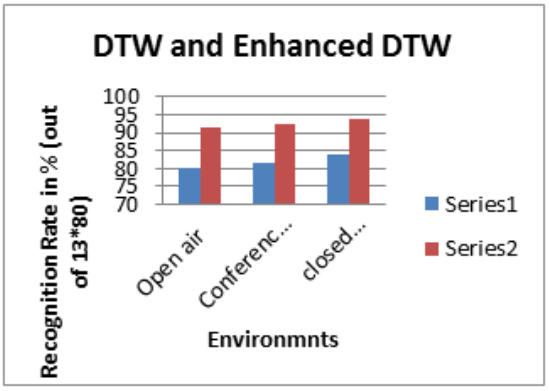


Fig. 4. Result analysis for both techniques

5 Conclusion

This paper presents a speech recognition system for Hindi language that can work efficiently for varying speaking rate speech utterances and for closely similar utterances. We design a unified model of DTW and HMM in order to combine template based approach and statistical approach. This model work efficiently for objective and we test it on a Local area Network application of query-system for rate of agriculture goods. It is tested by rural people. Enhanced DTW increased results from 84% to 94% by increasing 10% in recognition rate.

References

1. Kumar, K., Agarwal, R.K.: Hindi speech recognition system using HTK. International Journal of Computing and Business Research 2(2) (May 2011) ISSN (Online): 2229-6166
2. Aggarwal, R.K., Dave, M.: Design and Modeling of A Speech Understanding System for Hindi Language, Deptt. of Computer Engg. N.I.T., Kurukshetra
3. Ranjan, S.: A Discrete Wavelet Transform Based Approach to Hindi Speech Recognition. In: 2010 International Conference on Signal Acquisition and Processing (2010)
4. Abdulla, W.H., Chow, D., Sin, G.: Cross-words Reference Template for DTW-based Speech Recognition Systems. In: TENCON 2003. Conference on Convergent Technologies for Asia-Pacific Region, October 15-17, vol. 4, pp. 1576–1579 (2003)
5. The online encyclopedia of writing systems & languages, `http://www.omniglot.com/index.html`
6. Furui, S.: Cepstral analysis technique for automatic speaker verification. IEEE Trans. ASSP-29(2), 254–272 (1981)
7. Furui, S.: Speaker-independent isolated word recognition using dynamic features of spectrum. IEEE ASSP-34(1), 52–59 (1986)
8. ETSI ES 201 108, V1.1.3.: ETSI standard: speech processing, transmission and quality aspects (STQ); Distributed speech recognition; Front-end feature extraction algorithm; Compression algorithms; Sect. 4, pp. 8–12, (September 2003)
9. Pruthi, T., Saksena, S., Das, P.K.: Swaranjali: Isolated word recognition for Hindi language using VQ and HMM, Hughes Software Systems and IIT Guwahati (2000)
10. Sharma, K.K., Kapoor, P., Chakraborty, P., Nandi, G.C.: Dynamic Spectrum Derived MFCC and HFCC Parameters and Human Robot Speech Interaction. In: International Conference on Advances in Computer Engineering–ACE 2011 (2011)
11. Hocine Bourouba, E., Bedda, M., Djemili, R.: Isolated Words Recognition System Based on Hybrid Approach DTW/GHMM. Informatica 30, 373–384 (2006)
12. Yaniv, R., Burshtein, D.: An Enhanced Dynamic Time Warping Model for Improved Estimation of DTW Parameters. IEEE Transactions on Speech and Audio Processing 11(3) (May 2003)

Feedforward Feedback(FFFB) Method for Dereferencing the Non Blind Algorithms in Adaptive Beamforming

Vaibhav Narula and Rajat Srivastava

Amity School of Engineering and Technology, New Delhi, India
{vaibhav3589,razerrajat}@gmail.com

Abstract. This paper presents a novel method of de-referencing (Blinding) the Non-Blind algorithms employed in the application of Adaptive Beamforming developed for smart antenna applications. The method (primarily based on Adaptive Controller)involves taking a FeedForward channel from receiver antenna output and its summation(subtraction) with the system o/p FeedBack can actually replace the role of reference signal, conventionally required to calculate the feedback error. The summation/subtraction results in error signal which is then inserted into the Adaptive Algorithm. LMS and RLS Algorithms are used in this paper for experimentation. System input was a DS-SS QPSK digital signal . This Beamforming scheme was also successfully observed using Agilent ADS by doing basic level simulations for micro-strip patch array.

Keywords: Adaptive Beamforming, Feedforwar-, Feedback, Adaptive Controller, LMS, RLS, Microstrip-patch Antenna, Agilent ADS.

1 Introduction

Smart Antennas is a system of antenna array which identifies spatial signal signatures with the help of Adaptive Beamforming Algorithms. A process of adapting magnitude and phase of the signal from every antenna element to form narrow beams in order to concentrate the reception in specific direction while neglecting the interference , is known as Beamforming[3,4]. For fixed DOA(direction of arrival), simple beamforming is used whereas for time variant DOA, Adaptive Beamforming is employed [5].

Adaptive array systems as shown in figure 2(a) use antenna arrays controlled by robust adaptive algorithms, to dynamically vary the radiation pattern in accordance with the varying environment of the signal. There are practiced many performance criteria for such algorithms namely MMSE(min. mean square error), Maximum signal to interference and noise ratio (SINR), maximum likelihood (ML), Minimum noise variance and Maximum gain [6]. Our paper is concerned with MMSE based algorithms only.

The two major types of adaptive algorithms are blind and non-blind. Non-blind algorithms require a pilot(reference) signal to detect the desired signal and update their complex weights. These algorithms include LMS(Least Mean Square), RLS(Recursive

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 207–214, 2013.

Least Square), SMI(Simple Matrix Inverse) and CGM. In contrast to non-blind algorithms, blind algorithms do not need a pilot signal to find the complex weights. This paper examine the functioning of LMS and RLS algorithms, conventionally accepted as Non-Blind algorithms, with means of proposed method, as Blind algorithms.

1.1 Microstrip Patch Antenna

Microstrip patch antennas are planar structures which resonate from a strip of metal embedded over a dielectric substance and a metal plate which together constitute the ground plane. Common antenna shapes are a polygon, circular and elliptical, however any continuous shape can be used. The default microstrip antenna is a rectangular patch of metal over a ground plane. By varying its dimensions and altering its physical structure, several variations are possible. They are small, versatile, cheap and quite lightweight.

1.2 Related Work

Retrospective considerations will guide us through the immense amount of work already done in this field [2],[3],[4],[8],[9] but never before the conventional beamformer models were viewed in this light. This novel approach will take out two of the most efficient and popular beamforming algorithms from the less pragmatic Non-Blind category and present them with a new modular approach. Untill present , LMS and RLS Algorithms have always been used as Non-Blind beamforming applications. This new method of dereferencing can also lead to the exploration of adaptive beamforming based on Adaptive (MRAC) based controllers towards new directions.

2 Problem Statement

In Blind algorithms ‘d(n)’(reference signal) is not used where as in Non-Blind algorithms, ‘d(n)’(or some other closely correlated training string) has to be known both to the transmitter and receiver during the training period. Acquiring a training sequence all the time is very difficult when High Freq switching antennas deployed for high Bandwidth data exchange. Main approach of this paper is to address the issue hereby mentioned.

2.1 Feedforward-Feedback (Adaptive) Control Mechanism

In real life complex control systems such as Neural system functions efficiently only with a coordinated inputs from its Feedforward(FF) was well as Feedback(FB) motor faculties. [10]Feedforward : by identification, it anticipates the relation between system and environment to determine the course of action, Feedback : monitors discrepancies which can be used to refine actions of the controlled system. The controller, when coupled to an identification procedure, is precisely what is often referred to as an 'adaptive controller': it adapts its control strategy to changing estimates of the dynamics of the controlled system. Such system illustrated in Figure 1 contains both controller plus identification procedure.

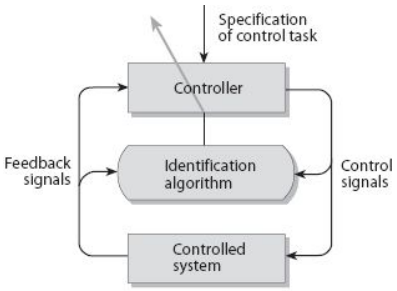


Fig. 1. To render a controller adaptive, an identification algorithm monitors control signals and feedback signals to provide the controller with updated estimates of the parameters that describe the controlled system

2.2 Adaptation of FF-FB to Adaptive Beamforming

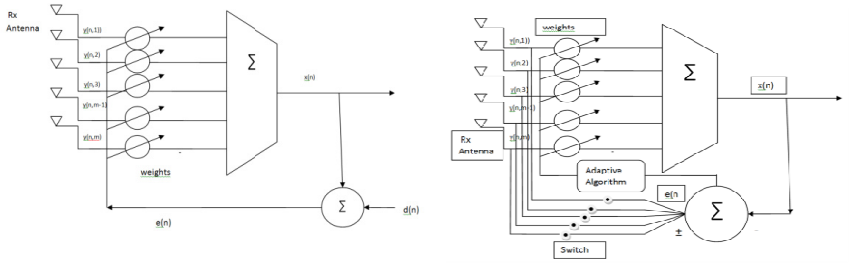
The proposed method is basically an adaptation of Feedforward-Feedback based Adaptive Controller into the scheme of Adaptive Beamforming .

The only addition in this model over the conventional one is that of a feedforward(FF) path. This FF highway('m' number of lines) consist of paths from each of the antenna output channels and is extracted prior to the weight coefficient multiplication. Signals from these lines are added (to)/subtracted(from) the Feedback(FB) from system output x(n). The resultant is then feedback into the adaptive system as error signal e(n) to initiate beamforming in the DOA(direction of arrival) of primary signal. Since it's a narrow band operation therefore only one weight is associated with each antenna channel.

A special characteristic of this model is that, the requirement of reference signal is completely surpassed. So following this model, all the earlier Non-Blind algorithms can now properly function even in the Blind format.

The model shown in Figure .2(b) is called FFFB(feedforward-feedback) beamformer. A special characteristic of this beamformer is that, it incorporates FF path to compensate for the reference signal. FF signal is extracted from the antenna output channels prior to weight coefficient multiplication. This FF highway of 'm'(number of array, antenna elements) channels are subtracted/added to the feedback from system output x(n). The resultant is then fed back into the adaptive system as error signal e(n) to initiate beamforming in the DOA(direction of arrival) of primary signal. Since it's a narrow band operation therefore only one weight is associated with each antenna channel.

Distinct channels of the FF highway are selectively switched on/off . The pattern of selective switching till present is entirely based on (best hit adjustments of) different case simulations and has yet not been embedded on a general mathematical formulation. Multiplication of the values(+1,-1,0) with the respective FF channel lines is a simple switching mechanism considered here for practical simulations.



(a) (b)

Fig. 2. (a) Conventional Beamformer ; (b) Feedforward-Feedback(FFFB) Beamformer

The expression of signal received at the antenna(m=8) is

$$\sum_{m=1}^{8} y_m(n) = \sum_{m=1}^{8}[S_m(n)p(\theta_0) + \sum_{i=1}^{Nu} U_{im}(n)p(\theta_i)] + N(n) \quad (1)$$

Here, y_m= received signal, S_m= Primary Signal, θ_0= DOA of Primary Signal, U_{im}= Interfering Signals, Nu= Number of Interfering Signals, θ_i=DOA of Interfering Signal, N=Ambient Noise, $p(\theta)$= steering vector. Weight update equations for LMS and RLS Algorithms are, (here $\alpha_i = \pm 1,0$)

$$w_{LMS}(n+1) = w(n) + \mu y_m(n)e^*(n) \quad (2)$$

$$w_{RLS}(n+1) = w(n) + \frac{\delta^{-1}y_m(n)e^*(n)}{\lambda + y_m(n)\delta^{-1}I} \quad (3)$$

$$e(n) = \sum_{i=1}^{8} \alpha_m y_m - y_m^H(n)w(n) \quad (4)$$

3 MATLAB Simulations

3.1 Case 1. (Directional Interference Signals + Ambient Noise)

Number of array elements (m)=8 ; Carrier Frequency = 2.53 GHz (within CDMA bandwidth) ; D(spacing) = $\lambda/2$(half of wavelength) ; Input Data : Direct Sequence Spread Spectrum (DS-SS) data stream modulated using BPSK scheme. ; y_1-y_8 are Individual receiver channel lines.

3.1.1 Least Mean Square(LMS) Algorithm

Step Size = 0.8 ; Direction(Angle) of interferences (in degrees) : 0,20,40,60,80 + ambient AWGN ; Signal to Interference Ratio(SIR)= -1 dB ;[for SIR worst case scenario has been considered, below -1dB the MSE becomes very large and beam pattern is destabilised ; and viceversa, increasing the SIR also increases the accuracy of beam patterns proportionally and thus smoothens the beams].

Table 1. Switch Values for FF channel

DOA	10^0	30^0	50^0	70^0
Y1	0	-1	-1	-1
Y2	-1	-1	0	1
Y3	-1	0	0	-1
Y4	0	1	-1	0
Y5	0	0	1	-1
Y5	0	0	0	1
Y7	0	0	0	-1
Y8	1	1	1	1

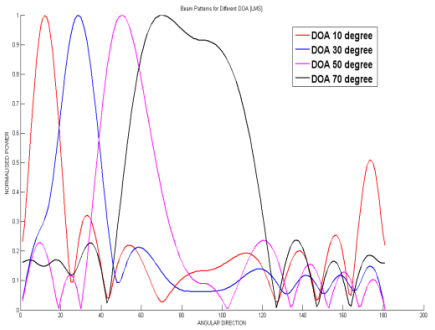


Fig. 3. Beam Pattern Plot for FFFB beamformer

3.1.2 Recursive Least Squares(RLS) Algorithm

Forgetting Factor = 0.9(λ) ; Direction(Angle) of interferences (in degrees) :10,30,50,70,90 + ambient AWGN ; Signal to Interference Ratio(SIR)= -70 dB ; [for SIR worst case scenario has been considered, below -70dB the MSE becomes very large and beam pattern is destabilised ; but viceversa, increasing the SIR doesn't much increases the accuracy of beam patterns in RLS Algorithm].

Table 2. Switch values for FF Channel

DOA	20^0	40^0	60^0	80^0
Y1	-1	0	-1	-1
Y2	1	1	0	1
Y3	1	0	-1	-1
Y4	0	-1	0	1
Y5	-1	1	-1	-1
Y5	-1	0	0	1
Y7	-1	-1	-1	-1
Y8	1	1	0	1

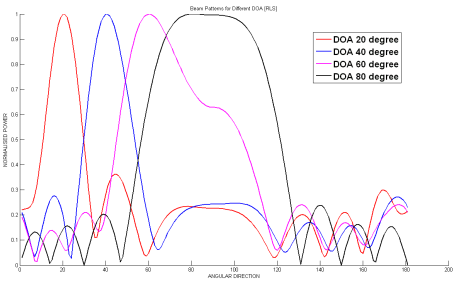


Fig. 4. Beam Pattern Plot for FFFB beamformer

One more important effect observed about the behaviour of conventional beamformer (both for LMS and RLS) was that its beamforming capabilities are seriously hampered for angles near zero degree and became better as the DOA moves away from 0 degree. Following Figure 5. simulation demonstrates this.

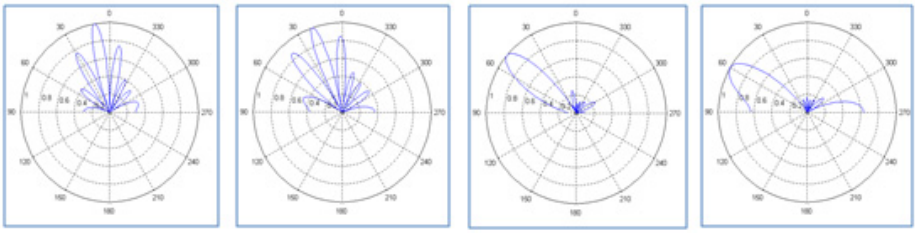


Fig. 5. Polar Graphs [conventional beamformer]; (from left to right) DOA=10^0 (LMS);DOA=20^0 (RLS); DOA=50^0 (LMS); DOA=60^0 (RLS)

3.2 Case 2 (Ambient Noise)

For the conditions having presence of ambient or un-directional AWGN, The FFFB beamformer does not need multiple Feedforward lines. Only employing last two channels produces directed beam as accurate as in case of reference based conventional beamformer.

$$e(n) = y(m) - y(m-1) - x(n) \tag{5}$$

here ,m=8(number of antenna channels/weights) where e is error , y(m) and y(m-1) are FF channels and 'x' is feedback signal. The following comparative simulations were carried on using LMS algorithm for SIR = -1, DOA=10^0 for non directional interference.

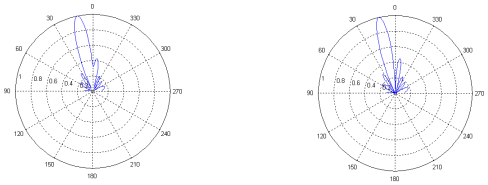
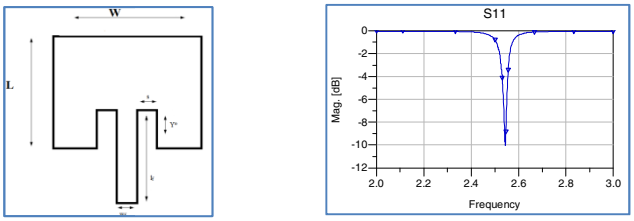


Fig. 6. (DOA=100,LMS);left plot is FFFB beamformer & rightplot is conventional beamformer

4 Agilent ADS(Advanced Design Systems) Simulations

In order to design the antenna, we required only the resonant frequency, height of the substrate and the dielectric constant. Based on these values the antenna was designed from the equations mentioned in [7]. The feeding method of the antenna was been selected as inset feed. The final dimensions of the antenna were Width=43.45mm ,length=35.2 mm, effective permittivity (ϵ_{reff}) =2.9. The length of the feedline (l_f) was 27.5mm and width(w_f) was 10.42mm with separation (s) as w_f /2, or 5.21mm and inset width (Y^o)as 9.33mm. To design the array the effective distance between two antennas was kept as λ_c /2, where λ_c is the resonant wavelength. The final design of a single antenna and the output of the simulated array is shown in Figure 7



(a) (b)

Fig. 7. (a) Design of a microstrip patch antenna ; (b) Return loss of the array

Now the coefficient values from FFFB beamformer system , for all the tested DOA's was put for the weighted multiport excitation (voltage) values of constructed microstrip array and Momentum Vis. beam pattern O/Ps were observed for E-plane beam analysis of the array. Following were the observations

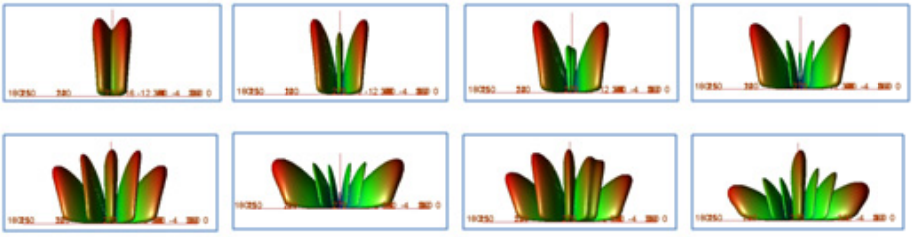


Fig. 8. E-Plane beam pattern of the array. Top Row (left to right) :DOA=10^0,20^0,30^0,40^0 Bottom Row (left to right) :DOA=50^0,60^0,70^0,80^0

5 Conclusion

Thus the proposed Feedforward-Feedback Beamformer can be successfully employed in implementing reference-less (Blind) LMS and RLS Algorithms. This new model also facilitates further improvements over the conventional beamformer other than dereferencing it. ADS simulations also show satisfactory results for optimum beamforming. Although beam patterns are not as sharp and perfect compared to MATLAB, but they still do it enough to indicate promise for further research on this platform also.

6 Future Scope

Till present switching decision control of this model rest purely on Brute Force method focussed on exhaustive simulation results and is therefore lacking any proper mathematical formulation. So future developments in this regard should first aim at constructing a rigorous mathematical foundation for its justification and progressive development. Moreover there also needs to be meticulous testing of this model on a

whole variety of applications wherever beamforming concept is utilised . Aggressive examination by scholars can also be initiated for adaptive beamforming using Agilent ADS and not restrict only to popular softwares like MATLAB.

References

1. Chryssomallis, M.: Smart antennas. IEEE Antennas and Propagation Magazine 42(3), 129–138 (2000)
2. Rani, S., Subbaiah, P.V., Reddy, K.C., Rani, S.S.: LMS and RLS Algorithm for Smart Antennas in a W-CDMA Mobile Communication Environment. ARPN Journal of Engineering and Applied Sciences 4(6) (August 2009)
3. Islam, M.T., Rashid, Z.A.: MI-NLMS adaptive beamforming algorithm for smart antenna system applications (July 2006)
4. Ali Qureshi, M., Aziz, A., Bakhsh, H.: CIR Improvement using Antenna Beamforming and Power Control. In: Proc. 2010 International Conference on Networking and Information Technology, ICNIT, Manila Philippines, pp. 294–297 (2010) ISBN: 978-1-4244-7577-3
5. Gross, F.: Smart Antennas For Wireless Communication. Mcgrawhill (September 14, 2005)
6. Balanis, C.A., Panayiotis: Introduction to Smart Antennas (Synthesis Lectures on Antennas). Morgan & Claypool (2007)
7. Balanis, C.A.: Antenna Theory. John Wiley & Sons, Inc. (1997)
8. Ramasamy, V.: Beamforming for MC-CDMA. Master Desertion, Virginia Polytechnic Institute and State University (2003)
9. Majeed, H., Umar, R., Basit, A.A.: Smart Antennas MIMO, OFDM & single carrier FDMA for LTE. Master Desertion, Lennaeus University (2011)
10. Feedforward and Feedback (Tutorial by Arbib, M. S.), http://www.answers.com/topic/feedback-and-feedforward

Extended Queue Management Backward Congestion Control Algorithm

V. Sinthu Janita Prakash[1], D.I. George Amalarethinam[2], and E. George Dharma Prakash Raj[3]

[1] Dept of Comp Science, Cauvery College, Tiruchirappalli, India
[2] Jamal Mohammed College, Tiruchirappalli, India
[3] Computer Science, Bharathidasan University, Tiruchirappalli, India
{sinthujanitaprakash,georgeprakashraj}@yahoo.com,
di_george@jmc.edu

Abstract. Network Congestion is a major problem with respect to the Internet. The Active Queue Management algorithms (AQM) are used to reduce congestion, and in this paper, a new AQM algorithm is proposed called as Extended Queue Management Backward Congestion Control Algorithm (EQMBCCA). This is an extended version of the QMBCCA, which has been proposed earlier by us. QMBCCA makes use of Internet Control Messaging Protocol (ICMP) Source Quench (ISQ) notifications and also the CE (Congestion Experienced) bit whenever the average queue size crosses minimum threshold value. EQMBCCA introduces a configurable intermediate threshold value IntThres between the minimum and maximum values and generates ISQ signals only if the congestion crosses the threshold value. Therefore the generation of ISQ messages is significantly reduced here. There is also reliability since CE bits are set in the packets once congestion occurs. The proposed algorithm is compared with the Queue Management Backward Congestion Control Algorithm (QMBCCA) in terms of Gain in Goodput, Loss percentage and ISQ traffic for FTP flows. It is found that the performance of EQMBCCA is almost equal to that of QMBCCA and there is a significant reduction in the ISQ traffic in the reverse direction.

Keywords: Congestion Control, Active Queue Management, RED, ECN, QMBCCA, Packet Loss, Goodput, Fairness, Delay.

1 Introduction

When multiple input streams arrive at a router whose output capacity is less than the sum of the inputs, congestion occurs [1]. The dropping of incoming packets in the absence of the buffer space is the usual method that a router reacts to congestion. This method is called as “Tail Drop” or “Drop Tail”.

TCP is the most leading transport protocol today [2]. It is assumed that persistent TCP flows account for most of the buffer demand and short-lived (e.g. web-traffic) TCP or non-elastic (e.g. streaming) flows just add a little bit more to the buffer requirement [3]. Whenever congestion occurs in the network, an intermediate router begins to drop packets. In response to this instance of congestion, TCP invokes its congestion control algorithms [4] so that congestion is eased.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 215–221, 2013.

Active Queue Management (AQM) aims at minimizing queuing delay while maximizing the bottleneck link throughput. [5] reveals that the likelihood of congestion grows exponentially with queue occupancy, suggesting that drop rates ought to increase accordingly

This paper is organised as follows. Section 2 gives a description about the related work on AQM Algorithms such as RED and its variants and QMBCCA. Section 3 proposes the new algorithm EQMBCCA. The network scenario, simulation parameters, performance metrics and the corresponding comparative graphs are given in Section 4. Section 5 concludes the paper.

2 Related Work

Active Queue Management algorithms basically detect congestion before the queue overflows and it also gives an indication of the congestion that has occurred to the end nodes [6]. [7] proposes a fair active queue management (AQM) mechanism based on a well-known economic model called supply and demand.

Random Early Detection [6] follows the concept of Early Random Drop by introducing the measure of average queue size and also dynamically changing the drop probability. A new AQM algorithm is proposed in [8] that exploits Round Trip Time (RTT) explicitly by introducing a passive RTT estimation technique in a router. Some of the variants of RED are Loss Ratio-based RED (LRED), PERED (Pre-Estimation RED) and BO-ARED (An Adaptive RED Algorithm Combined with Buffer Occupation) [9], [10], [11]. RED parameter tuning in response to changing network conditions like traffic load, link capacity and round-trip time is used in [12]. An algorithm is given in [13] which enhances RED as an add-on patch that makes minimal changes to the original RED. The article [14] investigates the influence of weighted moving average on the performance of the RED mechanism. Effective RED (ERED) [15] aims to reduce packet loss rates in a simple and scalable manner. A solution is proposed in [16] to overcome the difficulties of configuring the RED parameters by using a Kohonen neural network model. A discrete-time dynamical feedback system model of TCP/RED is used in [17].

QMBCCA is an alternative approach to the current Explicit Congestion Notification (ECN) algorithm involving the use of an existing Internet ISQ signaling mechanism in addition to the CE notification. By using ISQ, Congestion notification is kept at the IP level. Both the CE (Congestion Experienced) and ECT(ECN Capable Transport) bits are maintained as in ECN. The performance analysis of QMBCCA with ECN for web transfers was done by us in [18].

3 EQMBCCA

QMBCCA makes use of both ECN and Backward Explicit Congestion Notification (BECN) mechanism. In case of congestion, it generates ISQ and also sets the CE bits. There is an overhead in the generation of ISQ messages even for lighter congestion

and it affects the performance of internet routers. ISQ generation is carried by the control plane of the routers which is compared to be slower than the data plane. The transfer of packets between the control plane and data plane should be reduced in high speed routers.

ECN notifications are more reliable compared to ISQ notifications. When the destination receives the ECN notification, it generates ECN Echo acknowledgements continuously until the sender notifies the receiver. The sender reduces the congestion window and the threshold value and activates the CWR (Congestion Window Reduced) mechanism. Additional reverse traffic will be introduced when ISQs are sent to the source and is preferable to minimize such control traffic. On the other hand the RTT taken by the CE mechanism will result in the delay of notifying the congestion leading to higher queue variance and reduced throughput.

EQMBCCA tries to solve these issues by using CE mechanism for low congestion and ISQ in addition to CE for high congestion. In order to differentiate between high and low congestions, an intermediate value is chosen between the minimum and maximum values of RED called IntThres. IntThres is configurable and can be set anywhere between the boundary values. The router behaves as an ECN when the queue is below IntThres and as QMBCCA otherwise. There should be a reduction of the window size of the source for whichever notification reaches the sender first.

The EQMBCCA uses a single bit in the 32 bit unused field in the ICMP source quench packet to differentiate between an ISQ due to a marked packet and an ISQ due to a dropped packet. If the bit is set, it is due to a marked packet and the sender detecting it should wait a full RTT before increasing the window according to the TCP algorithm. If the bit is not set, it is due to a dropped packet and the source halves its congestion window and starts increasing the window according to the TCP algorithm. The algorithm for the proposed EQMBCCA is as follows.

```
For each packet arrival
Calculate the average queue size avg
       If (avg > maxth)
          {
            if (ECT  bit  marked)
              {
              drop the packet;
              send an ISQ2  back to the source;
              }
            Else drop the packet;
        }
       else if (avg < minth ) queue the packet;
       {
             else  if   (minth < avg < IntThres)
          {
            if (ECT  bit  marked)
                 {
      mark the CE bit if not already marked;
                     else drop the packet;
                 }
            }
```

```
        Else if (ECT bit marked)
            {
             mark the CE bit if not already marked;
             Send ISQ1 to the Source;
            }
         Else Drop the Packet;
       Send ISQ2 to the source;
}
```

The flowchart for the proposed EQMBCCA is given in Fig.1.

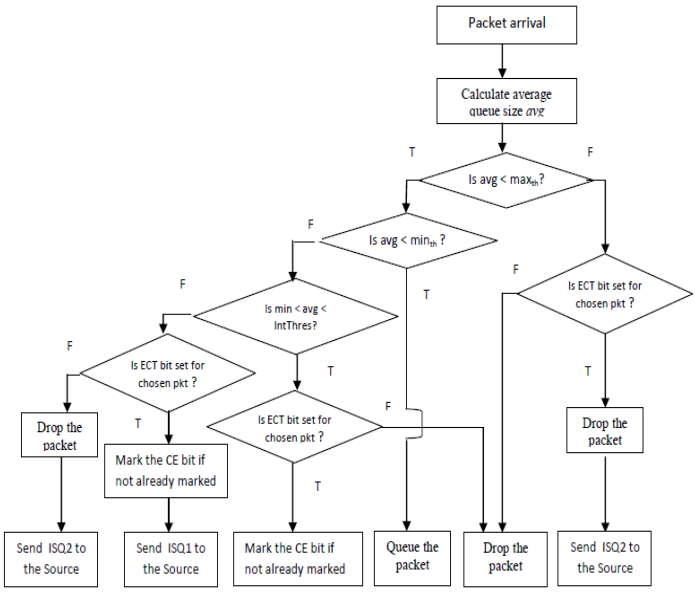


Fig. 1. Flowchart for EQMBCCA

4 Experimentation and Analysis

Simulation is done by Network Simulator (NS ver 2.1b8a). Fig.2. gives the network topology used for comparing the two algorithms QMBCCA and EQMBCCA. FC(1).. FC(n) serve as FTP clients with nodes FS(1)..FS(n) as corresponding FTP servers. Traffic flows from servers to clients. TCP used is New Reno with a data packet size of 1000 bytes, Acknowledgement packet size of 40 bytes and window size set to 100KB. The RED parameters were minth = 15KB, maxth = 3*minth, buffer size = 2*maxth, IntThres = 30KB, maxp = 0.1, wq = 0.002 and byte based dropping was used. . The total number of competing flows are varied from 20, 40, 60, 80, 100, 120, 140, 160, 180 and 200. The start time of all the flows are varied between 0s and 5s.

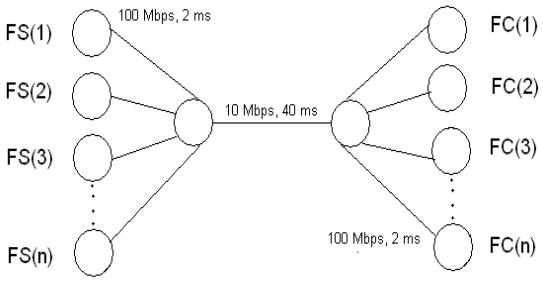


Fig. 2. Network Topology

The performance metrics taken into consideration for this topology are Gain in Goodput, Loss percentage and ISQ traffic in the reverse direction. Goodput for a TCP flow is computed based on the number of data packets received by the receiver.

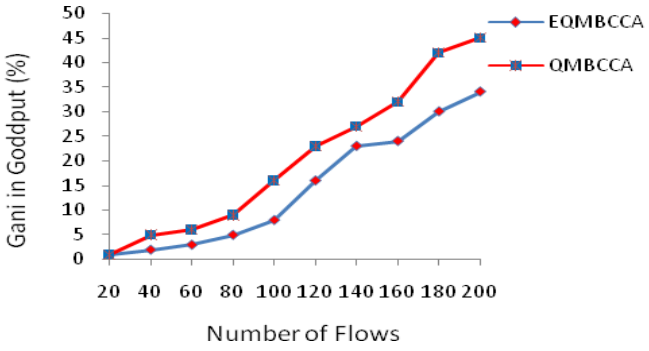


Fig. 3. Average Goodput Gain

Fig. 3. Shows the percentage of the goodput gain by varying the number of TCP flows. Due to early congestion notification by QMBCCA, it has a gain of 43.5% compared to 33.4% of EQMBCCA who generates ISQs only after it crosses the IntThres value.

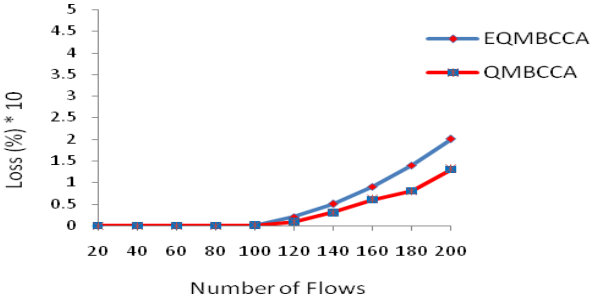


Fig. 4. Percentage Loss

Loss percentage measures the ratio of the number of packets dropped at the bottleneck link to the number of packets injected into it. QMBCCA experiences lesser loss than EQMBCCA (Fig. 4) during high congestion. The average goodput and loss percentage show no significant improvement over QMBCCA.

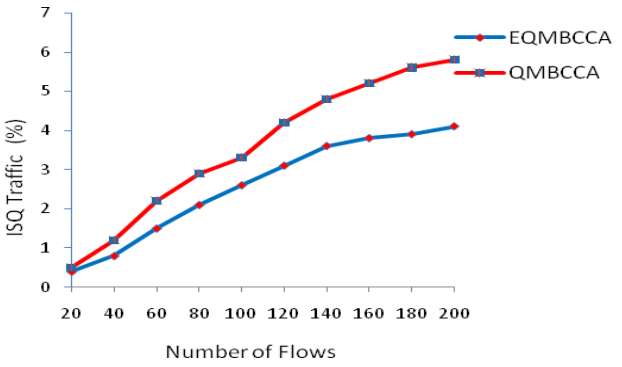


Fig. 5. ISQ Reverse Traffic

Percentage of ISQ traffic is computed as the ratio of the number of ISQs generated in the reverse direction of data flow to the total number of packets in the reverse direction. In Fig. 5. we observe that the percentage of ISQ reverse traffic is significantly reduced in EQMBCCA compared to QMBCCA since ISQs are generated only when the average queue size exceeds IntThres.

5 Conclusion

This paper proposes a new AQM algorithm called Extended Queue Management Backward Congestion Control Algorithm (EQMBCCA) which is an extended version of the QMBCCA proposed earlier. QMBCCA makes use of ISQ notifications and also the CE bits whenever the average queue size crosses minimum threshold value. EQMBCCA introduces a configurable intermediate threshold value between the minimum and maximum values and generates ISQ signals only if the congestion crosses the threshold value IntThres. Therefore the generation of ISQ messages is significantly reduced here. There is also reliability since CE bits are also set in the packets once congestion occurs. IntThres value is configurable and hence when it is fixed to a value nearer to minth, it behaves like a QMBCCA and when it is fixed nearer to maxth, it behaves like ECN algorithm. The algorithm tries to rectify the drawback of QMBCCA by significantly reducing the number of ISQs generated and also the drawback of ECNs RTT delay. The proposed algorithm is compared with the Queue Management Backward Congestion Control Algorithm (QMBCCA). The comparison is made in terms of Gain in Goodput, Loss percentage and ISQ traffic in the reverse direction for FTP flows. It is found that the performance of EQMBCCA is almost equal to that of QMBCCA and there is a significant reduction in the ISQ traffic in the reverse direction.

References

1. Curran, K., Woods, D., Mc Dermot, N., Bradley, C.: The Effects of Badly Behaved Routers on Internet Congestion. International Journal of Network Management 13, 83–94 (2003)
2. Mascolo, S., Casetti, C., Gerla, M., Lee, S.S., Sanadidi, M.: TCP Westwood: Congestion control with faster recovery. UCLA CSD Technical Report #200017 (2000)
3. Avramova, Z., De Vleeschauwer, D., Wittevrongel, S., et al.: Dimensioning Drop-tail and AQM (RED) buffers at access networks for optimal performance with bulk data TCP traffic. Computer Communications 33, S58–S70 (2010)
4. Ryu, S., Rump, C., Qiao, C.: Advances In Internet Congestion Control. IEEE Communications Surveys 5(1) (2003)
5. Barrera, I.D., Arce, G.R., Bohacek: Statistical approach for congestion control in gateway routers. Computer Networks 55(3), 572–582 (2011)
6. Özekes, S.: Evaluation of Active Queue Management Algorithms. Istanbul Commerce University Journal of Science 4(7), 123–139 (2005)
7. Saadatfar, H., Yaghmaee, M.H., Mashhadi, H.R.: A fair active queue management approach based on supply and demand model. In: 5th International Symposium on Telecommunications (IST), Tehran, December 4-6, pp. 261–265 (2000), doi:10.1109/ISTEL.2010.5734034
8. Hoshihara, H., Koga, H., Watanabe, T.: A New Stable AQM Algorithm Exploiting RTT Estimation. In: 31st IEEE Conference on Local Computer Networks, pp. 143–150 (2006)
9. Wang, C., Liu, J., Li, B., Sohraby, K., Thomas Hou, Y.: LRED: A Robust Responsive AQM Algorithm Using Packet Loss Ratio Measurement. IEEE Transactions on Parallel and Distributed Systems 18(1), 29–43 (2007)
10. Zhen-zhen, C., Yang, X.: PERED: A new AQM Algorithm. In: International Conference on Wireless, Mobile and Multimedia Networks, China, pp. 1–4 (August 2009)
11. Luo, Y., Chen, H., Yang, X., Hao, Y.: BO-ARED: A new AQM algorithm with adaptive Adjustment of parameters. In: 8th World Congress on Intelligent Control and Automation (WCICA), China, pp. 1852–1857 (August 2010)
12. Chen, Yang, S.H.: The mechanism of adapting RED parameters to TCP traffic. Computer Communications 32(13), 1525–1530 (2009)
13. Masoumzadeh, S.S., Meshgi, K., Ghidari, S.S., Taghizadeh, G.: FQL-RED: an adaptive scalable schema for active queue management. International Journal of Network Management 21(2), 147–167 (2011)
14. Domańska, J., Domański, A., Augustyn, D.R.: The Impact of the Modified Weighted Moving Average on the Performance of the RED Mechanism. In: Kwiecień, A., Gaj, P., Stera, P. (eds.) CN 2011. CCIS, vol. 160, pp. 37–44. Springer, Heidelberg (2011)
15. Abbasova, B., Korukoglub, S.: Effective RED: An algorithm to improve RED's performance by reducing packet loss rate. Journal of Network and Computer Applications 32(3), 703–709 (2009)
16. Lochin, Talavera, B.: Managing Internet routers congested links with a Kohonen -RED queue. Engineering Applications of Artificial Intelligence 24(1), 77–86 (2011)
17. Chen, J.Y., Hu, C.Y., Ji, Z.: Self-Tuning Random Early Detection Algorithm to Improve Performance of Network Transmission. Mathematical Problems in Engineering Article Number: 872347 (2011)
18. Prakash, S.J., George Amalarethinam, D.I., George Dharma Prakash Raj, E.: Fairness and Percentage Loss Analysis for Short Lived Web Transfers using QMBCCA. The International Journal of Computer Applications, IJCA 16(2), 40–43 (2011)

Electronic Guiding Stick to Assist the Visually Challenged

K. Divya, P. Dhivya, R. Gayathri, and P. Govindaraj

Alpha College of Engineering,
Department of Biomedical Engineering,
Chennai - 600 124, Tamilnadu, India
{divyammu10,spdhivya10,ragavithiyabme08,
govindvit07}@gmail.com

Abstract. Nowadays, Blind people use a stick as a tool for guiding their pathway when they walk or move away from their places. This paper introduces obstacles avoidance by using an 'electronic stick' with Infra Red (IR) sensors. The sensor has experimented successfully and which has detected the obstacles present along the pathway of the blind people in the range of approximately 100 – 550 cm. This designed system will have longer distance of obstacles detection unlike existing systems. The IR beams will be reflected back from the obstacles and detected by photo detector of the sensor, which has activates the buzzer and vibrator to produce an alarm and vibration respectively apart from these stick has mode selector switch to select either an alarm, vibration or both by this way the blind person will choose the safer path. This guiding stick has designed small, reliable and light weight to carry, in order to make blind people more comfortable and to improve their quality of life.

Keywords: Electronic stick, Infra Red sensor, Obstacles, Distance measurement, Blind people.

1 Introduction

Blindness is frequently used to describe severe visual impairments with or without residual vision. The extent of vision loss is described using an international scale [1]. According to the World Health Organization (WHO) in 2011fact sheet shows, there were about 284 million people are visually impaired worldwide, 39 million are blind and 245 have low vision, about 90% of the world's visually impaired live in developing countries, globally, uncorrected refractive errors are the main cause of visual impairment but in middle and low-income countries cataracts remain the leading cause of blindness, the number of people visually impaired from infectious diseases has greatly reduced in the last 20 years and 80% of all visual impairment can be avoided or cured [2]. The application of ultrasonic ranging scheme for producing electronic walking stick for blind [3]. The use of embedded system for developing an excellent stick using ultrasonic sensors for helping blind people instead of using a

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 222–227, 2013.

white stick. However, ultrasonic sensors are not a suitable technology because they are too large, have high weight, and consume a lot of power [4]. The application of Infra Red sensors for obstacle avoidance scheme, which can be employed for producing electronic walking stick [5]. Infrared sensors seems suitable for distance measurement as they have a small package, very little current consumption, a low cost, and a variety of output options [6].

2 Basic Concept

The basic concept of this work, to design an electronic guiding stick with obstacle avoidance system by using Infra Red sensor. The ultrasonic sensors are not a suitable due their too large size, high weight, and consumes much power, this motivated us work towards new sensor based stick with enhanced compliance. The designed electronic stick which has embedded with IR sensor to measure the distance in range of approximately 100 – 550 cm. This desined system has more advantages over existing systems like which detects obstacles on the longer distance precisely, uses lesser electronic components and easy to carry etc. The mounted IR sensor of the stick which senses the obstacles with longer ranges on the pathway of the blind person. The sensed IR beam from the obstacles has detected by the photo detector and the signal has fed to an alarm and vibrator, then immediately these two will alerts the blind person to guide the right pathway without involvement of any microcontroller unlike existing systems. The analog output voltage can be determined by using the following equation [6].

$$\text{Analog output voltage} = \text{IR Value} / 208.8555$$

GP2Y0A710K0F is a distance measuring sensor unit composed of Position Sensitive Detector (PSD), Infra Red Emitting Diode (IRED) and signal processing circuit. This device output voltage was corresponding to the detected distance likewise above sensor can also be used as a proximity sensor.The analog output voltage with distance to reflective object of the Infra Red sensor shown in Fig.1 [7].

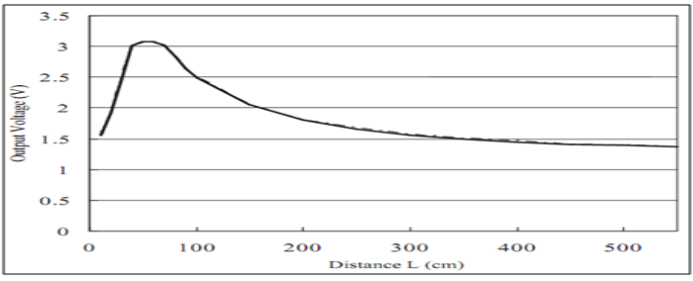


Fig. 1. Analog output voltage vs Distance to reflective object from GP2Y0A710K0F Infra Red sensor

The above graph shows, the sensor output voltage values are proportional to the distance. If the the obstacles present within the ranges of 100 – 550 cm, to be detected immediately and alerts the blind person by giving either vibration or alarm or both. The Earlier proposed system which has capable to detect only shorter distances about 80 cm and needs a microcontroller to process the IR signal. But this proposed system will have an advantage over earlier system by detecting the obstacle of longer distance and also consumes lesser time, lesser power and quicker responses to the obstacles. The designed stick was successfully tested with variety of common obstacles like stones, stones, tiles, clothes, tree, metal, wood, glass, concrete, plastic, human, vehicles etc. The mode selector switch of the stick enables easy use of the stick even person with vision impairment as well as hearing problem or both by choosing the modes like vibration with alarm or vibration without alarm.

3 Design and Operation

Design of this proposed system taken consideration about various parameters essential for guiding stick like, it should be small, light weight, reliable, easy to carry, consumes less power and performs precisely with suitable range of conditions. This system has utilized the hardware components like Infra Red (IR) sensor GP2Y0A710K0F (transmitter and receiver), Transistor, 555 Timer IC, buzzer, vibrator and mode selector switch (vibration and alarm or both).

Fig. 2. GP2Y0A710K0F IR sensor

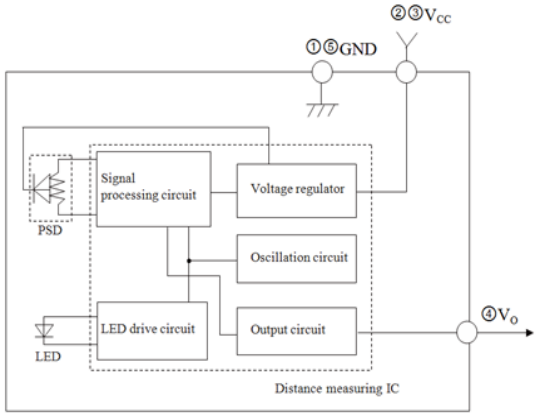


Fig. 3. Block diagram of GP2Y0A710K0F IR sensor

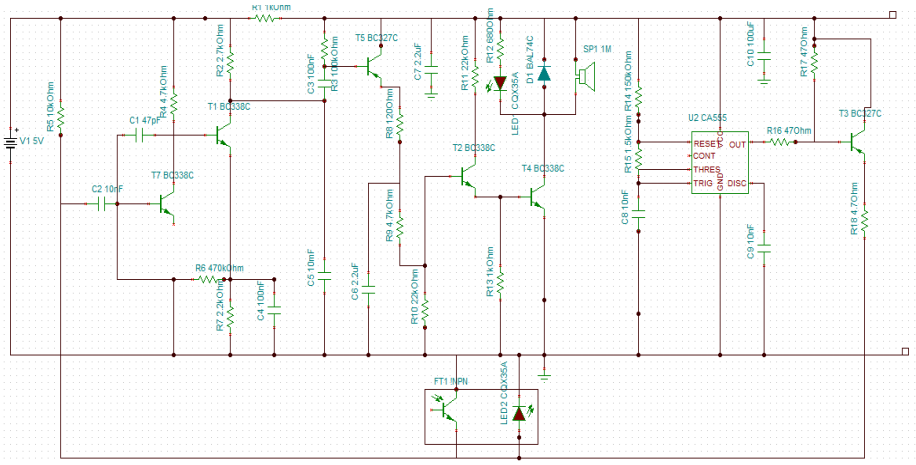


Fig. 4. Circuit design of the Guiding stick

3.1 Algorithm

i) IR signals are transmitted range of 100-500 cm to detect the obstacle(s)
ii) Obstacle(s) are detected by the IR sensor within the specified distance
iii) The received signal are transferred through the transistors for amplification of the received signals from IR sensor
iv) Amplified signals are fed to the controlling switch located in the handle of the stick to activate the buzzer and vibrator

3.2 IR Sensor

This system mainly utilized the IR sensor, model GP2Y0A710K0F and it's block diagram as shown in above Fig. 2 and Fig. 3 respectively, manufactured by SHARP CO [7].The sensor has small in size 58x17.6x22.5 mm, Long distance measuring ranges between 100 to 550 cm and with less power consumption (4.5 to 5.5 V, 30 mA). The Fig. 5 shows the graphical illustration of the designed guiding stick. The working of the proposed system as described as follows the IR sensor consists of both transmitter and receiver in combined manner.The power amplifier gives the optimum power supply to IR sensor containing Light Emitting Diode (LED) Transmitter. Fig. 6 shows the block diagram of the proposed smart guiding stick.

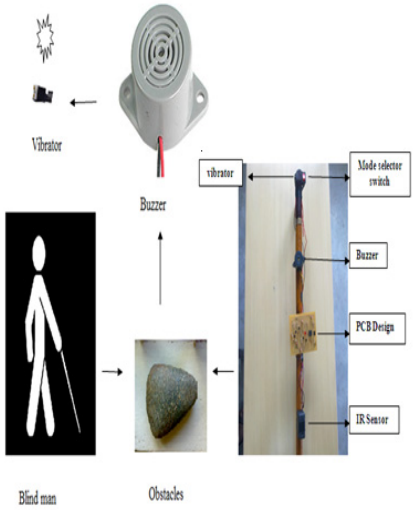


Fig. 5. Prototype of the Guiding stick

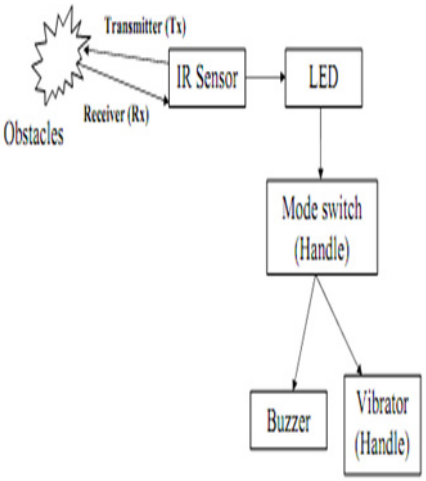


Fig. 6. Block diagram of Guiding Stick

The IR beams are transmitted from IR LED Transmitter within a range of 100 to 550 cm. The transmitted beam senses the obstacles and reflected back to photodiode receiver. The reflected beam from the obstacles as light energy and it will be converted into electrical energy. The output of the Receiver feeds the signals with minimal energy so we need the maximum current to operate the buzzer and vibrator for alert. The electrical signals are amplified by using Single Pin Double Throw (SPDT) switch. This switch which also controls the both buzzer and vibrator placed in the handle of the guiding stick.

4 Conclusions

We have designed the electronic smart guiding stick to detect the obstacles in the longer distance from 100 and 550 cm, by using GP2Y0A700K0F Infra Red sensor. The designed guiding stick which detects precisely the all type of obstacles like stones,tiles,clothes,tree,metal,wood,glass,concrete,plastic,human and vehicle etc. The detection sensitivity less in the case of longer distance obstacles when compared with nearer one. The placement of the sensor has an important role in the guiding stick which determines the sensitivity, angle and distance of the sensor over an obstacle. This proposed Infra Red sensor based smart guiding stick is more helpful to the person with any status of blind and it will improves the quality of life to those people.

References

1. International Council of Ophthalmology. International Standards: Visual Standards Aspects and Ranges of Vision Loss with Emphasis on Population Surveys; ICO: Sydney, Australia (2002)
2. World Health Organization (WHO) media centre fact sheet (2011), http://www.who.int/mediacentre/factsheets/fs282/en
3. Batarseh, D.T., Burcham, T.N., McFadyen, G.M.: An Ultrasonic ranging system for the blind. In: Proceedings of the Sixteenth Southern Biomedical Engineering Conference, vol. 4-6, pp. 411–413 (1997)
4. Palee, R., Innet, S., Chamnongthai, K., Eurviriyanukul, C.: Point-to point Distance Measurement Using Ultrasonic for Excellent Stick. In: International Technical Conference on Circuits/Systems, Computers and Communications Conference, Sendai, Japan (July 2004)
5. Yuzbasioglu, C., Barshan, B.: A new method for range estimation using simple infrared sensors. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1066–1071 (2005)
6. Innet, S., Ritnoom, N.: An Application of Infrared Sensors for Electronic White Stick. In: International Symposium on Intelligent Signal Processing and Communication Systems. IEEE Press, Thailand (2008)
7. SHARP GP2Y0A710K0 Optoelectronic Device, http://acroname.comGP2Y0A710K0-DATASHEET.PDF

PMH Stepper Motor Magnetic Circuit Design Using PDE Toolbox Simulations of Matlab for Different Topologies

E.V.C. Sekhara Rao[1] and P.V.N. Prasad[2]

[1] Department of EEE,
Chaitanya Bharathi Institute of Technology (CBIT), Hyderabad, India
chandrasekharev@yahoo.co.in
[2] Department of EE, UCE, Osmania University, Hyderabad, India
polaki@rediffmail.com

Abstract. This paper discusses about simulation of permanent magnet hybrid stepper motor (PMH) magnetic circuit using finite element model using Partial Differential Equation (PDE) toolbox of Matlab. It concludes mmf distribution during excitation due to permanent magnet and excitation coil for eight topologies. Topologies are designed with different airgaps and with different stator teeth.

Keywords: FEM, PDE toolbox, PMH stepper motor.

1 Introduction

Stepper motor is a special motor used for discrete torque without any interface. It is classified into three major types as permanent magnet stepper motor, Variable reluctance stepper motor and Permanent magnet hybrid (PMH) stepper motor. Among all PMH motor is widely used in many applications such as solar array tracking system in satellites, robotics and CNC machines due to its micro-stepping quality with high torque capacities. But its magnetic circuit analysis is complicated due to the presence of permanent magnet and double slotting. Its magnetic circuit is explained earlier with equivalent circuit model with linear assumptions. But getting operating point mmf with permanent magnet is complicated [1], finite element model is used for getting accurate operating point of permanent magnet. In this paper the FEM analysis of the PMH stepper motor is investigated using PDE toolbox of Matlab. PDE toolbox has both command mode and GUI mode. In this paper GUI mode is used to develop the required geometry of the PMH stepper motor then required boundary conditions are applied then solution is obtained through FEM relations for eight topologies.

2 FEM of PMH Motor

Tooth layer unit (TLU) [2] is a rectangle area that has a tooth pitch width and two parallel lines behind the teeth of stator and rotor. The area is shown in Fig.1. The

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 228–235, 2013.

factors of the nonlinear material and the non-uniform distribution of magnetic field in the teeth of stator and rotor are taken full consideration in this computation model. There are two basic assumed conditions in the computation model of TLU.

(i) The lines AB and CD of the TLU in Fig.1 are thought as iso-potential lines.
(ii) The magnetic edge effect of stator pole is ignored, which is assumed that the distribution of the magnetic field for every tooth pitch width is the same.

Fig 1 shows stator and rotor tooth geometry for one tooth pitch. In Fig 1, u_s, and u_r are scalar quantities of the iso-potential lines AB and CD. The magnetic potential difference F is given by

$$F = U_s - U_r \tag{1}$$

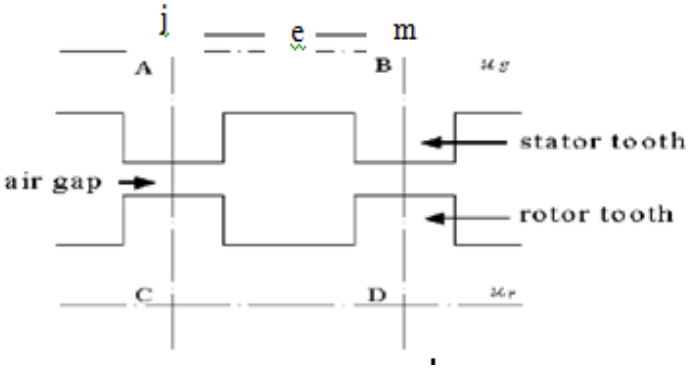


Fig. 1. The Tooth layer unit of PMH stepper motor

If Φ (α) is assumed as the flux in a tooth pitch width and per axial unit length of iron core, α is the relative position angle of stator and rotor. The specific magnetic conductance G of TLU is then given by

$$G = \frac{\Phi(a)}{F} \tag{2}$$

Apparently, G is related to the saturation extent of iron core and is changed with F and the relative position angle a. G can be got by the numerical computation on the magnetic field of TLU shown in Fig. 1. The lines AC and BD are the periodic boundary lines because the distribution of the magnetic field is considered as the same for every tooth pitch width. The magnetic field in TLU is irrational field and the magnetic equations [2] for the field are given in the rectangular coordinates by

$$\left.\begin{array}{c} \frac{\partial}{\partial x}\left(\mu\frac{\partial \boldsymbol{\varphi}}{\partial x}\right)+\frac{\partial}{\partial y}\left(\mu\frac{\partial \boldsymbol{\varphi}}{\partial y}\right) \\ \varphi_{CD}=0 \\ \varphi_{AB}=0 \\ \varphi\,(x,y)_{AC}=\varphi\,(x+\lambda,y)_{BD} \end{array}\right] \tag{3}$$

where φ is the scalar quantity, μ is the magnetic permeability and λ is the tooth pitch. For a certain position angle α and a magnetic potential difference F, the distribution of the magnetic field of TLU can be calculated by the 2D finite element analysis. The flux per axial length of TLU is given in (4).

$$\varphi(\alpha, F) = \sum B_e \, (\overset{\frown}{jm})_e \tag{4}$$

Here the nodes j and m are on the border AB as shown in Fig1. $(\overset{\frown}{jm})$ is the length of unit e from node j to m and B_e is the flux density. The specific magnetic conductance G will be used in the calculation of the whole nonlinear network equations of the motor.

3 Rationality Analysis Model

Though TLU, the key calculation model of the method is simple in drawing, its calculation model contains some approximate factors due to the basic assumed conditions. It is the main issue to be studied in this paper whether the assumed conditions are of engineering rationality or whether the assumed conditions are in agreement with the practical situations and whether the errors from the assumed conditions can be ignored from the point view of engineering. The two basic assumed conditions (i) and (ii) mentioned in 2 are analyzed below. A practical PMH stepper motor was chosen to analyze the rationality. It has 4 poles in the stator and 6 sections in the rotor with disc of NdFeB magnet axially magnetized. The main structure parameters of the motor are shown in Table 1.

As the motor is symmetrical, pair of poles of the motor has been chosen as the magnetic numerical calculation area, the magnetic field contains the nonlinear material, iron core, and the current area. The surface arc effect of the iron core and the pole edge effect of the stator are considered in the calculation model, by which the magnetic potential values and their differences on the lines behind the teeth could be found out and the engineering rationality of the assumed conditions could be verified. There is only axial current in Fig. 1 and the magnetic potential vector A_z suits Poisson's equation [2] as

$$\left.\begin{aligned} \frac{\partial A_z}{\partial x} &= \frac{b_i A_{zi} + b_j A_{zj} + b_m A_{zm}}{2\Delta} \\ \frac{\partial A_z}{\partial y} &= \frac{C_i A_{zi} + C_j A_{zj} + C_m A_{zm}}{2\Delta} \\ B_x &= \frac{\partial A_z}{\partial y} \\ B_y &= \frac{\partial A_z}{\partial x} \end{aligned}\right] \tag{5}$$

Δ is the area of the split triangle unit, b_i, b_j, b_m and c_i, c_j, c_m are the parameters of the split triangle unit.

Table 1. Data of PMH stepper motor

Stator poles	Tooth per stator pole	Outside diameter of stator (cm)	Inside diameter of stator (cm)	Outside diameter of stator shell (cm)
04	10	10.108	5.936	10.652
No.of tooth on rotor disk	No. of turns per phase	Section length of rotor	Outside diameter of rotor	Inside diameter of rotor
50	102	10.26	4.2	1.74

4 Simulation Model Using PDE Toolbox of Matlab

4.1 Geometry Design

Using graphical user interface (GUI) of PDE toolbox of Matlab, geometry of motor is designed for one pole pitch due to symmetry. Rotor is designed with 23 slots with equal spacing and then stator poles are designed with 05 slots for each pole for the above mentioned dimensions. Current coil is designed on stator pole. The obtained geometry is fig.2 for uniform airgap and fig.3 for non uniform airgap. Then stator outer shell is designed and the boundary conditions are provided with set formula as ((R1+R3) - (C3+SQ1+SQ2+SQ3+SQ7+SQ8+SQ9+SQ25+SQ26+SQ27+SQ28+SQ29 + SQ30)) + (C2-SQ4-SQ5-SQ6-SQ10-SQ11-SQ12-SQ13-SQ14-SQ15-SQ16-SQ17-SQ18-SQ19-SQ21-SQ20-SQ22-SQ23-SQ24-SQ37-SQ38-SQ39-SQ40-SQ41) + ((R2 + R4)-C4)+R5+SQ42+C6+C5+C1)*SQ43'), where R1, R2 are designed for stator pole 1 and R5 is designed as current coil on pole1. R4, R5 are designed for stator pole 2 and SQ42 is designed as current coil on pole2. SQ1, SQ2, SQ3, SQ7, SQ8, SQ9, SQ25, SQ26, SQ27 are designed as teeth on stator poles. SQ4-SQ5-SQ6-SQ10-SQ11-SQ12-SQ13-SQ14-SQ15-SQ16-SQ17-SQ18-SQ19-SQ21-SQ20-SQ22-SQ23-SQ24-SQ37-SQ38-SQ39-SQ40-SQ41are created equally to provide 23 rotor teeth. SQ43 is to provide the required boundary. R6 is the extra teeth on stator for smooth performance of the motor. From the mentioned set formula the geometry is obtained as shown in Fig.2 for uniform airgap with extra teeth on stator.

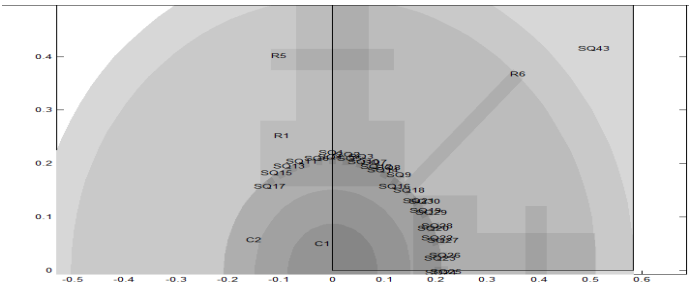


Fig. 2. Stator rotor geometry of PMH stepper motor for uniform narrow air-gap (0.137 mm) with extra teeth on stator

4.2 Motor Boundary Conditions

Motor boundary conditions are obtained using the following relations provided in magneto-statics application of PDE toolbox. The "statics" implies that the time rate of change is slow, so Maxwell's equations for steady cases,

$$\nabla \times H = J \tag{6}$$

$$\nabla \cdot B = 0 \tag{7}$$

$$B = \mu H \tag{8}$$

where 'B' is the magnetic flux density in tesla,' H' is the magnetic field intensity in AT/m, 'J' is the current density A/m^2, and μ is the material's magnetic permeability.

$$\text{Since } \nabla \cdot B = 0,$$

There exists a magnetic vector potential '**A**' such that

$$B = \nabla \times A \tag{9}$$

and

$$\nabla \times \left(\frac{1}{\mu} \nabla \times A\right) = J \tag{10}$$

The plane case assumes that the current flows are parallel to the *z*-axis, so only the *z* component of **A** is present,

$$\mathbf{A} = (0, 0, A),\ J = (0, 0, J) \tag{11}$$

And the preceding equation can be simplified to the scalar elliptic PDE

$$-\nabla \cdot \left(\frac{1}{\mu} \nabla A\right) = J \tag{12}$$

Where

$$J = J(x, y) \tag{13}$$

For the 2-D case, the magnetic flux density B is calculated as

$$B = \left(\frac{\partial A}{\partial y}, -\frac{\partial A}{\partial x}, 0\right) \tag{14}$$

The interface condition across sub-domain borders between regions of different material properties is that (H x n) be continuous. This implies the continuity of $\left(\frac{1}{\mu}\frac{\partial A}{\partial n}\right)$ does not require special treatment, since the variation formula of the PDE problem is used. In ferromagnetic materials, μ is usually dependent on the field strength $|B| = |\nabla A|$, so the nonlinear solver is needed. The Dirichlet boundary condition specifies the value of the magneto-static potential 'A' on the boundary. The Neumann condition specifies the value of the normal component of n $\left(\frac{1}{\mu}\nabla A\right)$ on the boundary. This is equivalent to specifying the tangential value of the magnetic field 'H' on the boundary. Visualization of the magneto-static potential 'A', the magnetic field 'H', and the magnetic flux density 'B' is available. 'B' and 'H' can be plotted as vector fields. The magnetic permeability μ is 1 in the air and in the coil. In the stator and the rotor, μ is defined by

$$\frac{\mu_{max}}{1+c[\nabla A]^2} + \mu_{min} \tag{15}$$

Where 'c' is coercive force which is the applied reverse magnetic field strength H required to force the net magnetic field back to zero after magnetization. The SI unit for H is A/m, and 1 A/m = 0.01257 oersteds. For getting magnetic potential at the border Dirichlet boundary condition is considered as (1, 0) and Neumann condition specified as (0, 0). Fig.3 shows geometry for one pole pitch for uniform airgap with extra teeth on stator after Dirichlet boundary condition is implied on the design.

4.3 PDE Model for Solution

Permeability for different core materials is calculated using (15). In this investigation two core materials are used for investigation as iron (99.8%) and iron (99.95%). These permeability values are given for stator and rotor core material portion in the geometry shown in Fig.2. Current density in these portions considered as zero.

Permanent magnet portion permeability is also calculated using (15) and current density is calculated [3]. Two types of permanent magnetic materials are used in this investigation like NdFeB, Se_2Co_{17}.Current density is considered zero and permeability as one for airgap and current coil to investigate permanent magnet operating point. Current density is considered for current coil to investigate total mmf under exciting condition. Considering all these conditions PDE solution is obtained as shown in Fig.4 which is magnetic potential diagrams for PMH stepper motor under permanent magnet excitation. Fig.5 is magnetic potential diagrams for PMH stepper motor under permanent magnet and current coil excitations. This magnetic potential is multiplied with axial length (10.26 cm) and air-gap lengths (0.137 mm, 0.93 mm) to get total mmf. PDE toolbox of Matlab for FEM analysis of PMH stepper motor is used to investigate mmf due to permanent magnet for different topologies which is difficult to investigate by mathematical model.

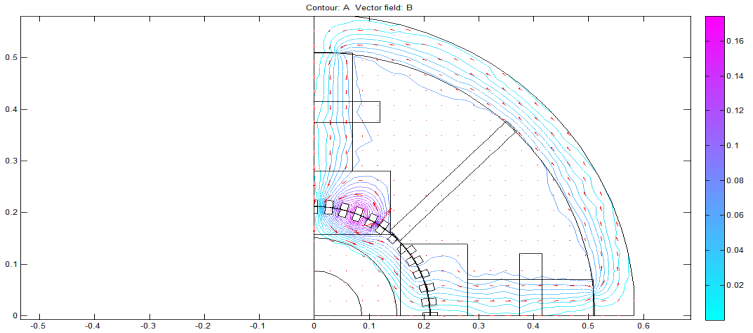


Fig. 3. PDE solution for Iron (99.8%) core material with NdFeB permanent magnet with extra teeth on stator without excitation

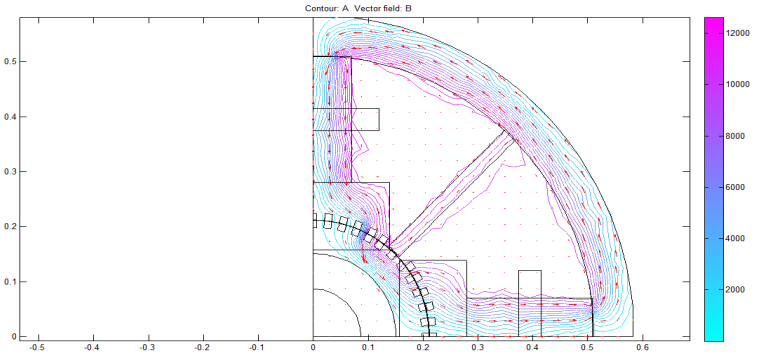


Fig. 4. PDE solution for Iron (99.8%) core material with NdFeB permanent magnet with extra teeth on stator with excitation for current density of 17056A/m^2

Different topologies considered are 1) Uniform air-gap (0.137 mm) with extra teeth on stator 2) Uniform air-gap (0.137 mm) without extra teeth on stator 3) Non-uniform air-gap (0.137 mm) with extra teeth on stator 4) Non-uniform air-gap (0.137 mm) without extra teeth on stator 5) Uniform air-gap (0.93 mm) with extra teeth on stator 6) Uniform air-gap (0.93 mm) without extra teeth on stator 7) Non-uniform air-gap (0.93 mm) with extra teeth on stator 8) Non-uniform air-gap (0.93 mm) without extra teeth on stator. Mmf is calculated for eight topologies due to permanent magnet excitation alone, due to stator coil excitation for two current densities (0.5A, 1A for 36SWG) and tabulated in Tab 2.

Table 2. MMF for different topologies

Topology	Current density (A/m^2)	MMF due to Permanent Magnet AT×10^{-4}	MMF due to excitation, AT	Topology	Current density (A/m^2)	MMF due to Permanent Magnet AT×10^{-4}	MMF due to excitation, AT
1	170648	4.578	23.6352	5	170648	0.970	8.1150
	341296	4.578	30.7940		341296	0.970	14.607
2	170648	6.159	92.3820	6	170648	0.811	4.6890
	341296	6.159	184.764		341296	0.811	9.7380
3	170648	4.578	46.4910	7	170648	0.649	2.4340
	341296	4.578	92.3820		341296	0.649	6.4920
4	170648	3.849	61.5880	8	170648	0.568	3.2460
	341296	3.849	107.779		341296	0.568	8.1150

5 Conclusions

FEM analysis of PMH stepper motor is done for eight topologies using PDE toolbox of Matlab. The results obtained for different topologies conclude that mmf

distribution is uniform and mmf interaction with stator and rotor teeth is effective with uniform airgap. When airgap is increased mmf is reduced. More mmf is interacted when an extra stator teeth is provided on the stator. This analysis is done without using any commercial FEM software first time for FEM analysis of Stepper motor. Time required for FEM analysis is comparatively low when compared with FEM commercial softwares.

References

1. Gieras, J.F., Wing, M.: Permanent Magnet Motor Technology Design and Applications. Marcel Deckker, Inc., New York, ISBN: 0-8247-0739-1
2. Dou, Y.: The Engineering Rationality of a Hybrid Stepping Motor Calculation Model. Proceedings of the CSEE 19(5), 35–38 (1999)
3. Pyrhönen, J., Jokinen, T., Hrabovcová, V.: Design of Rotating Electrical Machines. John Wiley & Sons, Ltd. (2008) ISBN: 978-0-470-69516-6

Harnessing Maximum Power from Solar PV Panel for Water Pumping Application

Mrityunjaya Kappali[1,*], R.Y. Uday Kumar[2], and V.R. Sheelavant[1]

[1] Faculty, Electrical & Electronics Engineering,
SDM College of Engg. & Tech., Dharawd -580002 – India
mrkappali@rediffmail.com, sheel125@gmail.com
[2] Faculty, Electrical & Electronics Engineering,
National Institute of Technology Karnataka, Surathkal, Mangalore -575025, India
udaykumarry@yahoo.com

Abstract. The problem of fast depleting fossil fuels is triggering exploration of alternate sources of electricity. Among such sources, Solar Photo Voltaic (PV) energy is gaining prominence due to its plentiful availability. Water pumping is an important application of solar PV power. However people are not opting for it in large numbers as the 'cost per watt' for solar pumping systems is high. The cost component can be reduced by harnessing more power per unit installed capacity of the solar panel. One method of realising this is by Maximum Power Point Tracking (MPPT) wherein a Power Electronic (PE) converter is used to match pump with the PV panel. Present paper deals with the MATLAB based simulation study of solar PV driven Permanent Magnet (PM) DC motor (brushed) Pump System. Two cases are considered: a) System without MPPT b) System with MPPT. It is shown that by varying duty cycle of the converter at different radiation levels, the pump speed and hence the useful mechanical output can be enhanced with MPPT. The simulation study reveals that the output power becomes the maximum when the motor voltage becomes the maximum. Hence it is proposed that the motor voltage can be used as a control parameter for varying the duty cycle of the converter in achieving maximum output.

Keywords: Solar PV, Water Pumping, Maximum Power Point Tracking.

1 Introduction

Water pumping is an important application of solar PV power. However people are not opting for it in large numbers as the 'cost per watt' for solar pumping systems is high. This is evident from the fact that around the world only 60000 solar pumping systems are installed [1]. This number is just 5000 in India [2]. Hence it is necessary to devise the means of reducing the cost component. This can be realised by harnessing more power per unit installed capacity of the solar panel.

* Corresponding author.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 236–241, 2013.

Solar PV panel exhibits typical Voltage vs. Current (V-I) (Fig.1) and Power vs. Voltage (P-V) (Fig.2) characteristics [3] as a function of solar radiation. At each radiation, represented proportionally by the panel short circuit current Iph, there exists a particular operating point at which the output power of the panel becomes the maximum. The process of controlling the operating point of solar PV panel so that it always corresponds to Maximum Power at the corresponding radiation is referred to as Maximum Power Point Tracking (MPPT). This needs matching between the Load and PV panel and can be accomplished by connecting a Power Electronic (PE) converter with variable duty cycle (D) as the interphase between the PV panel and the load [4]. Presently different control parameters are employed to vary D viz.:

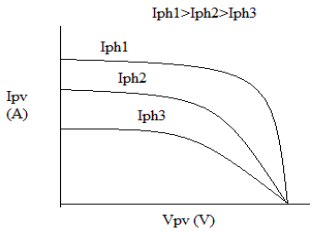


Fig. 1. 'V-I' plot for PV Panel

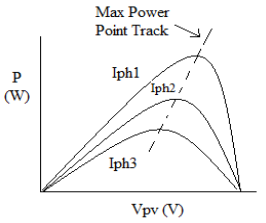


Fig. 2. 'V-P' plot for PV Panel

i) Power: In this method, the PV panel power (Ppv) is monitored continuously and D is varied till Ppv becomes maximum. But for computing power, it is necessary to perform multiplication as P = Vpv x Ipv. However, the need for multiplication operation makes the controller comparatively complicated.

ii) Speed: In this method speed of the pump (w) is monitored and D is varied till w becomes maximum which automatically means output power also becomes maximum. But speed being a mechanical parameter, monitoring it requires a transducer making the control system comparatively complicated.

Present paper deals with the simulation of solar PV driven water pump system. Two cases are considered: a) System without MPPT b) System with MPPT. Simulation is done using MATLAB - Simulink (version 7.5).

2 Solar Water Pumping

A typical Solar PV water pumping system with and without the provision of MPPT is shown in Fig.3 and Fig.4 respectively. Details of different components considered for simulation and testing purpose are given below.

PV Panel: The PV generator selected is parallel combination of two panels, each of rating: 74Wp (totalling 148Wp), Vm = 16.4 V, Im = 4.5A, Sun Technics Make. PV panel can be represented by a simple equivalent circuit [5] with a current source having a diode in parallel and resistance Rs in series (Fig.5). The current and power are given by the equations (1) & (2).

$$Ipv = Iph - Id \tag{1}$$

$$Ppv = Ipv \times Vpv \tag{2}$$

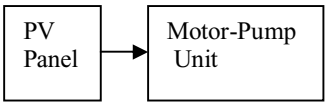


Fig. 3. Solar Pumping without MPPT

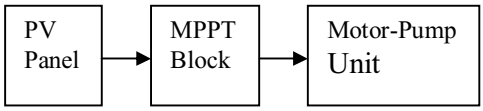


Fig. 4. Solar Pumping with MPPT

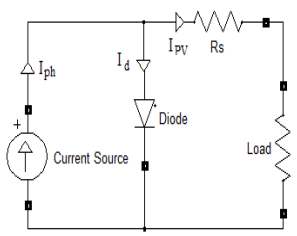


Fig. 5. PV Panel Equivalent Circuit

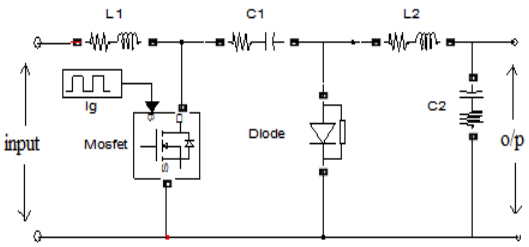


Fig. 6. Buck-Boost Converter

where Iph: short circuit current of the PV panel (A); Ipv: Load Current (A); Id: Diode Current (A); Vpv: Output Voltage (V); Ppv: Output Power (W). From experiment, Rs is found to be = 2.13 ohm. Iph, being directly proportional to the radiation, is used as a measure of radiation.

Motor-Pump Unit: It is a monoblock of PMDC Motor (brushed) and Centrifugal Pump. The specifications are: 12V, 70-100W, total head: 9m, Tata BP Solar make, Motor Inertia: 6.83 x 10^{-3} Kgm^2, Armature Resistance: 0.7 ohm, Armature Inductance: 0.12 x 10^{-3} H, C: voltage constant = 0.033 V/rad/sec. The DC motor is modelled representing the permanent magnet as separate excitation with constant voltage source.

Load Equation: The pump is used to lift water with delivery head, Hd = 5m and suction head, Hs = 1m. The characteristic 'Torque (T, Nm) vs. Speed (w, rad/sec)' is experimentally found yielding the load equation (4).

$$T = 4.8 \times 10^{-6} w^2 + 0.00019\, w + 0.092 \quad (4)$$

Converter (MPPT Block): Buck-boost converter (Fig.6) is employed as MPPT block [6]. The switching is done by MOSFET operating at 20 kHz. The converter is designed for a duty cycle range of 0.3 to 0.6.

3 Case I: System without MPPT

In this case, PV panel is connected directly to the motor-pump unit (Fig.7). Input required for this block is the information of radiation and the same is given in terms of Iph. The output variables are: current & voltage for panel as well as Motor (Ipv= Ia, Vpv= Va); Motor speed & torque (w, T). The simulation is run for different Iph and plots of speed and output power as a function of Iph are obtained (Fig.9 & 10).

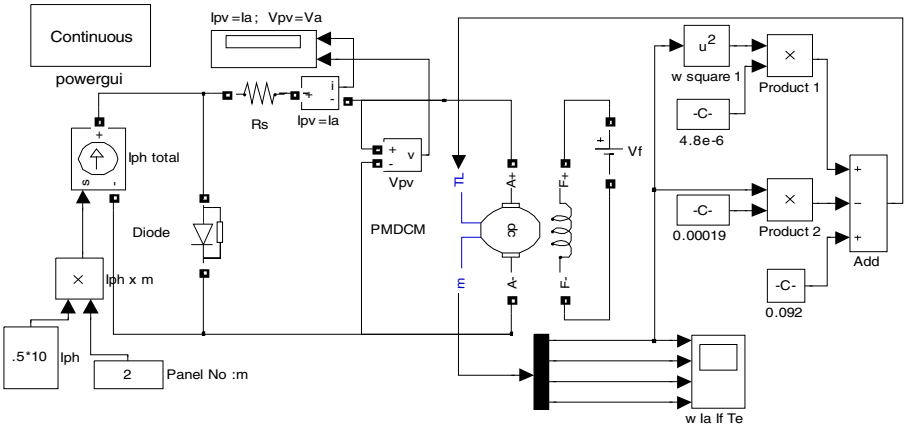


Fig. 7. Panel Pump Model without MPPT

4 Case II: System with MPPT

The complete model of PV panel, the converter and the motor-pump unit is realised as shown in Fig.8. Inputs required for this module are: a) Radiation level (given in terms of Iph). b) Duty cycle D for converter. The output variables are: panel current & voltage (Ipv, Vpv), Motor current & voltage (Ia, Va), Motor speed & torque (w, T). The simulation is run for different Iph and at each radiation for different duty cycles D. The plot of maximum power and speed as a function of Iph is obtained (Fig.9 & 10).

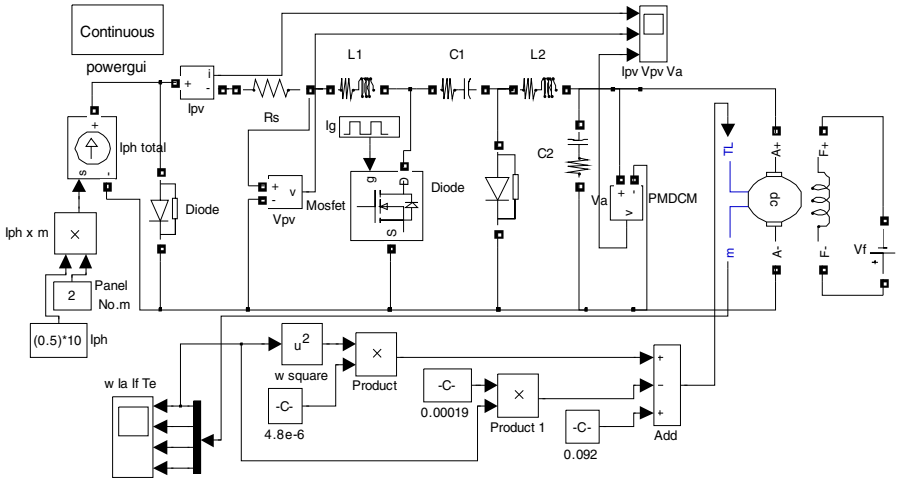


Fig. 8. PV Panel - Pump Model with MPPT

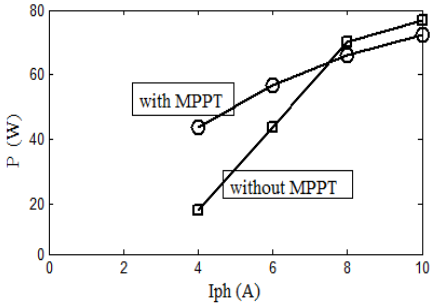


Fig. 9. 'Iph vs. P' comparison

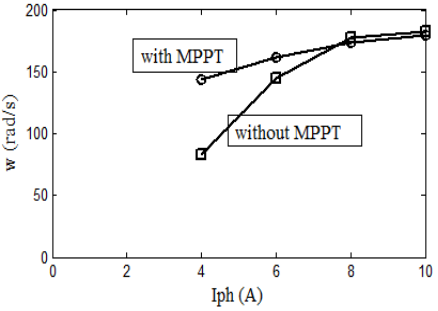


Fig. 10. 'Iph vs. speed' comparison

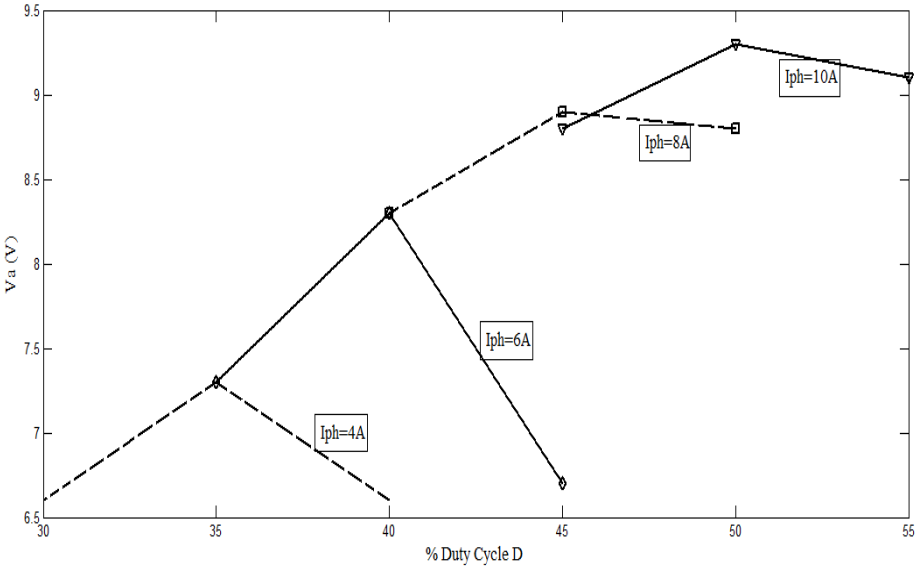


Fig. 11. '%D vs. Va' for pump with MPPT

5 Critical Observations and Discussions

Some important observations are made from the simulation study and results presented in the above sections 4 & 5:

a) On comparing the two schemes, it is found that in the scheme with MPPT there is an increase in the output power P and hence the speed w of the pump. This feature is quite considerable at lower radiation values and decreases as radiation increases (Fig.9 & Fig.10).

b) In the case of scheme with MPPT, it is observed that at a particular radiation, the output power becomes the maximum when the motor voltage Va becomes the maximum (Fig.11). This means for a particular radiation, there exists a specific motor

voltage (Vamax) at which output power becomes maximum and this occurs at a specific duty cycle. Hence the motor voltage Va can be used as a control parameter for varying the duty cycle of the converter in achieving MPPT. Va can be continuously monitored and D continuously varied so as to realise Maximum Va which automatically assures maximum output power at the corresponding radiation. Monitoring Va is simple as it is an electrical parameter and using it in a controller is simple as there is no multiplication operation involved. This method to realise MPPT is simple than the other methods which use either power or speed as control parameters.

6 Conclusions

Water pumping is an important application of solar PV power. Present paper has dealt with the simulation of solar PV driven water pump system. Two cases are considered: a) Motor pump unit connected to the panel directly and b) the Motor pump unit connected to the panel through a PE Converter for MPPT. It is shown that by varying duty cycles at different insolation levels, the pump speed and hence the useful mechanical output power can be enhanced with MPPT. An important observation made is that the output power becomes the maximum when the motor voltage becomes the maximum. Hence it is proposed that the motor voltage can be used as a control parameter for varying the duty cycle of the converter for the system with MPPT in achieving maximum output.

Acknowledgements. The authors thank their college authorities, NITK Surathkal & SDMCET Dharwad, for the support extended in carrying out this work.

References

1. Short, T.D., Mueller, M.A.: Solar powered water pumps: problems, pitfalls and potential. In: International Conference on PEMD, pp. 280–285 (2002)
2. Surendra, T.S., Subbaraman, S.V.V.: Solar PV water pumping comes of age in India. In: IEEE Photovoltaic Specialists Conference, pp. 1485–1488 (2002)
3. Vongmanee, V., et al.: Vector control of induction motor drive system supplied by pv arrays. In: IEEE Int. Conf. on Commn., Circuits and Systems, vol. 2, pp. 1753–1756 (2002)
4. Van Der Merwe, L., et al.: Maximum power point tracking-implementation strategies. In: IEEE Int. Symp. on Ind. Electronics, vol. 1, pp. 214–217 (1998)
5. Ghoneim, A.A.: Design Optimization of Photovoltaic Powered Water Pumping Systems. Elsevier Energy Conversion and Management 47, 1449–1463 (2006)
6. Rashid, M.H.: Power Electronics, 3rd edn. PHI Ltd.Pearson Publication (1988)

Implementation of Model Predictive Control for Closed Loop Control of Anesthesia

Deepak D. Ingole, D.N. Sonawane, Vihangkumar V. Naik,
Divyesh L. Ginoya, and Vedika Patki

Department of Instrumentation & Control Engineering,
College of Engineering, Pune- 411 005
Maharashtra, India
dingole21@yahoo.in, dns.instru@coep.ac.in,
{naikvihang,dlginoya007,patki.vedika}@gmail.com

Abstract. This paper focuses on the design and implementation of model–predictive controller (MPC) for the close loop control of anesthesia for a patient undergoing surgery. A single input (Propofol infusion rate) single output (bispectral index (BIS)) model of patient has been assumed which includes variable dead time caused by measurement of bispectral index. The main motivation for the use of MPC in this case relies on its ability in considering, in a straightforward way, control and state constraints that naturally arise in practical problems. The MPC can take into account constraints on drug delivery rates and state of the patient but requires solving an optimization problem at regular time intervals. We proposed a modified active set method algorithm using Cholesky factorization based linear solver to solve online convex quadratic programming (QP) problem, to reduce complexity of the algorithm, eventually to accelerate MPC implementation. Experimentation shows that excellent regulation of bispectral index (BIS) is achieved around set-point targets.

Keywords: Three Compartmental PK-PD model, Bispectral index (BIS), depth of anesthesia (DOA), Model–predictive control (MPC), Active Set Method (ASM).

1 Introduction

Closed loop control of anesthetic drug delivery has been a major issue of research for past many decades. Traditionally, an anesthesiologist used to decide the initial dose of the drug (bolus) and the variation of the drug during the period of surgery, based on different patient characteristics, such as age and weight. With the advent of several new EEG indexes, efforts have been directed towards designing of control strategy for automatic control of anesthesia drug delivery.

Anesthesia drug delivery involves the continuous administration of a combination of drugs with frequent adjustments to achieve loss of consciousness while maintaining normal vital signs during surgical stimulation. Anesthesia generally consists of

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 242–248, 2013.

several components: loss of consciousness and lack of awareness (hypnosis), lack of nociceptive reactivity (analgesia), hemodynamic stability and sometimes immobilizations (neuromuscular blockade) [1].

Model Predictive Control (MPC) is by far the most successful advanced control technology used in industry for solving process control problems with slow dynamics. Due to its profit potential, nowadays, its applicability is increased to real time and safety critical applications such as portfolio management, electrical drives, glucose control, anesthesia control, etc. [2]. MPC requires repeated unsupervised online solution of dynamical optimization problem. Consequently, numerical optimization technologies and efficient algorithms tailored for MPC optimization applications are very important to implement MPC successfully. One of the main reasons for its application in the anesthesia automation is that it can take account of physical and operational constraints, which are often associated with the patient safety. The proposed MPC uses the approximate linear pharmacokinetic-pharmacodynamics (PK-PD) model of the patient in the controller design, which regulates patient's BIS by manipulating the infusion rate of Propofol. An extensive simulation conducted to validate the robustness of the proposed MPC controller, by considering parameters variations in the selected model to account for change of patient's consciousness.

The remainder of this paper is organized as: Section 2 describes the three compartmental mathematical model of a patient. Section 3 explains a MPC problem formulation, controller design strategy and the modified active set algorithm for QP solver. Section 4 illustrates active set method and proposed method. In Section 5, clinical system and the results of the clinical trials are presented. Evaluation of the controller performance and discussion are given in Section 6.

2 Pharmacokinetic and Pharmacodynamic Model of Patient

The description of the BIS dynamics has been done mainly with physiological models. This model consists of a Pharmacokinetic part to describe the absorption, distribution, metabolism and excretion (ADME) of bioactive compounds in a higher organism and Pharmacodynamic part to describe the drug effect on the physiological variables of interest. The more used PK model is the Marsh model that has been widely studied and represents the patient as a set of compartments: central, fast and slow [3, 4, 5].The state space representation of the pharmacokinetic model is given as

$$\left.\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned}\right\}. \tag{1}$$

where, $x(t)$ is the vector of propofol concentrations in the compartments, u(t) is the infusion rate of propofol, y(t) is the effect site concentration of propofol, and A, B and C are the pharmacokinetic parameters. The pharmacokinetic parameters are given as

$$A = \begin{bmatrix} -\frac{k1+k2+k3+k4}{v1} & \frac{k2}{v1} & \frac{k3}{v1} & \frac{k4}{v1} \\ \frac{k2}{v2} & -\frac{k2}{v2} & 0 & 0 \\ \frac{k3}{v3} & 0 & -\frac{k3}{v3} & 0 \\ \frac{k4}{v4} & 0 & 0 & -\frac{k4}{v4} \end{bmatrix}, B = \begin{bmatrix} \frac{1}{v1} \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$C = [0 \quad 0 \quad 0 \quad 1], D = [0].$$

here, the numbers 1, 2, 3 and 4 correspond to central compartment, shallow peripheral compartment, deep peripheral compartment and the effect site compartment, respectively.

The Pharmacodynamic model is the static relation between the effect site concentration of propofol and the BIS value given by the following Hill's sigmoid Emax model [6]

$$E(t) = E0 - Emax\frac{y(t)^{\gamma}}{y(t)^{\gamma}+C50^{\gamma}}. \tag{2}$$

where, E0 is the BIS value before starting propofol infusion, Emax is the change of the BIS corresponding to infinite Propofol concentration, C50 is the effect site concentration for E = Emax/2, and γ is the Hill's coefficient. In the following, we assume Emax = E0.

3 Model Predictive Control (MPC) Problem Formulation

Linear Model Predictive Control is an optimization-based approach has been successfully applied to a wide variety of control problems. The basic idea is to select a sequence of $\mathbf{N_u}$ future control moves to minimize an objective function **J** (usually sum of square of predicted errors) over a prediction horizon of $\mathbf{N_p}$ sample times. Although a horizon of $\mathbf{N_u}$ control moves are calculated, only the first move is implemented, a correction for plant-model mismatch is made, and the optimization is performed again, this concept of MPC is called as receding horizon control-MPC strategy [7].The constrained MPC problem is to minimize the cost function

$$\min_{y,u}(J) = \frac{1}{2}\sum_{i=1}^{N_p}\|\hat{y}(t+i|t) - y_{ref}(t+i|t)\|^2 Q_y + \frac{1}{2}\sum_{i=0}^{N_u-1}\|\Delta u(ti+|t)\|^2 R_y \tag{3}$$

Subject to input and state constraints

$$u_{min} \le u(t) \le u_{max}, \Delta u_{min} \le \Delta u(t) \le \Delta u_{max}, y_{min} \le y(t) \le y_{max} \tag{4}$$

here, y_{ref} is the set-point; N_p and N_u are the prediction and control horizons while Q_y and R_y are the positive definite weight matrices. To get optimal solution, the MPC problem (3) is formulated as a standard quadratic programming (QP) problem [8]

$$\min_{U \in \mathbb{R}^{n_v}} J = \left\{ \frac{1}{2} U^T H U + F^T U \right\} \quad . \tag{5}$$

Subject to

$$AU \leq b. \tag{6}$$

where, $H = \Gamma^T Q_y \Gamma + R_y$ is an $(lN_u \times lN_u)$ positive definite Hessian matrix, and $F = \Gamma^T Q_y - \Gamma^T Q_y + Mu_{-1}u_{-1}$ is a $(lN_u \times 1)$ vector.

4 The Active Set Method

Interior Point Methods [8] and Active Set Methods [9] are the most commonly used approaches to solve standard QP problems .Compared to active set methods; interior point methods are often preferred for large problems since their computational performance is less sensitive to the number of constraints. On the other hand active set method gain more advantages for problems with small number of variables and fewer states. According to the primal active set framework introduced by (Fletcher, 1997) the QP can be solved iteratively as follows:

Step 1: Initialization and input
Compute a feasible initial point x^0, set $\mathcal{W}^0$ the initial working set of active constraints at x^0 Set $k = 0,1,2 \ldots Kmax;$

Step 2: Solving linear system
Solve

$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \lambda \end{bmatrix} = \begin{bmatrix} -F - Hx_k \\ 0 \end{bmatrix} = \begin{bmatrix} d \\ b \end{bmatrix}. \tag{7}$$

for $(\Delta x, \Delta\lambda)$, given $\mathcal{W}_k$, compute $\hat{x}_{k+1}$, $d = -F - Hx_k$, $b = A^T x$

Step 3: Determine the most negative Lagrange multiplier and remove the corresponding active index (constraint) from $\mathcal{P}_k$

$$\mathcal{P}_k = \left\{ i \notin \mathcal{W}_k : J_i \Delta_{xk} > 0 \ \frac{b_i - J_i x_k}{J_i \Delta x_k} < 1 \right\} \tag{8}$$

$$x_{k+1} = x_k + \Delta x \text{ and } \mathcal{W}_{k+1} = \mathcal{W}_k .$$

Step 4: Calculation of step length

Compute $\alpha_k = \min_{i \in \mathcal{P}_k} \left\{ \frac{b_i - J_i x_k}{J_i \Delta x_k} \right\}$ and $x_{k+1} = x_k + \alpha_k \Delta x$

It is clear that step 2 of the active set algorithm is where most of the computation occurs and hence optimizing the computational efficiency of step 2 should be the most fruitful.

4.1 Proposed Cholesky Factorization Based Linear Solver

In order to solve a KKT system as fast and reliable as possible, Cholesky and QR factorizations can be used. But according to (5, 6) the KKT matrix is indefinite, and therefore it is not possible to solve it using either of the two factorizations. In this paper, we present Cholesky factorization method for solving the KKT system by using range-space procedure which can be divided into subproblems. The KKT system (7) is solved based on range space procedure, corresponding to the convex equality constrained QP. The Hessian matrix in (7) $H \epsilon \mathbb{R}^{n\,X\,n}$ must be Symmetric Positive Semidefinite, because the procedure uses the inverted Hessian matrix H^{-1}. Table 1 shows steps to calculate solution of linear system in (7)

Table 1. Steps to compute solution of linear system using Cholesky factorization

Step 1: Compute Cholesky factorization $H = LL^T$	**Step 6 :** Compute Cholesky factorization $\bar{H} = MM^T$
Step 2: Compute K by forward substitution, solving $LK = A^T$	**Step 7 :** Compute q by forward substitution, solving $Mq = z$
Step 3: Compute G by backward substitution $LG = d$	**Step 8 :** Compute λ by backward substitution, solving $M^T\lambda = q$
Step 4: Compute $\bar{H} = K^TK$	**Step 9 :** Compute $r = K\lambda - G$
Step 5: Compute $z = K^TG + b$	**Step 10:** Compute x by backward substitution solving $L^Tx = r$

5 Experimental Results

The rigorous experimentation was carried out for 8 patients having age 25±2 years and weight 63±2 kg, who underwent minor surgery. Their individual parameters were calculated and inputted to the model. The controller has to maintain BIS between 40 and 60 during the surgery. Firstly, it is assumed that the patient is in a fully awake state (BIS≈ 100) and then the controller is tuned as the set-point is changed from 100 to 50.The MPC algorithm is developed in MATLAB as ***m-file*** to evaluate the influence of several parameters (γ and k4) on the PK/PD model with sample time 1s. In Fig. 1 the response of proposed active set method and MATLAB's QP solver **"quadprog"** can be compared, as it is shown that the response of proposed method is faster with lower settling time. In Fig. 2 the response of BIS with time is shown. The prediction horizon is 10 and control horizon is 3. Fig. 3 illustrates the performance of the controllers in presence of the noise with 0.1% magnitude and disturbance (t=300 to 400 sec). In both the cases controller is robust and successfully works in presence of noise.

The coefficient matrices of continuous-time state space PK-PD model are given as [4]

$$A = \begin{bmatrix} -0.3944 & 0.1362 & 0.0935 & 0.0215 \\ 0.0304 & -0.0304 & 0 & 0 \\ 0.0033 & 0 & -0.0033 & 0 \\ 0.2154 & 0 & 0 & -0.2154 \end{bmatrix}, \quad B = \begin{bmatrix} 0.1077 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$
$$C = [0 \quad 0 \quad 0 \quad 1], \qquad D = [0] \qquad (9)$$

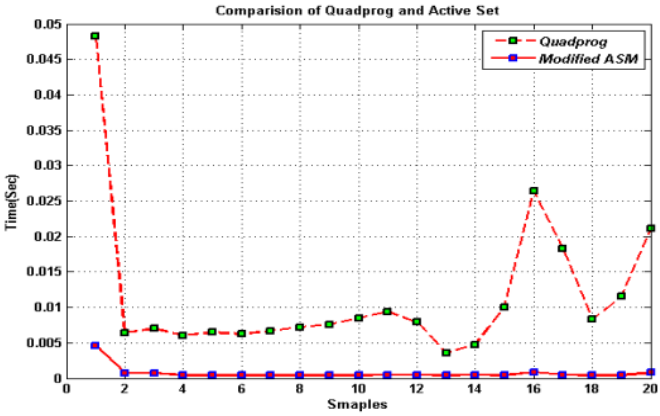


Fig. 1. Comparison of Quadprog and Modified Active Set method for MPC

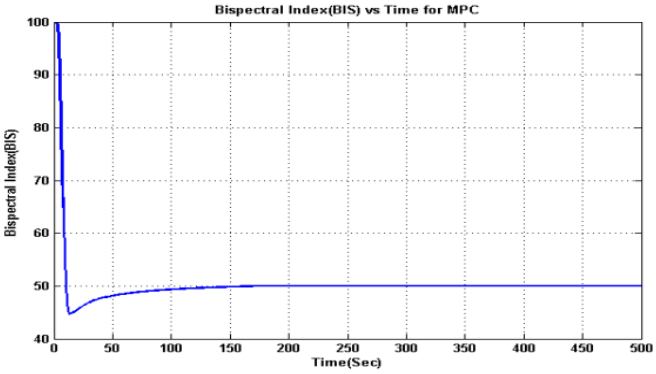


Fig. 2. BIS vs Time for MPC

6 Conclusion

In this paper, a model predictive control, for regulation of anesthesia using BIS as the controlled variable have been developed and evaluated thoroughly. The performance of proposed active set method algorithm with MATLAB's QP solver (quadprog) is compared and tested. By using such controllers the anesthesiologist will be more

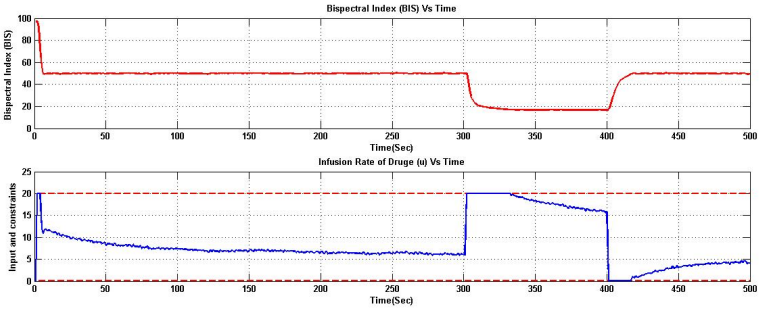


Fig. 3. BIS vs Time for MPC in case of Noise and Disturbances

confident about the safety of patient. Since, during surgery the presence of noise and disturbances are inevitable, using MPC for controlling the DOA which is a robust strategy, is a good choice for safety critical applications.

References

1. Dumont, G.A., Martinez, A., Ansermino, J.M.: Robust control of depth of anesthesia. International Journal of Adaptive Control Signal Process., 435–454 (2009)
2. Garcia, C.E., Prett, D.M., Morari, M.: Model Predictive Control: Theory and Practice - A Survey. Automatica 25(3), 335–348 (1989)
3. Schüttler, J., Ihmsen, H.: Population pharmacokinetics of propofol. Anesthesiology 92(3), 727–738 (2000)
4. Furutani, E., Tsuruoka, K., Kusudo, S., Shirakami, G., Fukuda, K.: A Hypnosis and Analgesia Control System Using a Model Predictive Controller in Total Intravenous Anesthesia During Day-case Surgery. In: SICE Annual Conference 2010, The Grand Hotel, Taipei, Taiwan, August 18-21 (2010)
5. Sawaguchi, Y., Furutani, E., Shirakami, G., Araki, M., Fukuda, K.: A Model-Predictive Hypnosis Control System Under Total Intravenous Anesthesia. IEEE Transactions on Biomedical Engineering 55(3), 874–887 (2008)
6. Kenny, G.N.C., Mantzaridis, H.: Closed-loop control of propofol anaesthesia. British Journal of Anaesthesia 83(2), 223–228 (1999)
7. Wright, S.J.: Applying new optimizaiton algorithms to model predictive control. In: Chemical Process Control-V, CACHE, AIChE Symposium Series, vol. 93 (316), pp. 147–155 (1997)
8. He, M., Ling, K.V.: Model Predictive Control On A Chip. In: 2005 International Conference on Control and Automation (ICCA 2005), Budapest, Hungary, June 27-29 (2005)
9. Fletcher, R.R.: Practical Methods of Optimization, 2nd edn. Wiley (1987)

A Technique to Reduce Crosstalk Delay on Data-Bus in DSM Technology

Anchula Sathish[1,*], M. Madhavi Latha[2], and K. Lal Kishore[2]

[1] RGMCET, JNT University Anantapur,
Nandyal, Andhra Pradesh-518501, India
[2] Jawaharlal Nehru Technical University, Hyderabad, Andhra Pradesh 500085, India
{sathish_anchula,mlmakkena}@yahoo.com

Abstract. Propagation delay through interconnects and data buses limits the performance of a highly integrated circuits. This delay is due to crosstalk. This crosstalk is due to coupling transitions occurring on the data bus and interconnects when data is transmitted. Hence reducing the crosstalk is one of the main challenges in DSM and VDSM technology. Reduction of crosstalk increases the reliability and performance of the system. One of the favorable techniques to reduce the crosstalk delay is to reduce the coupling transitions. Data Bus encoding technique is the promising method to reduce the crosstalk delay by encoding the data on the data bus. Hence an efficient Crosstalk delay reduction data bus encoding technique is proposed which can reduce the total crosstalk delay by around 43%, to 50% compared to unencoded data and 23% to 35% compared to others techniques for 8-bit,16-bit, 32-bit and 64-bit data buses.

Keywords: Crosstalk delay, Coupling transitions, interconnects, reliability, data bus, Bus encoding.

1 Introduction

Crosstalk effects are dominating the current technology which causes signal delays and signal noise on interconnects and on data buses because the wires are packed ever closer to each other and the inter wire coupling capacitances dominates the portion of total capacitance as the technology moves into towards nano regime. The crosstalk has become a major concern because of continuing decrease in dimensions of wires sizes and corresponding increase in length. The propagation delay of interconnects and data bus caused by self capacitance, coupling capacitance and resistance is becoming prominent than gate delay [1]. The performance and reliability of the chip depends more on interconnects and data buses than gate delays. The characteristics of data buses and long interconnects such as wire spacing [2], coupling length, wire length, wire width, wire material, driver strength, and signal transition time, etc. influences the coupling effect. Crosstalk delay faults can be reduced by reducing the coupling transitions [3].The total energy consumption and delay which determines maximum speed of the bus depends on crosstalk as given in [4-5].

* Corresponding author.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 249–254, 2013.

Crosstalk delay results due to charging and discharging of a coupling capacitance of data bus. Reducing the transition activity on the on-chip data buses is the one of the attractive way of reducing the crosstalk which intern reduces crosstalk delay. On of the simplest method to eliminate crosstalk is by using passive shielding [6]. However it requires twice the number of wires which results to a 100% area overhead. However instead of inserting shield wire between every pair of wires the spacing between the wires can be increased which can decrease the coupling transitions. Even then the area overhead is 100% [7]. Crosstalk preventing coding (CPC) can able to eliminate some of the crosstalk classes. For 32-bit bus it requires 46-bit bus [8] hence large area overhead. A bus encoding technique is proposed which can avoid forbidden patterns i.e the patterns of 010 and 101. Avoiding forbidden patterns can eliminate class 6 crosstalk. But this technique requires 52-bus lines for 32-bit data bus [5]. Selective shielding technique is proposed [9] which eliminates opposite transitions on adjacent bus lines. It also requires 48-bit bus for 32-bit data bus. A bus encoding technique is proposed for SoC buses to eliminate opposite transitions and to minimize power. It requires 55-bit bus for 32-bit data bus [10]. Dual-rail coding technique is proposed in [11] in which sends both the original as well as duplicated data bits are placed adjacently. This technique also requires 100% area overhead. In recent days it is discovered that encoding the data bus can eliminate or reduces the some classes of crosstalk with much low area overhead compare to the shielding techniques and others. Transition activity on the data bus can be reducing by employing bus encoding techniques. Several bus encoding techniques have been proposed to reduce power consumption during bus transmission in literature. These techniques mainly relay on reducing the data bus activity. Reducing power consuming transition by encoding the data on the data buses leads to reducing the bus activity hence overall power consumption is reduced [12-17]. However these techniques are not evaluated their performance for the crosstalk delay. The proposed technique not only reduces the power consuming transitions [18] but also delay due to crosstalk. It requires only 2 extra bus bit for any data bus width.

2 Crosstalk Delay

One of the important effects of coupling capacitances is that it may induce unwanted voltage spikes in neighboring bus wire. This is knows as Crosstalk. A wire on which a switching transition occurs is termed an *aggressor* and the wire on which it produces a noise spike is termed as a *victim*. Typically, an aggressor wire is physically adjacent to a victim wire and they may be modeled as being connected by a distributed coupling capacitance. Hence, a switching event in the aggressor wire while the victim wire is silent can result in the injection of current into the victim wire, causing an electrical spike. However, a large coupling capacitance relative to the self-capacitance of the wire can cause a large inadvertent spike on the victim that may cause a spurious switching event, potentially leading to errors on victim wire and increased delay due to charging and discharging. The analytical delay on data buses in deep sub-micron has been proposed by Sotiriadis et al.[19]. The crosstalk can be classified into six types 1C, 2C, 3C, 4C, 5C and 6C according to the C_C of two wires in 3-bit interconnect bus models. Table.1 defines the crosstalk classes and delay normalized to C_sR_w [20].

Table 1. Crosstalk Classes and their delays

Crosstalk Classes	Transition Patterns	Relative Delay on the middle wire
1	---,--↑,↑--,--↓,↓--,↑-↑,↑-↓,↓-↑,↓-↑, ↓-↓	0
2	↑↑↑,↓↓↓	C_sR_w
3	-↑↑,↑↑-,-↓↓,↓↓-	C_sR_w (1+λ)
4	-↑-,-↓-,↓↓↑,↑↓↓,↑↑↓,↓↑↑	C_sR_w (1+2λ)
5	-↑↓,-↓↑,↓↑-,↑↓-	C_sR_w (1+3λ)
6	↓↑↓,↑↓↑	C_sR_w (1+4λ)

3 Crosstalk Delay Reduction Coding

The proposed crosstalk delay reduction technique called Bus regrouping with hamming distance (BRG-HD) is based on reduction of the number of coupling transitions occurring on the data bus when a new data is to be transmitted. By implementing the following algorithm crosstalk delay can be reduced. The proposed algorithm for 16-bit Data bus (Db) is given as follows:

Let 16-bit data bus is represented by Db [0:15]

- Calculate the number of CT (coupling transitions) of the present bus data with the previous bus data.
- Calculate 6C, 5C, 3C, 2C and 1C type crosstalk transitions and its delay between present bus data with the previous bus data.
- If CT >= (n/2) then
- Consider the grouping of the present bus data. Now arrange the data on the data bus as
  Odd Group: $Db_0\ Db_2\ Db_4\ Db_6\ Db_8\ Db_{10}\ Db_{12}\ Db_{14}$
  Even Group: $Db_1\ Db_3\ Db_5\ Db_7\ Db_9\ Db_{11}\ Db_{13}\ Db_{15}$
- The Hamming Distance between odd group bits of present data and odd group bits of previous data is calculated. This is represented as OPHD = Odd position bits Hamming Distance
- The Hamming Distance between even group bits of present data and even group bits of previous data is calculated. This is represented as EPHD = Even position bits Hamming Distance
  Transmit the data by following the below conditions:
  If OPHD > EPHD, flip the data in odd bit positions and append bit '1' on the left and bit '0' on the right side of the encoded data.
  If EPHD > OPHD, flip the data in even bit positions and append bit '0' on the left and bit '1' on the right side of the encoded data.
  If OPHD = EPHD, flip the entire data and append bit '1' on the left and bit '1' on the right side of the encoded data.
- If CT<n/2 is true then transmits the data as it is, append bit '0' on the left and bit '0' on the right side of the encoded data.
- Calculate total normalized crosstalk delay.

Calculate the efficiency of reduction of normalized crosstalk delay for 8-bit,16-bit, 32-bit and 64-bit data busses for all crosstalk classes.

Table 2. Crosstalk delay efficiency for 16-bit data bus

Method	Class 6	Class 5	Class 4	Class 3	Class 2	Class 1
BINV	43.809666	8.2297806	34.299941	55.018762	73.57513	-28.85911
DYNAMIC	89.23031	24.717745	-22.634739	-68.925891	-22.797928	-8.0378076
BRG	92.4821	58.69488	-32.293497	-176.87617	-241.45078	-26.282668
NOVEL	90.214797	55.733864	13.246924	-228.89306	-399.87047	-31.734163
BRG-HD	93.153341	71.40524	19.478617	-220.23921	-314.37824	-42.306458
SHINV	60.396778	16.814599	14.960457	-32.973734	-52.720207	-41.082402
EESCT	78.475537	29.056309	-3.04628	-70.473734	31.476684	-39.917574

Table 3. Total Delay efficiency of proposed technique for different data bus widths

Method	8-bit	16-bit	32-bit	64-bit
BINV	41.9207014	27.07901	22.338939	12.444777
DYNAMIC	39.7110136	23.25034	26.158091	22.6029
BRG	44.720216	29.597859	32.619086	25.324946
NOVEL	42.3417405	37.609417	43.245249	41.936444
BRG-HD	43.7803451	47.18538	50.483031	46.771505
SHINV	39.8388607	24.834163	15.473742	13.091752
EESCT	34.2245716	27.804206	32.181245	22.790648

4 Performance Evaluation

The performance of proposed data bus encoding technique for crosstalk delay reduction is evaluated by using a VHDL code. The simulations are performed on 64-bit, 32-bit, 16-bit and 8-bit data bus by applying 10000 data vectors. Crosstalk delay due to 6C type, 5C type, 4C type, 3C type, 2C type, and 1C type crosstalk are shown in Table.2. The proposed technique performance is compared with the other existing six techniques [12-17]. Table.2 shows that proposed technique (BRG-HD) can able to reduce the crosstalk delay by around 93%, 71% and 19% for 6C, 5C and 4C respectively. It increases delay around 220%, 314% and 42% for 3C, 2C and 1C respectively. Hence it can be concluded that the proposed technique can effectively converts the critical delay crosstalk to non critical delay crosstalk. Table.3 shows the performance of the different techniques by varying the data bus widths. It shows that the performance decreases with the increase of bus width for all techniques except for Novel and proposed technique. In overall the proposed technique can able to reduce the total crosstalk delay by around 43% to 50% compared to unencoded data and 23% to 35% compared to others techniques for 8-bit, 16-bit, 32-bit and 64-bit data buses from Table.3. Hence this technique is the effective technique to reduce the total crosstalk delay in DSM and VDSM technologies.

5 Conclusion

The Proposed crosstalk delay reduction technique for data bus is based on reducing coupling transitions. The main aim of the proposed technique is to reduce total delay due to crosstalk. The reduction in worst case crosstalk reduces the overall energy consumption and delay on data bus to transfer data. The main advantage of the proposed technique is it requires only two extra bus line for any data bus width. Hence area overhead due to more bus lines is drastically reduced. The simulation results show that the proposed technique reduces the crosstalk delay effectively. The reduction in the crosstalk delay reduces the delay faults hence reliability of the system increases. Hence the proposed technique effectively overcomes crosstalk effect which is one of the challenges of DSM and VDSM technology.

References

1. Caignet, F., et al.: The challenge of signal integrity in deep-submicron meter CMOS technology. IEEE 89(4), 556–573 (2001)
2. Petrov, P., Orailoglu, A.: Low-Power instruction Bus Encoding for Embedded Processors. IEEE Transaction on Very Large Scale Integration Systems 12(8), 812–826 (2004)
3. Khan, Z., Arslan, T., Erdogan, A.T.: Low power system on chip bus encoding scheme with crosstalk noise reduction capability. IEEE Proceedings-Computers and Digital Techniques 153, 101–108 (2006)
4. Sotiriadis, P.P., Chandrakasan, A.: Low power bus coding techniques considering inter-wire capacitances. In: Custom Integrated Circuits Conference (2000)
5. Duan, Tirumala, A., Khatri, S.P.: Analysis and Avoidance of Cross-talk in On-Chip Buses. In: Hot Interconnects, pp. 133–138 (2001)
6. Ma, J., He, L.: Formulae and application of interconnect estimation considering shield insertion and net ordering. In: Proc. ICCAD, pp. 327–332 (2001)
7. Arunachalam, R., et al.: Optimal shielding/sapcing metrics for low power design. In: Proceeding of International Symposium VLSI, pp. 167–172 (2003)
8. Victor, B., Keutzer, K.: Bus encoding to prevent crosstalk delay. In: Proceedings of International Conference Computer-Aided Design, pp. 57–63 (2001)
9. Mutyam, M.: Selective shielding: A crosstalk-free bus encoding technique. In: Proceedings of the International Conference Computer-Aided Design, pp. 618–621 (2007)
10. Khan, Z., et al.: A Dual low power and crosstalk immune encoding scheme for system-on-chip buses. In: Proceedings of the International Workshop, Power and Timing Modelling Optimization and Simulation, pp. 585–592 (2004)
11. Rossi, D., et al.: Coding scheme for low energy fault tolerant bus. In: Proceedings of International Workshop On-line Testing, pp. 8–12 (2002)
12. Stan, M.R., Burleson, W.P.: Bus-Invert coding for low-power I/O. IEEE Transaction on Very Large Scale Integration Systems 3, 49–58 (1995)
13. Madhu, M., Srinivas Murty, V., Kamakoti, V.: Dynamic coding Technique for Low-Power data bus. In: Proc. IEEE Computer Society Annual Symposium on VLSI (2003)
14. Samala, N.K., Radhakrishnan, D., Izadi, B.: A Novel deep submicron Bus Coding for Low Energy. In: Proceedings of the International Conference on Embedded Systems and Applications, pp. 25–30 (June 2004)
15. Natesan, J., Radhakrishnan, D.: Shift Invert coding (SINV) for low power VLSI. In: IEEE Conference on Digital System Design, pp. 190–194 (2004)

16. Ravindra, J.V.R., Chittarvu, N., Srinivas, M.B.: Energy Efficient Spatial Coding Technique for Low Power VLSI Applications. In: Proceedings of the 6th International Workshop on System-on-Chip for Real-Time Applications, pp. 201–204 (December 2006)
17. Sathish, A., Subba Rao, T.: Bus Regrouping method to optimize Power in DSM Technology. In: Proc. IEEE-International Conference on Signal Processing, Communications and Networking, pp. 432–436 (January 2008)
18. Sathish, A., Madhavi Latha, M., Lal Kishore, K.: A technique to Reduce Transition Energy for Data-Bus in DSM Technology. IJCSI International J. of Computer Science 8(4(2)) (2011)
19. Sotiriadis, P.P., Chandrakasan, A.: A bus energy model for deep submicron technology. IEEE Transaction on Very Large Scale Integration Systems 10(3), 341–350 (2002)
20. Macchiarulo, L., Macii, E., Poncino, M.: Wire placement for crosstalk energy minimization in address buses. In: Proceedings Design, Automation and Test in Europe Conference and Exhibition, March 4-8, pp. 158–162 (2002)

Design and Implementation of Interior-Point Method Based Linear Model Predictive Controller

Vihangkumar V. Naik, D.N. Sonawane, Deepak D. Ingole, Divyesh L. Ginoya, and Neha S. Girme

Department of Instrumentation & Control Engineering,
College of Engineering, Pune- 411 005
Maharashtra, India
{naikvihang,dlginoya007,nehagirme}@gmail.com,
dns.instru@coep.ac.in, dingole21@yahoo.in

Abstract. Linear model predictive control (MPC) assumes a linear system model, linear inequality constraints and a convex quadratic cost function. Thus, it can be formulated as a quadratic programming (QP) problem. Due to associated computational complexity of QP solving algorithms, its applicability is restricted to relatively slow dynamic systems. This paper presents an interior-point method (IPM) based QP solver for the solution of optimal control problem in MPC. We propose LU factorization to solve the system of linear equations efficiently at each iteration of IPM, which renders faster execution of MPC. The approach is demonstrated practically by applying MPC to QET DC Servomotor for position control application.

Keywords: Model Predictive Control, Interior-Point Method, LU factorization, DC Servomotor.

1 Introduction

We are considering interior-point algorithm used on-line with linear Model Predictive Control (MPC). Linear MPC assumes a linear system model, linear inequality constraints and a convex quadratic cost function [1]. MPC can be formulated as a quadratic programming (QP) problem and solved at each sampling interval. It thus has the natural ability to handle physical constraints arising in industrial applications [2].

At each sampling instance MPC solves an online QP optimization problem, computes the sequence of optimal current and future control inputs by minimizing the difference between set-points and future outputs predicted from a given plant model over a finite horizon in forward time. Then, only the current optimal input is applied to the plant. The updated plant information is used to formulate and solve a new optimal control problem at the next sampling instance. This procedure is repeated at the each sampling instance and the concept is called as receding horizon control MPC strategy**.** Since this quadratic program can be large depending upon control problem,

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 255–261, 2013.

MPC requires long computation times at each sampling instants, therefore it is usually restricted to systems with slow dynamics and large sampling intervals, such as chemical processes [3]. The ability to solve the QP problem online become critical when applying MPC to systems with fast response time and/or embedded applications where computational resource may be limited [4].

Recently, many reports in the literature address applying MPC to control applications with short sampling intervals, by adapting fast optimization algorithms. These algorithms solve the resulting QP by exploiting the special structure of the control problem the MPC sub problem. The interior-point method (IPM) approaches the solution of Karush-Kuhn-Tucker (KKT) equations by successive decent steps. Each decent step is Newton's like step and the solution is obtained by solving system of linear equations using appropriate numerical methods in order to determine search direction. Matrix decomposition and factorization algorithm provides a means to simplify the computation involved in linear solver. Different decomposition and factorization methods are used such as Gauss elimination, QR, LU, and Cholesky [3], [5], [6].

In this paper, we present an interior-point algorithm to solve MPC problem, which utilizes Mehrotra's predictor-corrector algorithm [7], and the linear system at each IPM iteration is solved efficiently by a LU factorization based Linear Solver. This remainder of the paper is organized as follows: section 2 describes MPC problem formulation. Section 3 presents the interior-point method. MPC implementation on Quanser QET DC Servomotor is presented in section 4. Section 5 shows experimental results. Section 6 is conclusion.

2 MPC Problem Formulation

A discrete time linear time-invariant model of a system in a state space form is given as,

$$\left.\begin{aligned} x(t+1) &= Ax(t) + B\,u(t) \\ y(t) &= Cx(t) \end{aligned}\right\}. \tag{1}$$

where, $y(t) \in \mathbb{R}^m$ are output vectors, $u(t) \in \mathbb{R}^l$ are input vectors and $x(t) \in \mathbb{R}^n$ are internal states vectors. The objective function is defined as,

$$\min_{y,u}(J) = \frac{1}{2}\sum_{i=1}^{N_p}\|\hat{y}(t+i|t) - y_{ref}(t+i|t)\|^2 + \frac{1}{2}\sum_{i=0}^{N_u-1}\|\Delta u(t+i|t)\|^2 . \tag{2}$$

where, $\Delta u(t+i|t) = u_t - u_{t-1}$

subjected to linear inequality constraints on system inputs, where N_p is prediction horizon, N_u is control horizon [8]. In turn, the MPC problem is formulated as a QP problem,

$$\min_U J = \left\{\frac{1}{2}\,U^T H\,U + f^T U\right\} \text{ subjected to } AU \le b\,. \tag{3}$$

where, H is (Nu X Nu) Hessian matrix, f is (Nu X 1) column vector [8], [10].

3 Interior-Point Method

Consider the following standard QP problem,

$$\min_{x} \quad \frac{1}{2}x^{T}Qx + C^{T}x \text{ subjected to } Ax = b, Lx \leq k\,. \tag{4}$$

Except equality constraints in (4), it resembles (3). The KKT Conditions for optimality are given by,

$$\left.\begin{array}{c} Qx + A^{T}y + L^{T}z = -C \\ Ax = b \\ Lx + s = k \\ z^{T}s = 0 \\ z, s > 0 \end{array}\right\}. \tag{5}$$

where, y and z represents the Lagrange multipliers for the equality and inequality constraints respectively. The slack variable s is introduced for converting inequality into equality. By applying Newton's method search direction $(\Delta x, \Delta y, \Delta z, \Delta s)$ is obtained as shown in (6), where $J(x,\ z, \lambda, s)$ denotes Jacobian of $f(x,\ z, \lambda, s)$,

$$J(x,\ z, \lambda, s)\begin{bmatrix}\Delta x \\ \Delta y \\ \Delta z \\ \Delta s\end{bmatrix} = -\begin{bmatrix} r1 \\ r2 \\ r3 \\ r4\end{bmatrix} \tag{6}$$

If we consider only inequality constraints and augmented form of original system is given as,

$$\Delta s = Z^{-1}(r4 - S\Delta z)\,. \tag{7}$$

$$\begin{bmatrix} Q & L^{T} \\ L & -Z^{-1}S\end{bmatrix}\begin{bmatrix}\Delta x \\ \Delta z\end{bmatrix} = -\begin{bmatrix} r1 \\ r3 - Z^{-1}r4\end{bmatrix}. \tag{8}$$

We use Mehrotra's Predictor-corrector algorithm [7] to solve (8) for $(\Delta x, \Delta z)$. Predictor or affine and corrector-centering steps uses $(r1, r3, r4)$ values from (9) and (10) respectively to form the right-hand side of (7) and (8) which in turn finds $(\Delta x_{aff}, \Delta z_{aff}, \Delta s_{aff})$ and $(\Delta x_{cc}, \Delta z_{cc}, \Delta s_{cc})$ [1], [11], [12], [13].

$$\begin{bmatrix} r1 \\ r3 \\ r4\end{bmatrix} = \begin{bmatrix} Qx + C + L^{T}z \\ Lx + s - k \\ ZSe\end{bmatrix}. \tag{9}$$

$$\begin{bmatrix} r1 \\ r3 \\ r4\end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ diag(\Delta z_{aff})diag(\Delta s_{aff})e - \sigma\mu e\end{bmatrix}. \tag{10}$$

where, σ is centering parameter and μ is complementarity measure or duality gap. Thus, solution of linear system in (8) is to be calculated twice at each iteration of IPM, which puts major computational load on overall performance of IPM, and

eventually on MPC. Linear system in (8) uses two different right-hand side components but the same coefficient matrix, so just one factorization of this matrix is required per iteration. Such system can be solved using Gauss elimination, QR, LU, Cholesky factorization methods.

The coefficient matrix in a linear system (8) involved in IPM is symmetric and indefinite so conventional and more numerically stable Cholesky factorization cannot be used to solve the linear system. Compared to computational cost of QR decomposition ($2/3n^3$) and the problem of associated pivoting with ill condition of Gauss elimination, LU factorization proves to be more effective and accurate to solve system of linear equations [5], [6]. We propose LU factorization based Linear Solver for the same. The pseudo code of LU factorization Linear Solver for (8) is given as follows:

Pseudo Code-LU factorization

$A = [A_{11} \quad A_{12}; \quad A_{21} \quad A_{22}]$; where, $A_{11} = Q$; $A_{12} = L'$; $A_{21} = L$; $A_{22} = -Z^{-1} * S$
$B = [B_{11} \; ; \quad B_{22}]$; where, $B_{11} = r1$; $B_{22} = (r3 - Z^{-1} * r4)$
$A = LU$;
Solve LY=B // forward sustitution
UX=Y // backward substitution
$X = [X_1; X_2]$;
$\Delta x = X_1$; $\Delta z = X_2$;

4 MPC Implementation on Quanser QET DC Servomotor

In this paper the efficacy of the proposed method is implemented and demonstrated practically on Quanser QET DC Servomotor plant. The experimental setup has DC Servomotor with Q2 USB based data acquisition card and a PC equipped with proprietary QuaRc 2.1 software. QuaRc provides hardware-in-loop simulation environment [15]. Motor shaft position θ(t) is measured by an encoder. High pass filter is designed to get motor shaft velocity ω(t) numerically. The proposed real-time MPC implementation strategy is shown in Fig. 1.

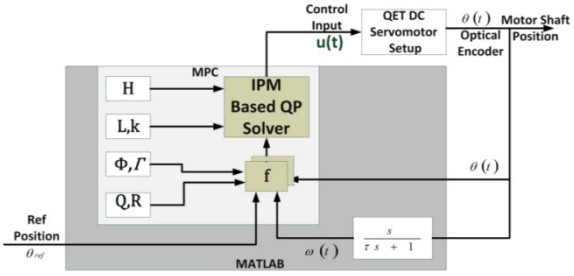


Fig. 1. Real-time MPC Implementation strategy for position control of QET DC Servomotor

The state space representation of DC Servomotor model is given as,

$$A = \begin{bmatrix} 0 & 1 \\ 0 & -10 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 210 \end{bmatrix}, C = [1 \quad 0], D = [0]. \tag{11}$$

5 Experimental Results

The proposed algorithm is coded in MATLAB. Fig. 2 shows per iteration computation time comparison of proposed method with QR factorization and MATLAB's QP solver *quadprog*.

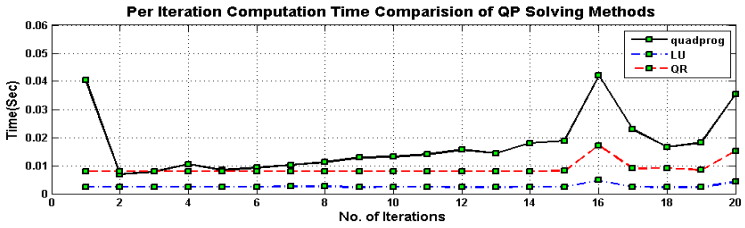


Fig. 2. Computation Time Comparison of QP Solving Algorithms

The performance of QP solver for linear MPC is tested for position control of DC Servomotor with 0.01s sample time. Fig. 3 shows the response of the system output for reference tracking and effect of change of prediction horizon (N_p) by keeping control horizon (N_u) constant. Step response of controller to load disturbances and system uncertainties is shown in Fig. 4. The constraints on control input are -10V and +10V, form Fig. 3 and Fig.4, the constraint handling capability of MPC is verified.

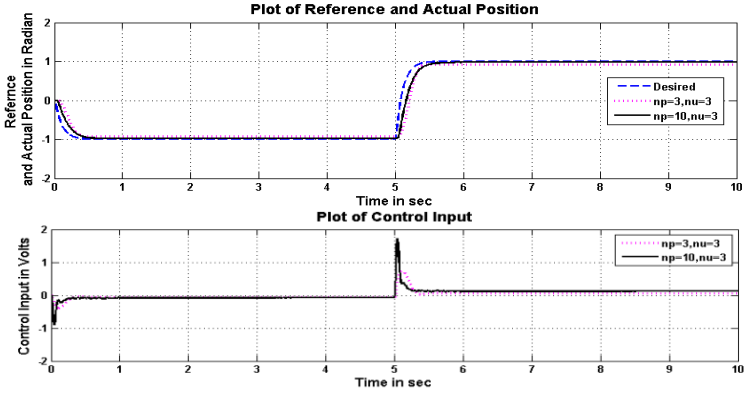


Fig. 3. MPC response to reference position for different prediction horizon

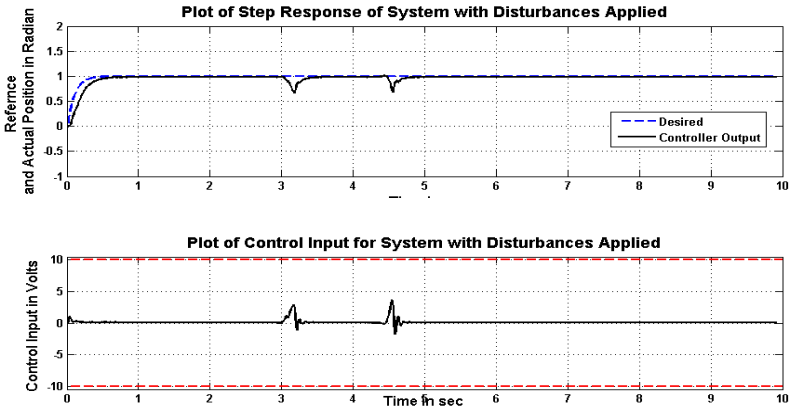


Fig. 4. MPC response to load disturbances

6 Conclusion

In this paper, we presented LU factorization based linear solver for solution of IPM, which is used for real time implementation of constrained linear Model Predictive Controller for position control of DC Servomotor. From experimentation results, it is observed that, the proposed method for MPC solves QP problem efficiently within the specified sample period and performs well for tracking a reference position of DC Servomotor which is relatively a fast dynamic system. It is also observed that, the controller is robust to handle parameter uncertainties and load disturbances.

Acknowledgment. We would like to thank Rashmi More from Department of Mathematics, College of Engineering, Pune, for her sustained contribution towards this work.

References

1. Wills, A.G., Heath, W.P.: Interior-Point Methods For Linear Model Predictive Control. In: Control 2004. University of Bath, UK (2004)
2. Wills, A., Mills, A., Ninness, B.: FPGA Implementation of an Interior-Point Solution for Linear Model Predictive Control. In: Preprints of the 18th IFAC World Congress, Milano, Italy, pp. 14527–14532 (2011)
3. Rao, C.V., Wright, S.J., Rawlings, J.B.: Application of interior point methods to model predictive control. Journal of Optimization Theory and Applications, 723–757 (1998)
4. Ling, K.V., Yue, S.P., Maciejowski, J.M.: A FPGA Implementation of Model Predictive Control. In: Proceedings of the 2006 American Control Conference, Minneapolis, Minnesota, USA, pp. 1930–1935 (2011)
5. Sudarsanam, A., Hauser, T., Dasu, A., Young, S.: A Power Efficient Linear Equation Solver on a Multi-FPGA Accelerator. International Journal of Computers and Applications 32(1), 1–19 (2010)

6. Chai, W., Jiao, D.: An LU Decomposition Based Direct Integral Equation Solver of Linear Complexity and Higher-Order Accuracy for Large-Scale Interconnect Extraction. IEEE Transactions on Advanced Packaging 33(4), 794–803 (2010)
7. Mehrotra, S.: On Implementation of a primal-dual interior-point method. SIAM Journal on Optimization 2(4), 575–601 (1992)
8. Maciejowski, J.M.: Predictive Control with Constraints. Pearson Education Limited
9. Wright, S.J.: Applying new optimization algorithms to model predictive control. In: Chemical Process Control-V, CACHE, AIChE Symposium Series, vol. 93(316), pp. 147–155 (1997)
10. Wills, A.G.: EE04025 - Notes on Linear Model Predictive Control. Technical Report (2004)
11. Kruth, T.R.: Interior-Point Algorithms for Quadratic Programming. IMM-M.Sc-2008-19, Technical University of Denmark (2008)
12. Nejdawi, I.M.: An Efficient interior Point Method for Sequential Quadratic Programming Based Optimal Power Flow. IEEE Trasactions on Power Systems 15(4), 1179–1183 (2000)
13. Wills, A.G., Heath, W.P.: EE03016 Inerior-Point Methods for Linear Model Predictive Control. Technical Report, University of Newcastle, Australia (2003)
14. Stinga, F., Roman, M., Soimu, A., Bobasu, E.: Optimal and MPC Control of the Quanser Flexible Link Experiment. In: 4th WSEAS/IASME International Conference on Dynamical Systems and Control (Control 2008), Greece, pp. 175–180 (2008)
15. Introduction to QuaRc 2.0 & DCMCT Instructor manual and Hardware Guide. Quanser Inc., Markham, ON, Canada (2010)

Design and Implementation of Discrete Augmented Ziegler-Nichols PID Control

Vedika Patki, D.N. Sonawane, and Deepak D. Ingole

Department of Instrumentation & Control Engineering,
College of Engineering, Pune- 411 005
Maharashtra, India
patki.vedika@gmail.com, dns.instru@coep.ac.in,
dingole21@yahoo.in

Abstract. Ziegler Nichols is oldest and widely accepted PID tuning method. Due to excessive overshoot in Ziegler-Nichol tuned PIDs (ZNPID), their performance is usually not acceptable for applications where small error tolerance band and precise control is required. To overcome this problem, we propose the gain updating method called as Augmented Ziegler-Nichols PID (AZNPID). We proposed an algorithm, which uses fourth order Runge-Kutta method to solve differential equations involved in PID, which eventually improves the peak overshoot of AZNPID compared to ZNPID. The proposed AZNPID is tested for third order and fourth order linear processes and its performance is compared with ZNPID.

Keywords: Discrete PID controller, Auto-tuned PID controller, MATLAB.

1 Introduction

In last fifty years, the proportional–integral–derivative (PID) controller has been most commonly used feedback control strategy. PID is used in wide range of applications like automotive, robot motion control, process control, aeronautics etc. [1] It is also integrated as software block in programmable logic controller (PLC), distributed control system (DCS), even an advanced control strategy like model predictive control (MPC) also requires PID at their basic level [2]. PID deals with past (I), present (P) and future (D) errors with a simple control strategy. Some important features of PID includes ease of implementation and good performance in absence of precise analytical model of system under control, which makes PID the most preferred choice of control engineers. In spite of all these advantages, the practical issues like actuator saturation, integral wind-up, anti-reset wind-up and skilled manpower required for accurate tuning restricts the applicability of PID for higher order and precise control systems.

To overcome these drawbacks many auto-tuning methods were proposed [3]-[8]. In [3] feed forward action is used to achieve dynamics of control. Complex control

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 262–268, 2013.

process is combined with separation integral PID and incomplete derivation PID to weaken the influence of swash disturbance and load disturbance.

In [4] PID is combined with General Predictive Control (GPC),along with lightening computational cost of GPC by exploring its internal mechanism and puts forward an improved algorithm PID-GPC. Based on the fractional-order PID control algorithm and dynamic matrix control (DMC) algorithm, the fractional order PID dynamic matrix control (FOPID-DMC) algorithm is given in [5].PID algorithm is combined with Model Predictive Control(PIDMAC) and described in [6].

In this paper, we propose an Augmented Ziegler-Nichol PID (AZNPID) tuning method to reduce system overshoot by auto tuning methods. The AZNPID modifies Ziegler-Nicole PID tuning rule with simple equations to obtain the gain values of Proportional (P), Integral (I) and Derivative (D) terms, tuned after each sample, based on error (e) and change in error (Δe) [9].

The rest of the paper is organized as: In section 2, an overview of PID and Auto-tuned PID controller is given. Section 3 illustrates experimental results. Section 4 discusses the conclusion.

2 Auto-Tuned PID Controller

2.1 Discrete PID Controller

The continuous time PID controller can be expressed as:

$$u(t) = K_p \left[e(t) + \frac{1}{t_i} \int e(t)dt + t_d \frac{d}{dt} e(t) \right] \quad . \tag{1}$$

where, u(t) is the control output, K_p is the proportional gain, t_i is integral time or reset time , t_d is derivative time or rate time, and e stands for the error between the reference and process output Y.

The values of K_p, K_i and K_d are decided using any standard rule. For this study, we have used Ziegler-Nichol's Tuning rule [11]. The values of critical gain K_{cr} and corresponding time period P_{cr} are determined experimentally or can be calculated using bode plot. $K_{p,}t_i$ and t_d are obtained as given in Table 1.

Table 1. Ziegler-Nichol tuning rule

Gain Coefficient	PID
Proportional Gain (K_p)	$0.6K_{cr}$
Integral time (t_i)	$0.5P_{cr}$
Derivative time (t_d)	$0.125P_{cr}$

In order to implement PID on dedicated hardware platform a discrete PID algorithm is considered by replacing derivative term and integral term using a backward difference method and sum using rectangular integration method. Discrete form of position algorithm for PID is given as:

$$u(n) = K_p[e(n) + \frac{t_s}{t_i}\sum e(j) + \frac{t_d}{t_s}[e(n) - e(n-1)]] \ . \quad (2)$$

where t_s is sampling period, n is index and j refers to time constant. Discrete time integral gain K_i and derivative gain K_d obtained as

$$K_i = K_p \frac{t_s}{t_i} \ , \ K_d = K_p \frac{t_d}{t_s} \ . \quad (3)$$

For saving hardware resources, alternate recursive algorithm is used which is given as,

$$u(n) = u(n-1) + a_0 e(n) + a_1 e(n-1) + a_2 e(n-2) \ . \quad (4)$$

where, $a_0 = K_p + K_i + K_d, a_1 = -K_p - 2K_d, a_2 = K_d$.

2.2 Auto-tuned Discrete PID Controller

Fig.1 shows an improved auto-tuning scheme of AZNPID. It is a gain updating method, based on values of α(k) which is function of error (e) and change in error (Δe).

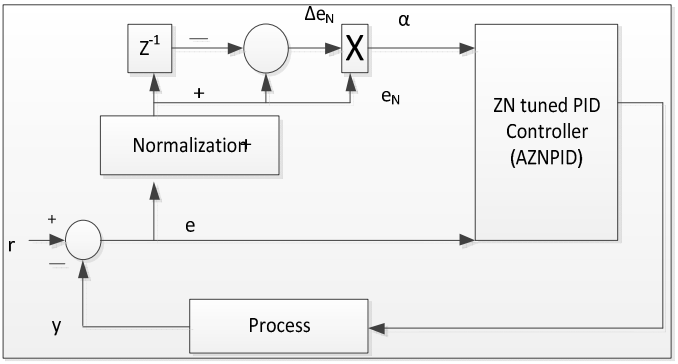


Fig. 1. Implementation of AZNPID

$$e(k) = ref^n - y(k) \ . \quad (5)$$

$$e_N(k) = \frac{e(k)}{abs(ref^n)} \ . \quad (6)$$

$$\Delta e_N(k) = e_N(k) - e_N(k-1) \ . \quad (7)$$

$$\alpha(k) = e_N(k) * \Delta e_N(k) \ . \quad (8)$$

Values of $K_{p,}K_i$ and K_d are modified depending on value of $\alpha(k)$ after every sample instant as [9]

$$K_p^m = K_p(1 + k11|\alpha(k)|) . \quad (9)$$

$$K_i^m = K_i\big(0.3 + k12\alpha(k)\big) . \quad (10)$$

$$K_d^m = K_d(1 + k13|\alpha(k)|) . \quad (11)$$

$$u_{auto} = K_p^m(k)e(k) + K_i^m(k)\sum_{i=0}^{k} e(i) + K_d^m(k)\Delta e(k). \quad (12)$$

where, $\Delta e(k) = e(k) - e(k-1)$& $e(j) = e(k) - e(k+1)$.

For implementing AZNPID on hardware (11) can be written in recursive algorithm form as,

$$u_{auto} = u_{auto}(k-1) + a_0 e(k) + a_1 e(k-1) + a_2 e(k-2). \quad (13)$$

where, $a_0 = K_{p\_m,} + K_{i\_m} + K_{p\_m}$, $a_1 = -K_{p\_m,} - 2K_{p\_m}$, $a_2 = K_{p\_m}$.

In (8)-(10) k11, k12, k13 are three positive constants which gives variation in process value to achieve desired output. We observed that, the constant values k11=1; k12=1; k13=12 are best suitable for the selected processes. From (4)-(7) it was observed that, when process output $y(k)$ is less than the set point, the value of K_{p_m} is greater than K_p so as to reach set-point quickly, K_{d_m} is greater than K_d to avoid oscillations in process output. At the same time K_{i_m} is less than K_i to correct offset errors, which plays a minor role when the process is reaching the set point. As the process output reaches the set-point, values of $K_{p\_m,}$ K_{i_m} K_{d_m} are approximately equal to K_p, K_i, K_d. As the process output increases beyond the set point, the values of $K_{p\_m,}$ K_{i_m}, K_{d_m} increases, this eventually increases control u ($u_{auto} > u$).This brings process back to the desired set-point.

Steps for implementation of AZNPID algorithm:

Step 1: Obtain gain margin and phase margin of system either by using bode plot method or by manual way, where t_i is set to infinity and t_d equal to zero. Increase gain K_p up to the value where system shows sustained oscillations (marginally stable system). Note down corresponding period of oscillations.

Step 2: Above K_p gives critical gain of the system K_{cr} .The period of oscillation obtained in step 1 represents P_{cr} .

Step 3: Applying Zeigler-Nichols formula, the values of K_p, K_i, K_d is obtained from Table 2 and (3).

Step 4: Using the values of constants k11, k12 and k13, AZNPID is applied to the system using (8)-(11).

3 Experimental Results

To test the performance of proposed AZNPID, a third order linear model of system has been considered. The transfer function of system is given as

$$\mathrm{Gp} = \frac{1 - 0.1\mathrm{s}}{(1 + \mathrm{s})^3} \ . \tag{14}$$

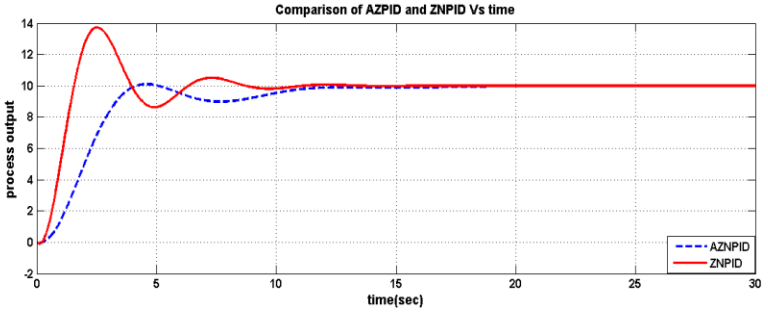


Fig. 2. Comparison of AZPID and ZNPID for third order linear system

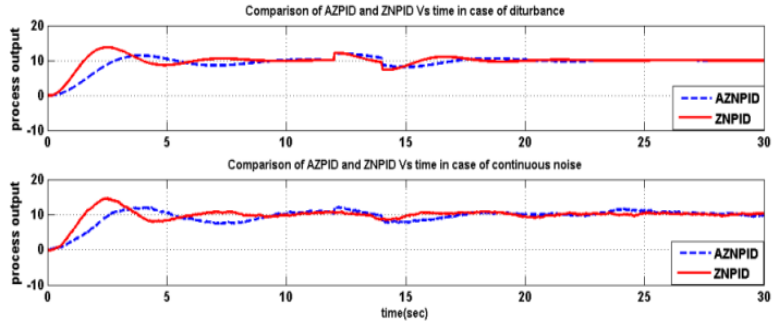


Fig. 3. Comparison of AZPID and ZNPID with disturbance and noise

In fig. 2 the response of proposed AZNPID and conventional ZNPID is compared, it is shown that, the response of proposed method improves the peak overshoot than ZNPID. In fig.3 illustrates the performance of the controllers in presence of the noise with 0.1% magnitude and disturbance (t=12 to 14 sec). In both the cases controller is robust and successfully works in presence of noise.

Similarly, the comparison of proposed AZNPID is noted for fourth order system. The transfer function of fourth order system is considered as

$$\mathrm{Gp} = \frac{1}{(1+\mathrm{s})^4} \tag{15}$$

The comparison of ZNPID and AZNPID are shown in fig.4 and fig.5.

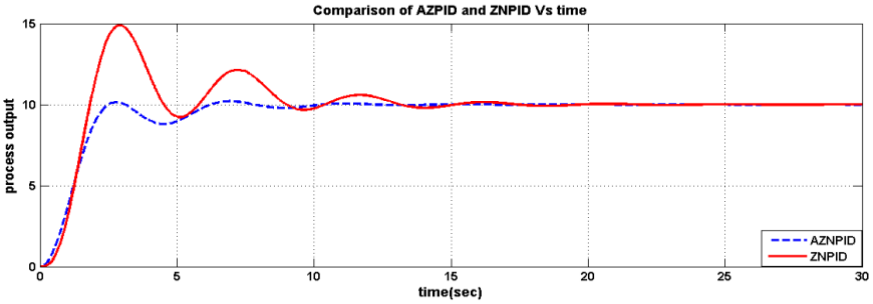


Fig. 4. Comparison of AZPID and ZNPID for fourth order linear system

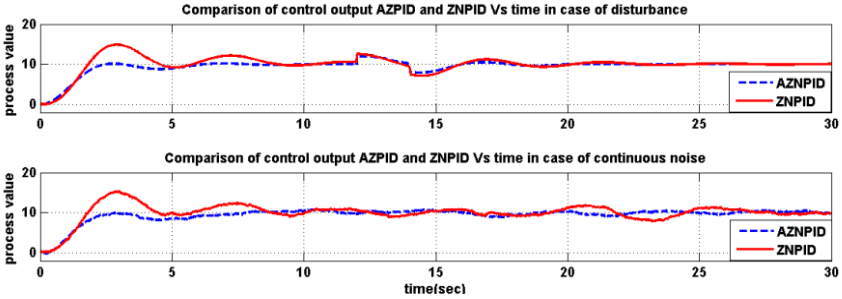


Fig. 5. Comparison of AZPID and ZNPID with disturbance and noise

The processes are recorded for different performance indices and summarized in Table 2.

Table 2. Performance analysis of third order linear model

System	Performance Parameter	ZNPID	AZNPID
$Gp = \frac{1-0.1s}{(1+s)^3}$	% overshoot	37.2	1.1
	Rise Time(sec)	1.43	3.3
	Settling Time(sec)	9.83	11.15
$Gp = \frac{1}{(1+s)^4}$	% overshoot	48.8	1.4
	Rise Time(sec)	1.7	2.04
	Settling Time(sec)	17.57	9.14

4 Conclusion

In this paper, a simple model-independent, discrete auto-tuned scheme of PID controller is presented. It has been tested for third order and fourth order linear systems. Using proposed AZNPID algorithm, it is observed that, the time domain specifications of the system remains stable, even if there is a change in the order of the system. The scheme adjusts gain of P, I, D after every sample instant based on gain updating factor alpha. Although variation of alpha is non-linear, original structure of PID remains intact. Moreover, the scheme can be applied for controlling the fast dynamic applications in dedicated hardware.

Acknowledgments. We would like to thank Mr. Rajesh Bhalerao, Senior Design Engineer, Honeywell Automation India Limited (HAIL) Pune, for her help and support towards this work.

References

1. Desbourough, L., Miller, R.: Increasing customer value of industrial control performance monitoring-Honeywell's experience. In: International Conference on Control Performance Chemical Process Control. AIChE Symposium, Prague, vol. 98, pp. 870–878 (2002)
2. Åström, K.J., Hugglund, T.: The future scope of PID Control. Control Engineering Practice 9, 1163–1175 (2001)
3. Du, S., Feng, Z., Fang, Z.: Design of Improved PID Algorithm for Position Control of Servo Unit. In: International Conference on Computer and Communication Technologies in Agriculture Engineering 2010, pp. 333–335. IEEE Press, Chengdu (2010)
4. Guo, W., Chen, X., Qiu, X.: Application of Improved PID Model Algorithmic Control Algorithm. In: International Conference on Intelligent Computation Technology and Automation, pp. 309–312. IEEE Press, Hunan (2008)
5. Guo, W., Wen, J., Zhou, W.: Fractional-order PID Dynamic Matrix Control Algorithm based on Time Domain. In: 8th World Congress on Intelligent Control and Automation, pp. 208–212. IEEE Press, Jinan (2010)
6. Chen, G.-Y., Ren, P., Pei, H.-L.: An improved PID Generalized Predictive Control Algorithm. In: Seventh International Conference on Machine Learning and Cybernetics, pp. 1877–1881. IEEE Press, Kunming (2008)
7. Vasyutynskyy, V., Kabitzsch, K.: A Comparative Study of PID Control Algorithms Adapted to Send-on-Delta Sampling. In: IEEE International Symposium on Industrial Electronics (ISIE 2010), pp. 3373–3379. IEEE Press, Bari (2010)
8. Chander, S., Agrawal, P., Gupta, I.: Auto-tuned, Discrete PID Controller for DC-DC Converter for fast transient response. In: International Conference on Power Electronics (IICPE 2010), India, pp. 1–7. IEEE Press, New Delhi (2011)
9. Dey, C., Mudi, R.J.: An Improved auto-tuning scheme for PID controllers. ISA Transition 48, 396–402 (2009)
10. Ang, K.H., Chong, G., Li, Y.: PID control system analysis, design, and technology. IEEE Trans. Control Syst. Technol. 13(4), 559–576 (2005)
11. Ogata, K.: Modern Control Engineering. Prentice-Hall of India Pvt. Ltd. (2001)

Harmonic Reduction in Three Phase Rectifier Using Current Injection

Binsy Joseph

Electronics Department
Fr. Conceicao Rodrigues College of Engineering
Bandra, Mumbai, India
bincyjosephy@yahoo.com

Abstract. This paper describes the Matlab simulation for harmonic reduction in three phase bridge rectifier using third harmonic current injection method. The third harmonic current injection is a method, to reduce the input current harmonic content of three-phase rectifier .Rectifier circuit without current injection is compared with that of current injection and power recovery, which proves that considerable reduction in harmonics takes place with current injection method.

Keywords: Harmonic reduction, Power recovery, Third harmonic current injection, THD.

1 Introduction

The application of power electronics in industry has led to the increase in number of non-linear loads. These loads act as harmonic current sources that pollute the main supply. When the generated harmonics are excessive, their effect can be reduced by the installation of harmonic filters. These range from relatively simple passive circuits that are tuned for particular frequencies to complex active filters which inject a compensating current into the supply lines depending on the load harmonic.

The third harmonic current injection is a method to reduce the input current harmonic content of three-phase rectifier. The method is based on the injection of the current whose frequency is equal to the triple of the line frequency at the rectifier input, while the injected current is taken from the rectifier output terminals.

2 Harmonic Reduction by Current Injection Methods

A novel method of reducing harmonic currents on the supply side of a three-phase bridge rectifier was proposed in [1]. The principle of the method is to modify the current waveforms in the dc side windings of the converter transformer by injecting a third harmonic current in to the neutral point of the transformer. Passive LC filters connected between the rectifier output and the secondary neutral point act as third harmonic current source. The major limitation of the method is the constant value of injected current for particular filter value and bridge voltage. With the change in load, the effectiveness of the compensation will be affected.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 269–274, 2013.

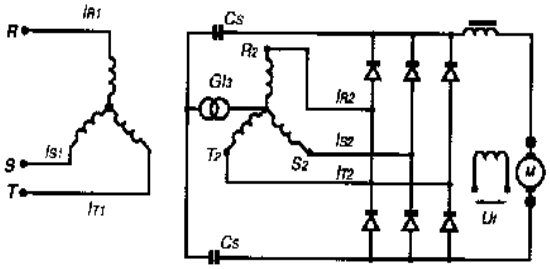


Fig. 1. Harmonic injection circuit using LC filters[1]

Another approach with controlled filters connected between the bridge rectifier output and star point of the transformer secondary, proposed in [2]. The current injection network for low harmonic rectifier proposed in the paper [3] consists of three-phase diode bridge, a current injection device and a current injection network. This transformer rejects the spectral components of the current taken by the network at even triples of the line frequency.

3 Analysis of Current Injection in Proposed Method

In a three phase diode bridge rectifier, in each time instant, the diode which is connected to the highest of the phase voltages and the diode connected to the lowest of the phase voltages conducts. The phase whose voltage is neither minimal nor maximal at the considered time instant remains unconnected. This results in discontinuities in the input currents. Patching of these discontinuities is the essence of the current injection method. The analysis of the current injection method is done on a three phase diode bridge rectifier having a constant load current Iout.

The input current waveform of the first phase for the interval – < ωt< is given by

$$i_{1(t)} = I_{out} \begin{cases} 1 \text{ ; for } -\pi/3 < \omega t < \pi/3 \\ 0 \text{ ; for } -2\pi/3 < \omega t < -\pi/3 \text{ and } \pi/3 < \omega t < 2\pi/3 \\ -1 \text{ ; for } -\pi < \omega t < -2\pi/3 \text{ and } 2\pi/3 < \omega t < \pi \end{cases} \quad (1)$$

The Fourier series of the above input current waveform is given by

$$i_1 = \frac{2}{\pi n}\left[\sin\frac{\pi n}{3} + \sin\frac{2\pi n}{3}\right] \quad (2)$$

The current injection method applies simultaneous injection in all three of the phases, which is achieved by applying a magnetic current injection device or injecting the third harmonic current at the input transformer secondary neutral point. Structure of the diode bridge rectifier applying simultaneous current injection is presented in Figure 2.

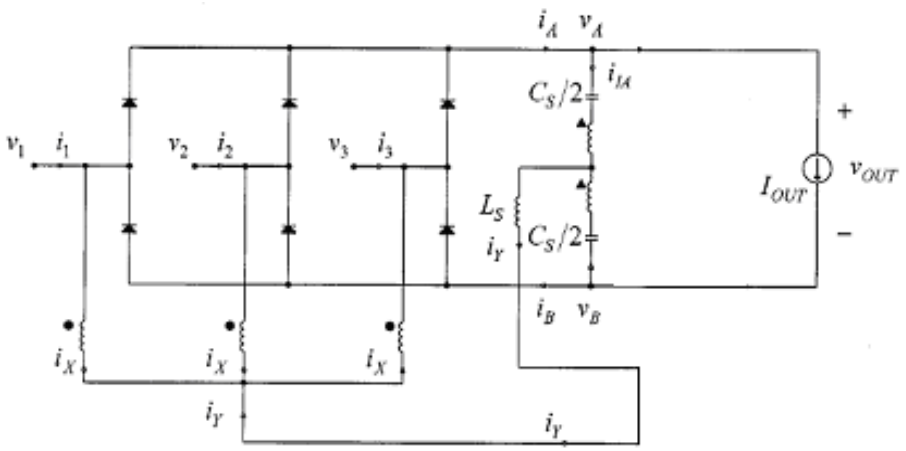


Fig. 2. Rectifier applying current injection

Output current of the current injection network i_y, is divided in three parts by the current injection device, and injected back at the rectifier input.
Input current of the first phase is given by

$$i1(t) = \begin{bmatrix} I_{out} + i_{1A} - i_x & for -\pi/3 \le \omega t \le +\pi/3 \\ -i_x & ; for -2\pi/3 \le \omega t \le -\pi/3; and +\pi/3 \le \omega t \le +2\pi/3 \\ I_{out} + i_{1A} - i_x & -\pi \le \omega t \le -2\pi/3; and +2\pi/3 \le \omega t \le \pi \end{bmatrix} \quad (3)$$

Input currents of the remaining two phases are characterized by the same equations, but with the phase displacement of $2\pi/3$ & $4\pi/3$ .The currents i_{1A} & i_{1B} taken by the current injection network from the bridge output terminals are the same, containing the spectral components at the triple of the line frequency. Then

$$i_{1A} = i_{1B} = \frac{1}{2} i_y = \frac{3}{2} i_x \quad (4)$$

To match the optimal input current waveform, current i_x is

$$i_x = \frac{1}{2} I_{out} \cos(3\omega t) \quad (5)$$

Resulting in the current injection network output current of

$$i_y = \frac{3}{2} I_{out} \cos(3\omega t) \quad (6)$$

Input current of the first phase is given by Equation (7).

$$i1(t) = Iout \begin{bmatrix} 1+\frac{1}{4}\cos\omega t & for\ -\pi/3 \le \omega t \le +\pi/3 \\ -\frac{1}{2}\cos 3\omega t & for\ -2\pi/3 \le \omega t \le -\pi/3; and\ +\pi/3 \le \omega t \le +2\pi/3 \\ 1-\frac{1}{4}\cos\omega t & -\pi \le \omega t \le -2\pi/3; and\ +2\pi/3 \le \omega t \le \pi \end{bmatrix} \quad (7)$$

The Fourier series expansion for the above input current waveform is

$$i_1 = \frac{1}{2\pi n}\frac{n^2-36}{n^2-9}\left[\sin\frac{\pi n}{3} + \sin\frac{2\pi n}{3}\right] \quad (8)$$

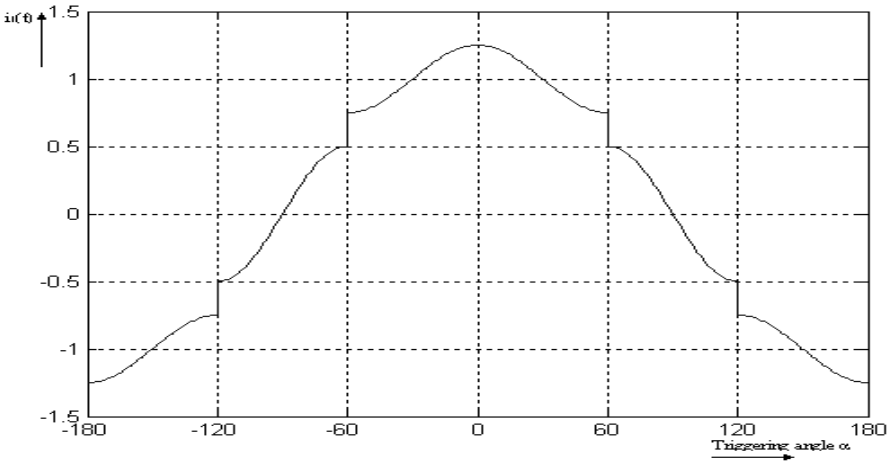


Fig. 3. Input current waveform after injection

The optimal third-harmonic current injection reduced the higher order harmonics for the factor

$$\text{Reduction factor} = \left|\frac{n^2-36}{4n^2-36}\right| \quad (9)$$

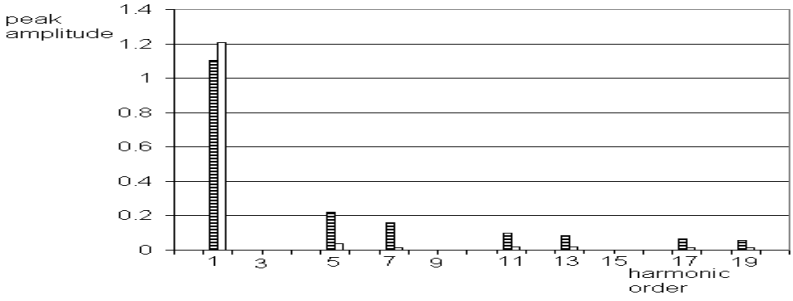


Fig. 4. Harmonic Spectrum ▤ -Without injection □ - With injection

4 Simulation Results

To simulate the performance of the third harmonic current injection, system-modeling package Matlab is used. The system consists of a three-phase rectifier, current injection network and passive resistance emulation network (power recovery circuit). The transformer at the rectifier input replaces current injection device. The current is injected at the neutral of the transformer secondary.

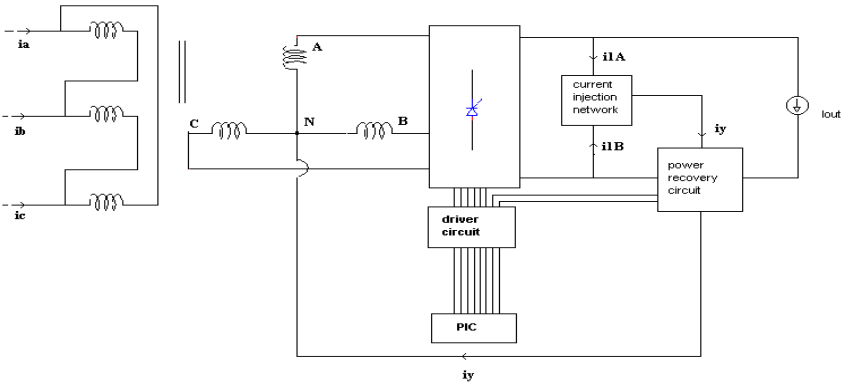


Fig. 5. Structure of three-phase thyristor rectifier with current injection

The main objective of the simulations carried out is to get an idea about the harmonic spectrum of input current of the rectifier circuit with and with out current injection. Simulations have done in three stages-

- Rectifier circuit without current injection and power recovery
- Rectifier circuit with current injection but no power recovery
- Rectifier circuit with both power recovery and current injection.

All the three cases were carried out for a set of firing angles. A constant current load was realized by an RL load.Figure 6 shows input voltage V1,input current i_{1p}waveform at the primary side of transformer and current waveform at secondary side if transformer$i_{1s.}$

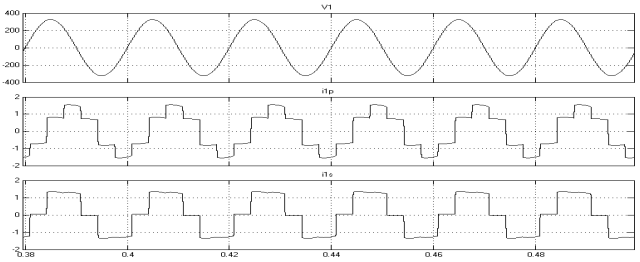


Fig. 6. Voltage and current waveforms without current injection for angle $\alpha=75^{o}$

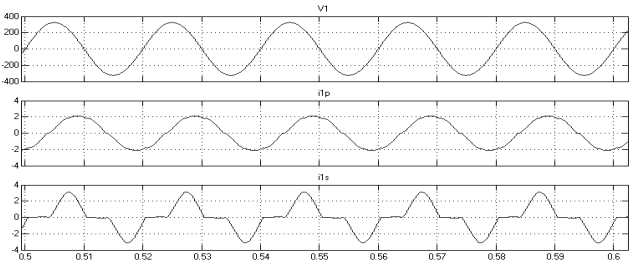


Fig. 7. Input voltage and current waveforms with current injection for angle $\alpha=75^{o}$

Comparison of THD for different firing angles with and without current injection is given in Figure 8.

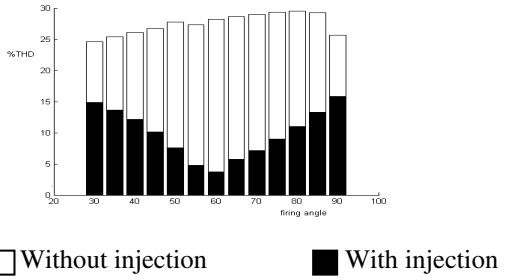


Fig. 8. Comparison of THD with and without current injection for various firing angles

5 Conclusions

The third harmonic current injection scheme for harmonic reduction in a three-phase controlled rectifier has been studied and analyzed. Simulation results shows that rectifier circuit with current injection and power recovery have considerable reduction in harmonics and have better output.

References

1. Lawrance, W.B., Mielczarski, W.: Harmonic current reduction in a three phase diode bridge rectifier. IEEE Transactions on Industrial Electronics 39(6) (December 1992); Han, B.: Single phase Active power filter using FFT with Harmonic Phase delay Compensation. Department of Electrical Engineering Myongji University, Yongin, Korea (Senior Member IEEE)
2. Mielczarski, W., Lawrance, W.B., Nowacki, R., Holmer, D.G.: Harmonic current reduction in three phase bridge rectifier circuits using controlled current injection. IEEE Transactions on Industrial Electronics 44(5) (October 1997)
3. Pejovic, P., Janda, Z.: Improved current injection network for three phase high power factor rectifiers that apply the third harmonic current injection. IEEE Transactions on Industrial Electronics 47(2) (April 2000)

Typing Pattern Recognition Using Keystroke Dynamics

Poonam Rangnath Dholi and K.P. Chaudhari

Marathwada Institute of Technology, Aurangabad
dholipoonam@yahoo.co.in, kiran_chaudhari@rediffmail.com

Abstract. Biometric authentication is individual characteristics that cannot be used by imposter to penetrate secure system. Keystroke dynamics based authentication verifies user from their typing pattern. To authenticate user based on their typing samples, it is required to find out he resemblance of a typing samples of user regardless of the text typed. Key event timing is extracted from key features Latency, Dwell time, Key interval, Up to up, Flight time and standard are measure in the form of FAR, FRR and ER. In this paper we introduces a k-nearest neighbor approach to classify users' keystroke dynamics profiles. For authentication, an input will be checked against the profiles within the cluster which has significantly reduced the verification load.

Keywords: Biometric, Keystroke dynamics, Identification, Verification.

1 Introduction

The increasing use of internet or automated information systems, together with our persistent use of computers has greatly simplified our lives, while making us completely dependent on computers and digital network [2]. In every development of technology the first and foremost step in preventing unauthorized access is user Authentication. User authentication is categorized into three classes: Knowledge based, Object or Token based, Biometric based. Currently, there are two major forms of biometrics: physiological and behavioral. Physiological based biometrics rely on biological attributes such as fingerprints, iris and retina patterns. Behavioral biometrics utilizes behavioral attributes such as voice, signature and keystroke dynamics.

Behavioral biometrics relies on the way we interact with the authentication system. The most popular systems are based on voice, signature and keystroke dynamics. They are more convenient than physiological methods because they generally do not require any special external hardware [3], [9].

In this paper we present a problem of unauthorized users. To overcome this problem various authentication techniques are introduced, we refer to the option of behavioral biometrics in particular through keystroke dynamics and propose a novel intruder classification approach based on the analysis of relation between intruders sessions. In section 2 we present science behind keystroke dynamics, which gives detailed information about keys timing automatic events. In section 3, working flow of keystroke dynamics authentication in terms of identification model and verification model. Section 4 suggests keystroke dynamics authentication algorithm based on k-nearest neighbor classification and section 5 concludes the paper.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 275–280, 2013.

1.1 Previous Works

The need of verifying user based on their keystroke dynamics, especially after the authentication process has ended. A new feature reduction method that is used for each user separately, based on similar di-graphs and their corresponding interval times [4]. In this paper three methods suggested for keystroke dynamics authentication problem. First method: The average and the standard Deviation, second method: Rhythm of striking and Third method: A comparison of time order and also it proposes new method i.e. fusion of above mention tree methods [6]. In this paper mean and standard deviation of latency, duration and di-graph is measured as a keystroke feature. Optimization techniques such as genetic algorithm and particle swarm optimization are used for feature extraction [7]. In this paper soft computing techniques are mentioned [8] to provide user authentication. In this paper k-means classification approach is given for authentication [10].

2 Science behind Keystroke Dynamics

Keystroke dynamics, is totally depends upon timing information of key press and key release. Pressing a keystroke generates three timing atomic events: key *down* event , key *up* event and key *press* event. Key down or key up occur when the key is depressed or released respectively, while key press is fired when an actual character is being inserted in. Keystroke dynamics can be captured via several different features extracted from the typing rhythm of the user including: latency between consecutive keystrokes, flight time, dwell time, based on the key down/press/up events (as shown in Fig. 1), overall typing speed, frequency of errors and control keys. Features extracted based on di-graph, tri-graph or n-graph segments of the entire text. In these, the latencies, intervals and flight time are measured for each sequence of keystrokes [4].

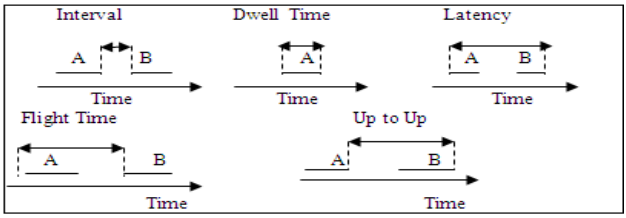


Fig. 1. The five temporal features: Interval, Dwell time, Latency, Flight time, and Up to Up [4]

2.1 Standard Measurements to Test the Approach

False Rejection Rate (FRR) - the percentage of times that a valid user is labeled as an adversary and denied access; also known as the false negative rate. False Acceptance Rate (FAR) - the percentage of times that an adversary gains access as a user: also known as false positive rate. Error Rate (ER) - This is actually an average of the FAR and FRR measures.

3 Working Flow of Keystroke Dynamics Model

In keystroke dynamics includes four modules i.e. keyboard monitoring, features extraction, classifier algorithm, database. Keyboard monitoring which monitors and captures the automatic events generated by the keyboard. Features Extraction which extracts a vector features (timing of key events) from the keystroke. Classifier Algorithms which consists of classification algorithms, which learns the user verification model based on her past data. Later this model is used to classify new sessions for verification. Database which contains user's typing sessions [4], [5].

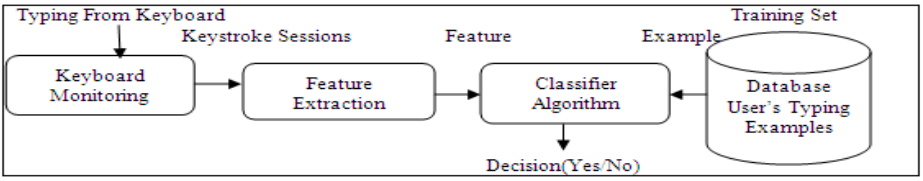


Fig. 2. Typing pattern verification-The user types a stream of keywords, which is monitored by keyboard monitoring then features are extracted by feature extraction and using classifier decision takes place

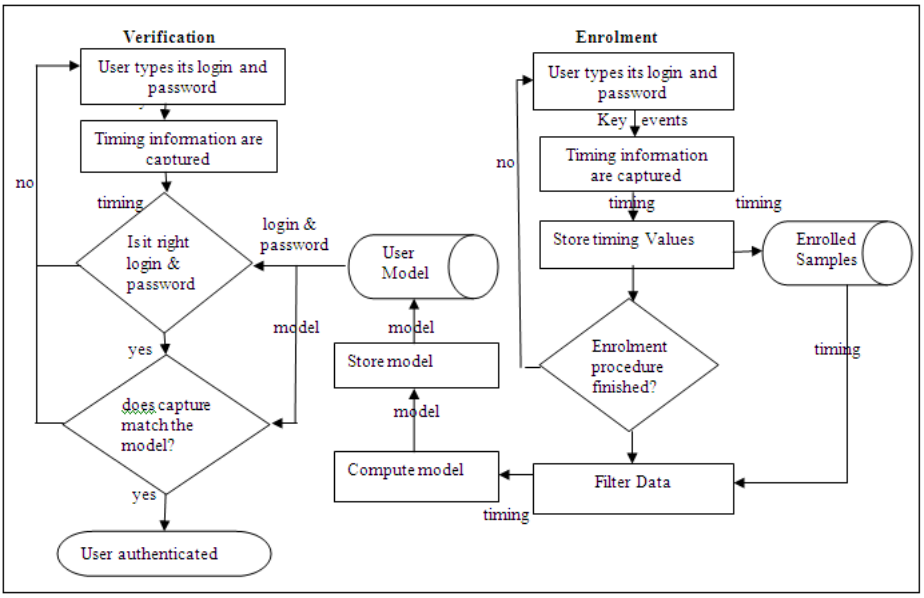


Fig. 3. Keystroke dynamics enrolment and authentication schemes: A password-based authentication scenario

In fig. 3 keystroke dynamics identification and verification scheme presents a password-based authentication. Keystroke dynamics system has different modes: Identification mode (Find) means One-to-many and Verification mode (Check) means One-to-one. In Identification (Enrolment) mode user types his username and password, timing information is captured by key events and this information is stored. In a training period user's typing information is stored into the database. When user login at verification mode that time then current timing information is match with stored timing information, then user is authenticated user.

4 Keystroke dynamics Authentication Based on K-Nearest Neighbor Classification

4.1 Preliminary

The duration of digraphs refers to the elapsed time between the first key pressed and the second key pressed. Same digraphs times in every session are aggregated using the median . Digraphs that are not shared between both sessions are filtered out. Remaining shared digraphs are sorted by its digraph times. For each digraph, its position in both sessions is compared and a distance value is obtained. In order to obtain one representative measure of distance, values obtained in previous step are added. This sum is called degree of disorder [10].For a general distance involving all *n*-graphs, the following formula is given [10]

$$d_{n,m,p}\,(t1,t2) = d_n(t1,t2) + d_m(t1,t2)\frac{M}{N} + d_p(t1,t2)\frac{P}{N}, \tag{1}$$

Where *n, m,* and *p* represent different *n*-graphs and *N, M*, and *P* represent the number of shared *n*-graphs with

N >*M*, *P*. A similarity measure is calculated for every pair of shared digraphs. A pair of digraphs, with times *t*1 and *t*2 respectively, is similar if

$$1 < \frac{\max\,(t1,t2)}{\min\,(t1,t2)} \leq\ t. \tag{2}$$

with constant t having a value of 1.25. Then, A-distance is calculated with the following equation:

$$A_n(t1,t2) < 1 - \frac{(no.of\ similar\ n-graphs\ shared)}{total\ no.of\ \ n-graphs\ shared}. \tag{3}$$

4.2 Proposed Classifier Authentication Method

To increase the matching efficiency of clustering or classification we used k-Nearest neighbor algorithm. K-Nearest neighbor approach is very simple classifier that can easily apply any distance measurement into classification mechanism. KNN is a method for classifying objects based on closest training examples in the feature space. The neighbors are taken from a sessions for which the correct classification is known.

The algorithm on how to compute the K-nearest neighbors is as follows: (i) Determine the parameter K is equal to number of nearest neighbors beforehand. (ii) Calculate the distance between the query-instance and all the training samples. (iii)

Sort the distances for all the training samples and determine the nearest neighbor based on the K-th minimum distance. (iv) Since this is supervised learning, get all the Categories of users training data for the sorted value which fall under K. (v) Use the majority of nearest neighbors as the prediction value.

This is suitable for characteristics of classification and distance based authentication problem. Therefore, we choose the k nearest neighbor approach for our clustering purpose.

Keystroke Dynamics Authentication Algorithm

Building representative user profiles (Enrolment phase): (i) Each legitimate user needs to provide several training samples. Each training sample will contain many *n*-graphs. (ii) A sorted *n*-graphs vector is formed by averaging corresponding groups of the graphs. (iii) A representative user profile is built by averaging all such vectors from all training Samples provided. For convenience, we use its representative mean m(A) to present the mean of the user A. Clustering process: (i) The k-nearest neighbor method is applied to cluster the representative profiles based on the distance measure. (ii) A threshold is set to control the size of the clusters. The optimal threshold is experimentally determined. For a user profile, it associates with a cluster and we call all other profiles associated with the list of this profile. Authentication Process (Verification phase): (i) When a keystroke dynamics sample *X* is claiming as user *A*, a positive authentication is confirmed if following conditions are satisfied: (a) *A* must be within the cluster that is closest to *X*, and (b) md(A,X) value must be the closest to m(A) within its cluster.

5 Conclusions

Keystroke Dynamics is used to verify authenticated user. Using k-nearest neighbor classification uses only its representative profile in the authentication process. The extra overhead of the KDAA algorithm is the need to cluster the database. Even though, this clustering work can be done offline without affecting the authentication process. This proposed k-nearest neighbor approach for user authentication will improve the authentication efficiency up to the 68% compared to the existing method. It will achieve extensively higher accuracy. Keystroke dynamics doesn't require any hardware, is low cost authentication software. It seems promising, still needs more efforts especially for identification.

Acknowledgement. We would like to thanks all pioneers mentioned in this paper for their contribution in the field.

References

1. Moskovitch, R., Feher, C., Messerman, A., Kirschnick, N., Mustafić, T., Camtepe, A., Lohlein, B., Heister, U., Moller, S., Rokach, L., Elovici, Y.: Identity Theft, Computers and Behavioral Biometrics. In: ISI 2009, Richardson, TX, USA, June 8-11 (2009)
2. Monrose, F., Rubin, A.D.: Keystroke Dynamics as a Biometric for Authentication. Courant Institute of Mathematical Science, New York (March 1999)

3. Revett, K.: A Bioinformatics Based Approach to Behavioral Biometrics. University of Westminster, Harrow School of Computer Science London, England HAI 3Tp (2007)
4. Shimshon, T., Moskovitch, R., Rokach, L., Elovici, Y.: Clustering Di-Graphs for Continuously Verify Users according to their Typing Patterns. Department of Information Systems Engineering, Ben Gurion University, Beer Sheva, Israel. 2 Deutsche Telekom Laboratories at Ben, Gurion University, IEEE (2010)
5. Rybnik, M., Tabedzki, M., Saeed, K.: A Keystroke Dynamics Based System for User Identification. University of Finance and Management Ciepla 40, 15-472, Bialystok, Poland, IEEE (2008)
6. Hocquet, S., Ramel, J.-Y., Cardot, H.: Fusion of Methods for Keystroke Dynamic Authentication. Universite Francois-Rabelais de Tours Laboratire d' Informatique (EA, 64 Avenue Jean Portalis 37200 TOURS, France (2010))
7. Karnan, M., Akila, M.: Identity Authentication Based on Keystroke Dynamics Using Genetic Algorith and Particle Swarm Optimization. ,Dept. of Computer Science and Engineering, Tamilnadu College of Engineering, Coimbatore, India, Research Scholar, Dept. of Computer Science and Engineering, Anna University Coimbatore, India, IEEE (2009)
8. Karnan, M.: Personal Authentication Based on Keystroke Dynamics using Soft Computing Techniques. Dept. of Computer Science and Engineering, Tamilnadu College of Engineering, Coimbatore, India, IEEE (2010)
9. Crawford, H.: Keystroke Dynamics: Characteristics and Opportunities. Department of computer Science Sir Alwyn Wiliams Building, University of Glasgow, United Kingdom, IEEE (2010)
10. Pedernera, G.Z., Sznur, S., Ovando, G.S., García, S., Meschino, G.: School of Engineering FASTA University, Mar del Plata, Argentina: Revisiting clustering methods to their application on keystroke dynamics for intruder classification, IEEE (2010)

Investigation on Series Active Filter with Small Energy Source for DC Voltage Control

K. Venkatraman, M.P. Selvan, and S. Moorthi

Dept. of EEE, National Institute of Technology Tiruchirappalli
Tamilnadu, India
venkatraman.kandadai@gmail.com, {selvanmp,srimoorthi}@nitt.edu

Abstract. — This paper presents an analysis of Series Active Filter (SAF) in conjunction with boost converter fed from a small capacity Renewable Energy Source (RES) for the control of DC voltage. Two control schemes like PI controller based Sinusoidal Pulse Width Modulation (SPWM) switching and hysteresis voltage controller based switching have been employed in the study. Boost converter is used to maintain the required DC voltage of SAF. The percentage total harmonic distortion of the compensated load voltage has been considered as a measure for comparison. It has been observed from the simulation results that under non sinusoidal supply and non liner loading conditions, the performance of SAF with hysteresis voltage controller is better.

Keywords: Power Quality, Voltage Quality, Series Active Filter (SAF), Voltage Source Inverter (VSI), Renewable Energy Source (RES), Sinusoidal Pulse Width Modulation (SPWM), Hysteresis Control, MATLAB/Simulink.

1 Introduction

Harmonic distortion in supply voltage is one of the main power quality disturbances frequently encountered by the utilities. The harmonic disturbances in the power supply are caused mainly by the non linear characteristics of the load. The presence of harmonics in supply voltage leads to transformer heating, electromagnetic interference and malfunctioning of solid state devices. Hence it is necessary to reduce the dominant harmonics to below 5% as specified in IEEE 519-1992 harmonic standards[1,2].

Conventionally, passive filters were used to eliminate current harmonics. However passive filters have demerits such as fixed compensation, bulkiness and occurrence of resonance with other elements. With the vast development in the field of power electronics engineering, active filters are introduced as custom power devices for mitigating the power quality problems. Various topologies of active filters have been proposed for harmonic mitigation. The Series Active Filter(SAF) based on voltage source inverter provides attractive solution for harmonic voltage problems[3-5].

A series active filter typically consists of three single phase Voltage Source Inverters (VSIs) connected to the distribution system through an injection transformer and required filters as shown in Fig. 1. It has the capability to inject harmonic voltage of the same amplitude but opposite phase than that of load[6,7].

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 281–286, 2013.

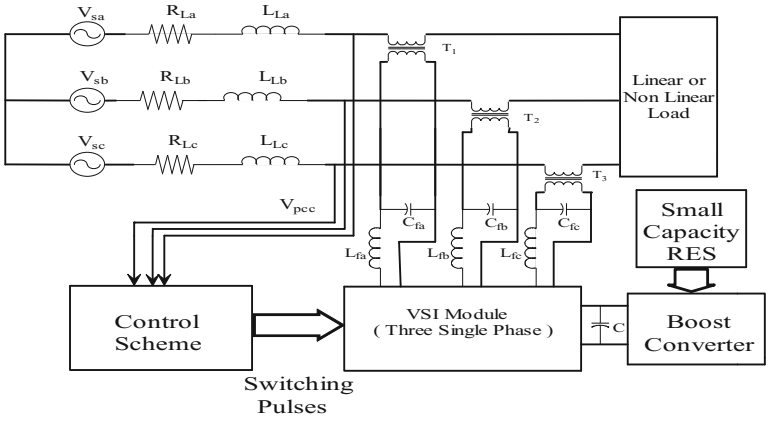


Fig. 1. Schematic diagram of SAF

Several control methods were proposed in the literature for elimination of voltage harmonics using series active filters such as instantaneous active and reactive power method or so called *'p-q' method, 'id-iq' method*, etc. [8, 9]. Different switching schemes for VSI like Sinusoidal Pulse Width Modulation (SPWM), Space Vector Pulse Width Modulation (SVPWM) and hysteresis PWM were also attempted in the literature [10-15].

Generally, SAF is operated to maintain load voltage balanced and sinusoidal. The system considered in this paper is a balanced three phase system with non linear loads. If SAF has to inject active component of harmonic power in to the system, it needs a source of energy. This energy has to be maintained continuously for successful operation of SAF. Hence, an energy source based on renewable energy with sufficient energy storage will be a better choice at the DC side of SAF. Usually small capacity renewable energy source such as PhotoVoltaic (PV) array is used for this purpose. Since the output voltage of the PV array is fluctuating one, in order to get required DC voltage for inverter, a boost converter with PI controller is employed. The present analysis utilizes abc-dq0 transformation [8,9] based reference voltage generation and employs either Proportional Integral (PI) controller with SPWM switching or hysteresis controller based PWM switching for VSI. The effectiveness of SAF with these schemes has been compared using MATLAB simulation under different operating conditions and the results are discussed.

2 PI Controller with SPWM Switching of SAF

The block diagram representation of the SAF control scheme employing PI controller with SPWM switching is shown in Fig. 2.In this control strategy, the reference voltage obtained using abc-dq0 transformation is compared with the actual injected voltage from the SAF, and the error signal is processed using a PI controller. The output of the PI controller augmented with the reference signal is compared with a

very high frequency carrier signal and the pulses are generated as shown in Fig. 3. The generated pulses are fed to the VSI of SAF.

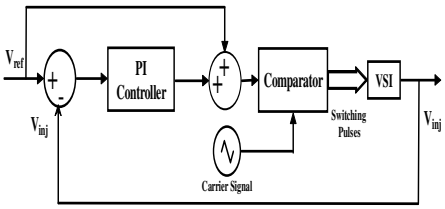


Fig. 2. PI Controller with SPWM

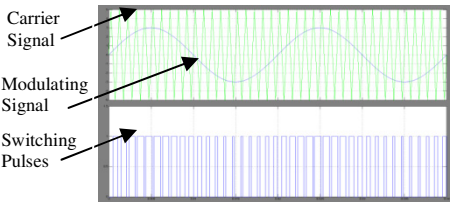


Fig. 3. Switching pulses employing SPWM

3 Hysteresis Switching of SAF

The block diagram representation of the SAF control scheme employing hysteresis controller is shown in Fig. 4. In this control strategy, the switching pulses are generated by comparing the injected voltage fed back from the terminals of VSI with the reference voltage and the associated hysteresis band. The pulse generation using fixed hysteresis band controller is shown in Fig. 5. This method is characterized by unconditioned stability, very fast response and good accuracy. On the other hand, the basic hysteresis technique exhibits several undesirable features, such as uneven switching frequency that causes acoustic noise and difficulty in designing input filters [13].

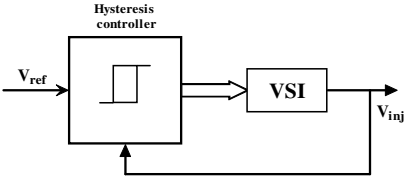


Fig. 4. Hysteresis based switching Controller

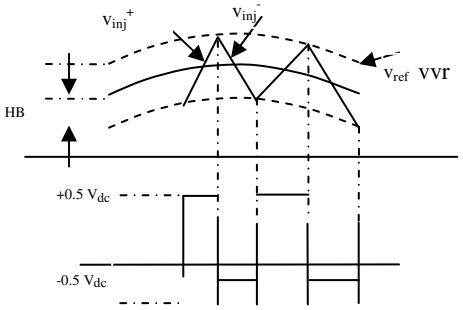


Fig. 5. Switching pulses employing Hysteresis

4 Dc Voltage Control of SAF

As the present system is designed for utilizing the small renewable energy source for controlling the DC voltage of inverter, boost converter is used to step up and maintain the DC voltage at the required level. A PI controller is used for controlling the boost

converter. If there is any change in input dc voltage of boost converter then PI controller adjusts the duty cycle accordingly and maintains constant dc voltage across inverter. The MATLAB Simulink implementation of boost converter with closed loop control is shown in Fig. 6.

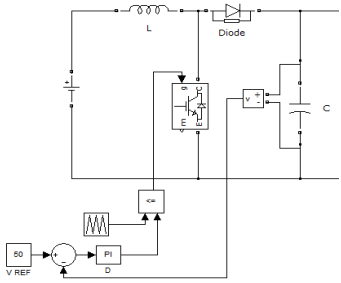


Fig. 6. Closed loop control of boost converter

Table 1. System Parameters

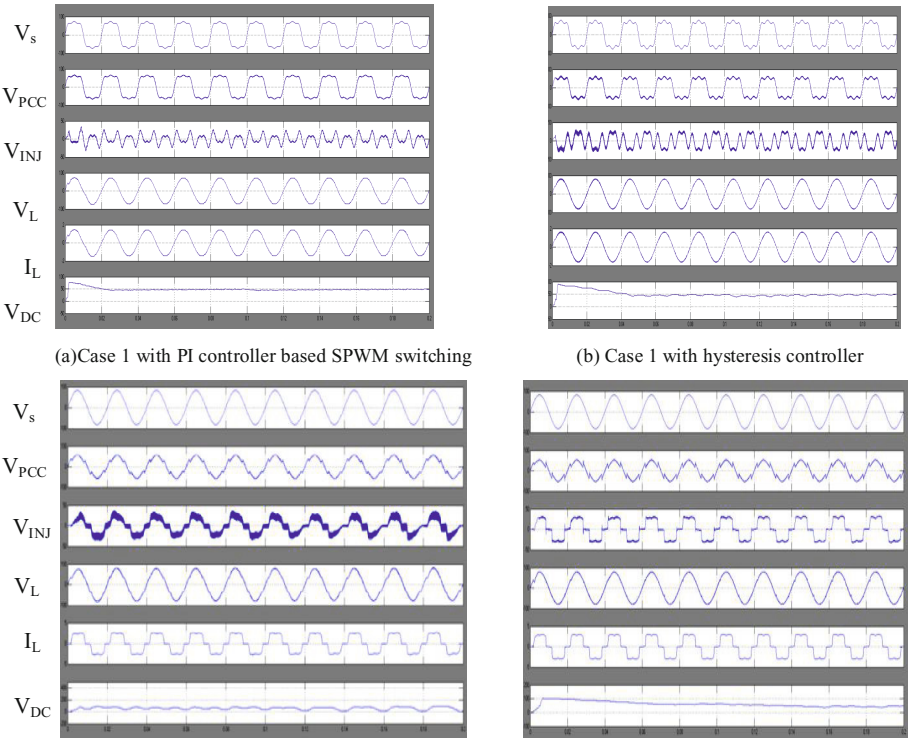
Component	Data
Source	100V, 50Hz
Line parameters	R=10 Ω ,L= 1mH
Load 1	200W
Load 2	Diode Bride Rectifier fed 50 Ω
DC capacitor	150μF- Case 1, 1000μF –Case 2
Injection Transformer	2 kVA, 3 phase, 100 / 100V

5 Simulation Results and Discussion

An intensive simulation study for the investigation of performance of series active filter in conjunction with small renewable energy source fed boost converter for DC voltage control is carried out using MATLAB/SIMULINK. The system parameters are furnished in Table I. Case studies with two different operating conditions as listed below have been carried out employing both the control strategies discussed in section 2 and section 3.

Case 1: R load with 3rd and 5th harmonics in supply – Non sinusoidal supply
Case 2: R load fed by uncontrolled rectifier sinusoidal supply – Non linear loading

The results of these case studies (supply voltage(V_S), voltage at common coupling point(V_{PCC}), injected voltage(V_{INJ}),load voltage(V_L),load current(I_L) and dc voltage(V_{DC})) are shown in Fig 7. In case 1, the supply voltage is having harmonics, as shown in Fig 7 (a) & 7 (b). Hence using the control strategy, an appropriate harmonic voltage is generated and injected into the system through a transformer in phase opposition so that the load voltage is maintained as sinusoidal as possible. Further, results of case 2 confirms due to non sinusoidal current drawn by the rectifier, a non sinusoidal voltage drop occurs because of which load voltage is distorted. This drop is generated and injected through transformer and load voltage is maintained as sinusoidal as shown in Fig 7(c)&(d).



(a)Case 1 with PI controller based SPWM switching (b) Case 1 with hysteresis controller

(c) Case 2 with PI controller based SPWM switching (d) Case 2 with hysteresis controller

Fig. 7. Simulation Results

Table 2. Comparision of Control Strategies

Case	% THD of Load Voltage Before Compensation	% THD of Load Voltage	
		PI Controller with SPWM Switching	Hysteresis Controller
Case 1	22.36	2.49	1.02
Case 2	16.45	4.85	3.42

6 Conclusion

This paper described the harmonic elimination performance of SAF with boost converter fed from small renewable energy source for the control of DC voltage. Two control strategies like PI control with SPWM switching and hysteresis control have been considered for the analysis of SAF. Different case studies have been carried out including 3rd and 5th harmonics in the supply voltage and introducing uncontrolled

rectifier loads. Based on the % THD of compensated load voltage, it can be inferred that the performance of series active filter is better in all the cases when hysteresis controller is employed. The small capacity renewable energy fed boost converter maintains the DC side voltage by delivering sufficient active power to SAF.

References

1. Chowdhury, B.H.: Power Quality. IEEE Potentials 20(2), 5–11 (2011)
2. IEEE recommended practice for evaluating electric power system compatibility with electronic process equipment, IEEE Standard 1346–1998 (1998)
3. Akagi, H., Tsukamoto, Y., Nabae, A.: Analysis and Design of an Active Power Filter using Quad Series Voltage Source PWM Convertors. IEEE Transactions on Industry Applications 5(1), 93–98 (1990)
4. Singh, Al-Haddad, K., Chandra, A.: A review of active filters for power quality improvement. IEEE Transactions on Industrial Electronics 46(5), 960–971 (1999)
5. Akagi, H.: Active harmonic filters. Proceedings of IEEE 93(12), 2128–2141 (2005)
6. Salmeron, P., Litran, S.P.: Improvement of electric power quality using series active and shunt passive filters. IEEE Transactions on Power Delivery 25(2), 1058–1067 (2010)
7. MdHasan, K.N., Romlie, M.F.: Comparative Study on Combined Series and Shunt Active Filters Using Two Control Methods. In: IEEE International Conference on Intelligent and Advanced Systems, pp. 928–933 (2007)
8. Routimo, M., Salo, M., Tuusa, H.: Comparison of Voltage-Source and Current-Source Shunt Active Power Filters. IEEE Transactions on Power Electronics 22, 636–643 (2007)
9. Bhattacharya, S., Divan, D.M.: Synchronous frame based controller implementation for a hybrid series active filter system. In: IEEE/IAS Annual Meeting, pp. 2531–2540 (1995)
10. Montero, M.I.M., Cadaval, E.R., Gonzalez, F.B.: Comparison of control strategies for shunt active power filters in three phase four wire systems. IEEE Transactions on Power Electronics 22(1), 229–236 (2007)
11. Moran, L., Godoy, P., Wallace, R., Dixon, J.: A new current control strategy for active power filters using three PWM voltage source inverters. In: IEEE Power Electronics Specialists Conference, June 20-24, pp. 3–9 (1993)
12. Holmes, G.D., Lipo, T.A.: Pulse Width Modulation for Power Converters-Principles and Practice. IEEE Press Series on Power Engineering. John Wiley and Sons, Piscataway (2003)
13. Bose, B.K.: An adaptive hysteresis band current control technique of a voltage fed PMW inverter for machine drive system. IEEE Transactions on Industrial Electronics 37(5), 402–408 (1990)
14. Zhang, X., Wang, J., Li, C.: Three Phase Four Leg Inverter Based on Voltage Hysteresis Control. In: IEEE Conference on Electrical and Control Engineering, June 25-27 (2010)
15. Srinath, S., Selvan, M.P.: A Combined Mode of Control for UPQC Connected to a Low Voltage Distribution System. Australian Journal of Electrical and Electronics Engineering 8(3), 257–270 (2011)

Modelling Analysis and Sensorless Operation of PMBLDC Motor

E. Kaliappan, C. Chellamuthu, and S. Rajkumar

Department of EEE
RMK Engineering College,
Chennai-601206, India
ekn.eee@rmkec.ac.in, sdrajkumarr@gmail.com

Abstract. This paper presents the simplified modelling and analysis of PMBLDC motor and for sensorless operation. The sensorless scheme employed is based on the backemf zero crossing detection method. PMBLDC motor is modelled using Matlab/Simulink. With the PMBLDC motor model the, dynamic characteristics of PMBLDC motor are monitored and controlled. The validity of sensorless operation is confirmed by simulation results. With small changes in the proposed model, Permanent magnet synchronous motor (PMSM) can also be analysed.

Keywords: PMBLDC, sensorless, modelling, backemf.

1 Introduction

Permanent magnet brushless direct current (PMBLDC) motor are rapidly gaining popularity and are widely used in a variety of applications ranging from small low power applications like fan and disk drives to a large industrial automation and aerospace application because of key features like better speed vs. torque characteristics, high efficiency, noiseless, long operating life and high power density. PMBLDC motors are normally powered by conventional 3Ø voltage source inverters (VSI) or current source inverters (CSI) which is controlled based on the rotor position information based on the rotor position information obtained from hall sensors, resolvers or absolute position sensors. But these position sensors have a lot of drawback like increase in cost, complexity in control, temperature sensitivity requiring special arrangements and reducing system reliability. Therefore sensorless techniques are becoming of great interest in recent times. A number of sensorless techniques have been developed for PMBLDC motor.

In back EMF zero detection crossing, it is difficult to extract the information about EMF, because the machine windings are carrying rapidly changing currents and experience substantial induced voltages from phase switching [1]. In third harmonic of motional EMF, the back EMF waveform has a three times that of fundamental component of any of the phase motional EMF [1]. In [2], difference of line voltages measured at the terminals of the motor and the commutation signals are obtained without neutral voltages. In back EMF integration method [3], the integration based on the absolute value of the open phase back EMF. Integration of the back EMF starts

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 287–291, 2013.

when the back EMF crosses zero. In sensorless BLDC motor, the commutation points of the inverter can be obtained by knowing the zero crossing points of the back EMF voltage of the open phase terminal voltage [4][6].

In this paper a Matlab/Simulink model of PMBLDC motor suitable for Sensorless operation is developed[4][5]. Using this model, the dynamic behavior of the motor is studied. The rest of the paper is organized as follows: Section 2 presents the Modelling of PMBLDC motor with sensorless control using Matlab/Simulink; Section 3 presents the simulation results of the proposed method, section 4 presents the conclusion is discussed.

2 Modelling of PMBLDC Motor

In the proposed scheme, a three phase star connected BLDC motor driven by a three phase inverter using six step commutations is used. The conducting interval of each phase is 120 electrical or two steps.The three phase star connected BLDC motor driven by a three phase inverter with the sensorless control and its ideal waveform are in the figure 1.The voltage equation of BLDC motor from Fig (1) can be derived as

$$V_a = RI_a + (L - M)\frac{dI_a}{dt} + E_a \tag{1}$$

$$V_b = RI_b + (L - M)\frac{dI_b}{dt} + E_b \tag{2}$$

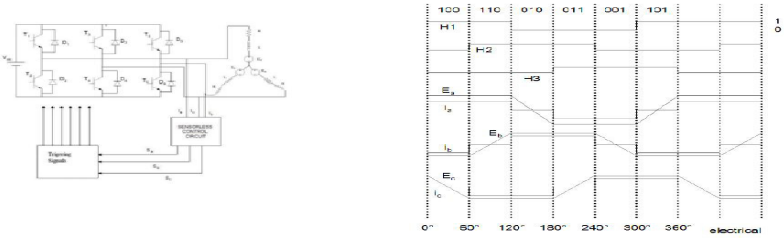


Fig. 1. Simplified BLDC Equivalent Drive and Ideal Waveform

$$V_c = RI_c + (L - M)\frac{dI_c}{dt} + E_c \tag{3}$$

Where,

R = Stator resistance per phase.

L = Stator inductance per phase.

i_a, i_b, i_c = Instantaneous stator phase currents.

V_a, V_b, V_c = Instantaneous stator phase voltages.

e_a, e_b, e_c = Instantaneous back EMFs in the respective phases.

The back-EMF is depending on magnetic flux in rotor because of permanent magnet with speed of rotor, given by

$$e_a = \begin{cases} \left(6E/\Pi\right)\theta_r & 0 < \theta_r < \Pi/6 \\ E & \Pi/6 < \theta_r < 5\Pi/6 \\ -\left(6E/\Pi\right)\theta_r + 6E & 5\Pi/6 < \theta_r < 7\Pi/6 \\ -E & 7\Pi/6 < \theta_r < 11\Pi/6 \\ \left(6E/\Pi\right)\theta_r - 12 & 11\Pi/6 < \theta_r < 2\Pi \end{cases} \tag{4}$$

From the equations (1) to (3), the line voltages are obtained as

$$V_{ab} = R\left(i_a - i_b\right) + L\frac{d}{dt}\left(i_a - i_b\right) + e_a - e_b \tag{5}$$

$$V_{bc} = R\left(i_b - i_c\right) + L\frac{d}{dt}\left(i_b - i_c\right) + e_b - e_c \tag{6}$$

The generated electromagnetic torque is given by

$$T_e = \frac{Z_p}{2\omega_r}\left(e_a i_a + e_b i_b + e_c i_c\right) \tag{7}$$

Where,

Z_p = Number of pole pairs.

ω_r = Electrical speed of rotor.

Electrical speed of the rotor is determined from the motion differential equation and it is given by

$$T_e - T_L = B.w_r + J\frac{d}{dt}\left(w_r\right) \tag{8}$$

Where,

J = Moment of inertia.

B = Friction coefficient.

T_L = Load torque.

3 Simulation Modelling of PMBLDC Motor

Figure 2 shows the simulation model of the PMBLDC motor drive consists of voltage and current block, terminal voltage block, generation of backemf block, sensorless control block i.e. proposed controller. The sensorless block gives the sensorless position information. Position of rotor position gives the incoming and outgoing phases.In BLDC block, the constant values of torque constant and backemf constant are 0.1418 Nm/A and 0.1V/rad/sec. The moment of inertia (J) is given by 0.1592kgm^2 and resistance value is 0.7 ohm. The inductance (L) and mutual inductance (M) value is given by trial and error method.

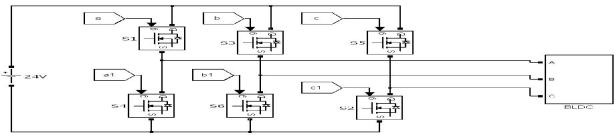


Fig. 2. BLDC Block

Figure 3-4 shows the subblock of BLDC motor which has voltage and current equation block, terminal voltage block, generation of EMF block, Sensorless Block.

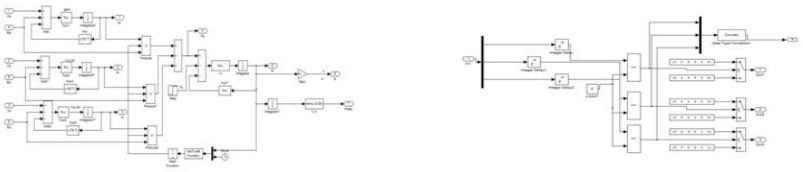


Fig. 3. Voltage and Current Equation and Sensorless Sub-block

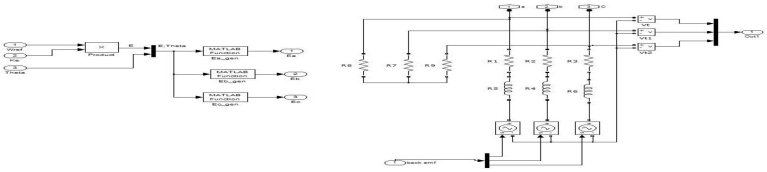


Fig. 4. Generation of back EMF and Terminal voltage sub-block

4 Simulation Results

The simulation results of simplified modelling and sensorless operation of permanent magnet brushless DC motor is obtained by MATLAB /Simulink. The version of MATLAB is 7.9.0(R2009) and the solver to run the model is ode45 (Dormand-Prince).The simulation model of PMBLDC motor is run for a period of 2 second.Load is applied at 0.5 seconds.Figure 10-17 shows the waveform for back EMF, phase current, rotor speed, electromagnetic torque under NO load and loaded conditions.

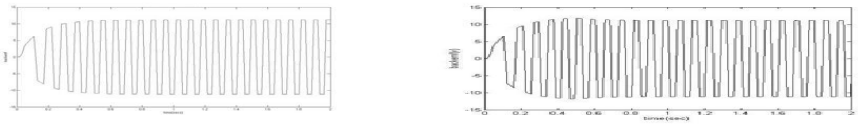


Fig. 5. Back EMF for NO Load and Loaded Condition

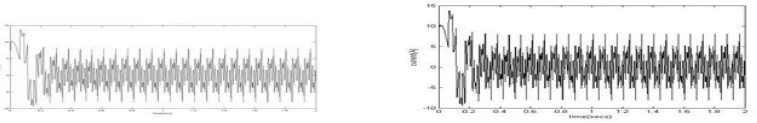


Fig. 6. Phase Current A for NO Load and Loaded condition

Fig. 7. Rotor Speed for NO Load and Loaded condition

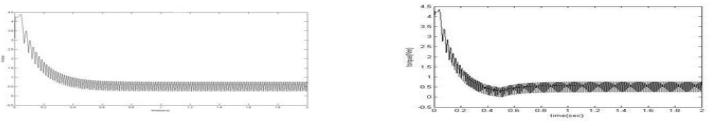


Fig. 8. Electromagnetic Torque for NO Load and Loaded condition

5 Conclusion

The simplified modelling is developed and simulation of sensorless operation of permanent magnet brushless magnet is developed by MATLAB /simulink simulation. The backemf zero crossing detection is used to detect the rotor position and backemf, current, rotor speed, electromagnetic torque waveform were obtained for both load and NO load condition without Hall sensors. This simplified model can be applied to other sensorless technique with the small change in the model.

References

1. Acarnley, Watson: Review of Position Sensorless operation of Brushless Permanent Magnet Machines. IEEE Trans. Ind. Appl. 53(2), 352–362 (2006)
2. Damodharan, P., Vasudevan, K.: Sensorless Brushless DC Motor Drive Based on the ZeroCrossing Detection of back EMF from the line voltage Difference. IEEE Trans. Energy Conv. 25(3), 661–668 (2010)
3. Lai, Y.-S., Lin, Y.-K.: A unified approach to zero-crossing point detection of back EMF for Brushless DC motor drives without current and hall sensors. IEEE Trans. Power Electron. 26(6), 1704–1713 (2011)
4. Lee, B.K., Ehsani, M.: Advanced simulation Model for Brushless DC Motor Drive. Electronic Power Components and Systems, 841–868 (2003)
5. Luk, P.C.K., Lee, C.K.: Efficient Modeling of Brushless DC motor drives. In: Industrial Electronics, Control and Instrumentation (1994)
6. Kaliappan, E., Chellamuthu, C.: A Simple Sensorless Control technique for PMBLDC motor using Back EMF Zero Crossing. European Journal of Scientific Research 60(3), 365–378 (2011)

Energy Loss Estimation: A Mathematical Approach*

Murali Krishna Talari[1], P. Sai Gautham[1], N.V. Ramana[2], and S. Kamakshaiah[3]

[1] EEE Dept, CVR College of Engineering, Hyderabad, AP, India
[2] EEE Dept, JNTU College of Engg, Jagityal, AP, India
[3] EEE Dept, Chairman of Electrical Science, JNT University, Hyderabad, A.P., India

Abstract. This work focuses on energy loss estimation in electrical sub transmission and distribution systems. The earliest empirical approaches, by Buller and Woodrow in 1928 and Hoebel in 1959, were modified by M.W. Gustafson in 1983, by changing the values of the coefficients and adding a constant loss term. It has been observed that this approach is not suitable for the present load scenario. In this paper, a new approach has been proposed and tested with real time data. It has been concluded that the relation ship between loss factor and load factor is not as complicated as perceived but easily understandable. By using exponential curve fitting, a relationship that is very close to reality can be obtained. An exponential curve has the ability to take on any shape between two given limits, based on the specified data. To verify the proposed approach data has been collected from a 33kV sub transmission line existing between 132kV/33kV Turkayamjal substation & 33kV/11kV Hayatnagar substation, APTRANSCO, Andhra Pradesh.

Keywords: Loss factor, Load factor, Load curve, Loss Curve.

1 Introduction

According to the fundamental laws of nature, energy cannot be converted from one form to the other, unless until there is an opposition to the conversion. An analogous statement exists for Electric Power Transmission and Distribution. According to this statement, power cannot be transmitted from one port to the other, without losses. They arise as power flows, to satisfy consumer demands, through the network. Electrical energy is converted from various forms of conventional and non conventional energy sources at suitable locations, transmitted at a high voltage over long distance and distributed to the consumers at a medium or low voltage. A distribution system consists of a substantial number of distribution transformers, feeders, and laterals, due to a large number of consumers spread over a wide geographical area. Some of the input energy is dissipated in the conductors and transformers along the delivery route. Total system loss is the difference between the energy sent and the energy received by the end users. Losses increase the operating cost, estimated to add *10%* to the cost of electricity and also approximately *25%* to the cost of delivery. These costs are levied on the consumers.

* We acknowledge Kiran Kumar Challa, M. E (IISc Bangalore) for his valuable discussion at each point of this research. We also thank V. Prameela, Assistant Engineer, APTRANSCO, Hyderabad for providing vital data.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 292–297, 2013.

Studies have indicated that losses in a distribution system contribute to a major part of losses in a power system (approximately about 70% of total losses). Therefore, the distribution system loss has become more and more of concern, because of the growth of load demand and the wide area it covers. It is important to estimate them accurately in order to take appropriate measures for reducing the losses. The evaluation and reduction of energy losses present the distribution companies with taxing challenges. The accurate estimation of electrical losses enables the supply authority to determine the operating costs, with greater accuracy, and abate the dissipation of power.

The loss factor can be used to calculate energy losses for those parts of electrical system where the current flowing is proportional to system load each hour, which would typically be the distribution and sub transmission systems. The near quadratic relationship that exists between load and loss factors has been used to develop empirical relations for estimation of loss. Attempting to make a lead in this aspect in 1928, Buller and Woodrow proposed a relationship between load factor and loss factor, by using a completely empirical approach [1]. This approach was improved by Hoebel in 1959 [2]. Martin W. Gustafson, a Transmission/Substation design engineer, modified the existing relationship between the two factors [3], [4]. He modified the constant coefficient and also added a new coefficient for losses [5]. He extended the same approach to evaluate transmission loss using special considerations [6]. Later on various methods were proposed to estimate sub transmission and distribution system losses, which include the use of artificial intelligence techniques like Fuzzy Logic and Neural Networks [7], [8]. The energy loss estimation techniques were depicted in [9].

Mathematics provides us curve fitting methods to establish the exact relationship between two sets of data. Previous approaches like that of Buller and Woodrow utilized curve fitting, but the assumptions made to find the relationship were not close to the actual conditions, but were ideal. In any relationship between loss factor and load factor, providing fixed limits to the values of the coefficients is important, as this helps in judging its accuracy.

Before discussing the new approach to estimate energy losses by using mathematical methods, it is necessary to understand the basic approach i.e. Buller and Woodrow Equation.

2 Buller and Woodrow's Approach

From two individual research works done by Buller and Woodrow in 1928 [1] and H. F. Hoebel in 1959 [2] an empirical equivalent hours loss factor equation was proposed.

$$EQ_f = (LD_f)^2(1-X) + (LD_f)X \quad (1)$$

Where, EQ_f = Loss Factor
LD_f = Load Factor
X = Constant Coefficient

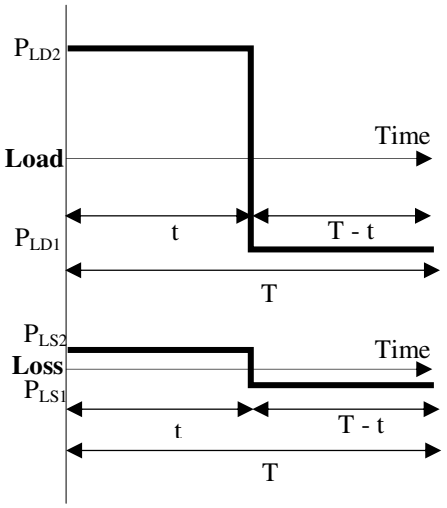


Fig. 1. Load Curve and a Loss Curve

Buller and Woodrow assumed an arbitrary and idealized case of a load curve, where it consisted of a peak load for a time t, and an off-peak load for a time (T-t), where 'T' was the complete time period taken into account as shown in Fig. 1. As the loss varies with square of the current and current directly proportional to the load, the loss varies with the square of the load. Therefore, the loss curve too is in the shape as that of the load curve [10].

$$LD_f = t/T + (P_{LD\,2}/P_{LD\,1})(1 - t/T) \tag{2}$$

$$EQ_f = t/T + (P_{LS\,2}/P_{LS\,1})(1 - t/T) \tag{3}$$

As loss varies with the square of the load,

$$EQ_f = t/T + (P_{LD\,2}/P_{LD\,1})^2(1 - t/T) \tag{4}$$

Buller and Woodrow then considered two boundary conditions, viz. complete peak (t→T) and complete off-peak (t→0). From these conditions, they came to a conclusion of the shape of the relationship between loss factor and load factor by fitting a curve between these two.

Even though, Buller and Woodrow's approach was considered to be useful, and it has been used, it consists of approximations far from reality. The load curve considered was an ideal one consisting of only peak and off peak as shown in Fig. 1. The average is a function of only peak and off peak. But the intermediate values affect the average. In the derivation of the equivalent hours loss factor it was considered that

$$P_{LSk} = C_k\, P_{LDk}\ [k=1,2,3\ldots] \tag{5}$$

with the consideration that $C_1=C_2$. This need not be right, as the value of the coefficient can be different at the off-peak, than at the peak.

3 Proposed Mathematical Approach

Many processes in nature have exponentia1 dependencies. The advantage of exponential equations is that with the right values of coefficients, any curve can be approximated in the form of an exponential equation i.e., within the limits of the given data. In exponential curve fitting, any function is approximated as

$$y = Ae^{Bx} \tag{6}$$

The advantage of such an equation is that 'A' decides the 'y' intercept of the curve, and 'B' gives the function its curvature. Taking the logarithm of both sides of (6)

$$\log y = \log A + Bx \tag{7}$$

Then, the best fit-values are

$$A = \frac{\sum_{i=1}^{n} \ln y_i \sum_{i=12}^{n} x_i^2 - \sum_{i=1}^{n} x_i \ln y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \tag{8}$$

$$B = \frac{n\sum_{i=1}^{n} x_i \ln y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} \ln y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \tag{9}$$

Where 'x_i' and 'y_i' are the sets of data, that must be fit in the required curve. Using (6) to fit a curve between the two factors, let the relationship between them be

$$LS_f = ae^{bLD\,f} \tag{10}$$

Taking the natural logarithm on both sides of (10)

$$\ln(LS_f) = \ln(a) + bLD_f \tag{11}$$

This can be plotted as a straight line between the natural logarithm of loss factor and load factor. Therefore, the general equation of the relationship can be shown as

$$\ln(LS_f) = A + B \times LD_f \tag{12}$$

The values of these coefficients can be determined by using (8) and (9). Due to its desired features of a given 'y' intercept value, and curvature, it can be used easily to estimate the loss factor of a system, from a given value of a load factor.

Let the data points that were collected be (LDf_1, LSf_1) (LDf_2, LSf_2)…….. (LDf_n, LSf_n). These data points can be obtained from the system, whose loss is to be estimated. Therefore, in the approximate exponential equation:-

$$A = \ln\left(\frac{\sum_{i=1}^{n}(\ln LSf_i \sum_{i=12}^{n} LDf_i^2) - \sum_{i=1}^{n}(LDf_i \ln LSf_i)}{n\sum_{i=1}^{n} LDf_i^2 - \left(\sum_{i=1}^{n} LDf_i\right)^2}\right) \tag{13}$$

$$B = \frac{n\sum_{i=1}^{n} LDf_i \ln LSf_i - \sum_{i=1}^{n}(LDf_i \sum_{i=1}^{n} \ln LSf_i)}{n\sum_{i=1}^{n} LDf_i^2 - \left(\sum_{i=1}^{n} LDf_i\right)^2} \tag{14}$$

These coefficients can be used for future estimation of losses, for a given system, by recording the required load factor.

4 Results

The proposed approach has been tested on a 33kV sub transmission line existing between 132kV/33kV Thurkayamjal substation and 33kV/11kV Hayathnagar substation, APTRANSCO, Andhra Pradesh. Annual load and loss data of the test system has been collected for the year 2010 and monthly load and loss factor were evaluated by plotting monthly load and loss curves.

Table 1. Monthly estimated loss factors

Month	Load Factor	Loss Factor	Buller & Woodrow	Proposed Approach
Jan	0.7515	0.6026	0.6849	0.5882
Feb	0.6598	0.4740	0.5797	0.4794
Mar	0.7380	0.5793	0.6690	0.5707
April	0.7547	0.5898	0.6886	0.5924
May	0.8480	0.7305	0.8020	0.7294
June	0.8516	0.7310	0.8065	0.7354
July	0.7973	0.6490	0.7396	0.6514
Aug	0.7360	0.5718	0.6667	0.5683
Sept	0.7763	0.6398	0.7144	0.6217
Oct	0.7881	0.6485	0.7285	0.6382
Nov	0.7996	0.6580	0.7424	0.6548
Dec	0.7063	0.5361	0.6323	0.5318

The constant coefficient 'X' of equation (1) has been calculated and value of the coefficient 'X' was 0.6432**.** It was further used to calculate the loss factors from respective monthly load factors. For the proposed method, the coefficients 'A' and 'B' were calculated from the average monthly load and loss factors. The values of the coefficients 'A' and' B' were -2.2073 and 2.231**.** These coefficients were again used to calculate the loss factors for the proposed approach. The summary of the analysis is shown in Table 1. From the results in the Table 1, it can be seen that the calculated values of the loss factor from the proposed approach are closer to the actual values than those obtained from Buller & Woodrow's approach. The function values are negative i.e. $\ln(LS_f)$ is negative. This owes to the fact that loss factor is always lesser than or equal to 1. Therefore, its natural logarithm is always negative. The above point limits the value of the first coefficient 'A'. The function exp(A) is also a loss factor. Therefore 'A' must be negative. This means that the graph between $\ln(LS_f)$ and load factor always has a negative y-intercept and lies in the fourth quadrant.

5 Conclusions

The method presented is a new tool capable of estimating sub-transmission and distribution system losses with greater accuracy, than what was previously available. It is also observed that the values of the coefficients are very important in estimating losses, but they cannot be generalized for any system. This is due to the fact that each system has a different load profile, which leads to different values of coefficients.

Hence, it is advised that a distribution or transmission company, use previous data, to calculate the values of the coefficients, for their respective load profile. Fitting an exponential curve to data does not need any tedious programming or simulation.

References

1. Buller, F.H., Woodrow, C.A.: Load factor equivalent hours values compared. Electr. World (July 1928)
2. Hoebel, H.F.: Cost of electric distribution losses. Electr. Light and Power (March 1959)
3. Gustafson, M.W.: Demand, energy and marginal electric system losses. IEEE Trans. Power App. Syst. PAS-102(9), 3189–3195 (1983)
4. Gustafson, M.W., Baylor, J.S., Mulnix, S.S.: Equivalent hours loss factor revisited. IEEE Trans. Power Syst. 3(4), 1502–1507 (1988)
5. Gustafson, M.W., Baylor, J.S.: Approximating the system losses equation. IEEE Trans. Power Syst. 4(3), 850–855 (1989)
6. Gustafson, M.W., Baylor, J.S.: Transmission loss evaluation for electric systems. IEEE Trans. Power Systems 3(3) (August 1988)
7. Hong, Y.Y., Chao, Z.T.: Development of energy loss formula for distribution systems using FCN algorithm and cluster-wise fuzzy regression. IEEE Trans. Power Del. 17(3), 794–799 (2002)
8. Hsu, C.T., Tzeng, Y.M., Chen, C.S., Cho, M.Y.: Distribution feeder loss analysis using artificial Neural Network. Electric Power Systems Research 34(2), 85–90 (1995)
9. Nagendra Rao, P.S., Deekshit, R.: Energy Loss Estimation in Distribution Feeders. IEEE Trans. Power Del. 21(3), 1092–1100 (2006)
10. Gönen, T.: Electric Power Distribution System Engineering. McGraw-Hill (1986)

Design and Development of Controller for Stand-Alone Wind Driven Self-excited Induction Generator

M. Sathyakala[1] and M. Arutchelvi[2]

Department of EEE, Saranathan College of Engineering, Trichy-620012,Tamilnadu, India
muthuvelsathya@yahoo.com, arutchelvi@gmail.com

Abstract. The 3-Φ self-excited induction generator driven by wind energy source is suitable for stand-alone applications. For such an application, the minimum excitation capacitance required for self-excitation of 3-Φ induction generator is taken up in this work and the detailed analysis is carried out to determine the range of wind speed variation and consumer demand for the designed capacitance value. An electronic load controller is designed to maintain the load voltage constant for these variations. The excess power resulting as a consequence of rise in load voltage due to variation in load is pumped to dump load. Exhaustive simulations have been carried out for such a scheme and the results have been presented in this paper.

Keywords: wind energy source, self-excited induction generator, electronic load controller, chopper, dump load.

1 Introduction

Today, the wind energy is the world's fastest growing energy source. Permanent magnet synchronous generator (PMSG) is particularly an attractive option in renewable energy applications, but the permanent magnets are quite expensive, demagnetization of permanent magnet at high temperature [1]-[3]. The induction generator (IG) is the most common generator in wind energy system applications in the stand-alone systems due to its simplicity, robust, small size, weight and ruggedness, more than 50 years life time [4]-[7],[9]. The interest in this topic is primarily due to the application of self-excited induction generator (SEIG) in isolated power systems where a capacitor bank is connected across the stator terminals, which provides the required excitation for SEIG [8]. The terminal capacitance on such machines must have a minimum value so that self-excitation is possible. Due to the brushless construction, low maintenance and low capital cost SEIG gives an important advantage over the synchronous generator. The SEIG requires minimum excitation capacitance for self-excitation which provides lagging reactive power of both IG as well as load. The design of excitation capacitance depends upon mainly on the magnetizing reactance and per unit frequency [10]-[15]. Although, all numerical results yields the solution for general R-L case, initial guessing of magnetizing reactance and per unit frequency is required, tedious and lengthy derivations are involved. A simple method and direct method for computing the minimum value of

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 298–304, 2013.

excitation capacitance required for initiating voltage build up in a 3-Φ SEIG under R-L load is presented in [15].

In this paper, the 3-Φ stand-alone wind-driven SEIG with an electronic load controller (ELC) feeding R-L load is undergone where the self-excitation capacitance is calculated by using direct method. In this paper, in section 2, the brief description of the system is explained. In section 3 and 4, the design of self-excitation capacitance and ELC has been developed. In section 5, by using MATLAB software the simulation results of SEIG with an ELC feeding 3Φ consumer (R-L) load and dump load is shown. The simulation for variation in wind speed of SEIG feeding constant 3Φ R-L load with the designed self-excited capacitance is shown.

2 Block Schematic Description

The schematic arrangement of the proposed stand-alone power generation scheme with SEIG is shown in Fig.1. The proposed scheme consists of the 3-Φ IG (3HP, 2.2kW, 380V, 5.5A, 50Hz) driven by a wind turbine, excitation capacitor bank along with ELC, which consists of a 3-Φ diode-bridge rectifier fed chopper circuit with dump load and the 3-Φ R-L load. Suitable values of self-excitation capacitors are connected across the SEIG such that it generates rated terminal voltage at full load. Since, the input power is nearly constant, the output power of the SEIG is held constant at varying consumer loads.

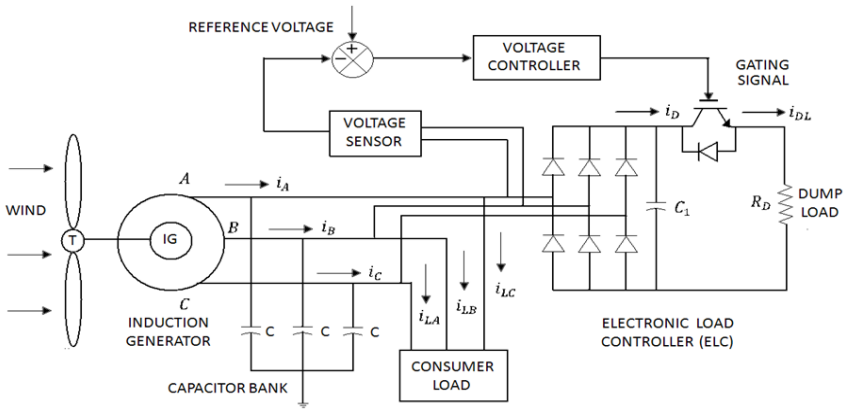


Fig. 1. Schematic diagram of the 3Φ SEIG with an ELC feeding three-phase consumer load

The power in surplus of the consumer load is dumped in a dump load through the ELC. The ELC consists of a 3-Φ uncontrolled rectifier with a chopper in series with a dump load. The ELC with chopper switch has the advantages of which, only one switching device is needed and so the driving circuit is required for the chopper switch alone and so, the scheme is very simple, cheap, rugged and reliable, low value of generation of harmonics and no reactive power demand and requires only one dump load and hence, it is inexpensive and compact. Thus, SEIG feeds two loads in

parallel such that the total power is constant, that is, P_{out}(generated power which should be kept constant)=P_c (consumer load power)+P_d(dump load power).

3 Design of Self-excitation Capacitance for 3-Φ SEIG

The typical parameters for 3-ΦIMsuch as R_s=4.8Ω,R_r=5.77Ω, X_{ls}=X_{lr}=15.66Ω, X_m=230Ω are estimated by conducting no-load test and blocked rotor test experimentally in the laboratory. The load parameters such as the resistance (R=25Ω) and the inductance (L=100mH) are calculated by considering the rated value of voltage and current of the 3-Φ IM. The per phase steady-state circuit of 3-Φ SEIG under R-L load is shown in Fig.2. where R_s, R_r, R, X_{ls}, X_{lr}, X, X_m, X_{smax} = per unit per phase stator, rotor (referred to stator) and load resistance, stator leakage and leakage (referred to stator) load and magnetising reactances (at base frequency) and maximum saturated magnetising reactance, C = per phase terminal-excitation capacitance , X_c = per unit per phase capacitive reactance (at base frequency) of the terminal excitation capacitance , F, v = per unit frequency and speed, N, f_b, Z_b= base speed in rev/min, base frequency, base impedance, V_g,V_0= per phase air-gap voltage and output voltage. Assumptions in the analysis of SEIG is reported in [15] and from Fig.2, the loop equation for the current I is given in (1) and Z is the net loop impedance given by,

$$IZ = 0. \tag{1}$$

$$Z = \left[\left(\frac{R_r}{(F-v)} + jX_{lr}\right)||X_m\right] + \frac{R_s}{F} + jX_{ls} + \left[-\frac{jX_c}{F^2} \parallel \frac{R}{F} + jX\right] \tag{2}$$

Since under steady-state excitation I ≠ 0, it follows from **(1)** that Z=0 or both real and imaginary parts of Z are zeros reported in [15]. For finding general solution for $C_{min,}$ all the coefficients are evaluated at X_m=X_{smax}, where

$$C = {}^{1}/_{2\pi f_b X_c Z_b}. \tag{3}$$

by simplifying (2), the obtained equation is given by

$$\alpha_4 F^4 - \alpha_3 F^3 + \alpha_2 F^2 - \alpha_1 F + \alpha_0 \ = 0. \tag{4}$$

Where,α_i, i=0, 1…, 4 are positive constants is reported in [15]. By solving (4) gives two real roots, the obtained C_{min} = 77μF and C_{max} = 152μF and (4) can be rearranged as

$$\alpha_4 F^4 + \alpha_2 F^2 + \alpha_0 = \alpha_3 F^3 + \alpha_1 F. \tag{5}$$

The two curves intersect in the range $F_{min} \leq F \leq F_{max}$, where $\mathbf{C_{min}}$ is estimated by using MATLAB programming.

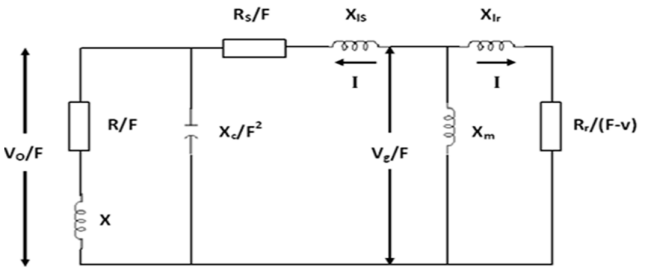


Fig. 2. Per unit per phase steady-state equivalent circuit of self-excited induction generator with R-L load

4 Design of Electronic Load Controller Components

The voltage rating of the uncontrolled rectifier and the chopper switch will be the same and dependent on the rms ac input voltage and the average value of the output dc voltage. The dc output voltage is calculated as,

$$V_{dc} = \frac{3\sqrt{2}V_{LL}}{\pi} = (1.35) * V_{LL} = 513.8V. \tag{6}$$

Where, V_{LL} is the rms value of the SEIG terminal voltage. By considering transient conditions the rms ac input voltage with the peak value is given by,

$$V_{peak} = \sqrt{2} * 10\% \text{ of rated voltage} = 580V \tag{7}$$

The ac input peak current (I_{peak})and the rating of the dump load resistance (R_D) is

$$I_{peak} = 2 * P/(\sqrt{3} * V_{LL} * 0.955) = 7A \tag{8}$$

$$R_D = V_{DC}^{\ 2} / P_{rated} = 119.7\Omega \tag{9}$$

$$C = (1 / 12 * f * R_D)(1 + (1/\sqrt{2} * RF)) = 210.83\mu F \tag{10}$$

where C is dc-link capacitance of the ELC (here RF=5%) is reported in [17].

5 Simulation Results

5.1 The Loss of Excitation and Rise in Voltage Due to Wind Speed Variation with Constant(R-L) Load

The stand-alone wind-driven 3-Φ self-excited induction generator is sensitive to wind speed variations. If the speed is decreased below the threshold of the self-excitation,

voltage at machine terminals will disappear. The simulation result of such a situation of which, variation in wind speed from 7m/s to 7.5m/s of 3-Φ wind-driven SEIG fed to the R-L load is shown in Fig.3(a), no voltage build-up for wind speed 7m/s from 0 to 1.5 seconds and the line current is shown in Fig.3(b). This shows that the designed excitation capacitance value of 77μF fails to excite. The increase in wind speed leads to the increase in generated voltage of SEIG. The Fig.3(c) and Fig.3 (d) shows the generated voltage and line current for increase in speed from 7.5m/s to 10m/s.

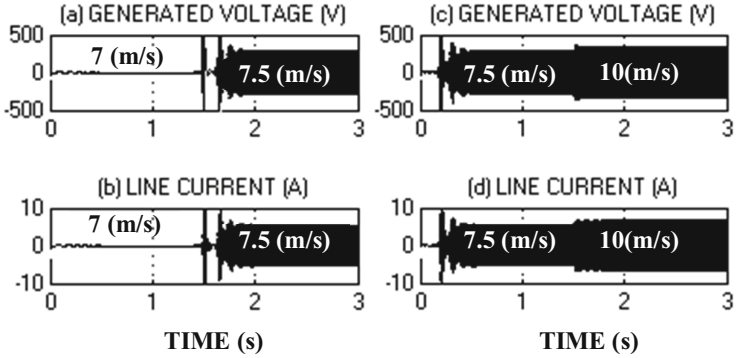


Fig. 3. Simulation results of range of wind speed at constant (R-L) load of SEIG (a) Generated voltage (V) (b) Line current (A) (c) Generated voltage (V) (d) Line current (A)

5.2 Simulation Results for 3-Φ Wind-Driven SEIG at Varying Consumer Load in Closed Loop Scheme

In closed loop control scheme, the generated ac output voltage of SEIG is converted into equivalent dc voltage by using 3-Φ uncontrolled rectifier which is taken as the reference voltage. When the consumer load is varied from rated load, then the increase in generated voltage is converted into equivalent dc voltage and is compared with the reference voltage, where the output obtained from the comparator is a positive value. A suitable PI controller is used and the output of PI controller and the triangular waveform is compared and the gate signal is given to the chopper switch. When the chopper switch is switched on, then the current flows through the dump load and the dump load consumes the difference power (generated power-consumer power). If the comparator output is zero, then the chopper switch is in switched off condition, where SEIG is operating at the rated (R-L) load.

The generated voltage, the line current, reactive power and real power of SEIG are shown in Fig.4(a), Fig.4(b), Fig.4(c), Fig.4(d), where SEIG is operated under rated R-L load from 0 to 2.5 seconds and from 2.5 to 5 seconds SEIG is operated under load variation of (R=25Ω and L=1H) with constant wind speed 10m/s. The gate pulse, dump load voltage and the dump load current of ELC from 2.5 to 5 seconds is shown

in Fig.5(e), Fig.5(f) and Fig.5(g).The Fig.4(a) shows the regulation of voltage at varying R-L load respectively, where the excess voltage which is increased due to R-L load change is fed to the dump load.

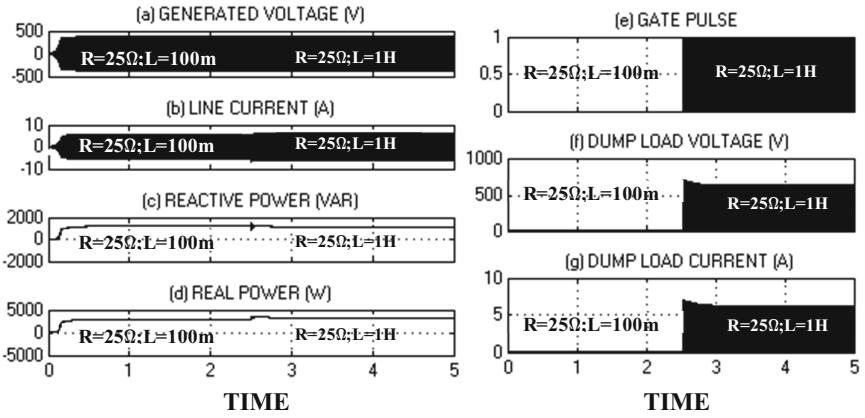


Fig. 4. Simulation results of (R-L)load variation of SEIG at constant wind speed (a) Generated voltage (V) (b) Line current (A) (c) Reactive power (VAR) (d) Real power (W) (e) Gate pulse (f) Dump load voltage (V) (g) Dump load current

6 Conclusions

In this paper, the design and development of controller for stand-alone wind driven SEIG with ELC in closed loop is developed and analyzed. The typical machine and load parameters of the scheme are estimated by the experimental tests undergone in the laboratory and the necessary self-excitation capacitance has been designed mathematically. The design of such a controller in closed loop has been taken up in this work. The range of wind speed variation of the estimated parameters of SEIG and the designed self-excitation capacitance at constant load is carried out. The regulation of voltage under the variation of consumer load at constant wind speed is achieved by the developed controller. The real and reactive power requirement for SEIG with ELC is investigated and all simulation results have been presented.

References

1. Enamul Haque, M., Negnevitsky, M., Kashem Muttaqi, M.: A Novel Control Strategy for a Variable-Speed Wind Turbine With a Permanent-Magnet Synchronous Generator
2. Van Niekerk, H.R.: Permanent magnet alternators for stand-alone electricity generation: Department of Electrical and Electronic Engineering, University of Pretoria, Pretoria, 0002
3. Mittal, R., Sandu, K.S., Jain, D.K.: Isolated Operation of Variable Speed Driven PMSG for Wind Energy Conversion System
4. Three-Phase Self-Excited Induction Generators, An Overview

5. Ofualagba, G., Ubeku, E.U.: The Analysis and Modelling of a Self-excited Induction Generator Driven by a Variable Speed Wind Turbine. Federal University of Petroleum Resources, Effurun, Nigeria
6. European Wind Energy Association (EWEA) and Greenpeace, WIND FORCE 12 Report (2004)
7. American Wind Energy Association (AWEA), http://www.awea.org
8. Bansal, R.C., Bhatti, T.S., Kothari, D.P.: A bibliographical survey on induction generators for application of nonconventional energy systems. IEEE Trans. Energy Convers. 18(3), 433–439 (2003)
9. Khan, P.K.S., Chatterjee, J.K.: Three-phase induction generators: A discussion on performance. Elect. Mach. Power Syst. 27, 813–832 (1998)
10. Rajakaruna, S., Bonert, R.: A technique for the steady-state analysis of a self-excited induction generator with variable speed. IEEE Trans. Energy Convers. 8(4), 757–761 (1993)
11. Levi, E., Liao, Y.W.: An experimental investigation of self-excitation in capacitor excited induction generators
12. Grantham, C., Sutanto, D.: Steady-state and transient analysis of self-excited induction generators
13. Malik, N.H., Mazi, A.A.: Capacitance requirements for isolated self-excited induction generators
14. Malik, N.H., Haque, S.E.: Steady state analysis and performance of an isolated self-excited induction generator
15. Jabri, A.K., Alolah, A.I.: Capacitance requirement for isolated self-excited induction generator
16. Juan Ramirez, M., Emmanuel Torres, M.: Electronic Load Controller for the Self-Excited Induction Generator
17. Singh, B., Murthy, S., Gupta, S.: Analysis and implementation of an electronic load controller for a self-excited induction generator. Inst. Electr. Eng. Proc.-Gener. Transm. Distrib. 151(1), 51–60 (2004)
18. Rajakaruna, S., Bonert, R.: A technique for the steady-state analysis of a self-excited induction generator with variable speed. IEEE Trans. Energy Convers. 8(4), 757–761 (1993)

A New High Step-Up Converter with Low Voltage Stress Using Coupled Inductor

K. Radhalakshmi, R. Dhanasekaran, S. Aiswarya, and S. Muthulakshmi

Syed Ammal Engineering College, Ramanathapuram -623502,
Tamilnadu, India
radhak10@yahoo.co.in, rdhanashekar@yahoo.com

Abstract. The Conventional Boost Converters are not able to provide voltage gain, A high – efficiency high Step-up converter is proposed, with low voltage stress on power switch. The circuit topology of the proposed converter consists of coupled inductor and the voltage lift technique to achieve high voltage gain. Additionally the voltage stress on power switch is clamped, and the energy stored in the leakage inductor is recycled in the proposed converter. Therefore the voltage stress on the active switch is reduced, and the conversion efficiency is improved.

Keywords: Step-Up converter, Soft switching, Coupled Inductor, Modified Non dissipative Snubber.

1 Introduction

Recently, DC-DC Converters with steep voltage ratio are usually required in many industrial applications. For example ,the front-end stage for clean- energy sources, the dc back-up energy system for an Uninterruptible Power Supply (UPS) and ,the telecommunications industry .The conventional Boost converters cannot provide such a high dc voltage gain ,even for an extreme duty cycle. In order to increase the conversion efficiency and voltage gain ,many boost converter topologies have been investigated in the past decade [1]-[6].Although voltage-clamped techniques are manipulated in the converter design to overcome the severe reverse-recovery problem of the output diode in high-level voltage applications, there still exists overlarge switch voltage stresses [1]. Zhao and Lee [2] introduced a family of high-efficiency, high step-up dc–dc converter. It can recycle the leakage energy and alleviate the reverse-recovery problem. Dumrongkittigule et al. [3] presented a high step-up DC–DC converter with integrated coupled inductor and common-mode electromagnetic interference reduction filter.Hyun-Lark Do [4] presented converter Consists of ZVS Boost Converter stage and a ZVS Half- Bridge converter stage.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 305–310, 2013.

2 Conventional Topology

2.1 Conventional Boost/Fly Back Converter with Coupled Inductor

By considering the coupled inductor, the boost converter (I) circuit, as shown in Fig.1.

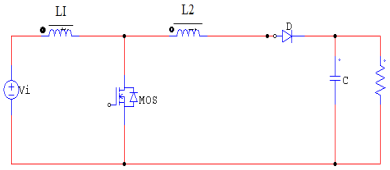


Fig. 1. Converter with coupled inductor

The voltage gain of the converter is

$$Gv = \frac{Vo}{Vs} = \frac{n}{1 - D} \tag{1}$$

Although a Conventional converter achieves high voltage step-up, it suffers from high voltage spike stress on the switching devices .Low efficiency occurs because of the leakage inductor and reverse recovery problems of the diode.

3 Proposed High Step-Up Converter

Fig. 2. Shows the proposed High step-up converter with voltage –lift technique and a coupled inductor. The diode D_1 is turned on when switch S is turned- off, then the voltage across the switch is clamped at a low voltage level, and the energy stored in the leakage inductance is recycled into C_1. Because the voltage across C_2 is constant, and voltage gain to be increased.Fig.3. Shows the typical waveform of the continuous conduction mode of converter.

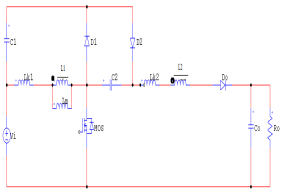


Fig. 2. Proposed High Step-up converter

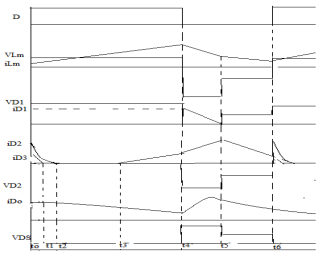
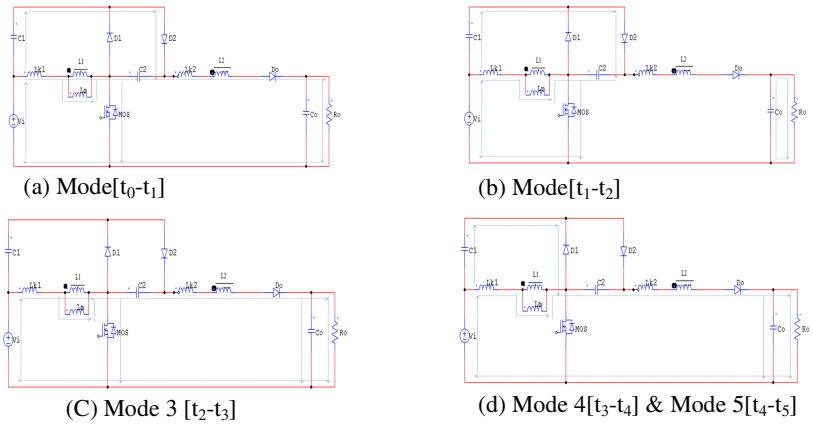


Fig. 3. Some typical waveforms under CCM operation

3.1 Continuous Conduction Mode

The operating mode of principle of converter under CCM is described. Fig.4. Shows mode of converter under CCM operation in the one switching period.



(a) Mode[t_0-t_1] (b) Mode[t_1-t_2]

(C) Mode 3 [t_2-t_3] (d) Mode 4[t_3-t_4] & Mode 5[t_4-t_5]

Fig. 4. Operating modes for CCM Operation

(a) Mode 1 [t_0-t_1]

At t=t_0, the switch S is turned on, diodes D_2 and D_3 are turned on and D_1 is turned off. Because the energy of Cs is discharged quickly. The voltage across C_2 is charged to Vs +V_{c1} as shown in Fig.4. (a)

(b) Mode 2 [t_1-t_2]

In this mode switch S is turned on, D_1 and D_3 are turned off and diode D_2 is turned on as shown in Fig.4. (b) . The inductor L_m and L_{k1} are charged from input voltage, the currents i_{Lm} and i_{Lk2} are equal and increased linearly with a slope of V_s/L_m. The capacitor C_2 is charged to V_S+V_{c1}.The load is supplied by C_0 at this mode.

(C) Mode 3 [t_2-t_3]

At t=t2, S is turned on, Diodes D1 & D2 is turned off and D3 is turned on .Fig.4. (c) shows the equivalent circuit of this mode.

(d)Mode 4 [t_3-t_4]

In this mode switch S is turned off, the diodes D_1 & D_3 are turned on and D_2 is turned off as shown in Fig. 4. (d). The energy stored in L_m and L_{K1} is released to capacitor C_1 and C_O.

(e) Mode 5 [t_4-t_5]

At this mode, V_{DS} =V_S+V_{C1}.S is still turned off. The equivalent circuit and operation of this mode is same as Fig.4. (d). In this interval, i_{D3} decreased linearly.

When the switch (Q) is turned on, the voltage across the magnetizing inductor (L_m) and V_{l2} are written as

$$VLm = Vs \tag{3}$$

$$VL\ 2 = nVLm = nVs \tag{4}$$

When the switch (Q) is turned off,

$$VLm = -Vc\ 1 \tag{5}$$

$$VL\ 2 = nVLm = -nVc\ 1 = Vc\ 2 + Vc\ 1 + Vs - Vo$$

Where the voltage across the capacitor (C_2) can be written as

$$Vc\ 2 = Vc\ 1 + Vs \tag{6}$$

By using voltage-second balance principle on N1 and N2 of the coupled inductor, the following Equations are written as

$$\int_0^{DT} Vs\,dt + \int_{DT}^{T} -Vc\,1\,dt = 0 \tag{7}$$

$$\int_T^{DT} nVs\,dt + \int_{DT}^{T} ((2Vc\,1 + Vs) - Vo)\,dt = 0 \tag{8}$$

From equation (7) and (8), VC1 and V_0 are derived as

$$Vc\ 1 = \frac{D}{1 - D} Vs \tag{9}$$

$$Vo = \frac{(2 + nD)}{1 - D} Vs \tag{10}$$

As a result, the voltage gain of the step-up converter can be represented as

$$Gv = \frac{(2 + nD)}{1 - D} \tag{11}$$

Equation (11) indicates that the proposed converter accomplishes a high voltage gain by using voltage –lift technique and increasing the turn's ratio of the coupled inductor.

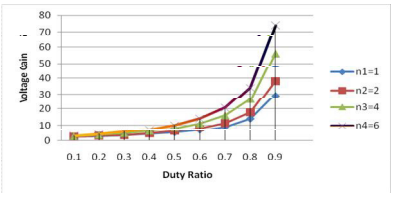


Fig. 5. Voltage gain against duty ratio under various turns ratio of the coupled inductor

3.2 Simulation

Fig.6. shows the simulation waveforms at CCM, which were measured to verify the performance of the proposed converter under output power Po=35w.this result indicate that the duty ratio must be raised to maintain a constant output voltage when output power is increased. The waveform agrees with the steady state analysis, and voltage stress of switch is effectively clamped by diode D_1 and capacitor C_1.Fig.7.shows illustrates the input current i_{Lm} in the primary side of coupled inductor. The input current i_1 is equal to the magnetizing current i_m is linearly charged by the input voltage Vs.

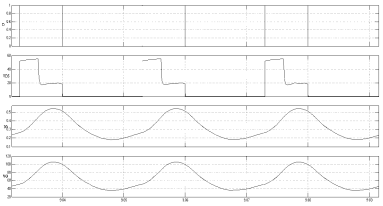


Fig. 6. Simulation of proposed Converter with V_O=100V and Po=35W

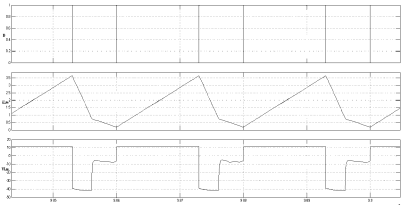


Fig. 7. i_{Lm} and V_{Lm} of proposed converter

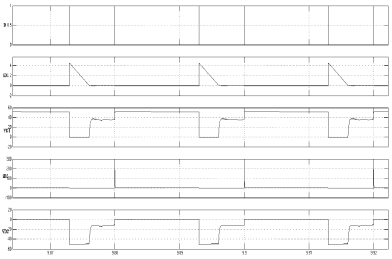


Fig. 8. i_{D1}, $V_{D1,}$ i_{D2} and V_{D2} of proposed converter

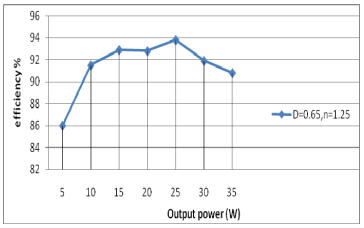


Fig. 9. Conversion efficiency of the proposed converter at different output powers

Fig.9. shows the experimental conversion efficiency of the proposed converter at condition D=0.65, n= 1.25. When D and n are equal to 0.42 and 6.67, the efficiency is 90.8% at full load and the maximum efficiency is 92.8% at Po = 20W.

Owing to the voltage stress on the switch during the switch turn- off period, the modified non-dissipative snubber is used. A new operating condition for the non-dissipative LC snubber is proposed in Fig.10.which is shown to reduce the voltage stress on the device to a great extent. In this mode, though the current stress on the device shall be more than that of the traditional RC snubber, it is shown that such stress can be reduced to a considerable extent by proper design of the snubber inductor.

Here, the frequency of oscillation between Cs and Ls is chosen to be much higher than the switching frequency. At the time of turn-on, the capacitor Cs is discharged through the loop of Ds, Ls and Q1 as shown in Fig. 10. Fig.11. Shows that, the voltage stress of the switch is reduced compare to proposed converter.

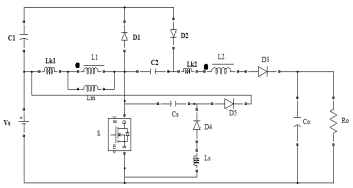


Fig. 10. Proposed Converter with Modified Non Dissipative Snubber

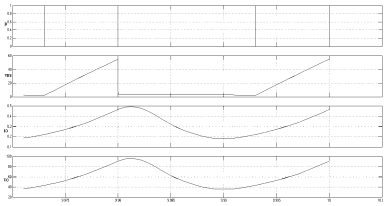


Fig. 11. Simulation of Proposed Converter with Non Dissipative Snubber

4 Conclusion

This paper presents step-up DC–DC converters that uses voltage-lift technique, adjusts the turn's ratio of the coupled inductor and achieve high step-up voltage gain. The proposed converter is highly efficient because it recycles the energy stored in the leakage inductor of the coupled inductor. Moreover, the voltage across the switch is clamped at the lower voltage level, enabling the converter to use the low rating switch to improve efficiency. This high-efficiency converter topology provides designers with an alternative choice to convert renewable energy efficiently, and it also can be extended easily to other power conversion systems for satisfying high-voltage demands. The proposed converter is appropriate for applications involving power conversion systems, such as small fuel-cell power conversion systems and small solar-cell power conversion systems.

References

1. Duarte, C.M.C., Barbi, I.: An Improved Family of ZVS-PWM Active Clamping DC-DC Converters. IEEE Trans. Power Electronics 17, 1–7 (2002)
2. Dumrongkittigule, T., Tarateeraseth, V., Khanngern, W.: Techniques for Commom Mode EMI Reduction in High Step-Up DC/DC Converter. In: Proc. IEEE 17th Int. Zurich Symp. on Electromagnetic Compatibility, pp. 541–544 (2006)
3. Do, H.-L.: A Zero-Voltage –Switching DC-DC Converter with High Voltage Gain. IEEE Trans. on Power Electronics 26(5) (2011)
4. Li, W., Zhao, Y., Deng, Y., He, X.: Interleaved Converter With Voltage Multiplier Cell for High Step-Up and High –Efficiency Conversion. IEEE Trans. on Power Electronics 25(29) (2010)
5. Kwon, J.-M., Kwon, B.-H.: High Step-Up Active –Clamp Converter With Input –Current Doubler and Output –Voltage Doubler for Fuel Cell Power Systems. IEEE Trans. on Power Electronics 24(1) (2009)
6. Dwari, S., Parsa, L.: An Efficient High Step-Up Interleaved DC-DC Converter with a Common Active Clamp. IEEE Trans. on Power Electronics 26(1) (2011)

Line Contingency Ranking Based on Detection of Accidental Islands for Autonomous Micro-grids

Pramod Kumar Muppidi and M. Venkata Kirthiga

Department of Electrical and Electronics Engineering
National Institute of Technology Tirchirappalli
Pramodkumar.muppidi@gmail.com, mvkirthiga@nitt.edu

Abstract. Introduction of Distributed Generator(DG) units in conventional distribution systems results in restructuring of the system. DG units impact line losses, bus voltage profile and stability of the existing system. This paper focuses on accidental island detection during line contingencies. This has been attempted by tracing the change in the configuration of the autonomous micro-grids. A comparison is done for radial and weakly meshed structures based on the voltage profile, indices, line losses and transient stability of the DGs on implementation of the proposed algorithm. Ranking of line contingencies is done with the database of the detected islands. MATLAB coding is developed to validate the proposed algorithm & the standard 33-node radial distribution system is considered for justification of the same.

Keywords: Intentional islands, accidental islands, autonomous micro-grids.

1 Introduction

Distributed Energy Resources (DER) are integrated an existing system to meet out the increasing demand. Distributed Generation brings many system and DG related advantages such as unloading transmission system, decreasing system losses, improving power quality and reliability. As deregulated operation of the distribution systems gains importance, the necessity towards optimizing the architecture of the system becomes inevitable for better service to consumers, achieving load balancing and loss reduction.

Active distribution network management improves the efficiency and reliability and hence alternatives like autonomously operated micro-grids are coming up. Island detection becomes important to ensure stability and reliability of micro-grids. A micro-grid[2] is an “Interconnection of small, modular generation units to low voltage distribution systems forming a self-sustained utility on isolation from main grid”.

2 Voltage Stability Indices

Two indices viz., NODAL VSI and LINE VSI [1] are used in this paper to determine the point of voltage instability as shown below:

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 311–316, 2013.

$$\text{Fast voltage stability index (FVSI)} = (4Z^2_{ij}Q_j / V_i^2 X_{ij}) \quad (1)$$

$$\text{Line stability factor (LQP)} = [4(X_{ij}/ V_i^2)(X_{ij} P_i^2/ V_i^2 + Q_j)] \quad (2)$$

Voltage collapse proximity index (VCPI)

$$VCPI_j(1) = Pj/Pj_{(max)} \quad \& \quad VCPI_j(2) = Qj/Qj_{(max)} \quad (3) \ \& \ (4)$$

where V_i, V_j = Voltage on sending and receiving buses
P_i, Q_i = Active and reactive power on the sending bus
P_j, Q_j = Active and reactive power on the receiving bus
S_i, S_j = Apparent power on the sending and receiving buses
$\delta_i - \delta_j = \delta_{ij}$, angle difference between sending & receiving bus voltages

All the above indices are expected to be less than unity for a stable system.

3 Island Detection and Line Contingency Analysis

An island[3] is "That part of a power system consisting of one or more power sources and load, for some period of time, separated from the rest of the system." Loss of Mains protection (LOM), loss of grid protection, anti-islanding protection and islanding protection are synonyms used for Island detection schemes.

The contingencies cause adverse effects on micro-grids where more DG units are connected. Hence when a part of the system forms an accidental island there is a risk for Brown out or Black out. Formation of accidental islands is more prominent in radial distribution systems but in interconnected systems this does not occur frequently. Other hazardous events viz., Line outage, Generator outage & Mal-operation of relays etc., are expected in a any autonomous system and hence island detection gains high priority in stability studies of micro-grids.

4 Proposed Algorithm

Ranking algorithm and island detection algorithm proposed in this paper are explained in this section.

4.1 Overall Algorithm Proposed

This algorithm proposes a scheme for ranking the line contingencies in an autonomous micro-grid. The proposed algorithm is unveiled in the following steps:-

1. Load flow analysis for base case of operation has been performed for Radial and Weakly meshed autonomous micro-grids and the indices are computed
2. A line is considered to undergo an outage (line no=1) and the architecture is traced for formation of islands.
3. Voltage indices, transient stability of DGs and bus/load isolation are checked

4. The above steps are repeated for all other lines to undergo outage.
5. The number of islands formed and the total load shed are tabulated to form database determining the complete information about the effects of line outages.

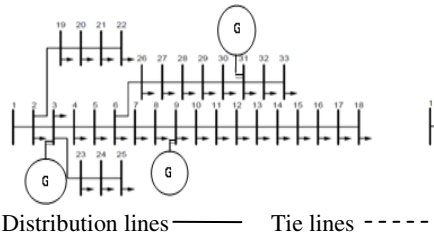
4.2 Proposed Algorithm for Detection of Islands

The main steps of island detection are as follows:-

1. A line contingency has been created between any two buses 'p' and 'q'.
2. Existence of closed loop from to-bus (q) to the from-bus (p) for the above line outage is traced. If true go to next step else skip the step 3 and go to step4.
3. The line data alone is updated, as no islands is formed and go to step5
4. Both the line & bus data are updated as islands have formed, and go to next step.
5. Newton Raphson load flow has been adopted for the updated data for both re-configured and the radial systems.
6. The voltage deviation is calculated as $\Delta V = V_{max}-V_{min}$
7. Results are tabulated and the line contingencies are ranked based on the minimum losses and voltage deviation calculated in step 5& 6.

5 Test System Details

The single line diagram of the autonomous micro-grid and reconfigured micro-grid considered in this work are shown in Fig. 2&3.



Distribution lines ——— Tie lines - - - - -

Fig. 1. Radial Autonomous Micro-grid

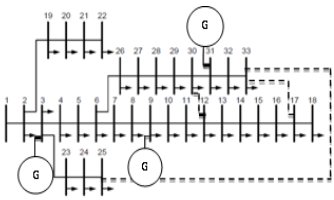


Fig. 2. Weakly meshed Autonomous Micro-grid

6 Results and Discussions

The standard 33 bus distribution system is transformed into a micro-grid and operated both in radial and reconfigured architectures.

6.1 Load Flow Results

The voltage profiles of the radial and reconfigured micro-grids are shown in table 1.

Table 1. Comparison of bus voltages for Radial and Weakly Meshed autonomous micro-grids

Bus No.	Bus voltage in pu Radial micro-grid	Bus voltage in pu Weakly meshed micro-grid	Bus No	Bus voltage in pu Radial micro-grid	Bus voltage in pu Weakly meshed micro-grid	Bus No	Bus voltage in pu Radial micro-grid	Bus voltage in pu Weakly meshed micro-grid	Bus No	Bus voltage in pu Radial micro-grid	Bus voltage in pu Weakly meshed micro-grid
1	0.9982	0.9982	10	0.9946	0.9940	19	0.9977	0.9977	28	0.9946	0.9911
2	0.9982	0.9982	11	0.9938	0.9933	20	0.9942	0.9942	29	0.9942	0.9898
3	1.0000	1.0000	12	0.9924	0.9921	21	0.9935	0.9935	30	0.9949	0.9898
4	0.9989	0.9985	13	0.9868	0.9887	22	0.9928	0.9928	31	1.0001	0.9927
5	0.9981	0.9973	14	0.9848	0.9876	23	0.9965	0.9970	32	0.9993	0.9911
6	0.9959	0.9944	15	0.9835	0.9872	24	0.9900	0.9913	33	0.9990	0.9900
7	0.9958	0.9942	16	0.9822	0.9870	25	0.9868	0.9889			
8	0.9967	0.9954	17	0.9804	0.9875	26	0.9957	0.9939			
9	0.9999	0.999	18	0.9798	0.9869	27	0.9956	0.9934			

6.2 Island Details for Weakly Meshed Autonomous Micro-grid on Line Contingency

Detection of accidental islands [3] and the load shed for the reconfigured structure of the micro-grid is done for the line outages at every location as shown in the table 2.

Table 2. Island detection for reconfigured micro-gird on occurrence of line contingency

Faul--ted Line	Load Shed/ Island	Isola--ted buses	V_{max} (pu)	V_{min} (pu)	P_{loss} (MW)	Faul--ted Line	Load Shed/ Island	Isola--ted buses	V_{max} (pu)	V_{min} (pu)	P_{loss} (MW)
1	Loadshed	1	1	0.98692	0.01940	18	Loadshed	22,21 20,19	1	0.98692	0.01764
2	Loadshed	2,1,20, 19,18,17	1	0.98692	0.01760	19	Loadshed	22,21,20	1	0.98692	0.01775
3	*		1	0.95634	0.04034	20	Loadshed	22,21	1	0.98692	0.01802
4	*		1	0.96573	0.03208	21	Loadshed	22	1	0.98692	0.01855
5	*		1	0.96942	0.02901	22	*		1	0.93920	0.05889
6	*		1	0.98691	0.01943	23	*		1	0.94572	0.05038
7	*		1.00653	0.98975	0.02146	24	*		1	0.97148	0.02501
8	*		1.01392	0.99038	0.02738	25	*		1	0.97887	0.02420
9	*		1.01107	0.97608	0.03522	26	*		1	0.98136	0.02271
10	*		1.00992	0.97714	0.03234	27	*		1	0.98373	0.02150
11	*		1.00887	0.97813	0.03000	28	*		1	0.98455	0.02061
12	*		1.00133	0.97426	0.02307	29	*		1	0.98625	0.01951
13	*		1.00077	0.97774	0.02142	30	*		1.001	0.98428	0.02293
14	*		1	0.98431	0.01949	31	*		1.007	0.97970	0.03145
15	*		1	0.98630	0.01934	32	*		1.002	0.98338	0.02420
16	*		1	0.98575	0.01947	33	*		1	0.97729	0.02137
17	Loadshed	18	1	0.98902	0.01765	34	*		1.002	0.98681	0.02009
						35	*		1.001	0.98728	0.02020

* - Indicates that there occurs no change in the system.

7 Transient Stability

Transient stability [4] – [5] of the DGs on fault occurrence is most inevitable for evaluating the system performance and stability. The power angle curves for the three generators numbered as G1, G2& G3 for all line contingencies are determined. Line fault is created at time t = 0 sec. and cleared at t= 0.1 sec. where the clearance of fault includes removal of faulty line from the system. The least impacting and worst affecting two line outages upon the generator stability are shown below in the figures.

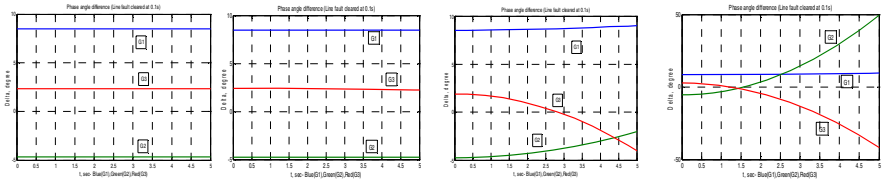


Fig. 4. Power angle curves of two stable and two unstable cases respectively

8 Ranking of Line Outages

As an attempt to create a database for sequence of events of line outage for the micro-grid under autonomous operation, all the line outages are ranked based on the system losses, voltage variations and number of islands as well as isolated buses formed as shown in table 3. The worst affecting contingency is ranked top.

Table 3. Ranking of contingencies for radial and reconfigured micro-girds

	Weakly meshed autonomous micro-grid						Radial autonomous micro-grid					
Rank	Line No.	ΔV (p.u)	Line No	P_{loss} (MW)	Line No	No. of Isolated Lines	Line No	ΔV (p.u)	Line No	P_{loss} (MW)	Line No	No. of Isolated lines
1	17	0.0110	2	0.0176	3	-	3	0.0132	3	0.0105	1	1
2	1	0.0131	18	0.0176	4	-	4	0.0132	4	0.0105	17	1
3	2	0.0131	17	0.0177	5	-	5	0.0132	5	0.0106	21	1
4	18	0.0131	19	0.0178	6	-	7	0.0132	6	0.0142	24	1
5	19	0.0131	20	0.018	7	-	8	0.0132	22	0.0143	32	1
6	20	0.0131	21	0.0186	8	-	6	0.0137	23	0.0143	16	2
7	21	0.0131	15	0.0193	9	-	17	0.0157	7	0.016	20	2
8	6	0.0131	1	0.0194	10	-	16	0.0174	24	0.0163	23	2
9	35	0.0134	6	0.0194	11	-	15	0.019	14	0.0191	31	2
10	15	0.0137	16	0.0195	12	-	25	0.0201	8	0.0191	15	3
11	29	0.0137	14	0.0195	13	-	26	0.0201	5	0.0192	19	3
12	16	0.0142	29	0.0195	14	-	27	0.0201	25	0.0194	22	3
13	34	0.0144	34	0.0201	15	-	28	0.0202	26	0.0196	30	3
14	28	0.0154	35	0.0202	16	-	1	0.0203	16	0.0197	14	4
15	14	0.0157	28	0.0206	22	-	2	0.0203	13	0.0199	18	4
16	27	0.0163	33	0.0214	23	-	18	0.0203	27	0.02	29	4
17	30	0.0166	13	0.0214	24	-	19	0.0203	28	0.0204	13	5

Table 3. (*continued*)

18	7	0.0168	7	0.0215	25	-	20	0.0203	17	0.0206	28	5
19	32	0.0185	27	0.0215	26	-	21	0.0203	12	0.021	12	6
20	26	0.0186	26	0.0227	27	-	22	0.0203	2	0.0212	27	6
21	25	0.0211	30	0.0229	28	-	23	0.0203	18	0.0212	2	6
22	33	0.0227	12	0.0231	29	-	24	0.0203	19	0.0213	11	7
23	13	0.023	32	0.0242	30	-	14	0.0205	20	0.0216	26	7
24	8	0.0235	25	0.0242	31	-	29	0.0218	29	0.0221	10	8
25	12	0.0271	24	0.025	32	-	32	0.0225	21	0.0221	25	8
26	31	0.0276	8	0.0274	33	-	13	0.0245	11	0.0223	9	9
27	24	0.0285	5	0.0290	34	-	12	0.0264	1	0.023	8	10
28	5	0.0306	11	0.0300	35	-	30	0.0278	32	0.023	7	11
29	11	0.0307	31	0.0315	1	1	11	0.0283	10	0.0236	6	12
30	10	0.0328	4	0.0321	17	1	31	0.0295	9	0.0252	5	21
31	4	0.0343	10	0.0323	21	1	10	0.0297	31	0.0255	4	22
32	9	0.0350	9	0.0352	20	2	9	0.0314	30	0.0368	3	23
33	3	0.0437	3	0.0403	19	3						
34	23	0.0543	23	0.504	18	4						
35	22	0.0608	22	0.0589	2	6						

9 Conclusions

This paper focused on detection of accidental islands formed in an autonomous micro-grid on line contingencies. This identification of formation of accidental islands has been utilized to rank the line contingencies in an attempt to maintain a database of the consequences of different line contingencies. The standard 33 bus distribution system has been transformed into an autonomous micro-grid in this work and both radial and reconfigured structures of the system have been considered for analysis. It is seen a weakly meshed architecture is best suited for an autonomously operated micro-grid from the voltage profile, indices and stability of the DGs. Also the weakly meshed structure is less vulnerable to accidental islands during line contingencies.

References

1. Suganyadevia, M.V., Babulal, C.K.: Estimating of Loadability Margin of a Power System by comparing Voltage Stability Indices. In: International Conference on Control Automation, Communication and Energy Conservation, pp. 1–4 (2009)
2. Gurunathan, S., Prakashkumar, V., Venkatakirthiga, M.: Optimal reconfiguration of Distribution system based on Ranking of buses. In: CDT National Conference on Efficient Energy Management in Urban and Rural Sector, New Delhi (2011)
3. Venkata Kirthiga, M., Arul Daniel, S.: Performance indices based controlled Islanding of Power system. In: Fourteenth National Power System Conference (NPSC 2006) (December 2006)
4. Barker, P., DeMello, R.W.: Determining the impact of DG on power system, radial distribution. In: Proc. IEEE Power Eng. Soc. Summer Meeting, pp. 1645–1656 (2000)
5. Baaran, M.E., Marakby, I.M.E.: Fault analysis on distribution feeders with distributed generators. J. IEEE Transactions on Power Systems 20(4), 1757–1764 (2005)

Investigations on Modern Self-defined Extinction Advance Angle Controller for CCC Based HVDC Systems

M. Rajasekaran and M. Venkata Kirthiga

Department of Electrical and Electronics Engineering,
National Institute of Technology Tiruchirappalli, India
rajasekarancareer@gmail.com, mvkirthiga@nitt.edu

Abstract. A Modern Self-defined Extinction Advance Angle Controller is proposed in this paper and suitability of the proposed controller for a Capacitor Commutated Converter(CCC) based HVDC system is investigated. The proposed controller is dedicated at the inverter end and hence the rectifier end converter is endowed with the standard CIGRE controller. The performance of the proposed controller is investigated under different operating conditions including post fault recovery and fault mitigation. The suitability of the proposed controller for a CCC HVDC system has been checked. CIGRE benchmark controller is adopted for comparison and it has been found that the proposed controller outperforms at post fault operating conditions. PSCAD/EMTDC simulation software has been used for justification of results.

Keywords: CCC HVDC system, Self-Defined Controller, CIGRE controller.

1 Introduction

High Voltage DC(HVDC) transmission is surpassing High Voltage AC transmission systems to cope with the increasing energy demand. As thyristors are used in conventional HVDC converters, the effectiveness of commutation of these devices is decided by the strength of the AC grid. A HVDC-topology, utilizing series capacitors for developing the voltage required for thyristor valve commutation was introduced in 1954 by Buseman [1]-[2].

The significant feature of HVDC transmission is the rapid controllability of transmitted power achieved by three inevitable controllers viz., Constant Current Controller(CC) at both the ends (Rectifier and Inverter ends), Constant Ignition Angle (CIA) Controller at the Rectifier end and Constant Extinction Advance Angle (CEA)Controller at the Inverter end. All the above mentioned controllers are assigned to control the firing delay angle of the switching devices taking into account the dc link voltage, device current and the ac network voltage for different control requirements [3].

A Modern Self – Defined Extinction Advance Angle Controller [4] is proposed in this paper for CCC based HVDC systems and its performance has been investigated for both steady state operation and for a DC link to ground fault. The behaviour of the

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 317–323, 2013.

proposed controller is compared with the standard CIGRE benchmark controller [5] under different operating conditions. A fault current mitigation technique is also introduced in this work, and the performance analysis has been discussed in detail.

2 Capacitor Commutated Converter

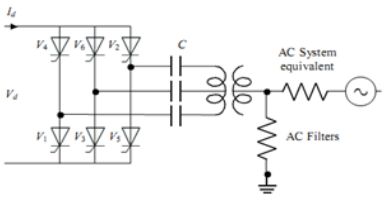


Fig. 1. Configuration of the CCC

A Capacitor Commutated Converter (CCC) is a modified conventional HVDC converter with additional capacitors between the transformer and valves to enhance smooth commutation, as shown in Fig. 1. This configuration is capable of supplying the reactive power requirement of a HVDC transmission system compared to the conventional Line Commutated Converter (LCC). Further, CCC gives dynamically stable performance at the inverter end, particularly when inverters are connected to weak AC systems and/or long DC cables. Capacitor in CCC configuration increases the commutation margin without any increase in the reactive power consumption of the converter station[2]. This improvement in commutation margin reduces the chances of commutation failure and hence CCC is dedicated to the inverter station, prone to commutation failure.

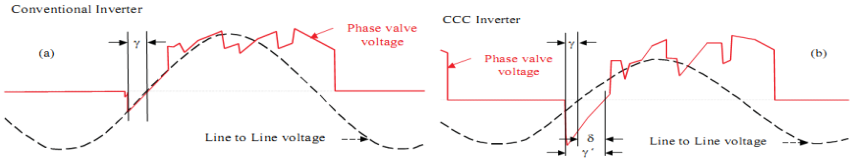


Fig. 2. AC bus voltage & valve voltage: (a) conventional inverter, (b) CCC inverter

As a matter of fact, it becomes inevitable to distinguish between the apparent extinction angle (γ′) and the real extinction angle (γ) as shown in Fig. 2. In the conventional HVDC converter, the extinction angle γ is defined as the electrical angle spanned between the instant at which the valve turns off and the positive zero-crossing of the line-to-line voltage at the ac converter bus, as given by (1)

$$\gamma = \prod - (\alpha + \mu) \tag{1}$$

where α is the inverter firing delay angle and μ is the overlap angle. However, in the case of the CCC this measurement does not take into account the capacitor voltage and does not measure the actual extinction angle. Therefore it is referred to as the apparent extinction angle (γ′) as given by (2)

$$\gamma = \prod - (\alpha + \mu) + \delta \tag{2}$$

where δ is the phase-lag angle between the AC bus voltage and valve voltage as shown in Fig. 2. The commutation margin angle γ′ in the CCC inverter is the angle between the instant of completion of commutation and the instant where the valve voltage crosses zero and grows positive there onwards.

3 Combined Control Characteristics of HVDC System

The above mentioned controllers are implemented at appropriate locations for effective and fast control in HVDC systems. A combined control operation is required to enable both forward and reversal of power flows and the characteristic of the converters is shown in figure 3. Under emergency control states viz., Reversal of power flow for control requirements, inversion operation of both converters to de-energize the line completely, emergency shutdown of the line etc., it is necessary to avail the CC and CIA controllers at the inverter end and CEA at the rectifier end. Thus each converter needs to be provided with all the three above mentioned controllers.

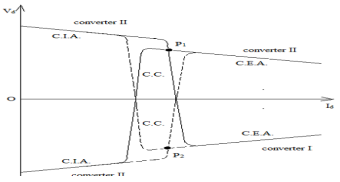


Fig. 3. Combined Control Characteristics

4 Modern Extinction Advance Angle Controller

As mentioned in the previous section, an effective CEA controller would help reduction in frequent commutation failures. In this paper a modern extinction advance angle controller [6], [7] is proposed, suitable for the inverter end station. The proposed controller acts at the device level and hence faster control on extinction angle is made possible. The schematic of the proposed controller is shown in Fig.4.

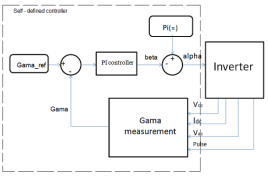


Fig. 4. Schematic of the proposed controller

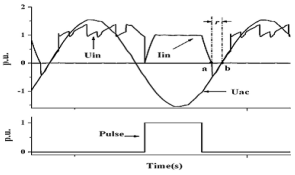


Fig. 5. Voltage waveform to measure γ

After the outgoing valve has completely de-ionized, when the voltage across the valve and AC commutating voltage pass through zero, the valve would exit conduction. This instant is denoted as node 'b' in Fig. 5. Hence, with firing pulse, zero-crossing instant of valve current and AC voltage of each valve as input signals, the nodes 'a' and 'b' are identified by the proposed controller and a unit time-pulse, the width of which is equal to the duration of extinction angle γ is generated. This time-pulse is fed to an integrator and the integrator generated an amplitude-pulse proportional to the extinction angle γ. This amplitude of the pulse is maintained until

the next extinction angle γ is measured. Through this method, the extinction angle γ of single valve in each cycle is measured and hence the same is maintained constant throughout the inverter operation of the bridge concerned.

The inputs of the measurement module are voltage, current, firing pulse and AC voltage of each valve and the outputs are measured extinction angle γ of each valve. The minimum γ value measured is selected as the input of closed-loop control of constant extinction angle control. This control scheme has been implemented using PSCAD /EMTDC software and has been shown in Fig. 6.

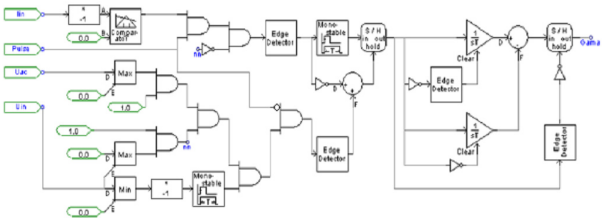


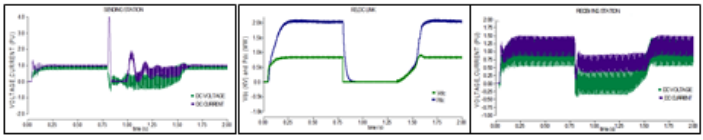
Fig. 6. Overall simulation diagram of the modern self-defined extinction angel controller

5 Results and Discussions

The efficacy of the proposed controller has been investigated for the standard CIGRE benchmark system with power rating of 2000MW, at 1100kV bipolar system with two six pulse bridges cascaded at each (rectifier and inverter) end of the link. The start-up characteristics of the system with the proposed controller are found to be satisfactory.

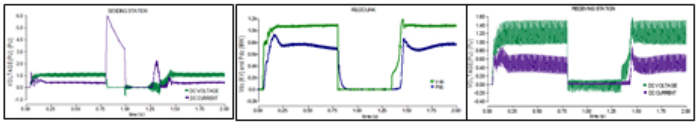
5.1 Response of the Proposed Controller for Post-fault Recovery

A DC link to ground fault has been simulated to occur at 0.8 s and allowed to persist for 600 ms. The efficacy of the proposed controller has been analyzed and the performance curves are shown in the Fig. 7 and Fig. 8.



a)Sending end station (V_{dc},I_{dc}) b) Receiving end station – P_{dc}, V_{dc} and I_{dc}

Fig. 7. Post fault recovery performance of the CIGRE controller at CCC HVDC system



a) Sending end station (V_{dc},I_{dc}) b) Receiving end station – P_{dc}, V_{dc} and I_{dc}

Fig. 8. Post fault recovery performance of the proposed controller at CCC HVDC system

Comparing the curves shown in Fig 7 a) and Fig. 8a) it is clear that the post fault recovery is faster and smoother when the proposed controller is implemented and the settling time is found to be just 102ms. Comparing the curves shown in Fig 7 b) and Fig. 8 b) it is evident that the DC link power, voltage and the inverter (receiving) end current recover back to the pre-fault values faster and hence resulting in an improved stability margin compared to the standard CIGRE benchmark controller.

5.2 Response of the Proposed Controller with Fault Mitigation Techniques

A fault mitigation technique shown in the Fig. 9 has been attempted in this work with two methods viz., 1. Isolating the inverter and rectifier operated at controlled firing angle 2. Isolating the inverter and rectifier operated at fixed firing angle(145°)

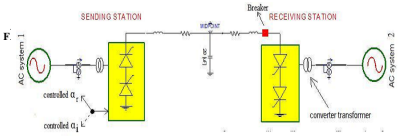
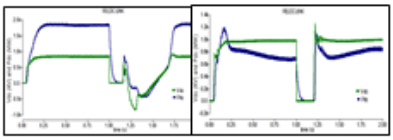
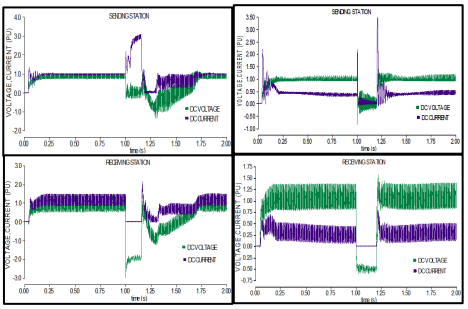


Fig. 9. Scheme for fault mitigation



a)cigre controller b) self defined controller

Fig. 10. Inverter isolated & controlled firing at rectifier(P_{dc},V_{dc})



a) cigre controller b) self defined controller

Fig. 11. Inverter isolated controlled firing at rectifier(V_{dc},I_{dc})

In the first method of mitigation the controller is let free to take control action and act upon the error signal so that to choose a suitable angle of firing to make the rectifier operate in inverter mode. On the other hand, the firing angle of the rectifier has been forcefully increased to 145° on occurrence of the fault on the DC link and

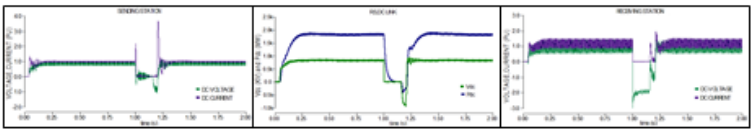


Fig. 12. Performance of the CIGRE controller inverter isolated and firing angle of rectifier forced to 145° system

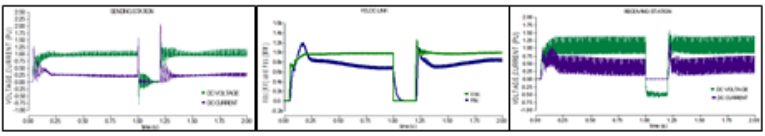


Fig. 13. Performance of the Self-Defined controller inverter isolated and firing angle of rectifier forced to 145° system

the circuit breaker is also operated at the same instant of the switching of the delayed firing angle in the second method. Both the converters are operated as inverters with an intension of evacuating the power from the link on fault occurrence.

A dc link to ground fault has been created at 0.8 s and fault is cleared at 1 s. The fault mitigation technique has been attempted for both the above mentioned cases. Performance of the proposed controller is compared with that of CIGRE controller for the same operating conditions and waveforms are shown in the Fig. 10-13.

It is concluded from the waveforms of the DC link voltage and current, that the settling time on post fault recovery with fault current mitigation is very much improved from (460ms to 434ms) on implementation of the proposed controller. Also the peak overshoot of the link current is found to reduce from (2 pu) to (1.25pu). It is clear that the proposed controller out performs the CIGRE benchmark controller.

6 Conclusions

A Modern self- defined extinction angle controller has been proposed for CCC based HVDC systems. It is inferred that the proposed controller is superior to the CIGRE benchmark controller at post fault recovery as the response of the controller is faster and smoother. Also a slight modification in the circuit has helped to limit the fault current to one third of that attained in the conventional control. A fault mitigation technique has also been suggested in this paper and performance of the proposed controller is found to be on par with the CIGRE controller when operated with the suggested fault mitigation technique for a CCC based HVDC system. PSCAD/EMTDC simulation software has been used for validating the proposed controller in this work. As a future work the proposed controller can be investigated for its suitability towards a Hybrid HVDC based system and a practical implementation of the same can also be attempted.

References

1. Buseman, F.: Economic supply of reactive power for Inverter stations. In: Direct Current, pp. 8–15 (June 1954)
2. Khatir, M., Zidi, S.A., Fellah, M.K., Hadjeri, S., Amiri, R.: Performance evaluation of an HVDC link with a capacitor commutated inverter connected to a very weak receiving ac network. Journal of Electrical Engineering 60, 209–214 (2009)
3. Padiyar, K.R.: HVDC power transmission systems. New Age International (P) Ltd. (2010)

4. Dong, M., He, J., Le, X., Huang, Y.: Realization of self-defined control system for constant extinction angle control using PSCAD/EMTDC. In: APPEC-Power and Energy Engineering Conference, pp. 1–5. IEEE Press (2009)
5. Faruque, M.O., Zhang, Y., Dinavahi, V.: Detailed modelling of CIGRÉ HVDC benchmark system using PSCAD/EMTDC and PSB/SIMULINK. IEEE Transactions o Power Delivery 21, 378–387 (2006)
6. Thahir, M.A., Kirthiga, M.V.: Investigations on a modern self-defined extinction advance controller for HVDC systems. In: International Conference on Process Automation, Control and Computing (PACC), pp. 1–6. IEEE Press (2011)
7. Ashfaq Thahir, M., Kirthiga, M.V.: Investigations on modern self-defined controller for hybrid HVDC systems. In: TENCON 2011 - 2011 IEEE Region 10 Conference, pp. 938–943. IEEE Press (2012)

A Novel Method for Modeling, Simulation and Design Analysis of SEIM for Wind Power Generation

R. Janakiraman[1] and S. Paramasivam[2]

[1] Sri Ramanujar Engineering College, Chennai, Tamil Nadu, India
janakiraman67@yahoo.co.in
[2] ESAB Engineering Services Ltd., Kanchipuram, Tamil Nadu, India
paramsathya@yahoo.com

Abstract. The objective of the research work is to propose a Novel Method for Modeling, Simulation and Design analysis of a Self-Excited Induction Machine for Wind Power Generation (SEIMWPG). The earlier models of such generators are facing huge mechanical losses, due to wear and tear in the tightly coupled mechanical gear systems and could not provide maximum efficiency. In this paper, the self-excited induction machine based wind electric generator design has been presented, it is possible to get maximum power with the variation of the mutual inductance of the stator and rotor windings, irrespective of the variation of wind velocities. The complete system is modeled and simulated in the MATLAB / SIMULINK environment.

Keywords: Doubly-fed induction generator (DFIG), Mutual Inductance, Self-Excited Induction Generator (SEIG), Wind Turbine (WT).

1 Introduction

Wind energy research has developed significantly over the past few years. in spite of this development, more technological advances are needed to make wind energy competitive with many other energy supply techniques. Modeling and simulation can be used to study the performance calculations of wind electric generating systems. In the model of SEIG, variation of the magnetizing inductance is the main factor in the dynamics of voltage build-up and stabilization [1]. A DFIG system has the effects of several parameters and grid strength (external line reactance value) on the system modes have been studied [2]. A complete analysis of a WT driven SEIG as a part of a supply system to an isolated load, has been carried out. To achieve this, a new simplified model has first been derived for a WT driven SEIG. [3]. A detailed DFIG wind turbine model including two-mass drive train, pitch control, nonlinear control and vector-control loops was developed [4],[5].

A new algorithm also has been presented in order to minimize the torque and flux ripples of the induction motor drive and to improve the performance of the basic direct torque control (DTC) scheme [6]. It is apparent that the calculation of all the machine inductances as defined by the inductances matrices is the key to the successful simulation of an induction machine [7]. The DQ-modeling approach for

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 324–330, 2013.

the three phase self-excited induction generator (SEIG) with a squirrel cage rotor and steady state performance are presented [8], [9].

2 Proposed Model of Self-Excited Induction Machine for Wind Power Generation

It is proposed that a wind turbine is coupled with a Self-Excited Induction Generator (SEIG) for electric power generation. In this case, no mechanical gearing system has been utilized; because of mechanical losses like wear and tear and mechanical friction losses, the mechanical gearing replaced by an electrical system, as discussed below in Fig.1, shows the flow diagram of the wind power generating system. Where, v_w is the wind speed in metres/second, L_m is the mutual inductance of the windings in henries, V_G is the voltage generated and V_k is the rated voltage output. The first step is to initialize the process of operation; the second step is to propose the variables involved in the process, such as wind speed, mutual inductance of the stator and rotor windings of the SEIG and voltage generation.

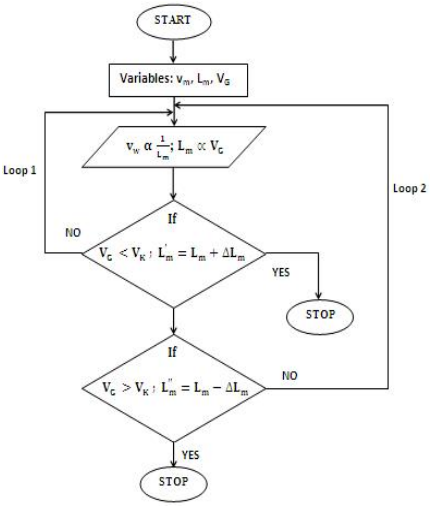


Fig. 1. Flow diagram of the proposed SEIMWPG

The main variable is the mutual inductance of the SEIG, in which it is dependent on the speed of the wind turbine, based on the wind velocity. The third step is the variation of the mutual inductance, in which it is inversely proportional to the wind velocity, and voltage generation is directly proportional to the mutual inductance of the windings of the SEIG. The fourth step is, if the voltage generation is less than the rated value, the loop.1, is directed to increase the mutual inductance of the SEIG and, to increase the voltage generation up to the rated value. The fifth step is, if the voltage generation is more than the rated value, loop.2, is directed to reduce the mutual

inductance of the SEIG and, to reduce the voltage generation up to the rated value. Thus, the wind power generation is always maintained rated value, irrespective of the variation of wind velocities.

3 Self-Excited Induction Generator (SEIG)

A Self- Excited Induction Generator is made by a three-phase induction machine, if its rotor is externally driven at a suitable speed, and a three-phase capacitor bank of a sufficient value is connected across its stator terminals. Fig.2, shows the schematic diagram of a SEIG system. Whenever the induction machine is driven at the required speed, the residual magnetic flux in the rotor will induce a small electro-motive-force (e.m.f) in the stator winding. The suitable capacitor bank causes this induced voltage to continue to increase until an equilibrium state is attained, which is due to magnetic saturation of the machine. The mutual inductance of the winding is made as variable with respect to the variation of the rotor speed. It can be varied so that, if the rotor speed is above the normal speed, the mutual inductance is made to reduce, and if the rotor speed is below the normal speed, the mutual inductance is made to increase; this arrangement is required to maintain the generated output at its rated value always.

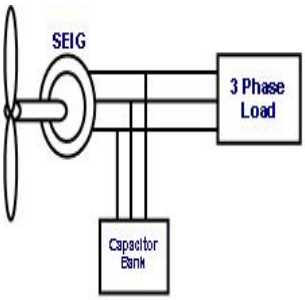


Fig. 2. Schematic Model of the SEIMWPG

3.1 Mathematical Modeling of a Self-Excited Induction Generator

The equation shown is used for developing the dynamic model of the SEIG,

$$[V_G]=[R_G][i_G]+[L_G]p[i_G]+\omega_{rG}[G_G][i_G] \quad (1)$$

Where $[V_G]$ and $[i_G]$ represents 4 x 1column matrices of voltage and which is given as
$[V_G]=[V_{sd}\ V_{sq}\ V_{rd}\ V_{rq}]^T$ and $[I_G]= [i_{sd}\ i_{sq}\ i_{rd}\ i_{rq}]^T$, [R], [L] and [G] represents 4×4 Matrices of resistance, generator inductance and capacitance as given. Further, L_m the magnetizing inductance, which can be obtained from the magnetizing curve of the machine.

$$[L]=\begin{bmatrix} L_{sd} & L_{dq} & L_{md} & L_{dq} \\ L_{dq} & L_{sq} & L_{dq} & L_{md} \\ L_{md} & L_{dq} & L_{rd} & L_{dq} \\ L_{dq} & L_{mq} & L_{dq} & L_{rq} \end{bmatrix} [G]=\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & L_m & 0 & L_r \\ -L_m & 0 & -L_r & 0 \end{bmatrix} [L]=\begin{bmatrix} L_s & 0 & L_m & 0 \\ 0 & L_s & 0 & L_m \\ L_m & 0 & L_r & 0 \\ 0 & L_m & 0 & L_r \end{bmatrix} \tag{2}$$

The relation between L_m and i_m is given as

$$\mathrm{L_m}=\frac{|\psi_\mathrm{m}|}{|\mathrm{i_m}|} \tag{3}$$

where $|\psi_m|$ and $\mathrm{i_m}$ are the magnetizing flux linkage and magnetizing current. The equation defining $\mathrm{L_m}$ $\mathrm{V_s}$ $|\mathrm{i_m}|$ used in the model is,

$$|\mathrm{i_m}|=1.447 * \mathrm{L^6_m}- 8.534* \mathrm{L^5_m}+18.174*\mathrm{L^4_m}-17.443*\mathrm{L^3_m} - 7.322*\mathrm{L^2_m}-1.329*\mathrm{L_m} + 0.6979 \tag{4}$$

$\mathrm{L_{dq}}$ used in matrix L represent the cross saturation coupling between all axes in space quadrature and is due to saturation

$$\mathrm{L_{dq}} = \mathrm{L_m}\frac{\mathrm{i_{md}}}{\mathrm{i_{mq}}}*\mathrm{L_{dq}} \tag{5}$$

$$\text{and} \quad \mathrm{L_{mq}} = \mathrm{L_m} + \frac{\mathrm{i_{mq}}}{\mathrm{i_{mq}}} * \mathrm{L_{dq}} \tag{6}$$

From the above equations representing, $\mathrm{L_{dq}}$, $\mathrm{L_{md}}$ and $\mathrm{L_{mq}}$ that under linear magnetic conditions, $\mathrm{L_{dq}} = 0$ and $\mathrm{L_{md}} = \mathrm{L_{mq}} = \mathrm{L_m}$, as expected. The two axes values of the total stator and rotor inductances are $\mathrm{L_{sd}} = \mathrm{L_{sl}}+ \mathrm{L_{md,}}$ $\mathrm{L_{sq}} = \mathrm{L_{sl}}+ \mathrm{L_{mq}}$ and $\mathrm{L_{rd}}=\mathrm{L_{rl}}+ \mathrm{L_{md,}}$ $\mathrm{L_{rq}} = \mathrm{L_{rl}}+ \mathrm{L_{mq}}$.The above equations $\mathrm{L_{sl}}$ and $\mathrm{L_{r1}}$ are the leakage inductances of the stator and rotor, respectively. Because of saturation, $\mathrm{L_{sd}}$ and $\mathrm{L_{sq,}}$ but it follows from previous arguments that under linear magnetic conditions, $\mathrm{L_{sd}} = \mathrm{L_{sq}}$. Hence $\mathrm{L_r} = \mathrm{L_{rl}} + \mathrm{L_{m..}}$

The electromagnetic torque developed by the generator is given by

$$\mathrm{T_e} = \left(\frac{3}{4}\right) * \mathrm{p} * \mathrm{L_m} * \left(\mathrm{i_{sq}}\mathrm{i_{rd}} - \mathrm{i_{sd}}\mathrm{i_{rd}}\right) \tag{7}$$

Thus, it is seen that Eq. (1) consists of four first order equations. An induction motor is hence represented by these four first order differential equations. Because of the non-linear nature of the magnetic circuit, the magnitude of magnetizing current, I_m is calculated as

$$\mathrm{I_m} = \left[\left(\mathrm{i_{sq}} + \mathrm{i_{rd}}\right)^2 + \left(\mathrm{i_{sq}} + \mathrm{i_{rq}}\right)^2\right]^{\frac{1}{2}} \tag{8}$$

Capacitor side equations are

$$[\mathrm{V_{sG}}]=(1/\mathrm{C})[i_C] \tag{9}$$

Also further,

$$[i_C]=[i_{sG}]+[i_L] \tag{10}$$

Where, $[V_{sG}]=[\ i_{sG}]+[i_L]$ Column matrices represent the direct and quadrature axis components of the capacitor current generator stator current and load current respectively. Load side equation is given as

$$[V_{sG}]=L_{Lp}[i_L]+R_L[i_L] \tag{11}$$

Thus, it is seen that the complete transient model of the SEIG in the d-q quasi stationary reference frame consists of equations from (1) to (11).

4 Modeling of Self-Excited Induction Machine for Wind Power Generation Using Mat Lab Simulink

Mat Lab Simulink is a powerful software tool for modeling and simulation. The equations from (1) to (11) have been implemented in the Simulink tool using various blocks in Fig.4.

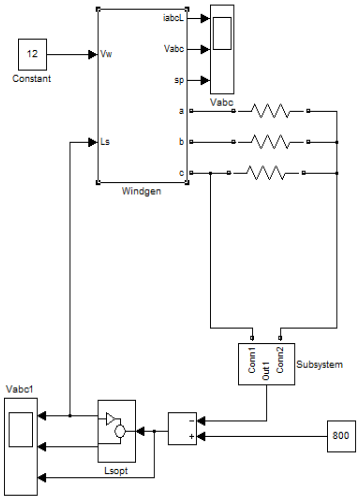


Fig. 3. Simulink model of proposed system

5 Simulation Results and Discussions

The model has been simulated using Mat Lab Simulink and the various results have plotted as below in Fig. 4(a) and Fig. 4(b).

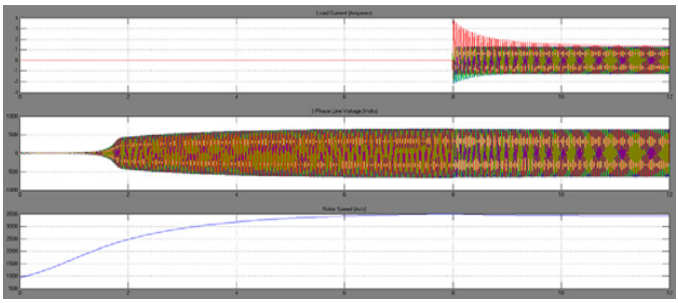


Fig. 4(a). Wind velocity (m/s), Generated voltage (v) and Load current (A)

In Fig. 4(a), shown that wind speed, voltage generated and load current. The voltage generated is maintained at rated value, irrespective of the wind speed variations with aid of machine windings mutual inductance variations, and the load current increases with respect to the increase in load connected.

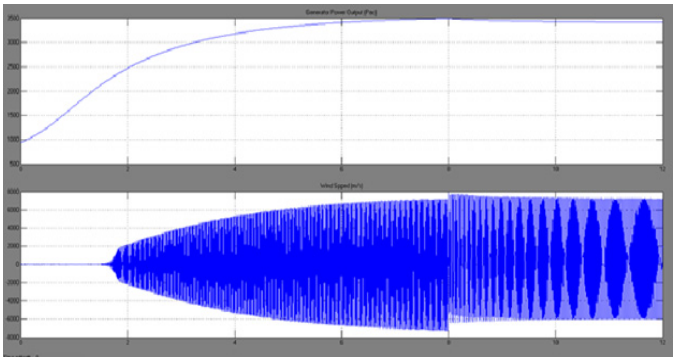


Fig. 4(b). Wind velocity (m/s) and Generator output Power, P_{ac}, (kW)

In Fig. 4(b)**,** shown that wind speed and the power output of the generator is met its rated value by the variation of the mutual inductance of the generator windings, irrespective of the various wind velocities.

6 Conclusion

This paper presented a novel method for the modeling, simulation and design analysis of a self-excited induction machine for wind power generation (SEIMWPG). The test model is designed and tested for various wind speeds and voltages. This test model will best suit for the power generations from the renewable energy source i.e., wind energy. The variation of the mutual inductance of stator and rotor windings has taken care of the wind speed fluctuations, in which it is said that the mutual inductance has varied, and to maintain the output voltage at its rated value. It is shown that the

simulated results for various wind speeds in this system should be useful to solve the power shortage and it will be enhanced to design for large scale generation and to meeting future power demands.

References

1. Debta, B.K., Mohanty, K.B.: Analysis on the Effect of Dynamic Mutual Inductance in Voltage Build-up of a stand-Alone Brushless Asynchronous Generator. In: NPEC, pp. 1–7 (2010)
2. Mei, F., Pal, B.: Model Analysis of Grid-Connected Doubly Fed Induction Generators. IEEE Transactions on Energy Conversion 23(3), 728–736 (2007)
3. Uqtug, Y., Demirekier, M.: Modelling, analysis and control of a wind-turbine driven self-excited induction generator. IEE Proceedings 135(4), pt. C, 368–375 (1988)
4. Yang, L., Xu, Z., Ostergaard, J., Dong, Z.Y., Wong, K.P., Ma, X.: Oscillatory Stability and Eigenvalue Sensitivity Analysis of a DFIG Wind Turbine System. IEEE Transaction on Energy Conversion, 1–12 (2011)
5. Tang, C., Guo, Y., Jiang, J.N.: Nonlinear Dual – Mode Control of Variable-Speed Wind Turbines with Doubly fed Induction Generators. IEEE Transactions on Control Systems Technology, 1–13 (2010)
6. Vasudevan, M., Atrumugam, R., Paramasivam, S.: Real time implementation of viable torque and flux controllers and torque ripple minimization algorithm for induction motor drive. ELSEVIER Energy Conversion and Management 47, 1359–1371 (2006)
7. Aroui, T., Koubaa, Y., Toumi, A.: Magnetic Coupled Circuits Modeling of Induction Machines Oriented to Diagnostics. Leonardo Journal of Sciences (13), 103–121 (2008)
8. Kishore, A., Prasad, R.C.: MATLAB SIMULINK Based DQ Modeling an Dynamic Characteristics of Three Phase Self Excited Induction Generator. In: Progress in Electromagnetics Research Symposium 2006, Cambridge, USA, March 26-29 (2006)
9. Sandhu, K.S., Jain, S.P.: Steady State Operation of Self-Excited Induction Generator with Varying Wind Speeds. International Journal of Circuits, Systems and Signal Processing 2(1), 26–33 (2008)

Modelling and Simulation of H-Bridge Topology for Z-Source Inverter Using Matlab/Simulink

S. Sabareswar and S. Lenin Prakash

Department of EEE, Saranathan College of Engineering, Trichy-620012, Tamilnadu, India
{sabareshwar,lenin.prakash.s}@gmail.com

Abstract. This paper presents the modelling of H-bridge topology for Z-source inverter (ZSI) feeding a RL Load with Sinusoidal PWM technique using Matlab/Simulink. The proposed model can be incorporated with any of the Z Source networks. The different modes of operation of single phase Voltage source inverter (VSI), and an additional mode when the same H-bridge topology is used as a ZSI have been analyzed. In the developed model the dynamic characteristics of load voltage, load current as well as switch voltage and current can be effectively monitored and analyzed. Therefore, it can be expected that the developed simulation model can be an easy-to-design tool for the development of ZSI based circuits including controller algorithms and topological variations with reduced computation time and memory size. Validations of simulation results of proposed model along with circuit simulation results have been presented.

Keywords: inverter, H-bridge, modelling, ZSI, VSI.

1 Introduction

To date, many methods for modelling of ZSI have been reported in the literature. One such method that presents the transient modelling and analysis of Voltage source type ZSI have been discussed in [1], in which, the ZSI is modelled by considering the inverter as a single switch. The limitation of this method is bidirectional power flow from source to load as well as load to source cannot be considered. In [2] state space average modelling of a three phase ZSI has been proposed. There is no sufficient literature available for modelling a single phase ZSI. In this paper modelling of single phase inverter which can be used with different kind of Z-Source networks.

There are numerous topologies available in ZSI such as Dual input Dual output ZSI [3], Trans Z-source inverter [4]. In the proposed method, the H-bridge inverter model developed for ZSI can be incorporated to all such topologies. Single phase VSI is modelled based on the output voltage and input current obtained for different possible modes operation occurring in VSI. From the output voltage obtained, the load current is derived. In case of an R load, the input current of the Single phase VSI is the absolute value of the load current obtained whereas in case of RL load it is not so. The input current fed to inverter is not the absolute value of load current. The relation between input current and load current is described in section 4. The load current is

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 331–338, 2013.

segregated as switch and diode currents which can be computed to generate the input current to the inverter. With this method of deriving input current, it is possible to include all modes of operation occurring in Single phase VSI as well as in ZSI. The simulation results by modelling of input current for a RL Load with all modes of operation in VSI and ZSI have been presented and from the results the advantages of ZSI by modelling H-bridge topology alone have been validated.

2 Different Modes of Single Phase VSI

The single phase VSI has three modes of operation based on the power flow between source and load. All the possible states are categorized in these three modes of operation. These are presented in Table 1. S_1, S_2, S_3, S_4 represents the switching state of the inverter switches. V_{dc} is the DC input supply voltage. When the switch is turned on, it is designated as logic '1' and when it is turned off it is designated with logic'0'. I_{D1}, I_{D2}, I_{D3}, I_{D4} are the diode currents. The power circuit diagram for H-bridge topology for ZSI is shown in Fig. 1.

2.1 Mode-I (Positive Power Mode)

In this mode of operation, the power flow will be from source to load. During this mode, the upper switch from one leg and lower switch of the other leg conduct together. From Table 1, the switches S_1-S_2 will be conducting for positive half cycle and S_3-S_4 will be conducting for negative half cycle. When S_1-S_2 conducts, the output voltage (V_o) and Load current (I_o) will be positive while the output voltage and Load current will be negative when S_3-S_4 conducts. Therefore in both half cycles the power generated will be positive. The input current to the inverter (I_{in}) for this mode is shown in Table 1.

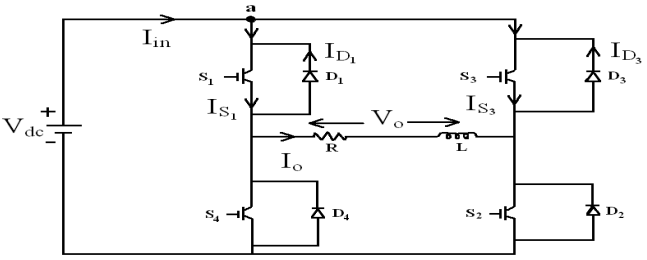


Fig. 1. Power circuit diagram of H-bridge topology for ZSI

2.2 Mode-II (Free Wheeling Mode)

This mode of operation occurs when the load doesn't receive power from source side, it simply freewheels within the switches and load. The condition for mode-II to occur is when the switches from upper leg or lower leg conduct simultaneously. The

switches S_1-S_3 (S_1-D_3 (or) S_3-D_1) (or) S_2-S_4 (S_2-D_4 or S_4-D_2) conduct together. This leads to switches and the load. The output voltage during this mode is zero and hence input current to the inverter (I_{in}) is also zero. The power flow from source to the load is zero during this mode.

Table 1. Switching States of Single Phase VSI

MODE	S_1	S_2	S_3	S_4	V_o	I_{D_1}	I_{D_2}	I_{D_3}	I_{D_4}	I_{in}
I	1	1	0	0	V_{dc}	0	0	0	0	I_o
	0	0	1	1	$-V_{dc}$	0	0	0	0	I_o
II	1	0	1	0	0	$-I_o$	0	I_o	0	0
	0	1	0	1	0	0	I_o	0	I_o	0
III	1	1	0	0	0	$-I_o$	$-I_o$	0	0	$-I_o$
	0	0	1	1	0	0	0	I_o	I_o	$-I_o$

2.3 Mode-III (Negative Power Mode)

In this mode of operation where the energy stored in the inductor of RL load is fed back by diodes D_3-D_4 (when S_1-S_2 is about to be gated on) or D_1-D_2 (when S_3-S_4 is about to be gated on) to the supply. This occurs prior to mode-I. At this instant, the diodes D_3-D_4 or D_1-D_2 will be conducting and the power flows from load to source. The output voltage is zero in this mode. The switching states for negative power mode are shown in Table 2.

3 Different Modes of Single Phase ZSI

All the three modes presented in Table 1 are applicable to both Single Phase VSI as well as ZSI. But ZSI has one more mode wherein the inverter switches of same leg are allowed to turn on simultaneously. This has been presented in Table 2.

3.1 Mode-I, Mode-II, Mode-III

The first three modes of operation are discussed in previous section wherein the power flow can be positive, zero and negative, which is applicable to H-bridge topology for ZSI also.

3.2 Mode-IV (Shoot through Mode)

This mode of operation is found only in ZSI wherein the switches of the same leg conduct together. For analysis purpose the switches of same leg are allowed to conduct is considered in H-bridge topology for ZSI. There are five possible combinations of H-bridge topology for ZSI based on the switching sequence of the switches.The output voltage and input current of the H-bridge topology for ZSI is

shown in Table 2. The input current magnitude in this mode is simply equal to input voltage by on-state resistance (R_{on}) of switches. By applying KVL to the Fig.1, the output voltage of the inverter is given by (1).

$$V_o = L\frac{dI_o}{dt} + RI_o \text{ (V)} \tag{1}$$

where R= Load Resistance (in Ω),L = Load Inductance (in H)

Table 2. Switching States of H-Bridge Topology for ZSI

MODE	S_1	S_2	S_3	S_4	V_o	I_{D_1}	I_{D_2}	I_{D_3}	I_{D_4}	I_{in}
I	1	1	0	0	V_{dc}	0	0	0	0	I_o
	0	0	1	1	$-V_{dc}$	0	0	0	0	I_o
II	1	0	1	0	0	$-I_o$	0	I_o	0	0
	0	1	0	1	0	0	$-I_o$	0	I_o	0
III	1	1	0	0	0	$-I_o$	$-I_o$	0	0	$-I_o$
	0	0	1	1	0	0	0	I_o	I_o	$-I_o$
IV	1	1	1	0	$V_{dc}/2$	0	0	0	0	$V_{dc}/2R_{on}$
	1	1	0	1	$V_{dc}/2$	0	0	0	0	$V_{dc}/2R_{on}$
	1	0	1	1	$-V_{dc}/2$	0	0	0	0	$V_{dc}/2R_{on}$
	0	1	1	1	$-V_{dc}/2$	0	0	0	0	$V_{dc}/2R_{on}$
	1	1	1	1	0	0	0	0	0	V_{dc}/R_{on}

4 Modelling of H-Bridge Circuit Considering All the Modes Occurring in VSI as Well as ZSI

Modelling of H-bridge Inverter circuit based on different modes of operation for VSI and ZSI has been presented here. By modelling of entire circuit it is possible to predict the exact nature of load voltages and currents. The output voltage of the inverter is obtained by considering all modes of operation occurring in VSI and ZSI. Here, the switches of the inverter S1, S2, S3 and S4 are considered as switching function. The output voltage equation is given by (2).

$$V_o = V_{dc}(S_1S_2\overline{S_3S_4}) - V_{dc}(S_3S_4\overline{S_1S_2}) + V_{dc}/2\left((S_1 - \overline{S_1}) * S_2S_3\right) + V_{dc}/2\left((\overline{S_3} - S_3) * S_1S_4\right) \tag{2}$$

Similarly, the input current to the inverter is obtained by applying KCL to the circuit at node 'a'. The input current obtained from equation (3) is applicable to all modes occurring in VSI and ZSI. The input current (I_{in}) equation is given by

$$I_{in} = \left(I_{S_1} - I_{D_1}\right) + \left(I_{S_3} - I_{D_3}\right) \text{ (A).} \tag{3}$$

where I_{S1}, I_{s3} are the current flowing through switches S1 and S3 during different operation modes. Similarly, I_{D1}, I_{D3} are the currents flowing through diodes during

freewheeling and negative modes of operation. The switch and diode currents expressions can be derived based on their conduction during individual modes of operation for both Single phase VSI and ZSI. From equation (1), the load current is computed. From the load current (I_o), the switch and diode currents are split. The switch current I_{S1}, I_{s3} and diode current I_{D1}, I_{D3} are given by expressions (4) to (7).

$$I_{S_1} = \Big(\big((S_1S_2I_{op}) > 0\big) - \big((S_1S_4I_{op}) > 0\big)\Big) * I_o + \Big(\big((S_1S_3I_{op}) > 0\big) + (S_1S_2S_3S_4) - (S_1S_2S_3)\Big) * I_o + \big((S_1S_4) * (V_{dc}/2R_{on})\big). \quad (4)$$

$$I_{S_3} = \Big(\big((S_3S_4I_{op}) < 0\big) - \big((S_2S_3I_{op}) < 0\big)\Big) * I_o + \Big(\big((S_1S_3I_{op}) < 0\big) + (S_1S_2S_3S_4) - (S_1S_4S_3)\Big) * I_o + \big((S_2S_3) * (V_{dc}/2R_{on})\big). \quad (5)$$

$$I_{D_1} = \Big(\big((S_1S_2I_{op}) < 0\big) + \big((S_1S_3I_{op}) < 0\big)\Big) * I_o + \Big(\big((S_1S_2S_3S_4) - (S_1S_4S_3)\big)\Big) * I_o. \quad (6)$$

$$I_{D_3} = \Big(\big((S_3S_4I_{op}) > 0\big) + \big((S_1S_3I_{op}) > 0\big)\Big) * I_o + \Big(\big((S_1S_2S_3S_4) - (S_1S_2S_3)\big)\Big) * I_o. \quad (7)$$

where I_{op} is the load current polarity which may be either positive or negative.R_{on} is the internal on state resistance of the switch. Similarly, the switch currents I_{s2}, I_{s4} and diode currents I_{D2}, I_{D4} can be derived.

5 Generation of PWM Signals

Here, the Sinusoidal PWM technique is used, which can be used for both Single phase VSI and ZSI [5], [6]. The modulation wave is modified to include the shoot through state of ZSI as well as the other states found in VSI. The expression for modulating wave considering the mode IV operation of ZSI is given by expression (8), (9), (10).

$$V_{am(S_1)} = V_{am} + T_o/T. \quad (8)$$

$$V_{am(S_2)} = V_{am} - T_o/T. \quad (9)$$

$$V_{am(S_3,S_4)} = V_{am}. \quad (10)$$

where T_o(in sec) is the period during which the shoot through state occurs,T (in sec)is the total switching period, $V_{am(S_1)}, V_{am(S_2)}, V_{am(S_3,S_4)}$are the amplitude of modultion wave for switches S_1, S_2, S_3-S_4 respectively.V_{am} is the amplitude of modulation wave.The amplitude modulation index is taken as m_a=0.88 and frequency

modulation index is considered as m_f=20. The switching frequency is taken as 1000Hz. The PWM pulses generated are given as input to the inverter switches.

6 Simulation Results

The results of modelling of H-Bridge inverter for both VSI and ZSI for RL Load are simulated with Matlab/Simulink. Fig.2 shows the simulation model of H-Bridge topology for ZSI. Fig. 3 shows the simulation results of Output voltage *(V_o)*, Load current *(I_o)* and Input current to the inverter *(I_{in})* for Single Phase VSI by modelling as well as circuit simulation. Fig.4 shows the simulation results of Output voltage *(V_o)*, Load current *(I_o)* and Input current to the inverter *(I_{in})* of ZSI considering mode IV operation by modelling as well as by circuit simulation.

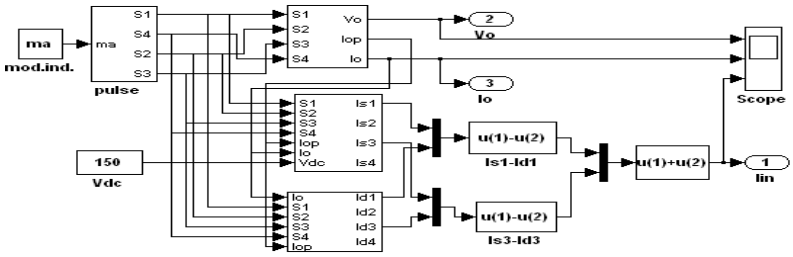


Fig. 2. Simulation model of H-bridge topology for ZSI using Matlab/Simulink

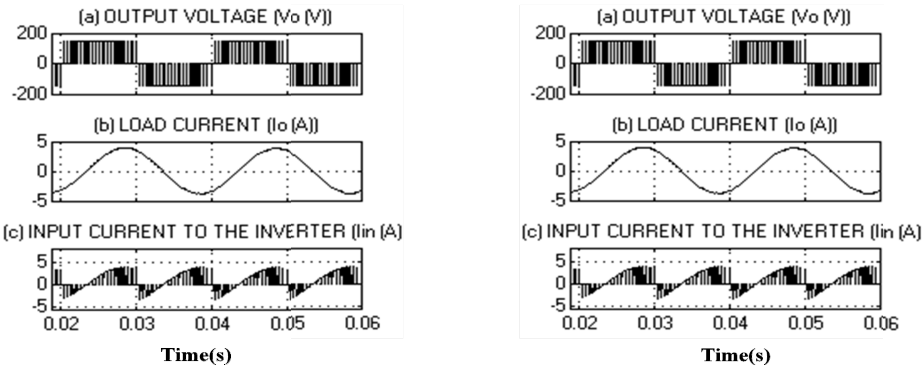


Fig. 3. Simulation results of (a) Output Voltage *($V_o(V)$)*(b) Load current *($I_o(A)$)* (c) Input Current to the Inverter *($I_{in}(A)$)*for Single phase VSI by modelling as well as by circuit simulation

The simulation parameters are:

$$R_{LOAD}= 14\ \Omega\ ;\ L_{LOAD}=0.1H,\ V_{dc}= 150\ V,\ R_{on} = 1m\ \Omega$$

For different modes of operation in Single phase VSI, it is found that the waveforms of Output voltage, Output current of the inverter and Input current to the inverter obtained by modelling are verified with Table 1 and hence the results are matched with circuit simulation. The fundamental component of output voltage is given by

$$V_{0,1}=m_a*V_{dc}=132V \tag{11}$$

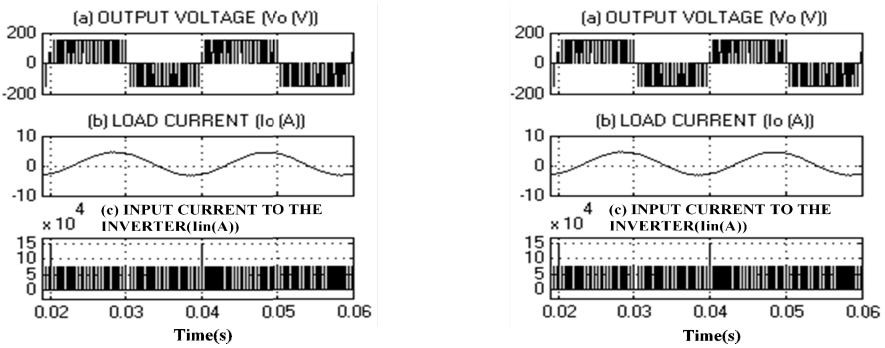


Fig. 4. Simulation results of (a) Output Voltage *($V_o(V)$)*(b) Load current *($I_o(A)$)* (c) Input Current to the Inverter *($I_{in}(A)$)*for Single phase ZSI considering Mode IV operation by modelling as well as by circuit simulation

The fundamental component of output voltage is computed with fourier block in MATLAB and is found to be similar to that of expression (11).Fig.5(c) and Fig. 6(c) shows the simulation results of Input current fed to the inverter by modelling as well as with circuit simulation. The magnitude of input current during mode IV operation of ZSI is given in Table 2, which is simply $V_{dc}/(2*R_{on})$ when either of three switches conduct (or) V_{dc}/R_{on} when all the switches conduct together. The simulation results of Output voltage of the single phase H-bridge topology for ZSI by modelling as well as circuit simulation are shown in Fig.5 (a) and Fig.6 (a).

7 Conclusion

In this paper, the modelling of H-bridge topology for ZSI has been presented. The simulation results show that the H-bridge inverter model can be incorporated for various topologies of ZSI. The modelling of output voltage followed by load current of the Single phase VSI based on different modes of operation are implemented using Simulink. The input current fed to the inverter is modeled based on operation modes of VSI and it is validated with circuit simulation. This concept has been extended to include Shoot through states of ZSI. Ultimately, the results for all modes of operation of ZSI and VSI are computed and hence the results are verified with simulation circuits.

References

1. Loh, P.C., Vilathgamuwa, D.M., Gajanayake, C.J., Lim, Y.R., Te, C.W.: Transient Modelling and analysis of pulse-width modulated Z-source inverter. In: Conf. Rec. 40th IEEE IAS Annu. Meeting, October 2-6, vol. 4, pp. 2782–2789 (2005)
2. Gajanayake, J., Vilathgamuwa, D.M., Loh, P.C.: Development of a comprehensive model and a multiloop controller for Z-source inverter DG systems. IEEE Trans. Ind. Electron. 54(4), 2352–2359 (2007)
3. Dehghan, S.M., Mohamadian, M., Yazdian, A., Ashrafzadeh, F.: A Dual-Input–Dual-Output Z-Source Inverter. IEEE Trans. Power Electronics 25(2) (February 2010)
4. Qian, W., Peng, F.Z., Cha, H.: Trans-Z-Source Inverters. IEEE Trans. Power Electronics 26(12) (December 2011)
5. Loh, P.C., Vilathgamuwa, D.M., Lai, Y.S., Chua, G.T., Li, Y.W.: Pulse-width modulation of Z-source inverters. In: Conf. Rec. 39th IEEE IAS Annu. Meeting, October 3-7, vol. 1, pp. 148–155 (2004)
6. Loh, P.C., Blaabjerg, F., Wong, C.P.: Comparative evaluation of Pulse width Modulation Strategies for Z-Source Neutral-Point-Clamped Inverter. IEEE Trans. Power Electronics 22(3) (May 2007)

Cascaded Multilevel Inverter for Industrial Applications

B. Amala Priya Shalini and S.S. Sethuraman

Department of Electrical and Electronics Engineering
Sri Venkateswara College of Engineering
Chennai, India
amalapriya_shalini@yahoo.co.in, sethuram@svce.ac.in

Abstract. This paper work presents certain Sequential Switching Hybrid Modulation strategies and compared for single phase multilevel inverters. Hybrid modulation represents the combination of fundamental frequency and multilevel sinusoidal modulations and designed for performance of the well-known alternative phase opposition disposition (APOD). The main characteristic of these modulations are the reduction of switching losses, achieve balanced power dissipation among the devices and equal load sharing.

Keywords: cascaded MLI, harmonics analysis, hybrid modulation, switching loss with EMI problem.

1 Introduction

The concept of multilevel converters has been introduced since 1975. Several multilevel converter topologies have been developed. Cascaded H bridge inverter (CHBI) is the most important topologies in MLI. In CHBI topology, single phase Capacitor, Battery and Renewable Energy can be used as multiple d.c voltage sources [2]. Multi-level inverters (MLIs) are finding increased attention in industries as a choice of electronic power conversion for medium voltage and high-power applications, because improving the output waveform of the inverter reduces its respective harmonic content and hence, the size of the filter used and the level of electromagnetic interference (EMI) generated by switching operation. The output of conventional two-level inverter is in the form of square wave ac power which usually contains undesirable harmonics [6]. When this output is fed to an electrical device such as an electrical motor it causes heating which in turn causes increased losses and finally resulting in decreased efficiency. The THD value is responsible for reducing the quality of output. From the survey, it has been found that cascaded multilevel inverter is more suitable for high power applications such as STATCOM in transmission lines, A.C Drive application in industries. The multilevel inverter can operate by Hybrid Modulation strategies. Then the performance of symmetrical CHB multilevel inverter with respect to harmonic content, number of switches and voltage stress across the switch is calculated by using MATLAB/Simulink Software. A detailed harmonics analysis is done on the CHB-MLI by considering 5-level operation.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 339–344, 2013.

2 Hybrid Strategy

Aim of the paper is to reduce the switching loss of multilevel sinusoidal modulation (MSPWM) schemes with low computational overhead. Also reduces its respective harmonic content and hence, the size of the filter used and the level of EMI generated by switching operation [1].

2.1 Need of Hybrid Modulation Techniques

In the recent 15 years, however, there has been growing awareness that electromagnetic interferences (EMI) were affecting the human body by cellphone usage and high power lines cause cancers and diseases. For 4 Billion years, life on earth has been in an EMI free environment. So EMI protection is essential for the Biosphere.

2.2 Review of MSPWM Schemes

The major MSPWM schemes: Alternative Phase Opposition Displacement (APOD), Single carrier Multiple Sinusoidal modulation (SCSPWM), Carrier-based PWM (CBPWM), Phase Shifted Carrier PWM (PSC) shows analysis of different PWM techniques in Fig 1.

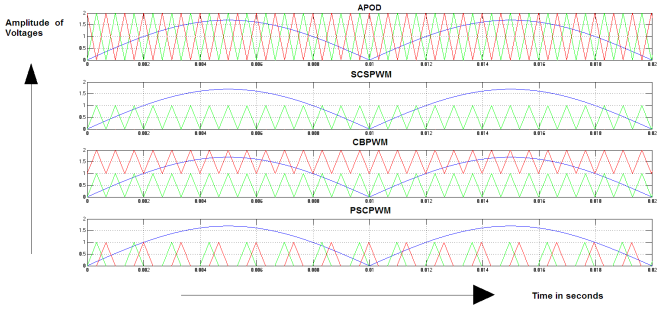


Fig. 1. Comparison of triangular carrier and modulating signal

Unipolar carrier-based N-level PWM operation consists of (N − 1)/2 different carriers, same as the number of FBI cells (K=N−1/2). Fig.1 shows comparison of sinusoidal reference and carrier signals for five-level CMLI operation. For APOD, all carriers are phase opposition by 180◦ from its adjacent carrier. SCSPWM is a result of sinusoidal-modulating signals with a fundamental frequency fo and carrier signal. For CBPWM, the reference waveform is sampled through a number of carrier waveforms displayed by contiguous of the reference waveform amplitude. For PSC all carriers are Phase Shifted by 90◦ from its adjacent carrier with single sinusoidal-modulating signals. From the review of MSPWM, APOD is an efficient than others [1].

2.3 Basic Principle of Hybrid Modulation

Hybrid modulation is the combination of fundamental frequency modulation (FPWM) and MSPWM for each inverter cell operation, so that the output inherits the features of switching-loss reduction from FPWM, and good harmonic performance from MSPWM. In this modulation technique, the four switches of each inverter cell are operated at two different frequencies; two being commutated at FPWM, while the other two switches are modulated at MSPWM, therefore the resultant switching patterns are the same as those obtained with MSPWM pulses.

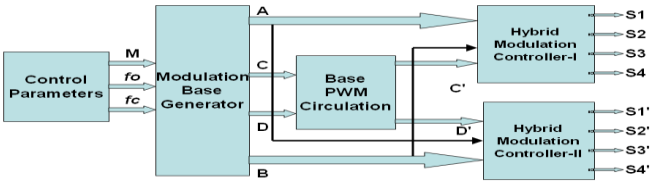


Fig. 2. Block diagram of proposed sequential switching hybrid modulation

(i) Control Parameters

Control parameters are M,f0,fc : M-modulation index(0.85), f0 –fundamental frequency(50HZ), fc -carrier frequency(1500HZ). The M for MSPWM is defined as M = Am/KAc, K = (N − 1)/2, the modulation frequency is given as mf = fc/fo, where N= number of levels, Ac - carrier amplitude and Am- fundamental modulated signal amplitude.

(ii) Base-Modulation Generator Design

In this modulation strategy, three base-modulation pulses are needed for each cell operation in a CMLI. A sequential switching pulse (SSP) (A) is a square-wave signal with 50% duty ratio and half the fo. This signal makes every power switch operating at MSPWM, and FPWM sequentially to equalize the power losses among the devices. FPWM (B) is a square-wave signal synchronized with the modulation waveform; B = 1 during the positive half cycle of the modulation signal, and B = 0 during negative half cycle. MSPWMs (C or D) for each cell, differs depends upon the type of carrier and modulation signals used shown in Fig 3. (a)

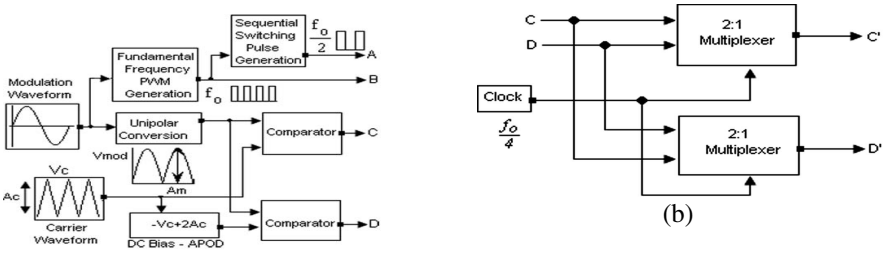


Fig. 3. Block diagrams : (a) Base modulator for APOD (b) Scheme of base PWM circulation for five-level cascaded multilevel inverter

(iii) Base PWM Circulation

The scheme of five-level base PWM circulation is shown in Fig 3 (b), consists of two 2:1 multiplexer, and selects one among the two PWMs based on the select clock signal. The clock frequency is fo/4, makes the time base for PWM circulation from one module to another. After two fundamental frequency periods, the order is changed so that the first module HPWM becomes the second module, the second becomes the third, etc., while the last module HPWM shifts to the first

(iv) Hybrid-Modulation Controller

HMC combines SSP, FPWM, and MSPWM that produces SSHM pulses. It is designed by using a simple combinational logic and the functions for a five-level HPWM are expressed as

$$\begin{aligned} S1 &= ABC' + \overline{A}B \\ S2 &= A\overline{B}C' + \overline{A}\,\overline{B} \\ S3 &= \overline{A}\,\overline{B}C' + A\overline{B} \\ S4 &= \overline{A}BC' + AB \end{aligned} \quad \text{and} \quad \begin{aligned} S1' &= ABD' + \overline{A}B \\ S2' &= A\overline{B}D' + \overline{A}\,\overline{B} \\ S3' &= \overline{A}\,\overline{B}D' + A\overline{B} \\ S4' &= \overline{A}BD' + AB \end{aligned}$$

Where A is an SSP, B is an FPWM, C_ is an MSPWM for cell-I and D_ is an MSPWM for cell-II. In Fig.4, it is shown that each gate pulse is composed of both FPWM and MSPWM. If SSP A = 1, then S1, S2, S1_, andS2_ are operated with MSPWM, while S3, S4, S3_, and S4_ are operated at FPWM. If SSP A = 0, then S1, S2, S1_, and S2_ are operated at FPWM, while S3, S4, S3_, and S4_ are operated with MSPWM. Since A is a sequential signal, the average switching frequency amongst the four switches is equalized. Voltage stress and current stress of power switches in each cell is inherently equalized with this modulation. After every two fundamental periods, the HPWM pattern is changed so that the first module (S1, S2, S3, and S4) becomes the second module (S1_, S2_, S3_,and S4_), and the second one shifts to the first, and is shown in Fig. 4.

3 Simulation Results

The waveforms of output voltage, the implementation of HPWM circulation makes the inverter modules operate at same average switching frequency with the same conduction period. As a result, all inverter cells operate in a balanced condition with the same power-handling capability and switching losses. To evaluate the quality of the output voltage waveforms, the values of total harmonic distortion (THD) calculated up to 50th order of harmonics, as suggested in the IEEE standard 519.

$$\mathrm{THD}(\%) = 100\sqrt{\frac{\sum_{n=2}^{\infty} V_{n,\mathrm{rms}}^2}{V_{1,\mathrm{rms}}^2}}$$

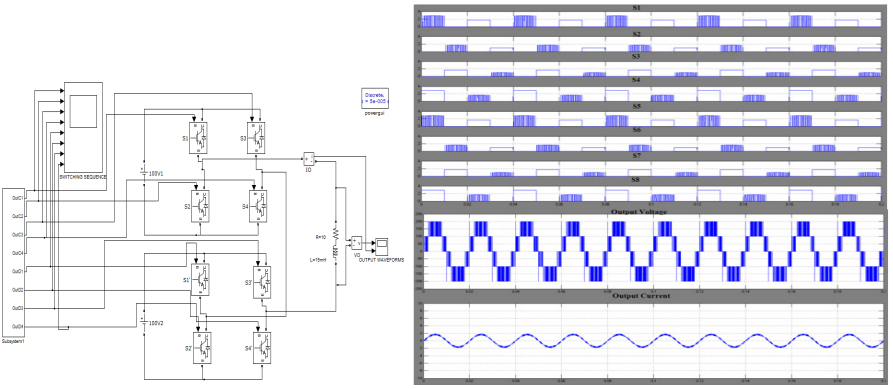


Fig. 4. Block diagram of five levels CMLI with hybrid switching states (s1-s8), output voltage and current

Table 1. Comparison of Spectrum Analysis in Different Mspwm Schemes

Types of MSPWM	THD (%)	Individual Harmonic Content (%)											
		2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th	13th
HAPOD	22.50	4.22	17.00	3.67	6.28	3.49	5.96	3.00	3.65	2.44	6.23	1.59	5.32
HSCMSP	41.43	7.75	31.75	7.30	17.27	6.55	10.64	5.64	6.01	4.56	4.45	3.50	3.09
HCBPWM	30.74	2.15	9.33	5.25	15.44	4.76	17.84	2.15	9.37	3.98	7.10	2.27	6.27
HPSCP	38.40	4.89	25.78	4.22	12.12	5.56	9.23	9.45	8.36	11.79	7.88	9.95	6.74

The harmonic spectrum of Vo and the %THD valve measured using FFT. From that, APOD is the efficient module among four MSPWM and its voltage harmonics (%THD of proposed module=22.50) is less when compared with others. It should also be observed that there in significant reduction in the THD content compared to the fundamental frequency techniques (%THD of exiting module).

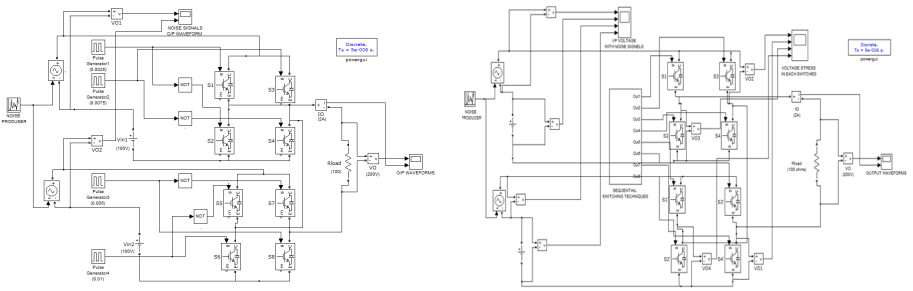


Fig. 5. Simulink model of CMLI: (a) Existing System, (b) Proposed System

From the Fig.5, the FFT analysis of proposed module (APOD), the voltage harmonics (%THD=25.19) is less when compared with existing module (fundamental frequency techniques) THD is 41.33%.

4 Conclusion

This paper has provided a brief analysis of cascaded multilevel inverter circuit topologies and a new family of SSHM techniques for CMLI, operating at a lower switching frequency is proposed. Compared to conventional MSPWM schemes, switching-loss reduction is obtained and the harmonic performance of the SSHM schemes are analyzed in the entire range of modulation index and it seems to be good. An early patent for the cascaded multilevel inverter can be traced back to 1975, but not available until the mid-1990s. Today, more and more commercial products are based on the MLI structure, and more and more worldwide research and development of multilevel inverter-related technologies is occurring.

Acknowledgement. I would like to thank my parents, Mr. C.N.Veeramani , friends, the teaching faculty and non-teaching staff members of Electrical and Electronics Department for their valuable support.

References

1. Govindaraju, C., Baskaran, K.: Performance analysis of cascaded multilevel Inverter with hybrid phase-shifted carrier modulation. Aust. J. Electr. Electron. Eng. 7(2), 121–132 (2010)
2. Rodriguez, J., Bernet, S., Wu, B., Pontt, J.O., Kouro, S.: Multilevel voltage-source-converter topologies for industrial medium-voltage drives. IEEE Trans. Ind. Electron. 54(6), 2930–2945 (2007)
3. Rodríguez, J., Lai, J.S., Peng, F.Z.: Multilevel inverters: A survey of topologies, controls and applications. IEEE Trans. Ind. Electron. 49(4), 724–738 (2002)
4. Gupta, R., Ghosh, A., Joshi, A.: Switching characterization of cascaded multilevel-inverter-controlled systems. IEEE Trans. Ind. Electron. 55(3), 1047–1058 (2008)
5. McGrath, B.P., Holmes, D.G.: Multicarrier PWM strategies for multilevel inverters. IEEE Trans. Ind. Electron. 49(4), 858–867 (2002)
6. Chithra, M., Bharathi Dasan, S.G.: Analysis of cascaded H-bridge multilevel inverter with photovoltaic arrays. In: Proceeding of the International Conference on Emerging Trend in Electrical and Computer Technology (ICETECH), pp. 442–447 (March 2011)
7. Zaragoza, J., Pou, J., Ceballos, S., Robles, E., Ibanez, P., Villate, J.L.: A comprehensive study of a hybrid modulation technique for the neutral point clamped converter. IEEE Trans. Ind. Electron. 56(2), 294–304 (2009)
8. Rashid, M.H.: Power Electronics: Circuits, Devices and applications, 3rd edn. Pearson Education (2004)
9. Mohan, N., Undeland, T.M., Robbins, W.P.: Power Electronics: Converters, Applications, and Design, 2nd edn (1995)
10. Bimbhra, P.S.: Power Electronics, 4th edn. Khanna Publishers (2010)

Efficient Path Finding Algorithm for Transmission of Data in Agricultural Field Using Wireless Sensor Network

Smitha N. Pai, K.C. Shet, and H.S. Mruthyunjaya

Manipal University, Manipal, N.I.T.K., Suratkal
India
{smitha.pai,mruthyuhs}@manipal.edu, kcshet@nitk.ac.in

Abstract. Transmission of data from the source to the sink requires an efficient path so that the network life is enhanced. Longer shelf life for battery is essential when it is difficult to replace the battery too often as in agricultural field. To achieve this, in this paper we have proposed an algorithm to find the efficient path between the sink (base station) and the source so that the network can be alive for one cropping season.

Keywords: Wireless sensor network, data transmission, agriculture.

1 Introduction

Sensors are responsible for capturing, analyzing and transmitting information using radio signals. Sensor networks contain a base station usually called a sink and enormous number of other sensors which carry out the task of sensing and transmitting, along with relaying information of other nodes. Sink is a high computation device used to gather information of the whole network and is normally connected to the Internet or other data processing equipment. Sensors are highly resource constrained, major being battery power, transmission range, data and program memory. Scarcity of water is a major issue faced by the world. An awareness of water consumption is very much essential to avoid depletion of water in the water bed. Knowledge of water usage if provided to a farmer could help them in significantly improving the way water is managed [1]. Sensors placed in the soil could help in detecting the soil moisture; and this data collected by the network could help in regulating the water flow to the fields [2]. Analysis of data collected could help in predicting the yield and any new strategies can be adopted to overcome any associated problems. In this paper new routing protocol is devised for enhancing the life of the sensor network placed in the agricultural field.

2 Related Work

Sensors are application specific and can be used to continuously sample physical data like humidity, temperature, soil moisture etc. COMMON-Sense is one such project

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 345–348, 2013.

associated with monitoring the regulation of water supply to the field [3]. This project uses AODV protocol for path finding, establishment and maintenance [4]. In this paper we try to improvise and make certain changes to AODV to suit the requirement for agricultural application using network simulator ns2.34 using a protocol named as NEWAODV.

3 Routing Algorithm

Reactive route finding algorithm maintains a routing table in each node. Though this application has no mobile nodes, routes are computed dynamically when the nodes lose their energy. Control packets are used to trace the efficient path and data packet for transmission of data.

3.1 Routing Table

The routing table is maintained in every node and has the following information. Previous node id, next node id, source address, the destination address (sink), pointer to neighbor's linked list of 4 byte each and sequence number (route discovery count), hop count from source to sink of 1 byte each.

3.2 Packet Type

NEWAODV_BEACON packet format is for route discovery, NEWAODV_PATH is for route establishment. Both packet format types have in common the source address (4 bytes), the time when the packet is sent (4 bytes), packet type (NEWAODV) of 1 byte. The route discovery contains the hop count (1 byte) which is the hop count from the source to the sink.

3.3 Path Discovery, Establish and Maintenance

The NEWAODV_BEACON packet is broadcast from the source to all its neighbors, the neighbors in turn broadcast the signal to its neighbor until it reaches the sink. Packets which are having higher hop count to any node from the source are dropped. Packets which come back to the source node after looping are dropped. If more than one packet reaches the sink with higher or same hop count, packets which reach later are dropped. Packets at any node having the same hop count, reaching through another route but arrive late are dropped.

NEWAODV_BEACON packet on reaching the sink waits for a time equal to 1.5 times the time taken by the first packet to reach the sink. This is required so that we wait for all packets with shorter route to reach sink. A unicast NEWAODV_PATH packet is then sent from the sink to the source establishing a reverse path on the way back. The path which is established between source and sink is used to continuously send the data from the source to the sink until any link failure occurs. On the occurrence of the link failure a new path is again established using the broadcast

signal NEWAODV_BEACON from the node where the link failure occurred. This process is continued until there is no path to the sink from the source.

3.4 Comparison of NEWAODV with AODV Protocols

In the case of AODV, during the process of path discovery, multiple RREP (Route Reply) packets, in response to a single RREQ (Route Request) packet lead to heavy control overhead. The periodic beaconing leads to unnecessary bandwidth consumption. There is high overhead when data packets gets delivered to too many nodes which are not destined to receive them [5]. In NEWAODV only one reply is sent for one route request. Periodic beaconing is removed and beacon signals are sent only when the route fails.

4 Simulation Results

Simulation is carried out using ns2.34. The topology used is hexagonal [6]. The initialization parameters are as per the requirement for the agricultural field, using Tiny node mote. Transmit and receive power for the radio link is set to 5dBm and-104dBm. Carrier sense and capture threshold is -104dBm and 10dBm. Height of the antenna is 1m. Two ray propagation model with the range of transmission 530m. and frequency of operation 915MHz. Transmit, receive and idle power for sensor as 19dBm, 15dBm, -20dBm. Initial energy in the battery is calculated as 20304J. Constant bit rate interval is 300secs [7]. Number of nodes used 448. Table 1 show the relative comparison of two routing protocol. In order to have similar features between the two routing protocols, the timers (beacon signals, neighbor purge, route purge) used in AODV were bypassed, but the data packets were not able to reach the destination.

Table 1. Relative comparison between two routing protocol

Variables used for comparison	NEWAODV	AODV without timer	AODV with timer
Total energy consumed	0.109710J	0.981866J	1.118299
Percentage energy consumed	0.000540J	0.004836J	0.005508
Idle	0.082880J	0.862066J	0.863075
Transmit energy	0.014588J	0.054941J	0.116819
Receive energy	0.012242J	0.065858J	0.138404
Packets reaching destination	960	0	960
Transmission and receptions	84866	612351	1787112
Drops with no route to destination	0	192	0

Using all the timers the results were calculated again, with data being able to reach the sink at the additional cost of energy. Source and sink were placed at extreme end opposite corners of the agricultural field to maximize the number of transmission and receptions.

The simulation results show that avoiding unnecessary transmission and reception of timer signals we are able to achieve an improvement in the amount of energy consumed in the case of NEWAODV. This is achieved by waiting for a time interval of 1.5 times the time required for the first packet to reach the sink and drop all other packets avoiding unnecessary multiple route reply packet from the sink to source. In the case of AODV removing the broadcast timers, neighbor purge, route cache purge, local repair timer we are not able to send the data to the sink. Introducing these had led to unnecessary overhead increasing the bandwidth and additional transmission and receptions consuming more energy.

5 Conclusion

The NEWAODV protocol is far efficient than the AODV for the current application. During link failure, currently new path is discovered from the location where the link is broken. Improvement to this can be obtained finding new path only upto the location where a path already exist.

References

1. Panchard, J., Rao, S., Prabhakar, T.V., Jamadagni, H.S., Hubaux, J.: COMMON-Sense Net: Improved Water Management for Resource-Poor Farmers via Sensor Networks. In: International Conference on Information and Communication Technologies and Development (2006)
2. COMMON Sense slides (2004), http://commonsense.epfl.ch/resources.html
3. Panchard, J.: PhD Thesis, Wireless Sensor Networks for Marginal Farming in India (2009)
4. Perkins, C.: Ad hoc On-Demand Distance Vector (AODV) Routing (July 2003), http://www.ietf.org/rfc/rfc3561
5. Rasheed, A., Mohammad, K.: Exploration and Comparison of Several AODV Implementations: A Survey. Communications of the ACS 2 (2009)
6. Pai, S.N., Shet, K.C., Mruthyunjaya, H.S.: Efficient Deployment Strategies for sensors placed in a sparse network. International Journal of Micro and Nano Systems 2, 11–14 (2011)
7. Pai, S.N., Shet, K.C., Mruthyunjaya, H.S.: Simulation Environment for Sensors placed in Agricultural Field. In: International Conference and Workshop on Emerging Trends in Technology (February 2012)

Analysis of a New Random Key Pre-distribution Scheme Based on Random Graph Theory and Kryptograph

Seema Verma[1] and Prachi[2]

[1] Department of Electronics, Banasthali University, Tonk, India
[2] Department of CSE & IT, ITM University, Gurgaon, India
seemaverma3@yahoo.com, prachi@itmindia.edu

Abstract. Wireless Sensor Networks (WSNs) vast myriad of futuristic applications makes it a matter of incessant research interest. Random key pre-distribution (RKP) seems to be best suited for WSN due to high security requirements and resource constrained nature. In this paper we present and implement a new RKP scheme on TinyOS. Later on, we perform a rigorous mathematical analysis of our scheme under ER (Erdos-Renyi) and kryptograph model. However, most of the previously presented schemes are based on ER model, our results prove that kryptograph model is more vital for secure WSNs.

Keywords: RKP, random graph theory, kryptograph.

1 Introduction

WSN compromises of huge number of ultra-small autonomous sensors capable of sensing physical and environmental conditions. They can be widely used in number of diversified applications, for instance, unattended surveillance, disaster management, nature reserves, in seismically threatened structures. For battery operated sensors, best optimization of communication and computation becomes mandatory. Further, when the thrust is miniaturization, provision of security features becomes challenging [1]. Especially, in hostile environment where privacy of data is very crucial and can be snooped easily, security becomes essential. To embark upon problem of security, viable and efficient cryptographic mechanisms are imperative, this escalate need for a key management algorithm efficient in all aspects like energy, memory etc. Generally, sensor networks rely on symmetric key algorithms to avoid the high computation cost of public key crypto-systems such as Diffie-Hellman key exchange [2].

We present a new static RKP approach to remove overhead of dynamic key generation and dependency on base station. Our scheme can be easily embedded in various pre-existing schemes to enhance their performance. Here, we discuss a few probabilistic approaches proposed earlier for WSN.

Eschenauer et al. in [3] pioneered first innovative RKP approach (basic scheme) for WSNs. Node exchange keys with neighbors. Nodes sharing key are securely linked. Else they establish path key through intermediate node sharing key with both nodes. Based on [3] Chan et al. [4] extended the idea and proposed q-composite key pre-distribution to increase network resilience at the cost of processing overhead. In

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 349–352, 2013.

q-composite, pairwise key is computed from at least q (q>=1) shared key between nodes. Hashed RKP was proposed by T. Shan et al. [5]. Keys are hashed different times for distinct nodes. Only first node get original key, j^{th} node receives (j-1) time hashed version. This also increases network resilience as hashed key of captured node get compromised, not original key and hashed versions. Two methods were proposed by Law et al. [6] key redistribution scheme, forward and reverse key distribution to avoid path key establishment. S. Zhu et al. introduced Pairwise Key Establishment protocol [7] to avoid communication overhead involved in Shared-Key Discovery. Instead of randomly distributing keys pseudo random function uses node's ID as seed. Any node can determine keys of another only by its ID. JianBo Fu et al. [8] modify q-composite RKP scheme based on the kryptograph. Jianmin Zhang et al. [9] dealt with the problem of small number of compromised nodes may affect a large fraction of keys. But, despite much research effort key agreement remains an open problem.

Rest of the paper is organized as: Our new scheme is presented in section 2. Section 3 elaborates mathematical analysis for both ER and kryptograph model. Simulation results for two previous models are followed by section 4. Finally, section 5 draws conclusion.

2 Our New Scheme

In this paper, we implemented our scheme using TinyOS mote simulator (TOSSIM) [10] to provide a high fidelity simulation for WSN. We also use TinyViz (GUI tool for TOSSIM) for visualization of our scheme and exchange of messages. Here, sensors are randomly distributed over a geographical area. In key pre-distribution phase, we generate key pool of (K) keys along-with their IDs. Afterwards, assign subset of k keys (k<=K) to each node. In shared key discovery phase, instead of broadcasting list of key IDs send key IDs only to nodes we actually want to communicate. Communication with desired neighbors makes our scheme communication efficient. For instance, suppose A and B are within communication range and want to communicate then A send its list of key IDs to B, if they share key then B send shared key ID to A. Key IDs doesn't give attack opportunity to adversary for traffic analysis attack. Now, A and B communicate by encrypting message through shared key. If don't share key B send its list of key IDs to A. During path key establishment, A broadcast list of key IDs of A and B. Intermediate node determine if it share keys with both or not. If not, it further broadcast message. Otherwise, if an intermediate node(C) has common key k_1, k_2 with A and B respectively then instead of issuing an unused key from C's key ring, C randomly select key from key pool. In basic scheme if C gets captured communication between A and B also gets compromised. However in our case, even if C gets captured communication remains intact. Further, after getting new key (k_3) encrypt k_3 with k_1 and k_2 to enhance security using RC5 algorithm. Now, send these encrypted values with key IDs of shared keys are sent to A. Now, firstly A determine which of its key from $k_1 \ldots k_k$ is common with C and decrypt $E_{k1}(k_3)$ and forward $E_{k2}(k_3)$ to B. Now, B decrypt $E_{k2}(k_3)$ to get key k_3 which from now on act as shared key between A and B. As path using C is used only once, effect of traversing hops to set up path key is negligible.

3 Analysis of Scheme Based on ER and Kryptograph Model

Here, we perform mathematical analysis of our scheme based on random graph theory proposed by ER and kryptograph model. In ER model [11], p will be provided as:

$$p = \ln(N)/N + c/N \tag{1}$$

where p=probability of connectivity, N=number of nodes and c=real constant.

As a result, key pool size K is computed as,

$$p = 1 - ((K - k)!)^2/((K - 2k)!P!) \tag{2}$$

Kryptograph is more realistic and powerful as compared to ER model since ER assumes that any two nodes can be directly linked, irrespective of geographical location. But, it's a well-known fact that WSN consider communication range (r) and so is kryptograph (perfectly goes with WSN). Also, ER model assume edges exist independently. However, to make WSN connected whether a new edge required to be inserted or not is dependent on connectivity provided by previously inserted edges.

Connectivity and resiliency always remain in conflict in WSN. So, to achieve them simultaneously with kryptograph [12] we determine connectivity in terms of N:

Probability that link exists between nodes is roughly equal to k^2/k.

$$\Pr[\text{link exists}] = 1-(1-k/K)^k \sim k^2/K \tag{3}$$

$$k^2/K \sim \log N/N \tag{4}$$

If K>=N and satisfies equation (4) then network is connected with high probability. A network become resilient if k/K~logN and we ensure all that for our network.

4 Simulation Results

Lastly, we carried out simulation to support theoretical results. Fig. 1 show graphs of connectivity for our scheme. In both graphs, modify k for various K to check p. Graph illustrate that connectivity based on kryptograph is much better than ER model.

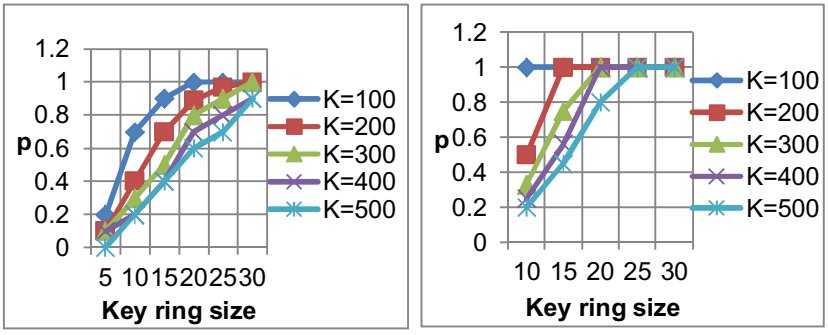


Fig. 1. Probability of connectivity of our scheme based on **ER model** and **Kryptograph**

5 Conclusion

In this paper, we implemented a new RKP scheme on TinyOS for WSN which is more secure and communication efficient as compared to previously presented schemes. Further, we discussed random graph theory through ER and kryptograph model in context of WSN. Later on, we performed rigorous mathematical analysis of our scheme based on ER and kryptograph model. Simulation results conclude that scheme based on kryptograph has more probability of connectivity and has better resilience and fault tolerance as compared to scheme based on ER model.

References

1. Prajapati, K., Nyathi, J.: An Efficient Key Update Scheme for Wireless Sensor Networks. In: Proceedings of International Conference on Wireless Networks, Las Vegas, USA, pp. 8–14 (2006)
2. Diffie, W., Hellman, M.: New directions in cryptography. IEEE Transactions on Information Theory, 644–654 (1976)
3. Eschenauer, L., Gligor, V.D.: A keymanagement scheme for distributed sensor networks. In: Proceedings of the 9th ACM Conference on Computer and Communications Security, New York, USA (2002)
4. Chan, H., Perrig, A., Song, D.: Random key predistribution schemes for sensor networks. In: IEEE Symposium on Security and Privacy, pp. 197–213. Carnegie Mellon Univ., PA (2003)
5. Shan, T., Liu, C.: Enhancing the key pre-distribution scheme on wireless sensor networks. In: IEEE Asia-Pacific Conference on Services Computing, pp. 1127–1131. IEEE Computer Society, Los Alamitos (2008)
6. Law, C.-F., Hung, K.-S., Kwok, Y.-K.: A novel key redistribution scheme for wireless sensor networks. In: IEEE International Conference on Communications, pp. 3437–3442. IEEE Computer Society, Washington, DC (2007)
7. Zhu, S., Xu, S., Setia, S., Jajodia, S.: Establishing pairwise keys for secure communication in ad hoc networks: a probabilistic approach. In: Proceedings of the 11th IEEE International Conference on Network Protocols, pp. 326–335. IEEE Computer Society, Washington, DC (2003)
8. Fu, J.B., Li, Q.L., Li, S., Ssanyu, L.: A Modified Q-Composite Random Key Predistribution Scheme Based on Kryptograph. In: International Conference on Computer Application and System Modeling, Taiyuan, pp. V15-240–V15-243 (2010)
9. Zhang, J., Cui, Q., Liu, X.: An Efficient Key Management Scheme for Wireless Sensor Networks in Hostile Environments. In: International Conference on Multimedia Information Networking and Security, Hubei, pp. 417–420 (2009)
10. Levis, P., Lee, N., Welsh, M., Culler, D.: TOSSIM: Accurate and Scalable Simulation of Entire TinyOS Applications. In: Proceedings of the First ACM Conference on Embedded Networked Sensor Systems (SenSys), New York, USA (2003)
11. Spencer, J.: The Strange Logic of Random Graphs. In: Algorithms and Combinatorics, vol. 22. Springer (2000)
12. Pietro, R.D., Mancini, L.V., Mei, A., Panconesi, A., Radhakrishnan, J.: Redoubtable Sensor Networks. ACM Transactions on Information and Systems Security 11(3), Article 13 (2008)

An Algorithm for Variable Precision Based Floating Point Multiplication

Rohit Sreerama, Satish Paidi, and Neelima Koppala

Sree Vidyanikethan Engineering College, Tirupathi
{rohit.sreerama,paidysatish,koppalaneelima}@gmail.com

Abstract. An algorithm for computing the variable precision multiplication is proposed in this paper. In the algorithm a parallel multiplier of size m is used to compute the multiplication of two numbers which are of n×m bits. Multiple precision floating point format is used to represent the numbers, only the mantissas are considered in the multiplication as the exponents are easily obtained by adding the exponents of the two operands to be multiplied. By using this algorithm the partial products obtained in the multiplication are added as soon as they are computed, resulting in the use of the lowest memory for storage of intermediate results, so bits size of the final result will be 2n×m.

Keywords: variable precision, floating point multiplication, parallel multiplier, Karatsuba algorithm.

1 Introduction

The speed of computation in computers has increased dramatically during the last decade. This increase in speed is due to the development of VLSI technology that has enabled the integration of millions of transistors on the same chip. The parallelism and the pipelining are used in operations at the hardware level to attain the current performances in terms of computation speed [1]. Even though the computation speed has increased the accuracy in the computations are not yet developed to that extent. The main problem of these computing capabilities is the accuracy of the results. Accuracy plays a vital role in the computation of large operands. In many applications, such as scientific computing, the large number of arithmetic operations and the reliance placed on the results makes it important to provide accurate and reliable results without errors. An error in applications can be very costly. For example, the destruction of the Ariane 5 rocket during its first flight in 1996 [7] or the overdoses radiation administered to patients in the National Oncology Institute of Panama in 2000 [8]. These are caused because of the accumulation of rounding errors and catastrophic cancellation that can lead quickly to results that are completely inaccurate. The variable precision increases the efficiency and accuracy of conventional arithmetic. The variable precision computing allows the precision of the computation to be varied based on the problem to be solved and the required result [3].This paper presents an algorithm for variable precision multiplication.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 353–357, 2013.

2 Variable Precision Floating Point Representation

The figure 1 represents the format for the variable precision floating point**.** Each variable precision number consists of 16 bit exponent field (E), a sign bit (S), a two bit type field (T) and a five bit length (L) and significant field (F) which contains L+1 significant words (F(0) to F(L)).The sign bit is zero if the word is positive and is one if the word is negative. The exponent is represented as two's compliment integer, the type field contains whether the number is normalized, infinite, zero or not a number. The length bit shows the number of m bit words present in the significant. The words in the significant are stored in the format of most significant F (0) to least significant F (L) [5].

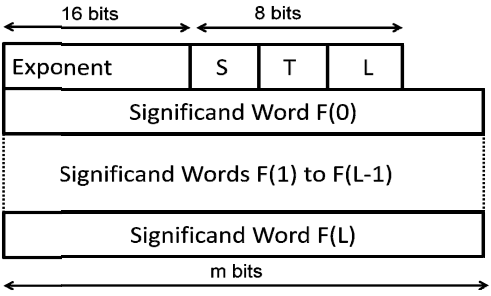


Fig. 1. Representation of variable precision

3 Classic Multiplication Method

In the classic multiplication all the partial products are generated first and then addition of the obtained partial products will take place. The classic method for variable precision multiplication is performed by using the multiplier, adder, and a long accumulator to store partial products [4][6]. Here we need large memory to store all the partial products that are generated before adding. If the size of the operand increases then the complexity of the classic multiplication also increases. To multiply two n word variable precision numbers, n^2 partial products are generated and accumulated.

4 Karatsuba Algorithm

It is a fast algorithm for multiplication. It reduces the multiplication by splitting the multiplier and the multiplicand into two parts that is least significant part and most significant part. Here the basic step of Karatsuba algorithm is the formula that allows us to multiply two large numbers X and Y using smaller multiplications [5]. Using this algorithm, n-bit multiplications are divided into n/2-bit multiplications. The Karatsuba algorithm is mainly based on the replacement of multiplication by two

additions. As it reduces the overall multiplication it is faster than the classic multiplication method. The complexity of the Karatsuba algorithm is low T(n) $=O(n^{log3})$ [9]. Though the overall complexity in the multiplication was reduced the Karatsuba algorithm is complex in hardware implementation and the complexity depends on the implementation devices. In the FPGA implementations the routing complexity increases with the increase in the size of the operand.

5 Proposed Algorithm

The proposed algorithm is based on classic multiplication and also has the advantages of Karatsuba algorithm. This algorithm reduces the memory that is used to store the partial products that are generated during computation in classic multiplication method by adding the partial products as soon as they are computed. This algorithm only uses the memory of (n x 2m) bits instead of (n^2 x 2m) bits that are used in the classic multiplication. This algorithm splits the operands A and B and the result into m bits. Depending on the value of the m the size of the multiplier and the memory are considered. The splitting of operand is shown in the figure 1.

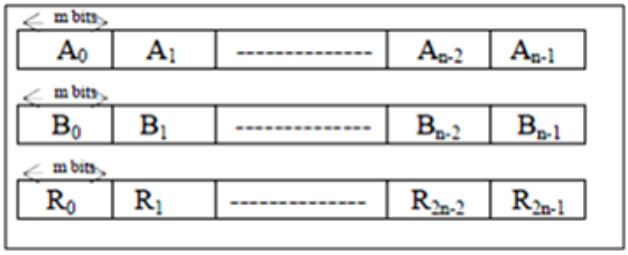


Fig. 2. Splitting of Operand in m-Bit Word

Variable Precision Multiplication Algorithm
Inputs A, B
Input m
Output R= A×B
Initialization R= 0
For j=0 to n-1 do
For i=0 to n-1 do
$R=R+ (A_i \times B_j) \times 2^{(i+j)}$
End for
End for
Return $R = (R_0, R_1, \ldots, R_{2n-1})$

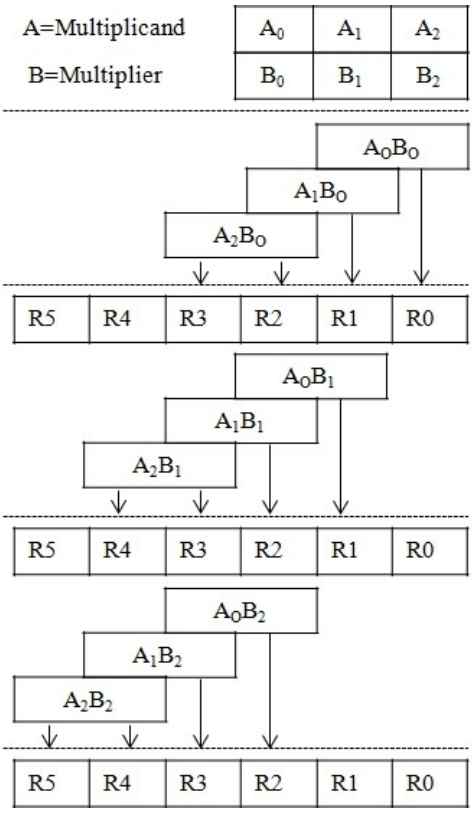


Fig. 3. The proposed method of computing the variable precision multiplication for 3 m bits operands

The routing problem in Karatsuba algorithm is solved here as we are using classic multiplication as base for algorithm but the theme of splitting the digits is taken from the Karatsuba algorithm so this proposed algorithm will be faster than the classic algorithm and will have all the advantages of the Karatsuba algorithm and can be implemented on hardware**.**

6 Implementation Results

The algorithm was implemented for 32 bit operands using Xilinx ISE 10.1. It has been mapped and routed on Xilinx FPGA circuit of family Spartan 3E XA3S250E. The implementation results are tabulated below.

Table 1. Occupation rate on XA3S250E FPGA Device

Logic Utilization	Occupation Rate	%
Slices	514/2448	20%
4 input LUTs	829/4896	16%
Bonded IOBs	128/172	74%
Multi 18x18SIOs	12/12	100%

7 Conclusion

In this paper an algorithm for variable precision multiplication was proposed and it has the advantages of both the classic multiplication and the Karatsuba algorithm. The design was implemented using Verilog HDL and verified using extensive directed-random vectors. This multiplier is suitable to be used in applications where a large dynamic range is required or in rapid prototyping applications where the required number range has not been thoroughly investigated.

References

1. Ali, H., Al-Hashimi, B.M.: Architecture Level Power-Performance Tradeoffs for Pipelined Designs. IEE Circuits, Devices and Systems Series 18 (2006)
2. Erle, M.A., Hickmann, B.J., Schulte, M.J.: Decimal Floating-Point Multiplication. IEEE Transactions on Computers Issue Date 58(7) (July 2009)
3. Smith, D.M.: Using Multiple Precision Arithmetic. Computing in Science & Engineering (July/August 2003)
4. Schulte, M.J., Swartzlander, E.E.: A Family of Variable-Precision Interval Arithmetic Processors. IEEE Trans. on Computers 49(5) (May 2000)
5. Erdem, S.S., Koç, Ç.K.: A Less Recursive Variant of Karatsuba-Ofman Algorithm for Multiplying Operands of Size a Power of Two. In: 16th IEEE Symposium on Computer Arithmetic (2003)
6. Siva Kumar, G., Srinivasa rao, B.K.N.: Design of Area and Power Efficient Floating-point Co-processor. (IJAEST) International Journal of Advanced Engineering Sciences and Technologies 10(2), 249–251
7. Arnold, D.N.: The Explosion of the Ariane 5 (2000), http://www.ima.umn.edu/~arnold/disasters/ariane.html
8. WISE News Communique, Radiological accident in Panama, http://www10.antenna.nl/wise/index.html, http://www10.antenna.nl/wise/549/5278
9. Karatsuba, A., Yu, O.: Multiplication of multiple numbers by mean of automata. Dokadly Akad. Nauk SSSR 145(2), 293–294 (1962)

Service Crawling in Cloud Computing

Chandan Banerjee[1,2], Anirban Kundu[2,3], Sumon Sadhukhan[1], Shoubhik Bose[1], and Rana Dattagupta[4]

[1] Netaji Subhash Engineering College, Kolkata 700152, India
{chandanbanerjee1,sumon.sadhukhan8,sbose78}@gmail.com
[2] Innovation Research Lab (IRL), Howrah, West Bengal 711103, India
anik76in@gmail.com
[3] Kuang-Chi Institute of Advanced Technology, Shenzhen 518057, P.R. China
anirban.kundu@kuang-chi.org
[4] Jadavpur University, Kolkata 700032, India
rdattagupta@cse.jdvu.ac.in

Abstract. The cloud service user of a single service relies on specific service capabilities to query relevant data. It can be difficult to consistently query since data is distributed across multiple services. In this paper, we are going to develop an algorithm to crawl the services within a cloud service inventory invoking fetch capabilities and following Internet protocol addresses among the resources assembling retrieved data in a centralized query enabled indexing service. Crawling techniques are being used for pre-cache information [1] that a user needs for post processing information at current available resource(s) for improving user latency to safe requests.

Keywords: Cloud Computing, Service Crawling.

1 Introduction

Web search engines [2] crawl the Web and update information since World Wide Web (WWW) is changing its shape based on information and geographically distributed locations. Search Engines [3] are often compared based on the indexing percentage of total Web sphere. Typical model of Search Engine does not scale as the network grows. Search Engines are now facing the challenge of intelligent storing of mass data within their own network for fast processing with scalable facilities. Cloud introduces mass data storage & processing. Convergence of cloud systems and Web bring in several new challenges. One of the important challenges is to generate service analytics in real time. A Web crawler [4] is an automated script for crawling through Web pages for creating index of searching data. Web crawlers typically work based on Unified Resource Locator (URL) to cache Web pages by recursive traversal [5]. Cloud crawler works with Internet Protocol (IP) addresses of a cache stored in a tree structure. Hosts are visited using specific threads for specific networks.

Section 2 shows our proposed framework and the corresponding approach. Experimental results are presented in Section 3. Conclusion is depicted in Section 4.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 358–361, 2013.

2 Proposed Approach

A cloud crawler is used to fetch the services for creating a framework of cloud service crawler engine using proper indexing methodologies. A crawler for a specific service is a program for extracting outward Web links (URLs) and further adding them into a list after processing. Thus, a cloud service crawler is a program which fetches as many relevant services as possible for the specific users. It uses the Web link structure in which the order of the list is important, because only high quality Web pages are considered as relevant. Fig. 1 shows the proposed service based cloud crawler. Here, an element insertion means that the element is inserted at the pointer location within the m-way tree. A special traversal technique is utilized for visiting all the nodes within each network or sub-network. Each node is selected twice. Second time it is actually popped from stack. An advantage of our algorithm is that data need not to be stored in the client node. The result is directly sent to the crawler server after scanning a single node.

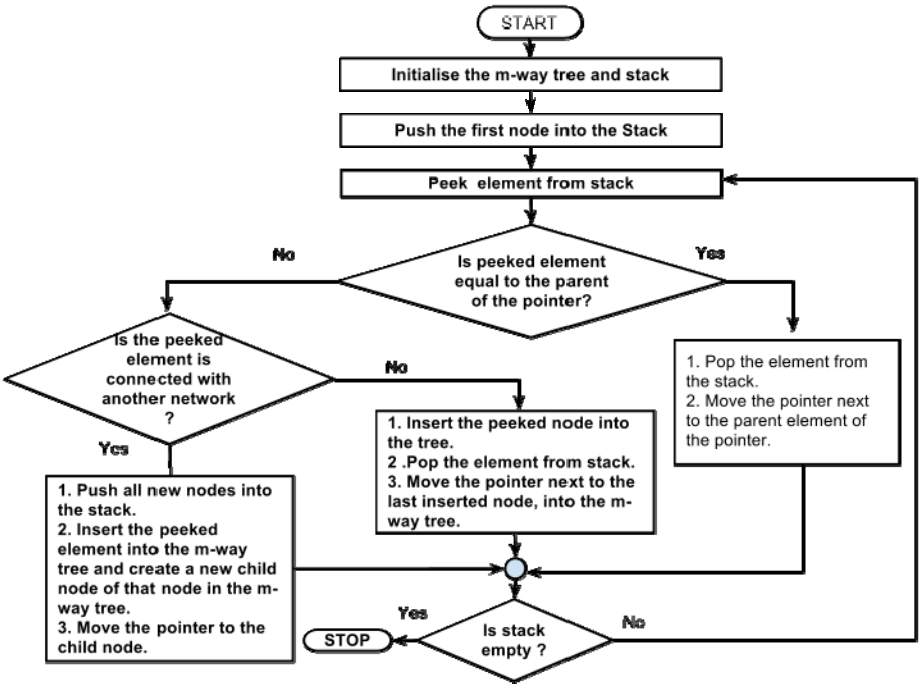


Fig. 1. Flowchart of Service based Cloud Crawler

Fig. 2(a) shows the basic framework used in cloud service crawler engine. Fig. 2(b) shows an arbitrary cloud cluster. There are total four network clusters within a cloud. Circular boxes indicate the clusters and rectangular boxes indicate the resources or slaves of each cluster network. Table 1 show the result which is based on our proposed approach as shown in Fig. 1.

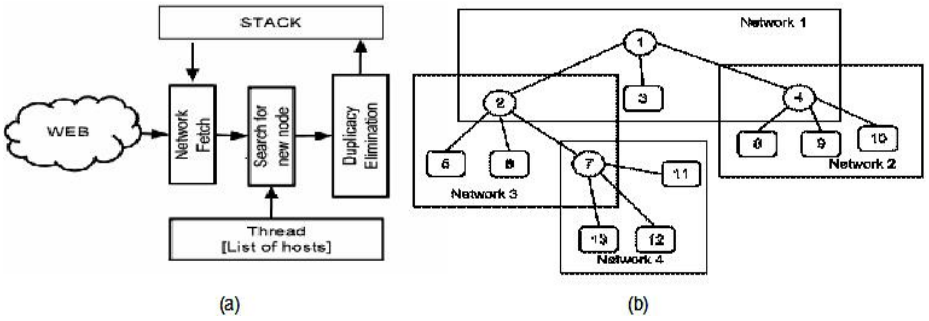


Fig. 2. (a) Cloud Service Crawl Engine; (b) Arbitrary Cloud Cluster Scenario

Table 1. Proposed approach based on Fig. 2(b)

Stack	Peek	Result	Stack	Peek	Result	Stack	Peek	Result
Empty			1			1,2,3,4	1	
1,2,3,4 ,8,9,10	4		1,2,3,4, 8,9,	10		1,2,3,4, 8	9	
1,2,3,4	8		1,2,3	4		1,2	3	
1,2,5,6 ,7	2		1,2,5,6, 7,11,12 ,13	7		1,2,5,6, 7,11,12	13	
1,2,5,6 ,7,11	12		1,2,5,6, 7	11		1,2,5,6	7	
1,2,5	6		1,2	5		1	2	
Empty	1				Final result as M-Way Tree			

3 Experimental Results

Four clusters have been used for experimental purpose using tree traversal as shown in Fig. 3 using cloud crawler based on IP addresses of cache. Threads have been utilized to visit distinct hosts in a concurrent manner.

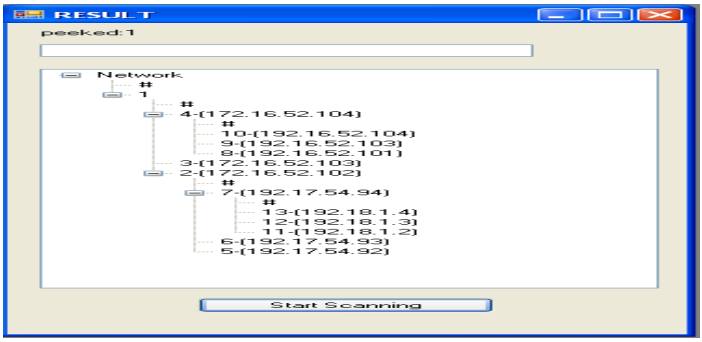


Fig. 3. Crawling results

4 Conclusion

In this paper, service crawling has been discussed. There is no need to store data into client node as result is directly sent to crawler server scanning each node. Cloud crawler works with IP addresses of a cache following an m-way tree structure.

References

1. Alonso, R., Barbara, D., Garcia-Molina, H.: Data caching issues in an information retrieval system. ACM Transactions on Database Systems 15(3), 359–384 (1990)
2. Brin, S., Page, L.: The anatomy of a large-scale hyper textual Web search engine. Computer Network ISDN Syst. 30, 107–117 (1998)
3. Yang, K.-H., Pan, C.-C., Lee, T.-L.: Approximate search engine optimization for directory service. In: Parallel and Distributed Processing Symposium, Dept. of Comput. Sci. & Inf. Eng., Nat. Taiwan Univ., Taipei, Taiwan (2003)
4. Lu, J., Wang, Y., Liang, J., Chen, J., Liu, J.: An Approach to Deep Web Crawling by Sampling. In: Web Intelligence 2008, pp. 718–724 (2008)
5. Gupta, P., Johari, K.: Implementation of Web Crawler. In: 2nd International Conference on Emerging Trends in Engineering and Technology, pp. 838–843 (2009)

Data-Retention Sleep Transistor CNTFET SRAM Cell Design at 32nm Technology for Low-Leakage

S. Rajendra Prasad[1], B.K. Madhavi[2], and K. Lal Kishore[3]

[1] Dept. of ECE, ACE Engineering College, Hyderabad, A.P., India
srprasad447@gmail.com
[2] Dept. of ECE, GCET, Keesara, Hyderabad, A.P., India
[3] Dept. of ECE, JNT University, Hyderabad, A.P., India

Abstract. In today's VLSI Circuits design, one of the key challenges is the increase in power dissipation of the circuits, which in turn shortens the service time of battery-powered electronics, reduces the long-term reliability of circuits due to temperature-induced accelerated device and interconnects aging processes, and increases the cooling and packaging costs of these circuits. Leakage power accounts for an increasingly larger portion of total power consumption in nanometer technologies. A CNTFET is the analogue of silicon MOSFET in which CNTs replace the silicon channel. This paper proposes a new circuit level technique called Data-Retention Sleep transistor method to reduce the standby leakage current in the CNTFET SRAM Cell Design. This technique reduces leakage power by significant amount compared to sleep transistor method with minimal area and delay overhead.

Keywords: SRAM Cell, Leakage-Power, CNTFET, Data-Retention Sleep Transistor, HSPICE.

1 Introduction

Integrated circuit (IC) design has always been driven by the demand for having more functionality integrated on a single chip. In Today's System-on-a-Chip (SoC) designs, this functionality includes multiple processor cores, on-chip memory, audio/video encoder/decoder, various I/O controllers, RF front-end, signal processing engines, and multiple voltage regulators. To meet this demand, the semiconductor industry has successfully followed the Moore's Law, resulting in tremendous advances in CMOS manufacturing processes. The accompanying down scaling of the minimum feature sizes has enabled us to double the number of transistors every 15-18 months. One side effect of technology scaling is that the Critical Dimension1 (CD) has become so small that the atomicity of the physical features and dopant levels is becoming assessable. This results in large variations in the physical and electrical characteristics of interconnect and transistors which in turn affect the performance and power consumption of the circuit. Traditionally, process variations have been modeled by considering the worst-case process corners in order to evaluate the performance of the design [1].

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 362–368, 2013.

Yet another consequence of technology scaling is that integrated circuit densities and operating frequencies are continuing to go up. The result is that chips are becoming larger, faster, and more complex, therefore, consuming ever larger amounts of dynamic power [2]. At the same time, CMOS scaling toward nano technologies requires very low threshold voltages and ultra-thin gate oxides to retain the current drive and alleviate the short-channel effects. The side effect of threshold voltage and oxide thickness scaling is an exponential increase in both subthreshold and tunneling gate leakage currents, which adds to total power consumption of the chip.

Given the importance of low-power design, this paper is proposes a new technique at circuit-level for low-power CNTFET SRAM cell design.

2 The Carbon Nanotube FET

Demand of scaling down semiconductor devices and integrated circuits into nanometers causes semiconductor industry faces to difficult challenges such as: increased short-channel effects, reduced gate control, exponentially rising leakage currents, severe process variations, and unmanageable power densities. As a consequence, using Carbon Nano-Tube Field Effect Transistors (CNTFET) is as an optimized Nano-scaled device for implementing high performance, very dense and low power circuits [3-4].

The core of a CNTFET is made of carbon Nano-tube. CNFETs, similar to MOSFETs, have Ptype and N-type devices, and 4 terminals. The great advantage of CNTFET devices is that the P-type and N-type CNTFETs with the same device size have the same mobility, which can significantly simplifies the process of transistor sizing, particularly in complex circuits with a large number of transistors. The most promising type of CNTFET for using as transistor is Single-Wall Carbon Nano-Tube (SWCNT) which consists of only one cylinder. This mentioned type can act as either a conductor or a semiconductor depending on the angle of the atom arrangement along the tube. This is called the Chirality vector and is represented by the integer pair (n, m). One of the most superiority of CNTFET is the ability of adjusting threshold voltage. The threshold voltage is defined as the voltage level required to switch on a transistor and can nearly be set according to our requirements. This capability helps us to alleviate the V_{th} loss during the implementation [5].

3 The Proposed CNTFET SRAM Cell Design

Aggressive CMOS scaling has resulted low threshold voltage and thin oxide thickness for transistors manufactured in nano technology regime. As the memories in current technologies occupy a large portion of chip area, their static power dissipation is one of the major components of power dissipation in chips because the leakage power is proportional to the area of a circuit. In the past much research has been conducted to address the problem of leakage power consumption in

SRAM's. Among the dynamic techniques, gating techniques [6-10] have been proven very effective in power reduction. Recently authors in [11] applied a sleep transistor circuit level leakage reduction technique to SRAM cell design based on new emerging technology called CNTFET. The key idea of the sleep transistor CNTFET SRAM cell is to disconnect the cell from the supply voltage V_{DD} and ground GND in the standby mode. This is done by using footer NMOS sleep transistor and Header PMOS sleep transistor. The drawback of this technique is that the virtual ground node may get charged to V_{DD} or the virtual Supply node may get discharged to '0'. Because of this the stored bit may get destroyed. So this technique cannot save the state during stand by mode. Possible solution for this drawback is to strap the virtual ground or virtual supply to a fixed voltage node, which is either virtual ground or virtual power supply node, to a voltage such that the voltage difference between the rails of CNTFET SRAM cell becomes greater than the data retention voltage (DRV) [10] so that it can retain the stored data.

This paper proposes a new circuit level leakage reduction technique called "Data-Retention Sleep Transistor" technique that retains its state in stand by mode. This new technique is applied to CNTFET SRAM cell as shown in fig. 1. This technique uses two extra transistors, one in the supply path and other in the ground path and by selecting appropriate voltage values for biasing the virtual ground and virtual supply nodes, we can achieve higher power saving. During the standby mode, if the voltage difference on the power supply rails of the CNTFET SRAM cell is fixed, the proposed data retention sleep transistor technique achieves significantly higher leakage power savings compared to sleep transistor technique. Leakage currents of the CNTFET SRAM cell vary with the exact values of its supply V_{DD} and ground GND voltages. As a result, optimum virtual ground VV_{GND} and virtual supply VV_{DD} voltage levels exist for which the SRAM cell leakage power is minimized.

3.1 Operation of Data-Retention Sleep Transistor CNTFET SRAM Cell

For the cell to be accessed for a read or write operation, the 'Sleep' signal becomes '0' and the 'SleepBar' signal becomes '1', so that sleep transistors M7 and M8 are ON, which causes the voltage of the virtual supply and virtual ground nodes to become V_{DD} and 0, respectively. As soon as the operation is completed, the WL goes to '0', which means that the 'Sleep' signal becomes '1' and the 'SleepBar' signal becomes '0', and the corresponding cell row enters the standby mode. In this state, the strapping transistors M9 and M10 turn ON and the voltage of virtual ground and virtual supply become VV_{GND} and VV_{DD}, respectively. The SRAM cell leakage power is lowered due to source body biasing of the pull-up, pull-down, and access transistors. In Data-Retention Sleep Transistor SRAM cell, like Sleep Transistor cell, having a smaller potential difference between the two rails of the cell, i.e., $\Delta V = VV_{DD} - VV_{GND}$, in the standby mode results in lower leakage; however, it also makes the cell more susceptible to noise in the standby mode.

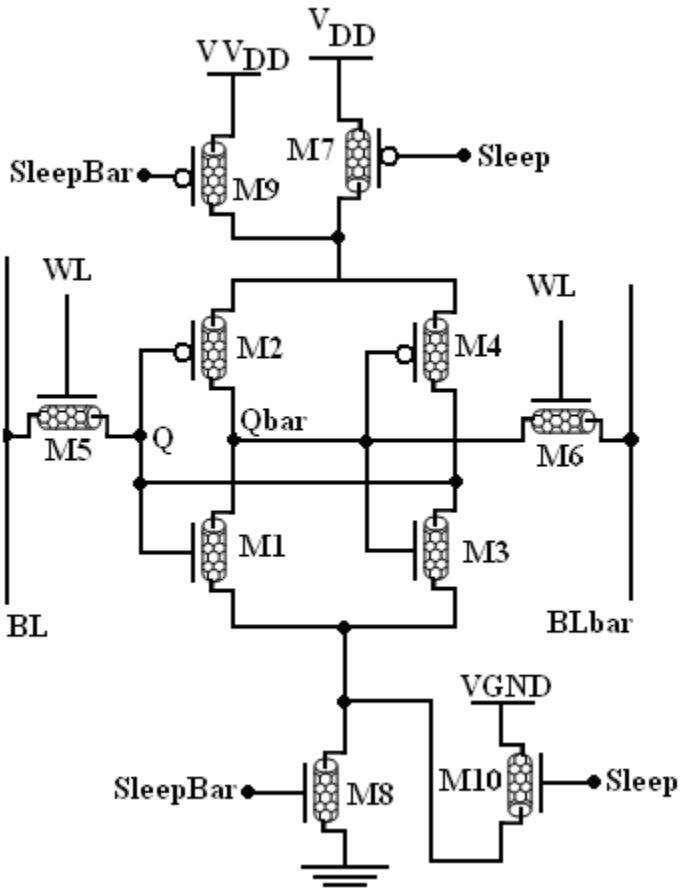


Fig. 1. The proposed CNTFET SRAM Cell based Cache

4 Simulation Results and Discussions

The CNTFET SRAM Cell is designed at 32nm technology by applying this proposed Data-retention sleep transistor technique. This circuit is simulated in HSPICE using Stanford CNTFET model at 32nm feature size with supply voltage V_{DD} of 0.9V [12]. The following technology parameters are used for simulation [13]: Physical channel length (L_channel) = 32.0nm; The length of doped CNT source/drain extension region (L_sd) = 32.0nm; Fermi level of the doped S/D tube (Efo) = 0.6 eV; The thickness of high-k top gate dielectric material (T_{ox}) = 4.0nm; Chirality of tube (m, n) = (19, 0); CNT Pitch = 10nm; Flatband voltage for n-CNTFET and p-CNTFET (V_{fbn} and V_{fbp}) = 0.0eV and 0.0eV; The mean free path in intrinsic CNT (L_{ceff}) = 200.0nm; The mean free path in p+/n+ doped CNT = 15.0nm; The work function of Source/Drain metal contact = 4.6eV; CNT work function = 4.5eV.

Table 1. Lekage Current through OFF transistors when Q=1

Sl. No.	Leakage Currents (A)	Conventional 6T CNTFET SRAM Cell	6T CNTFET SRAM Cell With Sleep Transistors	6T CNTFET SRAM Cell With Data-Retention Sleep Transistors
1	I (M1)	1.767e-06	1.130e-08	1.210e-09
2	I(M4)	7.133e-07	2.365e-08	1.184e-09
3	I(M6)	6.789e-06	3.168e-07	2.275e-08
4	Total Leakage	7.843e-06	3.292e-07	2.0268e-07

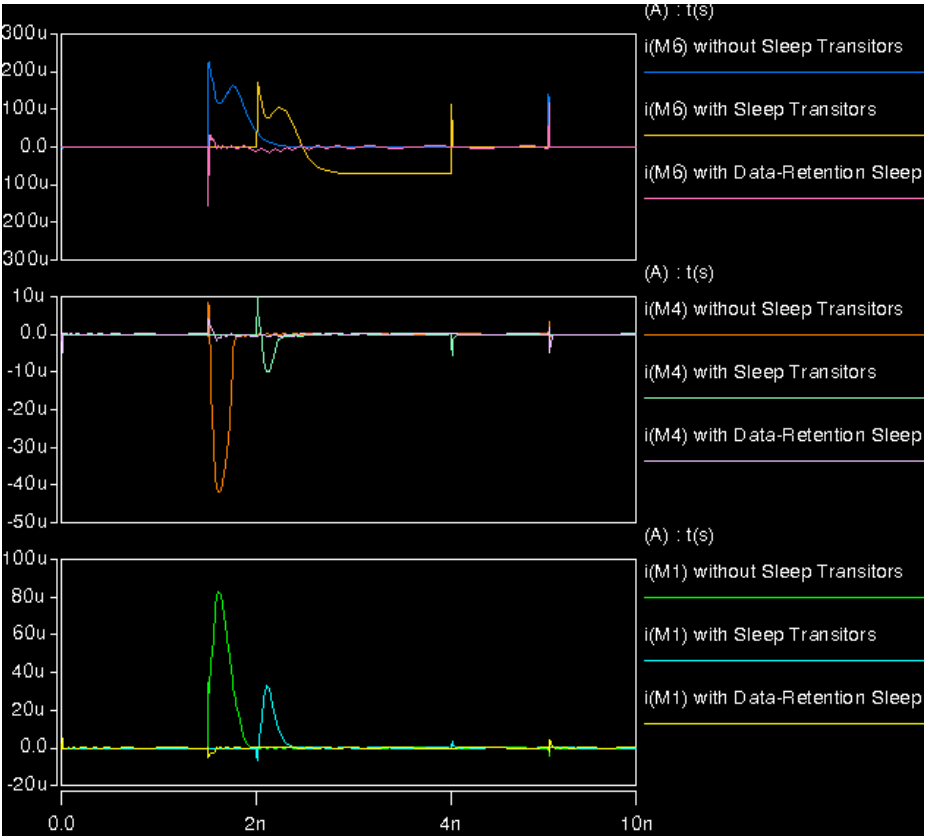


Fig. 2. Leakage Currents flowing through OFF Transistors when Q='1'

In data retention sleep transistor CNTFET SRAM cell having a smaller potential difference between the two rails of the cell, i.e., $\Delta V = VV_{DD}\text{-}VV_{GND}$, in the standby mode results in lower leakage; however, it also makes the cell more susceptible to noise in the standby mode. Notice that if ΔV is too low, all six transistors will work in the sub-threshold region, but as long as ΔV is greater than the Data Retention Voltage (DRV) of the cell, the data is retained. If $\Delta V = VV_{DD}\text{-}VV_{GND}$ is fixed, leakage currents through OFF transistors are lowered because of the stacking and the body effect. Analysis shows that when ΔV is fixed, there are optimum values for VV_{DD} and VV_{GND} for which the leakage power dissipation of the cell is at the minimum. In these simulations we have assumed both 'BL' and 'BLbar' are precharged to $V_{DD.}$ The CNTFET SRAM cell is simulated and simulated waveforms and results are shown in fig. 2 and Table 1. From this simulation results it can be seen that Data-Retention Sleep Transistor technique is more efficient compared to Sleep Transistor technique, especially for small values of ΔV and hence, small values of V_{DD}. This is useful for ultra low-power applications, which need lower noise immunity. With small delay and area overhead Data-Retention Sleep transistor technique results in a more robust and power-efficient CNTFET SRAM cell design.

5 Conclusions

CNTFETs are promising candidates to replace Si-MOSFETs due to its high current driving capability, tolerance to temperature and low leakage currents. CNFETs can also be considered as a potential alternative to CMOS in memory systems. In this paper a circuit-level leakage reduction technique called Data-Retention Sleep Transistor method is applied to CNTFET SRAM cell at 32nm Technology. This proposed Cell is simulated using HSPICE and simulated results show that in this technique the leakage power is reduced by significant amount compared to sleep transistor method but with small area and delay overhead. The results show that this technique is useful for standby leakage power reduction. The results show that using two sleep transistors is more beneficial than one sleep transistor.

References

1. International Technology Roadmap for Semiconductors by Semiconductor Industry Association (2009), `http://public.itrs.net`
2. Abdollahi, A., Pedram, M.: Power minimization techniques at the RT-level and below. In: Al-Hashimi, B.M. (ed.) SoC: Next Generation Electronics. IEEE Press, New York (2005)
3. Deng, J.: Device modeling and circuit performance evaluation for nanoscale devices: silicon technology beyond 45nm node and carbon nanotube field effect transistors. Doctoral Thesis. Stanford University (2007)
4. Appenzeller: Carbon Nanotubes for High-Performance Electronics—Progress and Prospect. Proc. IEEE 96(2), 201– 211 (2008)
5. Patil, N., Lin, A., Zhang, J., Wong, H.S.P., Mitra, S.: Digital VLSI logic technology using Carbon Nanotube FETs: Frequently Asked Questions. In: 46th ACM-IEEE Design Automation Conference, pp. 304–309 (2009)

6. Powell, M.D., Yang, S., Falsafi, B., Roy, K., Vijaykumar, T.N.: Gated-V_{DD}: a circuit technique to reduce leakage in cache memories. In: Proc. of International Symposium on Low Power Electronics Design, pp. 90–95 (2000)
7. Rajendra Prasad, S., Madhavi, B.K., Lal Kishore, K.: Design of Low Write-Power Consumption SRAM Cell based on CNTFET at 32nm Technology. International Journal of VLSI Design & Communication Systems (VLSICS) 2(4), 167–177 (2011)
8. Rajendra Prasad, S., Madhavi, B.K., Lal Kishore, K.: Design of 32nm Forced Stack CNTFET SRAM Cell for Leakage Power Reduction. In: IEEE International Conference on Computing, Electronics and Electrical Technologies (ICCEET 2012), India (accepted for publication, 2012)
9. Rajendra Prasad, S., Madhavi, B.K., Lal Kishore, K.: Reduction of Leakage-Power in CNTFET SRAM Cell using Stacked Sleep Technique at 32nm Technology. In: IEEE International Conference on Advances in Engineering, Science and Management (ICAESM 2012), India (accepted for publication, 2012)
10. Qin, H., Cao, Y., Markovic, D., Vladimirescu, A., Rabaey, J.: SRAM leakage suppression by minimizing standby supply voltage. In: Proc. of International Symposium on Quality Electronic Design, pp. 55–60 (2004)
11. Rajendra Prasad, S., Madhavi, B.K., Lal Kishore, K.: Low Leakage-Power SRAM cell design using CNTFETs at 32nm Technology. In: Springer 3rd International Conference on Communication, Network, and Computing (CNC 2012), Chennai, India, pp. 165–171 (2012)
12. Stanford University CNFET Model website, `http://nano.stanford.edu/model.php?id=23`
13. Deng, J., Philip Wong, H.S.: A Compact SPICE Model for carbon nanotube field effect transistors including non-idealities and its application-Part I: Model of the intrinsic Channel region and Part 2: Full device model and circuit performance benchmarking. IEEE Transaction on Electron. Device 54(12) (2007)

Intelligent Robust Router

Channamallikarjuna Mattihalli[1], Naveen Kolla[1], Bangi ChinnaSubbanna[1], M. Azath [2], and Sathish Kumar Konga[2]

[1] Department of Electrical and Computer Engineering
DebreBerhan University, Ethiopia
{Ckmattihalli,suveenabangi}@gmail.com, naveenkolla@yahoo.com
[2] Department of Computer Science
DebreBerhan University, Ethiopia
{azathhusssain,konga.knga}@gmail.com

Abstract. Through this paper our attempt is to give a onetime networking solution by the means of merging the VLSI field with the networking field as now a days the router is the key player in networking domain so the focus remains on that itself to get a good control over the network, Networking router today are with minimum pins and to enhance the network we go for the bridging loops which effect the latency and security concerns. The other are is of multiple protocols being used in the industry today. Through this paper the attempt is to overcome the security and latency issues with protocol switching technique embedded in the router engine itself. This paper is based on the hardware coding which will give a great impact on the latency issue as the hardware itself will be designed according to the need. In this paper our attempt is to provide a multipurpose networking router by means of Verilog code, by this we can maintain the same switching speed with more secured way of approach we have even the packet storage buffer on chip being generated by code in our design in the so we call this as the self-independent router called as the VLSI Based router. This paper has the main focus on the implementation of hardware IP router. The approach here is that router will process multiple incoming IP packets with different versions of protocols simultaneously and even it is going to hold true for the IPv4 as well as for IPv6. With the approach of increasing switching speed of a routing per packet for both the current trend protocols. This paper thus is going to be a revolutionary enhancement in the domain of networking.

Keywords: Robust Router, packets, FPGA, RTL, IP.

1 Introduction

Our approach here is to design a variable hardware router code by using Verilog and the same to be implemented till the SOC (System On Chip) level router. In this paper we are making a VLSI design for the implementation at the synthesizable level the same can be further enhanced to SOC level, but our main aim is limited to the NetList generation level which would give the result prediction and workable module vision. Our focus being in this is to make this router as much variable as we can which will

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 369–376, 2013.

give the robustness for the design to be called even as a Robust Router in which we can make the same router to not only go for N number of connections but also to detect all variety of packets and route the same. To do so we have to add the code with specific case's for every type of packets we want to add to our router to route, with this paper of hardware code our approach is to get the basic packets routing with multiple protocols starting with the IPv4 and IPv6.

2 Literature Survey

In this we are comparing the existing generic router architecture and our new robust router architecture. This will give the difference in the designing and would reflect our paper enhancements that we are upgrading in our robust router paper.

2.1 Generic Router Architecture

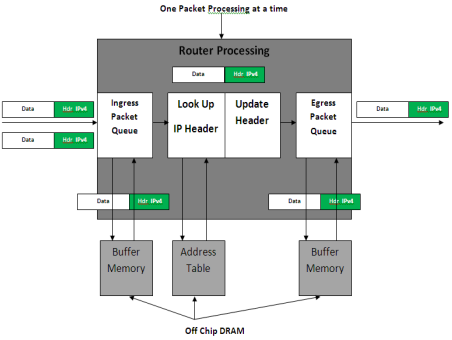


Fig. 1. Generic architecture

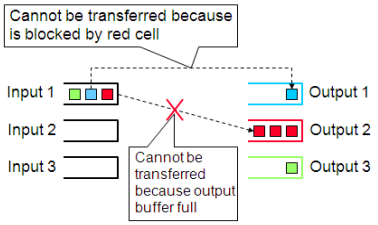


Fig. 2. Congestion flow

In the architecture we can look that the generic architecture is processing a single packet of a specific protocol at a given time and the output queue buffer also one for one egress channel ring by which there is the overloading of the queue buffer and will result in the congestion. The congestion flow is as shown below.

3 A New Robust Router Architecture

The architecture of robust router is totally based on the Verilog code which would enable our design in the implementation of parallel packet processing for N number of channels. This intern enables the multi packet processing at the same time. With the Verilog code being the base of design we have an option for the addition of protocol case and respective look-up table makes us go for the Multi-protocol processing at the same time. By which we are unable to provide the multi-packet multi-protocol routing at the same time with same speed.

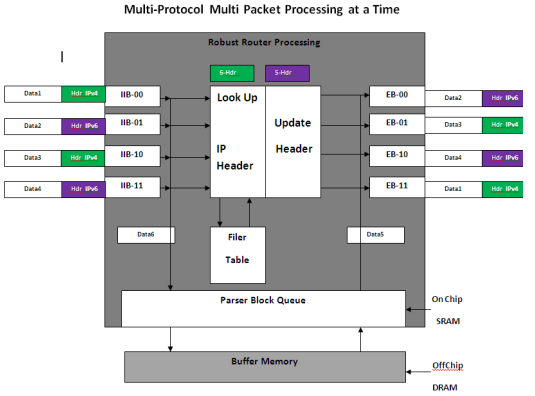


Fig. 3. Multi-protocal Multi packet processing at a tree

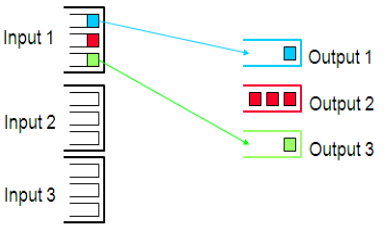


Fig. 4. Congestion flow with vertical queue

While designing the robust router a special concern is kept in the mind of the switching speed issue to give the maximum speed with parallelism being added. The egress output buffer queuing problem was also solved by providing a separate queue for every ingress channel in the egress channel with N vertical queue by which we can avoid the congestion to a remarkable level which is as shown below. The size issue is another special feature of our robust router which makes our robust router a unique system. As discussed earlier in the paper we are trying to make the Robust router on to the chip level design so we further advance it to the level of Ethernet based router which will make the router to be implemented on the standalone systems, which will be a revolutionary enhancement in size matter from room full of router to just the PCI slot operating Router and will make network work more faster. It looks something like this below.

Generally, the router can be interfaced with 'N' number of I/O devices. Block diagram given below.

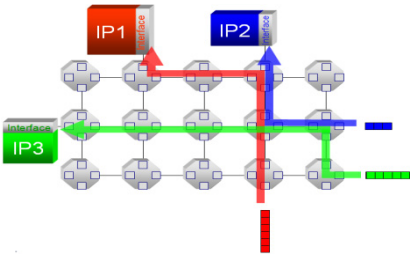


Fig. 5. Forwarding IP packet

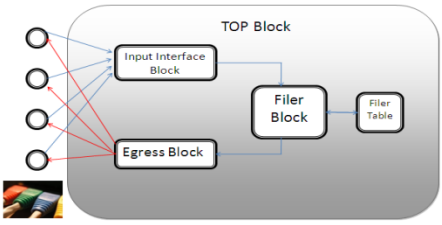


Fig. 6. Block diagram

The paper design has the following modules.

3.1 Input Interface Block

This block is mainly responsible for receiving the incoming IP packets over multiple input channels. This block asserts the necessary response signals in order to communicate with the IP packet driver modules. After receiving the IP packets, this block forward the same to the Ingress Block for the further process. This block forwards the same to packet store block as well as parser block.

3.2 Packet Store Block

This is responsible for storing the error free received packets. This module receives the packet contents from Ingress block and dispatches the same based on the request from Egress block.

3.3 Parser Block

This block is mainly responsible for parsing the complete packet into multiple set of data according to its field. The parsed contents will be inputted to the filer block. The above three blocks are merged all together as IIB in code to single file.

3.4 Filer Block

This block is responsible for selecting the egress ring. The block receives the parsed data from the parser block. The parsed data will be forwarded to the filer table. In response to this, the filer table provides the output ring number. Then, the received output from the filer table will be forwarded to the egress block.

3.5 Filer Table

It is a user configurable table. This table contains a set of data in its each slot, against which the data sent by the filer block will be compared. If the filer block inputted data matches with the data of any slot of filer table, then that slot's data will be used as egress ring through which received packet will be forwarded.

3.6 Egress Block

This block receives the data from filer block as egress ring number through which the received packet shall be forwarded. Upon receiving the egress ring number, this block initiates the communication with packet store block to fetch the packet to be forwarded. Then, the fetched IP packet will be forwarded to the output interface block with the output channel details, over which the packet has to be transmitted.

3.7 Output Interface Block

Upon receiving the packet with output port details from the egress block, this block forwards the IP packet over mentioned output channel. This block is also responsible for asserting all the necessary handshaking signals for the receiving device while transmitting the packets. The Egress Block and Output Interface Block are merged together in code as single file.

4 System Flow Diagram

The system flow diagram is as shown below which makes us to understand the flow of the signals through the system from each block by block and transaction carried between the blocks to accomplish the task of the robust router. The flow diagram described here is a brief one, which helps us to understand the flow of every block. Every block have the state machine cycle included in them to enhance the system logical transaction to the level of parallelism. The flow diagram is as shown below.

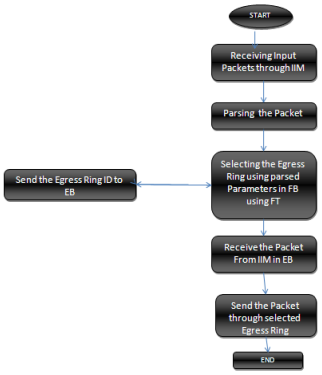


Fig. 7. System flow diagram

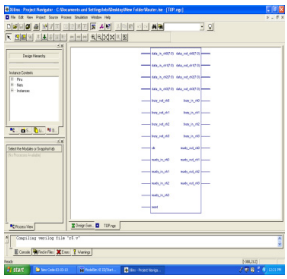


Fig. 8. Net List

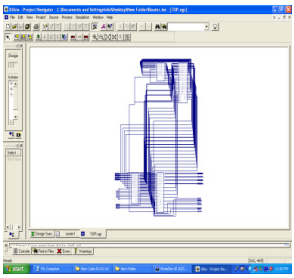


Fig. 9. View of thesystem at the RTL

First the packet is received from the ingress channel ring to the input interface block the packet is parsed to data packet and header packet, the data packet is stored in the parser queue and the header is sent to the filer block. The flier block then checks weather the packet is IPv4 or IPv6 and accordingly send the request to the filer table to router the packet to required destination. The filer table cross verifies the egress ring channel with it Dest-IP address and send the egress ring ID to the filer block. The filer block send and enables the particular egress ring in egress blocks and gives the command to the particular egress ring in egress block. Then in egress block the stored data packet in the parser queue is added back with header and is sent out with the specified egress ring channel. In this way the every packet is processed and routed in robust router.

5 Conclusions and Further Research

5.1 Net List of the Robust Router

The Net List is RTL level of the robust router system, which is syntasizable and can be extracted on the Xilinx tool. By which we can get preface look of the system and a transition from the frontend of the VLSI designing to backend of the VLSI designing. Which means the same can run on FPGA kit and test its robustness and errors of the system can be debugged before it is taken to SOC Level and to Fab-Labs. The snap below is the Pin configuration of the proposed Robust Router. The NetList level can further go after I/O Padding get the exact pin configuration which can be derived accordingly but will be similar one. The Snap below is the Top level system view of the system at the RTL.

5.2 SOC Designing

The further movements of the VLSI Designing will require the sharp knowledge of the VLSI backend designing and can be fabricated at the 45 nano technology using the Cadence Encounter Tool which will enable us to take the system to SOC level the steps are as shown below.

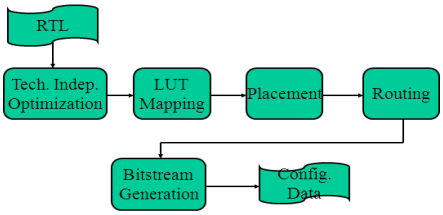


Fig. 10. System to SOC level the steps

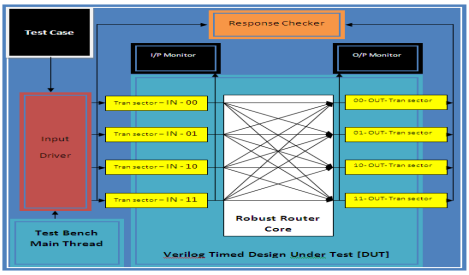


Fig. 11. DUT architecture

The RTL level of design which we get from the Net List of the system will have gate delay, propagation delay and wire delays included in them. These are all calculated and made into an optimization level. Then the design is fixed into LUT'S and the mapped between the LUT'S further the placement of the LUT'S are prissily done keeping mind the power utilization and the delay calculated earlier. Then the routing is done between the CLB'S. Further the bit-stream is generated to test the system and verification done across the Net List output to get the exact design. Then the system design is masked and made to the GDSSI Level further to be sent on to the Fab-Labs for fabrication.

5.3 Design under Test

The design under test [DUT] is made to test the system robustness under different cases. The DUT architecture includes the Test case which will define the test. The

input driver block will generate the test input signals for the system testing. The input and output transacted will make the system get the input and output according to the system core requirement. The input and out monitor are placed to compare the system testing. At last the Response checker is to give the system testing pass or fail. The DUT is as shown in figure 11.

For the testing of the robustness here we are mixing theIPv4 and IPv6 packets and we testing the system in different cases.

5.4 Advantages

- General purpose router
- Router hardware code is variable
- High switching speed with any no. of I\o pin connection
- More secured
- Robust router can handle all type of packets (Implemented on IPv4 and IPv6)

5.5 Application

- Can be used as public internetworking router
- Can be used as corporate router
- Software company private router (client to client, developer to client)
- Router for networking research
- In other words one point networking solution

5.6 System Test Analysis

Number of gates: - 59855
Total Run time: - 9.018 ns
Switching speed: - 662.5 Tbps
System Frequency: - 0.11 GHz
Can be fabricated: - by 45 nano technology with Cadence Encounter tool (cost 12000$)

This paper code is will give the output for the IPv4 and IPv6 packets put in the router at the same time, With this paper of Robust Router we will be able to route both the packets at the same time at the same speed. The Robustness is simulated with Model-Sim Tool with different Test Cases and the same code's NetList is extracted with Xilinx Tool for the synthesizable code. The same can be taken to the SOC (System on chip) level with Cadences Encounter Tool.

- The same Verilog code design can be taken to the implementation of MPLS (Multi-Protocol Label Switching).
- The some code design can be taken to the SOC (System on chip) level and can be implemented as the Ethernet Standalone System Router.
- The same code can be made variable with TCP and UDP Protocols.

References

1. Katsube, Y., Nagami, K., Esaki, H.: Toshiba's Router Architecture Extensions for ATM: Overview. IETF RFC 2098 (April 1997)
2. Rekhter, Y., Davie, B., Katz, D., Rosen, E., Swallow, G.: Cisco Systems' Tag Switching Architecture Overview. IETF RFC 2105, Amir.Palnitkar, 2nd edn. (February 1997)
3. Moy, J.: OSPF: Anatomy of an Internet Routing Protocol (1998)
4. Bjerregaard, T., Mahadevan, S.: A survey of research and practices of network-on-chip. ACM Comput. Surv. 38(1), 1 (2006)
5. Roth Jr., C.H.: Digital System Design Using VHDL
6. Jenkins, J.H.: Designing with FPGAs and CPLDs
7. Abromovici, M., Breuer: Digital System Testing and Testable Design
8. Balfour, J., Dally, W.J.: Design tradeoffs for tiled cmp on-chip networks. In: ICS 2006: Proceedings of the 20th Annual International Conference on Supercomputing, pp. 187–198 (2006)

Maude Specification Generation from VHDL

Fateh Boutekkouk

Department of Mathematics and Computer Science, University of Larbi Ben M'hedi, Route de Constantine, BP 358, Oum El Bouaghi, 04000, Algeria
fateh_boutekkouk@yahoo.fr

Abstract. In this paper, we present our flow that permits Maude specification generation from VHDL code. Firstly, a XML like Intermediate Format (IF) is created showing VHDL structures and statements in a hierarchical form. This format is an abstraction of the original VHDL code. Secondly, a Maude code is generated from this IF. Both Hardware system Maude specification and properties are then passed to the Maude model checker for verification purpose. Our idea is thus to combine between VHDL simulation and Maude based formal verification capabilities for hardware systems validation. The impetus behind this cooperation between simulation and formal verification is to enable hardware designers to discover errors that could not be detected by VHDL discrete event simulator.

Keywords: VHDL, Maude, Simulation, Formal specification, Formal verification.

1 Introduction

In order to increase productivity in electronic systems manufacturing, hardware engineers have developed Hardware Description Languages (HDLs) for describing the structure and function of integrated circuits at many levels of abstraction.

HDLs offer the necessary software and Hardware specific statements for sequenciality, concurrency, timing and synchronization expression. These languages are able to perform functional and timing simulations, design space exploration, performances estimation and synthesis. Among these languages, we find SystemC, Verilog, SystemVerilog and VHDL in particular. The latter is an industrial IEEE standard HDL [3]. It is designed to fill a number of needs in the hardware design process. Among these needs, ensuring the correctness of computer circuits is certainly the most difficult task. The most commonly used verification technique in HDLs is simulation. However, in many cases, simulation can miss important errors. A more powerful approach is the use of formal methods such as temporal logic model checking, which can guarantee correctness. In this context, we have developed an approach that permits Maude [4] code generation from VHDL code. Maude is based on rewriting logic [5]. In order to apply Maude model checker, we have to introduce in addition to system specification, the specification of the properties to be verified.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 377–380, 2013.

2 Our Flow

The entry of our proposed flow is a VHDL behavioural description. From this, a XML like Intermediate Format (IF) is generated. It captures the structure and the dynamic of VHDL design in an abstract manner. In other words, the IF abstracts the original VHDL code by taking into consideration only the statements having the impact on the global state of the system such as signals assignments and waits.

Before IF generation, designer should perform a functional and timing simulation to discover eventually errors. From the IF, a Maude code is generated following some correspondence rules. At this stage, we can simulate Maude code by executing rewriting rules. This step is very important to validate the correctness of the transformation from VHDL code to Maude code (i.e. execution of Maude code leads to same results as we simulate VHDL program). The last step of the flow is the formal verification that requires in addition to system specification, the properties specification. Figure 1 gives an example of a VHDL behavioural description. In this example, we have an entity named *exemple* with two bit ports: an input port *X* and an output port *Y*. The architecture (named *Description*) includes one process named *P* with signal *X* as a sensibility list. Here the process *P* tests the value after a mount of time. At the end, the process suspends on *X* (there is an implicit statement: *wait on X* at the end of the process). When an event occurs on *X*, the process *P* resumes its execution. The process *P* acts as an infinite loop. Figure 2 shows the corresponding IF description for the example of figure 1. Note that all implicit waits in the original VHDL code become explicit in the corresponding IF.

In order to transform the IF code to a Maude specification, a set of correspondence rules are defined. A VHDL process is declared as a Maude class with a set of attributes like the sensitivity list, the current state of the process and the name of the signal which is updated by the process. A VHDL signal is also declared as a Maude class. Each signal is characterized by its current value, its future value and the timeout. In order to preserve VHDL simulation semantic, we add two new attributes that are *isChanged* and *co*. *isChanged* is a Boolean to specify whether the value of the signal is changed or not. According to VHDL simulation semantic, the value of a signal can not be updated (modified) unless the process is blocked (i.e. it reaches a wait statement). If this is the case (i.e. the process is blocked) and if the current value of the signal differs from the future value then the signal is updated (i.e. the current value becomes the future value) and *isChanged* is set to true otherwise it is false. When updating a signal, all processes that are sensible to this signal resume their execution. We use *co* (counter) to specify the number of processes that are sensible to a specified signal. Whenever a process resumes its execution, *co* is decremented. When *co* is equal to zero (i.e. all processes sensible to this signal are waked up), *isChanged* is set to false (i.e. the end of a cycle) to start a new cycle. The progression of VHDL simulation time is specified by defining a class named *Horloge* with two attributes: a list to stock sorted signals timeouts and the current physical time. Using Maude, we define a sort named *Processstatevalues* to specify the set of a process states. For instance: *ops begin updatey1 waitX end : -> Processstatevalues* .

Let *P* be a process with a sensibility list *(X)*. In Maude, such a process is declared as: *< P: Process | listeSensibility: (X), state: begin >* .

Similarly, the specification of a signal *X* in Maude looks like:

< X: Signal | currentValue: N, futurValue: F, timeOut: T, isChanged: false, Co: m >

The horologe is declared as a Maude class named *Horloge*:
< H: Horloge | times : L, currentTime : C >
We use the functions *head* and *throw* to extract the first element and to delete this element from a list respectively. The elements of L are sorted in a descendant fashion.
Figure 3 gives an example of a rewriting rule.

```
0entity Exemple is
       port (X : in bit;
             Y : out bit );
 end Exemple;
architecture Description of Exemple is
  begin
    P:Process (X)
      begin
        if X ='1' then
           Y <= "1" after 10;
        else
           Y <= "0" after 5;
        endif;
    end process
 end Description;
```

Fig. 1. Example of VHDL behavioural description

```
< ENTITY Name : Exemple >
   < Port>
     < X: in   Type: bit >
     < Y: out  Type: bit >
   </Port>
 </ENTITY Exemple>
< ARCHITECTURE Name: Description    Entity: Exemple >
 < BODY>
   < Process Name: P  Param: X >
   <BODY>
     < IF (X='1') >
       < AFFEC Name:  Y  Value: "1" afterTimes: 10 ns >
    < ELSE >
      < AFFEC Name: Y Value: "0"  afterTimes: 5 ns >
    < /IF >
    <Wait on X>
     < /BODY>
    < /Process>
   < /BODY>
  </ARCHITECTURE>
```

Fig. 2. The generated IF code for the example of figure 1

```
crl [rl2] :
< P: Process | listeSensibilite : (X), state : updatey1, SIGNAL : Y>
< X: Signal | currentValue : N, futurValue : F, timeOut : T,
isChanged : false, Co : m >
< Y : Signal | currentValue : N1, futurValue : F1, timeOut : T1,
isChanged : false, Co : n >
< H : Horloge | times : L, currentTime : C >
=>
< P : Process | listeSensibilite : (X), state : waitX, SIGNAL : X>
< X : Signal | currentValue : N, futurValue : F, timeOut : T,
isChanged : false, Co : m>
<Y : Signal | currentValue : F1, futurValue : F1, timeOut : T1,
isChanged : true, Co : n >
<H : Horloge | times : throw C from L, currentTime : head (L) >
if T1 == head (L) and F1 =/= N1 .
```

Fig. 3. Example of a rewriting rule in Maude

3 Conclusion and Future Work

The process of formal verification passes by many steps: first an IF is generated from VHDL code then a Maude code is generated from this IF following a set of well defined rules. IF is an abstraction of the original VHDL code. In order to formally verify the correctness of the Hardware functionality, we have to specify in addition to hardware system specification a set of properties. At the end, we call Maude model checker. As a perspective, we plan to generate Maude code for VHDL structural and data flow styles. We also intent to discover some pertinent properties for verification.

References

1. Brayton, R.K., Clarke, E.M., Subrahmanyam, P.A.: Formal Methods in System Design. Special Issue on VHDL Semantics 7(1/2) (1995); Borrione, D. (Guest ed.)
2. Breuer, P.T., Fernndez, L.S., Delgado Kloos, C.: Clean Formal Semantics for VHDL. In: European Design and Test Conference, pp. 641–647. IEEE Computer Society Press (1993)
3. IEEE Standard VHDL Language Reference Manual. IEEE, IEEE Std 1076 (2000)
4. McCombs, T.: Maude 2.0 Primer, Version 1.0. International report. SRI. International (2003)
5. Meseguer, J.: Rewriting as a Unified Model of Concurrency. In: Baeten, J.C.M., Klop, J.W. (eds.) CONCUR 1990. LNCS, vol. 458, pp. 384–400. Springer, Heidelberg (1990)
6. Reetz, R., Schneider, K., Kropf, T., Verysys, H.: Formal specification in VHDL for hardware verification. In: Proceedings of Design, Automation and Test in Europe (DATE), February 23-26, pp. 257–263 (1998)
7. Wilsey, P.A., Davis, K.C.: Modeling VHDL's Execution Semantics in EXPRESS, Department of Electrical and Computer Engineering, University of Cincinnati (November 13, 1993)

Dynamic Probabilistic Packet Marking

K.P. Chaudhari[1] and Anil V. Turukmane[2]

[1] M.I.T. College of Engineering Aurangabad
Department of Computer Science & Engineering, Aurangabad [M.H], India
Kiran_Chaudhari@rediffmaill.com
[2] P.E.S. College of Engineering Aurangabad
Department of Computer Science & Engineering, Aurangabad [M.H], India
anilturukmane@gmail.com

Abstract. Most of the probability of packet marking (PPM) have existed many problems such as the lost of marking information, the difficulties to reconstruct attack path, low accuracy and so on. In this work, we present a new approach, called dynamic probabilistic packet marking (DPPM), to further improve the effectiveness of PPM. Instead of using a fixed marking probability, we propose to judge whether the packet has been marked or not then choose a proper marking probability. DPPM may solve most of the problems in PPM method. Formal analysis indicates that DPPM outperforms PPM in most aspects.

Keywords: Structured Network, secure data sharing.

1 Introduction

The denial-of-service (DoS) attack has been a pressing problem in recent years. DoS defiance research has blossomed into one of the main streams in network security. Various techniques such as the pushback message, ICMP trace back, and the packet filtering techniques are the results from this active field of research. The probabilistic packet marking (PPM) algorithm by Savage et al. Has attracted the most attention in contributing the idea of IP trace back. The most interesting point of this IP trace back approach is that it allows routers to encode certain information on the attack packets based on a predetermined probability. Upon receiving a sufficient number of marked packets, the victim (or a data collection node) can construct the set of paths that the attack packets traversed and, hence, the victim can obtain the location(s) of the attacker(s).

The Probabilistic Packet Marking Algorithm The goal of the PPM algorithm is to obtain a constructed graph such that the constructed graph is the same as the attack graph, where an attack graph is the set of paths the attack packets traversed, and a constructed graph is a graph returned by the PPM algorithm. To fulfill this goal, Savage et al. suggested a method for encoding the information of the edges of the attack graph into the attack packets through the cooperation of the routers in the attack graph and the victim site. Specifically, the PPM algorithm is made up of two separated procedures: the packet marking procedure, which is executed on the router side, and the graph reconstruction procedure, which is executed on the victim side.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 381–384, 2013.

The packet marking procedure is designed to randomly encode edges' information on the packets arriving at the routers. Then, by using the information, the victim executes the graph reconstruction procedure to construct the attack graph. We first briefly review the packet marking procedure so that readers can become familiar with how the router marks information on the packets.

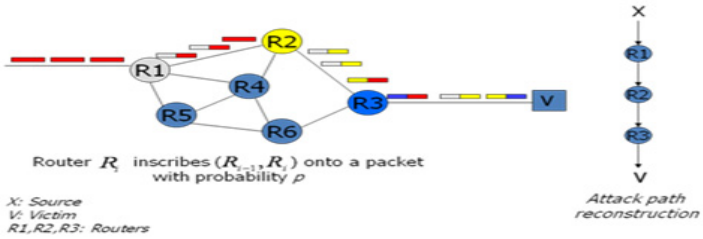


Fig. 1. Example – Probabilistic Packet Marking

A new probabilistic packet marking technology (called as P3M) based on path identification to defense serious distributed denial of service attacks and solves complex computation and other problem existed in traditional probabilistic packet marking (PPM) technologies. First contribution is constructing a new payload to carry router address and path identification. The second contribution is designing a new path identification scheme based on router addresses and hash algorithm. P3M is a practical technology to defense DDoS.To overcome above shortcomings; we design a new probabilistic packet marking technology -- P3M in this article. Comparing with the traditional PPM technologies, our first contribution is a new payload (called as P3M payload below) carrying router address and path identification to avoid influencing the normal running of recombining packets and Qos mechanism. Our second contribution is a new path identification scheme based on router addresses and hash algorithm. The use of path identification makes our probabilistic packet marking technology P3M simple when victim computes DDoS attack paths. And path identification also could be used by other network security equipments.

2 Principal of P3M Technology

As shown in Figure 2, pretending routers R1, R4, R5 have deployed P3M technology and the network path between terminals A and B is Path AB= {R1, R4, R5}, so the work flow between A and B is:

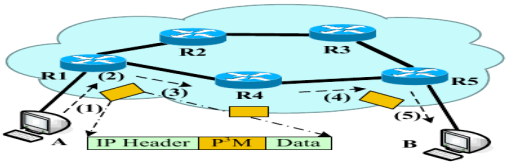


Fig. 2. P3M scheme

A sends a packet *P* whose source ip address is *IP*a and destination ip address is *IP*b.

After *P* arrives at router R1, firstly R1 will insert a P3M payload between P's ip header and transport layer data, alter the "Protocol "domain value of ip header to 256 and insert its raw value to P3M payload's "Protocol" domain. Secondly, R1 set P3M payloads "Hops" domain 1, and compute the "Path Identifier" value PI1 according to formula (1). Finally, R1 get p by computing according to formula (2) and alter the "Router Address" and "Locator" domain value by probability p. Notice that the "Locator" domain value is equal to "Hops" domain value.

$$P = 1 / h \tag{2}$$

In formula (2), h is the currently "Hops" value. After above processing, R1 transmits packet *P*.After receiving *P*, R4 will firstly check the value of "Protocol" domain in ip header. If it is unequal to 256, R4 will perform the same operation as R1; else, R4 will add "hops" value 1 and recalculate the "Path Identifier" domain value. Finally, R4 computes *p* as formula (2) and alter the "Router Address "and "Locator" domain value by probability *p*. After above processing, R1 transmits packet *P*.R5 will perform the same operation as R4.P reaches the destination B. If B finds DDoS attack, after receiving a certain amount of packets, it can reconstruct the DDoS attack paths by its trapped P3M payloads. The packet marking procedure aims at encoding every edge of the attack graph, and the routers encode the information in three marking fields of an attack packet: the start, the end, and the distance fields.

3 Simulation Result

In this section, we present the simulation results to show that the DPPM algorithm is able to guarantee the correctness of the constructed graph, independent of the marking probability and the structure of the attack graph. First, we describe the simulation environment. The Simulation n Environment Every simulation of the DPPM algorithm starts with a testing network rooted at the victim, that is, the attack graph. The configuration of the network follows the assumption stated in Section 2.1. In addition, the network has at least one leaf router, that is, a router with zero incoming edges. Each edge between two routers is directed and is assumed to have infinite capacity. Thus, no packet is lost under this environment. Next, we describe the properties of the simulated packets. All packets are homogeneous in terms of type, size, etc. Every packet's destination is set to the victim, and every packet starts its itinerary at one of the leaf routers of the testing network chosen at random. Further, the paths traversed by the packets are chosen at random.

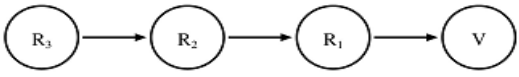


Fig. 3. An example linear network with three edged

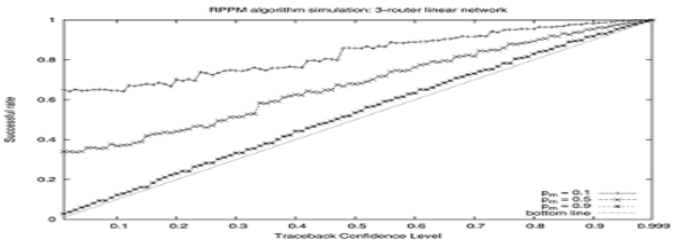


Fig. 4. The simulations shows that the larger the making probability is, the closer to the worst-case execution the simulation result becomes

4 Conclusion

In this work, we have shown that there are some problems in PPM algorithm: the overwritten problem, limited marking field, low accuracy and so on. Dynamic probabilistic packet marking has solve these problems by using dynamic probability and fragment-reassembly. Meanwhile, using the expected number of required marked packets E½X as the termination condition is not sufficientPath reconstruction is the fundamental goal of packet marking. Reduced false positives. High false positives are actively suppressed due to the above improvements. Effectiveness to handle large-scale DDoS attacks which is dominant in today's Internet.

Acknowledgments. The author is extremely thankful for the respected guide Prof.K.P.Chaudhari for their encouragement.

References

1. Pressman: Software Engineering – A Practitioners Approach, 7th edn. Booch, G., Rumbaugh, J., Jacobsan, I.: UML User Guide; Baker, F.: Requirements for IP Version 4 Routers. RFC 1812 (June 1995)
2. Savage, S., Wetherill, D., Karl, A., Anderson, T.: Network Support for IP Trace back. IEEE/ACM Trans. Networking 9, 226–237 (2001)
3. Song, D., Perrig, A.: Advanced and Authenticated Marking Schemes for IP trace back. In: IEEE INFOCOM 2001, Anchorage, AK (2001)
4. Year, A., Perrig, A., Song, D.: Fast Internet Trace back. In: IEEE INFOCOM (2005) (in press)
5. Peng, T., Leckie, C.: Adjusted Probabilistic Packet Marking for IP Traceback. In: Gregori, E., Conti, M., Campbell, A.T., Omidyar, G., Zukerman, M. (eds.) NETWORKING 2002. LNCS, vol. 2345, p. 697. Springer, Heidelberg (2002)
6. Rizvi, B., Fernandez-Gaucherand, E.: Analysis of adjusted Probabilistic Packet Marking. In: IP Operations & Management (IPOM 2003), Kansas City, MO (October 2003)

Advanced TeleOphthalmology for Blindness Prevention

Samit Desai, Poornima Mohanachandran, Pramod K. Singh,
Leena Devakumar, Sriram Kannan, and Amit Suveer

i2iTeleSolutions, Bangalore, Karnataka, India
{samit,poornima}@i2itelesolutions.com

Abstract. Healthcare practices are increasingly using complex medical imaging and data collection methodologies. A growing shortage of medical specialists as well as rising incidence of diseases complicates the issue further. Ophthalmology is one area where medical imaging plays a crucial role in diagnostic decisions. This paper describes an advanced system for TeleOphthalmology which supports sharing of patient information, medical data and images in a secure manner to experts without constraint of geography. Advanced workflow options, imaging processing tools for decision support and access from mobile platforms such as iPhone, iPAD and Playbook is supported.

Keywords: Medical Imaging, TeleOpthalmology, TeleMedicine, Decision Support, m-Health.

1 Introduction

Telemedicine is the use of electronic information and communication technologies to enable health care delivery when distance separates participants. Advances in medical instrumentation, medical imaging, telecommunication and signal processing have fuelled growth of Telemedicine.

Telemedicine may be as simple as a voice call between a patient and a medical expert or as complex as a TeleSurgery. An emerging trend is the use of mobile phones for Telemedicine, referred to as m-health. The increased display and processing capabilities of hand held devices in combination with advances in communication speeds makes m-health a reality. Concept of Telemedicine is three decades old and holds great future potential and offers many research opportunities.

2 TeleOpthalmology

TeleOphthalmology is Telemedicine applied to the area of Ophthalmology where patients' eye related problems are examined, diagnosed and treated, even though the patient and Ophthalmology expert are in different geographical areas.

Globally, every 5 seconds someone goes blind. The leading causes of blindness and low vision are age-related macular degeneration, cataract, glaucoma, diabetic retinopathy in adults, and Retinopathy of prematurity in infants. The majority of these are preventable if diagnosed early. [1]

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 385–389, 2013.

Diabetic retinopathy (DR) is retinopathy (disease of retina) arising due to diabetes. Early diagnosis of DR and timely treatment would significantly reduce the risk of vision loss. [6] Also retinopathy of prematurity (ROP) is an eye disease that affect retina of premature babies. All babies with a birth weight less than 2,000 g are at risk of ROP. They require screening by an ROP-trained ophthalmologist within the first 4 weeks of life, with subsequent screenings determined by the initial findings. If not diagnosed or untreated ROP can lead to blindness. With regard to disease burden in India, there are 27 million live births every year and 2 million of the new born are at risk of ROP and needs an eye examination. The challenge here is also that there are only a handful of pediatric retina experts to review and offer advice. [3]

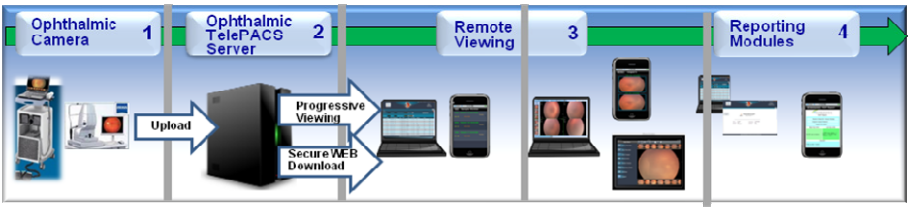


Fig. 1. TeleOphthalmology workflow

In both diabetic retinopathy and ROP the blood vessels in retina are affected and diagnosis is carried out by imaging the retina and reviewing the vasculature. Image based diagnosis makes TeleOphthalmology a viable option for diagnosis and follow up of these diseases. [1]

Critical features of the TeleOphthalmology system were developed taking into consideration the disease burden, acute shortage of medical experts and the diagnostic approaches employed. A TeleOphthalmology system should enable sharing of patient history (including family history & medical history), medical images and medical reports with an expert or set of experts in a secure manner. The next step would be review by an expert physician which includes, reviewing changes w.r.t previous visits if any and preparing a detailed report with diagnosis, suggested treatment and follow up plans. To enable easy access for expert physicians a sophisticated work flow along with secure access of patient studies from portable mobile devices was implemented TeleOphthalmology workflow implementation is shown in Fig 1. Images of the eye are captured using an ophthalmic camera. Ophthalmic cameras support imaging the anterior and posterior segment of the eye and presently non mydriatic portable cameras are also available specifically for the purpose of screening. Images and patient data uploaded to an internet based server make the studies available for review by authorized users anywhere in the world. Encryption and image compression technologies are used to ensure quick uploads and secure communication.

Standard formats are used for capturing patient information and data, thus minimizing user errors. In the case of medical images and reports DICOM format is used. However all available Ophthalmic cameras do not support DICOM format yet. In order to support non DICOM Ophthalmic cameras the system uses the “SC” (Secondary Capture) modality mapping as recommended by DICOM standard.

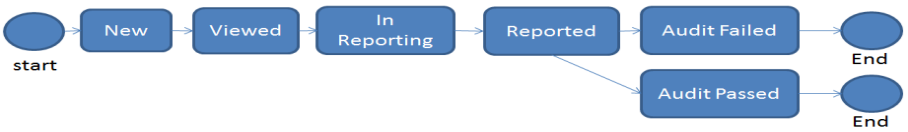


Fig. 2. TeleOphthalmology workflow- Study state machine

One important aspect of TeleOphthalmology [2] system would be quick transmission of data and images over the internet. This is addressed by the use of suitable compression algorithms. In order to provide a superior viewing experience and to avoid lenghthy downloads progressive streaming is employed.

The progress of a study through the TeleOphthalmology system is indicated by the study states. The various study states are indicated in Fig 2. The study states enable appropriate user access and activity. Assigning a study to an expert or a set of experts for review and reporting is supported by the system. Second opinion on a reported study is supported. The system also supports analysis of diagnosis variability across physicians and audit of reports.

In addition to reports in commonly used file formats, quick and error free reporting is supported using templates. Disease specific templates follow the international standards for disease classification. E.g. ICROP (International Classification for ROP). Reporting templates enable quick and comprehensive reporting while driving standardisation across physicians and ensuring completeness of a report.These reports are accessed at the patient end and handed over to the patient.

3 m-Health

Last decade saw significant advances in mobile technologies and advancement in hardware and software capabilities of mobile devices. Reports show increased penetration of mobile devices compared to that of computers. [7] Improved communication capabilities (e.g. 3G) also happened during this time. These advancements make the mobile platform an attractive option for healthcare delivery and 24/7 access to limited number of medical experts a reality.

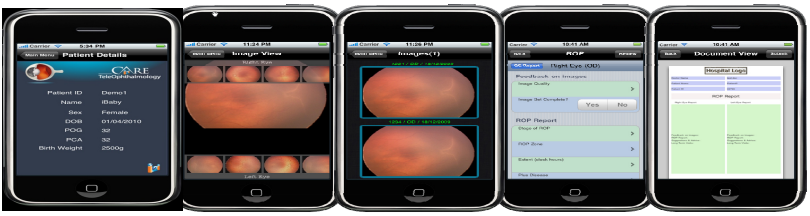


Fig. 3. Physician reporting via iPhone

The TeleOphthalmology system is supported on popular mobile platforms such as iPAD and iPhone. TeleOpthalmology image viewer enables the specialist to view all images of a particular study. Image manipulations such as zoom, pan are supported.

Reports are created post image review. Touch enabled mobile devices make image review and template based reporting, very easy for the user.Specificity and Sensitivity of use of iPhone for TeleOphthalmology has been published. [7]

4 Decision Support Using Image Processing

As described in introduction Diabetic Retinopathy (DR) and Retinopathy of prematurity (ROP) are diagnosed by visual examination of retina or by review of retinal images. Therefore, remote diagnosis using store-and-forward telemedicine is a promising strategy for improving the delivery and accessibility of care for these diseases. In order to improve efficiency of utilizing an expert's time the retinal images may be captured by a trained technician instead of an expert physician. [6]

A crucial preliminary step in the analysis of retinal images is the identification and localization of important anatomical structures, such as the optic nerve head (ONH), the macula, and the major vascular arcades. While many advances have happened in digital fundus imaging there are still challenges in diagnosis and analysis. Noises, Non uniform illumination, inconsistent image contrast, low contrast of vasculature are some of the challenges faced. Direct examination using an Ophthalmoscope or imaging using contrast agents are alternatives that are available to a physician. To circumvent these problems a new technique was developed that allows detection of features that are subtle or undetectable, significantly helping diagnosis and treatment planning. A few such examples are determining edges of vascular area in APROP images, determining occluded vessels (in neonates and adults), determining vessel growth post treatment etc.

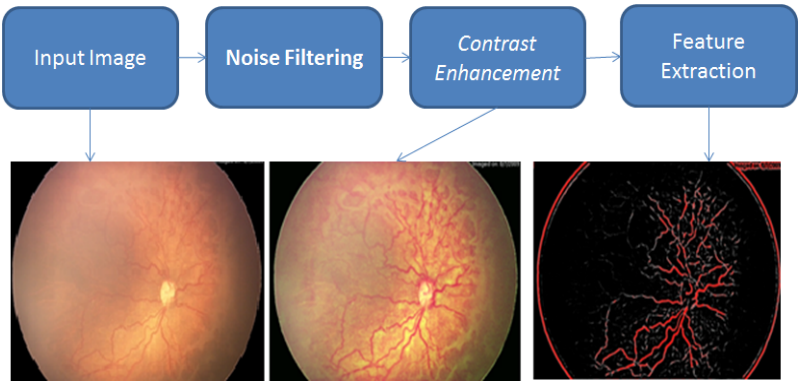


Fig 3a. original image **Fig 3b.** CE **Fig 3c.** VNM

The TeleOphthalmology system employs noise filtering, contrast improvement and vessel extraction algorithms to highlight features of interest for the Ophthalmologist. These algorithms are implemented and tuned to run taking local image statistics into account. Pseudo colouring algorithm is employed to improve visualization.

Currently three protocols are being validated “Grey Enhancement (GE)”, “Colour Enhancement (CE)” and “Vesselness Measure (VNM)”. On upload of a set of images to the system an expert physician is enabled review of the original images as well as the processed images. As an advanced feature the processed images are made available only to expert users of the system.

Intial clinical evaluation show that CE protocol enabled better visualization of capillary non-perfusion zones and edge of vascular loops in APROP. VNM protocol was found to be useful in determining if the vessels have reached the ora serrata. Both protocols helped in delineating zones of avascularity. [3]

5 Conclusions

An advanced TeleOphthalmology system is described. Secure and easy exchange of medical data and images is described. Upload and storage of anterior and posterior segment images is supported. The system is built on industry standards such as DICOM for image storage and management and uses reporting templates based on international classification of diseases. A sophisticated workflow is implemented and made available on portable mobile devices. Advanced image processing methods are implemented to aid the physician to review subtle features in the images.

Acknowledgment. The clinical evaluation was carried out under the guidance of Dr Anand Vinekar, Head of Paediatric Retina Narayana Nethralaya Postgraduate Institute of Ophthalmology, Bangalore. The system is in use by KIDROP (Karnataka Internet Assisted Diagnosis of ROP) team. The authors would like to thank Dr Anand Vinekar and the KIDROP team for their continued support.

References

1. Yogesan, K., Kumar, S., Goldschmidt, L., Cuadros, J.: TeleOphthalmology, pp. 3–7, 9–21 (2006) ISBN 978-3-540-70515-4
2. Desai, S.: Medical image transcoder for telemedicine based on wireless communication devices. In: Electronics Computer Technology (ICECT), pp. 389–393 (2011)
3. Vinekar, A.: The ROP challenge in rural India: preliminary report of a telemedicine screening model. In: International Experience with Photographic Imaging for Pediatric and Adult Eye Disease, Retina Physician, supp. 9-10 (2008)
4. Hani, A.F.M.: Analysis of foveal avascular zone in colour fundus images for grading of diabetic retinopathy severity. In: Engineering in Medicine and Biology Society (EMBC) IEEE, pp. 5632–5635 (2010)
5. Teleophthalmology: A Model for Eye Care Delivery in Rural and Underserved Areas of India. International Journal of Family, Medicine 2011, doi:10.1155/2011/683267
6. Hegde, K., et al.: Comparing the accuracy of a trained technician versus a Pediatric Ophthalmologist-in-training in diagnosing Retinopathy of Prematurity using wide-field digital imaging. In: World ROP Congress, Abstract book, p. 52 (2009)
7. Kumar, S., et al.: TeleOphthalmology Assessment of Diabetic Retinopathy Fundus Images: Samrtphone vs Standard Office Computer Workstation. TeleMedicine and e-Health, doi:10.1089/tmj.2011.0089

Performance Analysis of Ad Hoc On-Demand Distance Vector Routing and Dynamic Source Routing Using NS2 Simulation

Mayank Kumar Goyal, Yatendra Kumar Verma,
Paras Bassi, and Paurush Kumar Misra

Deptt. of CSE/IT, JIIT University, Noida, UP, India
{mayankrkgit,yatendra54,paras123,misrapaurush}@gmail.com
http://www.jiit.ac.in

Abstract. A mobile ad hoc network is a collection of wireless mobile nodes dynamically forming a temporary network without the use of any existing network infrastructure. A number of routing protocols like Dynamic Source Routing and Ad Hoc On-Demand Distance Vector Routing have been implemented. This paper presents the comparative performance analysis of two routing protocols for mobile ad hoc networks. A network simulator has been used for performance evaluation of Ad Hoc On-Demand Distance Vector Routing and Dynamic Source Routing protocol in this paper. To compare the performance of these routing protocols, the simulation results were analyzed using graphs and trace file based on Quality of Service metrics such as throughput, packet delivery fraction, packet loss and average end-to-end delay by varying network size up. The results presented in this paper clearly indicate that the protocols behave differently under different environments.

Keywords: Performance analysis, Ad Hoc Network, AODV, DSR, NS2.

1 Introduction

A mobile ad-hoc network is a kind of wireless ad-hoc network, and is a self-configuring network of mobile routers connected by wireless links - the union of which form an arbitrary topology [1][7]. The routers are free to move randomly and organize themselves arbitrarily thus, the network's wireless topology may change rapidly and unpredictably [2]. Routing is the process of moving a data packet from source to destination. Routing is usually performed by a dedicated device called a router. Part of this process involves analyzing a routing table to determine the best path.

1.1 Ad Hoc On-Demand Distance Vector Routing (AODV)

The Ad hoc On Demand Distance Vector routing algorithm is a routing protocol designed for ad hoc mobile networks [1][3].It is capable of both unicast and multicast

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 390–396, 2013.

routing [9]. It makes routes using a route request / route reply query cycle. When a source node wants a route to a destination for which it does not already have a route, it broadcasts a route request (RREQ) packet across the network. Nodes receiving this data packet update their information for the source node and set up backwards pointers to the source node in the route tables .A node receiving the RREQ may send a route reply (RREP) if it is either the destination or if it has number more than or equal to that contained in the RREQ [5]. It uses sequence numbers to ensure the freshness of routes. It is loop-free, self-starting, and scales to large numbers of mobile nodes [5].

1.2 Dynamic Source Routing (DSR)

Dynamic Source Routing is a routing protocol for wireless mesh networks [10]. It has only two major processes which are Route Discovery and Route Maintenance [4]. The key property of Dynamic source routing is the use of source routing instead of relying on the routing table at each intermediate device [5][6]. Each node receiving a route request packets, rebroadcasts it, unless it is the destination or it has a route to the destination in its route cache. Such a node replies to the RREQ with a route reply (RREP) packet that is routed back to the original source node. RREQ and RREP packets are also source routed. Table1 shows the comparison of Dynamic Source Routing, and Ad Hoc On-Demand Distance Vector Routing.

Table 1. Comparison between AODV and DSR

Protocol property	**DSR**	**AODV**
Loop Free	YES	YES
Multicast routes	YES	NO
Distributed	YES	YES
Unidirectional Link Support	YES	NO
Multicast	NO	NO
Periodic Broadcast	NO	YES
QoS support	NO	NO
Routes Maintained in	Route Cache	Route Table
Reactive	YES	NO

This paper is organized as follows: Section 2 describes the performance analysis of Ad Hoc On-Demand Distance Vector Routing and Dynamic Source Routing and presents the simulation result with analysis related to the objectives. Section 3 provides conclusion.

2 Performance Analysis

2.1 Performance Metrics

In order to evaluate the performance of ad hoc network routing protocols, the following metrics is considered:

End-to-End Delay: The average time interval consumed between the generation of a packet in a source node and the successfully delivery of the packet at the destination node.

Packet Delivery Fraction: The ratio of the number of data packets successfully delivered to all destination nodes and the number of data packets generated by all source nodes.

Number of Packets Dropped: The number of data packets that are not successfully delivered to the destination.

Throughput: The throughput metric measures how well the network can constantly provide data to the sink. Throughput is the number of packet arriving at the sink per millisecond.

2.2 Simulation Environment

The following simulation parameters have been taken for the simulation and are presented in table 2

Table 2. Simulation Parameters

Channel type	Channel/Wireless Channel
Radio-propagation model	Propagation/TwoRayGround
Network interface type	Phy/WirelessPhy
MAC type	Mac/802_11
Interface queue type	Queue/Drop Tail/PriQueue
Link layer type	LL
Antenna model	Antenna/Omni Antenna
Max packet in ifq	50
X dimension of topography	500
Y dimension of topography	400

2.3 Simulation Results

Table 3 shows simulation results for 25 nodes at varying simulation time. Similarly, simulation results for 5, 10, 15, and 20 nodes were computed to analyze the performance of AODV and DSR.

Table 3. Simulation Results for 25 Nodes

Performance metric	**AODV**	**DSR**	**AODV**	**DSR**
Start Time (s)	10	10	10	10
Stop Time (s)	50	50	100	100
Generated Packets	1899	2540	6001	9316
Received Packets	944	1056	2990	4442
Packet Delivery Ratio	49.7104	41.5748	49.825	47.6814
Total Dropped Packets	9	0	20	14
Avg. End-to-End Delay (ms)	168.334	105.797	179.814	114.896
Avg. Throughput(kbps)	1065.17	2114.17	4122.10	11426.70

Table 4. Simulation Results for 25 Nodes

Performance metric	**AODV**	**DSR**	**AODV**	**DSR**	**AODV**	**DSR**
Start Time(s)	10	10	10	10	10	10
Stop Time(s)	150	150	200	200	250	250
Generated Packets	14307	17401	22602	25482	30912	33572
Received Packets	7144	8486	11294	12528	15447	16573
Packet Delivery Ratio	49.9336	48.7673	49.969	49.1641	49.9709	49.3655
Total Dropped Packets	18	11	4	8	17	8
Avg. End-to-End Delay	125.247	94.768	103.29	86.4695	95.6824	82.0416
Avg. Throughput(kbps)	12034.7	25485.1	21572. 79	40393.7	31757.01	55648.06

2.4 Performance Analysis of Ad Hoc On-Demand Distance Vector Routing and Dynamic Source Routing

The graphs below show the behavior of AODV and DSR with respect to various metrics considered above. The X-axis shows the simulation time and Y-axis shows the metrics considered for simulation.

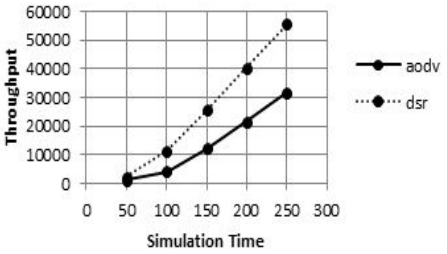


Fig. 1. Simulation Time Vs Throughput with 25 nodes

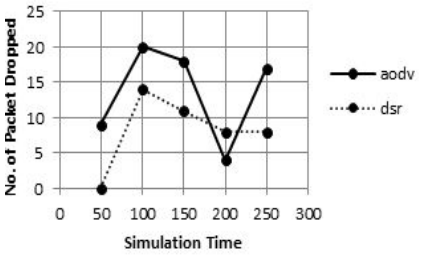


Fig. 2. Simulation Time Vs No. of Packet Dropped with 25 nodes

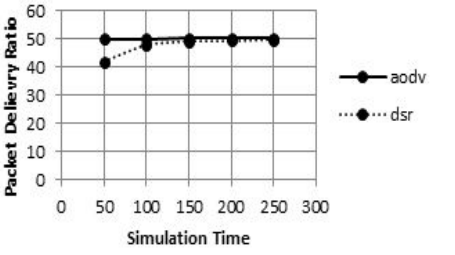


Fig. 3. Simulation Time Vs Packet Delivery Ratio with 25 nodes

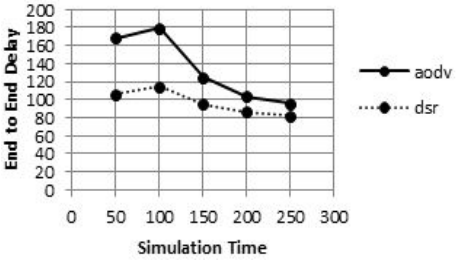


Fig. 4. Simulation Time Vs End to End Delay with 25 nodes

The graphs below show the performance of the two routing protocols with respect to various metrics considered above. The X-axis shows the no. of nodes and Y-axis shows the metrics considered for simulation. Simulations were carried out with the number of nodes as 5, 10, 15, 20 and 25.

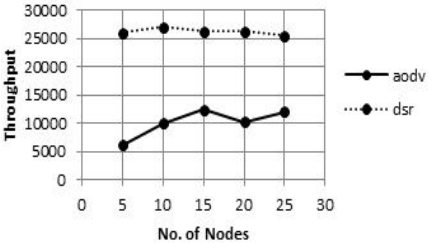


Fig. 5. No. of Nodes Vs Throughput with simulation time 150s

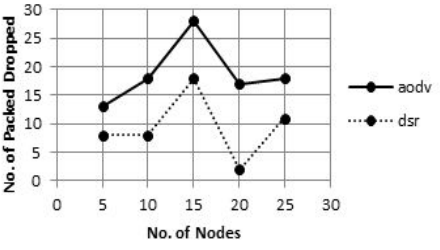


Fig. 6. No. of Nodes Vs No. of Packed Dropped with simulation time 150s

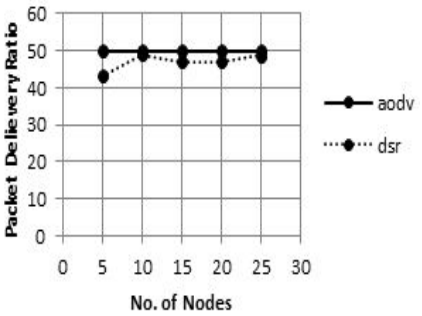


Fig. 7. No. of Nodes Vs Packet Delivery Ratio with simulation time 150s

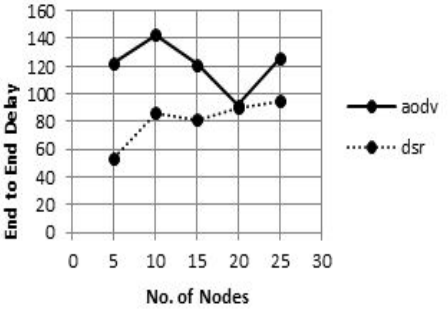


Fig. 8. No. of Nodes Vs End to End Delay with simulation time 150s

Throughput: Fig.5 shows that with increase in no. of nodes, throughput of DSR tends to be much constant as compared to AODV. So, the throughput of DSR outperforms AODV.

No. of Packed Dropped: Fig.6 shows that the behavior of AODV and DSR are same in case of no. of packet dropped, but the no. of packet dropped in case of AODV is more as compared to DSR.

Packet Delivery Ratio: Fig.7 shows the packet delivery ratio for AODV remains constant with increase in no. of nodes whereas the packet delivery ratio of DSR varies with increase in no.of nodes.

End to End Delay: Fig.8 shows the end to end delay for DSR increases with the increase in no. of nodes whereas for AODV it varies with the increase in no. of nodes.

3 Conclusion

This paper shows that when Mobile Ad hoc Networks are used for a longer duration of time then either Ad Hoc On-Demand Distance Vector Routing or Dynamic Source Routing can be used , because after some duration both the protocols show same behavior with respect to packet delivery ratio. The general observation from the simulation is that for application-oriented metrics such as packet delivery fraction Ad Hoc On-Demand Distance Vector Routing, outperforms Dynamic Source Routing in more “stressful” situations (i.e., smaller number of nodes and lower load and/or mobility), with widening performance gaps with increasing stress (e.g., more load, higher mobility). The result also indicates that as the number of nodes in the network increases Ad Hoc On-Demand Distance Vector Routing gives nearly constant throughput. Considering the overall performance of Ad Hoc On-Demand Distance Vector Routing it performs well with varying network size. Delay is high initially in AODV but after some time it is very low.

References

1. Abdullah, A., Ramly, N., Abdullah, M., Derahman, M.N.: Performance Comparison Study of Routing Protocols for Mobile Grid Environment. IJCSNS International Journal of Computer Science and Network Security 8(2), 82–88 (2008)
2. Murthy, C.S.R., Manoj, B.S.: Ad Hoc Wireless Networks: Architecture and Protocols: Routing Protocols for Ad Hoc Wireless Networks, p. 299. Prentice Hall Professional Technical Reference, New Jersey (2004)
3. Altman, E., Jimenez, T.: NS Simulator for Beginners. Lecture notes, Univ. de Los Andes, Merida, Venezuela and ESSI. Sophia-Antipolis, France (2003)
4. Staub, T.: Ad-hoc and Hybrid Networks: Performance Comparison of MANET Routing Protocols in Ad-hoc and Hybrid Networks. Institute of Computer Science and Applied Mathematics, University of Berne, Switzerland, pp. 1–38 (July 2004)
5. Alexander, Z.: Performance Evaluation of AODV Routing Protocol- Real-Life Measurements. In: SCC (June 2003)
6. Johnson, D.: Dynamic Source Routing for Mobile Ad Hoc Networks. IEFT MANET Draft (April 2003)
7. Nesargi, S., Prakash, R.: Configuration of Hosts in Mobile Ad Hoc Networks. In: 21st Annual Joint Conf. of the IEEE Computer and Communications Societies, vol. 2, pp. 1059–1068 (June 2002)
8. Johnson, D., Maltz, D., Broch, J.: DSR: The Dynamic Source Routing Protocol for Multihop Wireless Ad Hoc Networks. In: Perkins, C. (ed.) Ad Hoc Networking, ch. 5, pp. 139–172 (2001)
9. Perkins, C.E., Royer, E.M.: Ad Hoc On-Demand Distance Vector Routing. In: Proc. of the 2nd IEEE Workshop on Mobile Computing Systems and Applications, pp. 90–100 (February 1999)
10. Johnson, D.B., Maltz, D.A.: Dynamic source routing in adhoc wireless networks. In: lmielinski, T., Kmh, H. (eds.) Mobile Computing. Kluwer Academic (1996)

Modified Contrast Limited Adaptive Histogram Equalization Based on Local Contrast Enhancement for Mammogram Images

Shelda Mohan and M. Ravishankar

Department of Information Science and Engineering, Dayananda Sagar College of Engineering, Bangalore, India
{sheldamohan,ravishankarmcn}@gmail.com

Abstract. Optimal Contrast enhancement for detection of masses and micro calcification of mammogram images using Contrast Limited Adaptive Histogram Equalization (CLAHE) based on local contrast modification (LCM) is presented in this paper. The LCM-CLAHE is proposed to highlight the finer hidden details in mammogram images and to adjust the level of contrast enhancement. The proposed method is tested for mammographic images from MIAS database. The performance of the proposed method is obtained using Peak Signal to Noise Ratio (PSNR). The results are compared with other standard enhancement techniques such as Histogram Equalization, Unsharp Masking (USM) and CLAHE. The experimental results of proposed method show that this method provides better contrast enhancement with preserving all the local information of the mammogram images.

Keywords: Local Contrast Modification (LCM), Enhancement parameter, CLAHE, PSNR, USM, HE.

1 Introduction

In the present medical scenario detection of breast cancer in its early stage is a very immense challenge. Even with the advancement in medical technology it is complex to detect cancerous cells in its premature stage. Annual report on status of cancer reveals that one in eight women develops cancer in their lifetime and it is one of the major causes of death for woman in United States [1]. In breast cancer detection the critical part is a method to distinguish between normal tissues and cancerous tissues. Differentiating of this by human eye is very difficult. Mammography is the primary method in the detection of breast cancer, and it is an X-Ray imaging technique. Mammography is very effective method of finding breast diseases. Even with this effective method over 10 percent of the cancerous lesions are left undetected and also has the drawbacks like low contrast images [2].

In order to deal with the low quality or low contrast X-Ray images, Image Enhancement is used. Image Enhancement helps in improving the quality and appearance of the image. Enormous research is being done in the image enhancement in medical field and various techniques has been developed which improved the image quality to a certain extend [3].

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 397–403, 2013.

2 Related Works

Histogram Equalization (HE) is one of the popular methods for contrast enhancement which modify the gray level histogram of an image to a uniform distribution [4]. But in many cases it produces over enhancement in output image and loss of local information which leads to insufficient medical details during diagnosis. To overcome these drawbacks, many variants of HE have been proposed [5-8].

In medical imaging (such as mammogram enhancement) local contrast are more important than global contrast. In such type of applications Global Histogram Equalization (GHE) is insufficient because it cannot deal with local features of original image due to its global nature. Adaptive Histogram Equalization (AHE) method will perform throughout all pixels in the entire image and maps gray level using local histograms, but it takes more time [5]. Pizer has proposed AHE in which the input image is divided into blocks and then the mapping functions are computed for those blocks using CLAHE [7].

The standard CLAHE method produces over enhancement which results in the loss of some local information [9]. In order to overcome this limitation we have proposed LCM-CLAHE. This method will produce optimal contrast without losing any local information of the mammogram image which is most important for detection of breast cancer. The proposed method LCM-CLAHE consists of two stages of processing to increase the potentiality of contrast enhancement and to preserve the local details in the images. The details of the proposed method are presented in the next section.

3 Proposed Method (LCM-CLAHE)

Histogram Equalization uniformly distributes the intensity of the image .But it produces over enhancement in the output image which leads to the loss of local details in the mammogram images [9].AHE differs from the ordinary histogram equalization in the respect that HE generates only one histogram whereas adaptive method computes several histogram, corresponding to a distinct section of the image and uses that to redistribute the intensity values of image.AHE is a better enhancement method for medical images . However, it has a drawback of noise over amplification. CLAHE is a variant of AHE which reduces the noise amplification. Using CLAHE also we have found that it is also not so suitable for mammogram images of very fine details. In Histogram Modified (HM)-CLAHE the author have proposed global modification of histogram along with CLAHE [10].But in mammogram images local details are more important than global details for the detection of cancerous cells. So in the proposed method we have used a local contrast enhancement to highlight the fine details hidden in the mammogram image and an enhancement parameter to control the level of enhancement along with standard CLAHE. So incorporating LCM with CLAHE produces an optimal contrast enhancement with all local information of mammogram images which may not be obtained using Standard CLAHE.

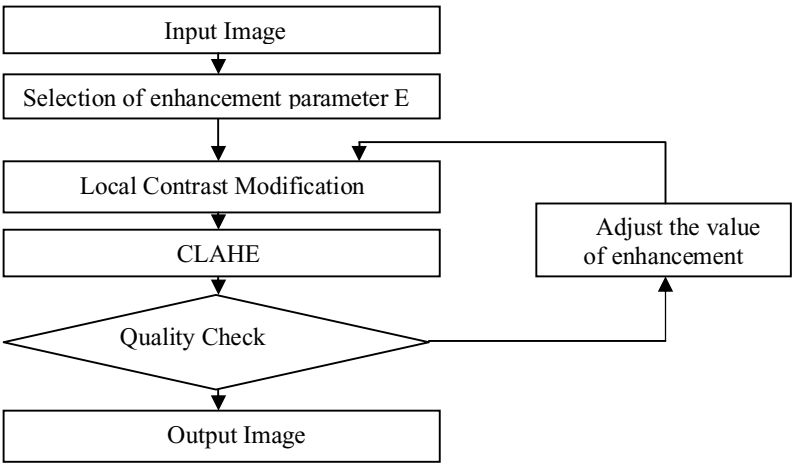


Fig. 1. Flow chart of proposed LCM-CLAHE

Figure 1 show the steps involved in the proposed method. The original image and the enhancement parameter is given as input to LCM .In LCM we modify the image to produce the finer details hidden in the mammogram image and that output image is give as input to CLAHE and CLAHE will further enhance the image with quality check. The quality check used in the proposed method is PSNR. The following sections will summarize in detail the theory of LCM, CLAHE and PSNR.

3.1 Local Contrast Modification (LCM)

The first stage in the proposed mammogram image enhancement method is applying local contrast enhancement on the input mammogram image. The function used in method is designed in such a way that it takes both global as well as local information to produce the enhanced image. The local information is extracted from a user defined window of size n×n pixels. The transformation function can formulated as given below

$$T = \frac{E \cdot M}{\sigma} \tag{1}$$

$$g = T * (f - m) + m \tag{2}$$

where g and f are LCM enhanced and input image respectively and E is the enhancement parameter, M is the global mean of the input image, m is the local mean and σ is the local standard deviation, E is a positive constant and value is in between 0 and 1. The expression for local mean and standard deviation for the user defined window is computed as follows

$$m(x,y) = \frac{1}{n \times n} \sum_{x=0}^{n-1} \sum_{y=0}^{n-1} f(x,y) \tag{3}$$

$$\sigma = \sqrt{\frac{1}{n \times n} \sum_{x=0}^{n-1} \sum_{y=0}^{n-1} \left(f(x,y) - m(x,y) \right)_2} \tag{4}$$

After getting the local mean and standard deviation of all windows, we calculate the average of this and that value is used in equation (1) and (2). This method will highlight the finer details of mammogram images .The enhanced image after this method is given as input to CLAHE.

3.2 CLAHE (Contrast Limited Adaptive Histogram Equalization)

The CLAHE method has good tractability in choosing local histogram mapping function. This method divides the image into appropriate regions and applies histogram equalization to them. It modifies the intensity values of the image by employing a nonlinear methodology in order to maximize the contrast for all pixels of the image. The clipping level selection of the histogram reduces the undesired noise amplification. In the proposed method Local Contrast Enhancement process is the first step and then CLAHE is applied to Local Contrast Modified image. By selecting optimal clipping level in CLAHE and the enhancement parameter in Local Contrast Enhancement the proposed method achieves optimum contrast enhancement for the mammogram images.

4 Experimental Results and Discussion

This section presents the experimental results of the proposed method LCM-CLAHE. In this paper, the most popular image enhancement techniques like HE, USM and CLAHE techniques are chosen in order to validate the proposed technique. The Local Contrast Enhancement in the LCM-CLAHE method preserves the local information. Determining the optimum contrast enhancement without losing fine details is a very b challenging in mammogram contrast enhancement. With the Enhancement parameter E=0.4 and clipping factor=0.015 the proposed method achieves optimum contrast enhancement for mammogram images as shown in Table 1.

Table 1. Comparison of PSNR values produced by HE, USM, CLAHE and proposed method LCM-CLAHE

Image	PSNR			
	HE	USM	CLAHE	LCM-CLAHE
mdb019	33.76	32.19	54.01	**46.09**
mdb213	29.68	35.15	53.68	**45.17**
mdb238	30.71	35.52	50.71	**43.95**
mdb076	31.01	34.58	55.34	**44.97**
mdb086	29.31	35.46	51.41	**45.97**
mdb201	34.78	30.10	50.68	**43.85**
mdb202	34.84	32.01	52.59	**43.43**
mdb218	33.04	34.00	55.04	**43.94**
mdb214	29.08	35.71	51.17	**44.39**
mdb220	33.52	32.91	50.20	**45.79**
mdb235	32.59	35.41	50.55	**43.11**
mdb239	34	31.75	53.28	**43.03**
mdb242	29.50	34.84	55.20	**45.60**
mdb243	31.26	33.74	52.99	**44.45**
mdb244	31.66	32.90	51.74	**43.51**
mdb249	31.34	33.35	52.79	**43.94**
mdb254	34.16	32.68	55.97	**44.14**
mdb280	47.23	29.11	55.73	**44.46**

5 Performance Measure

An image is said to be enhanced if it allows the viewer to better perceive the desirable information in the image. The performance measure used here is Peak Signal to Noise Ratio (PSNR) [11]. The PSNR value of an image G with respect to the original image F, both of size M×N pixels, is calculated as shown below

$$M_s = \sum_{m,n} \left[F(m,n) - G(m,n) \right]^2 \tag{5}$$

$$PSNR = 10 * \log 10 \frac{255^2}{M_s} \tag{6}$$

Where Ms is the mean Squared Error given by Equation (5).

From experimental result, it is clear that LCM-CLAHE provides an optimal contrast enhancement for image mdb235 (PSNR=43.11) without losing the finer information of original image where as for Unsharp Masking (PSNR=36.09), HE (PSNR=29.09) is not enhanced properly and CLAHE (PSNR=52.34) which shows over enhancement.

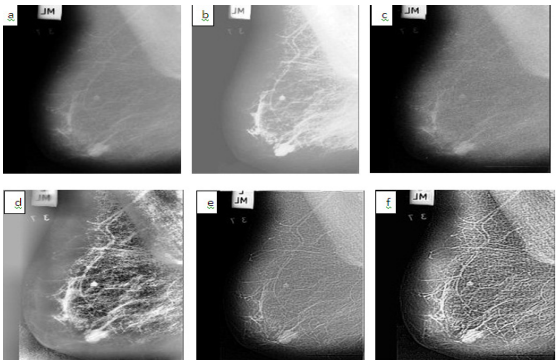


Fig. 2. Enhancement results for fatty mammogram image (mdb005) (a) original mammogram image (b) Image Enhancement using Histogram Equalization (c) Image Enhancement using USM (d) Image Enhancement using CLAHE (e) LCM Enhanced Image (f) Proposed method

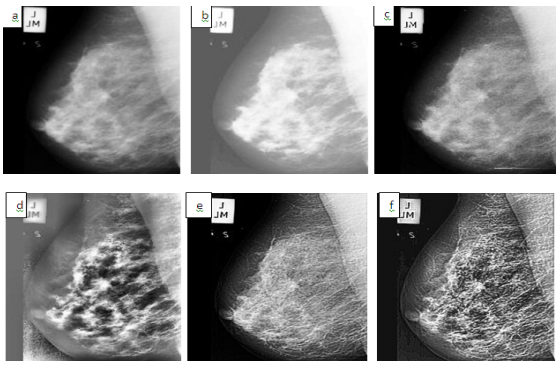


Fig. 3. Enhancement Results of Fatty Glandular Mammogram images (mdb219) (a) original image (b) Image Enhancement Using Histogram Equalization (c) Image Enhancement Using USM (d) Image Enhancement Using CLAHE (e) LCM Enhanced Image (f) Proposed Method

6 Conclusion

A new approach for image enhancement using Modified CLAHE based on local contrast enhancement has been presented in this paper. The proposed method provides a better contrast enhancement and information preservation of input mammogram image. The mammogram images enhanced using proposed method are satisfying both subjective and objective evaluation. The experimental results of LCM-CLAHE show that this method is more effective without compromising contrast as well as original information. All types of mammogram images like fatty, fatty-glandular and dense glandular can be enhanced effectively using this method. Currently, work is under progress for amending the proposed method for other types of medical images.

References

1. Wingo, P.A., Ries, L.A., Giovino, G.A.: Annual report to the nation on the status of cancer. with a special section on lung cancer and tobacco smoking, J. Natl. Cancer Inst. 91, 675–690 (1999)
2. Moskowitz, M., Donegan, W.L., Spratt, J.S. (eds): Cancer of the Breast, Philadelphia, Saunders, pp. 206–239 (1995)
3. Lassó, E., Trucco: Vessel enhancement in digital X-ray angiographic sequences by temporal statistical learning. Comput. Med. Imaging Graph. 29, 343–355 (2005)
4. Gonzalez, R.C., Woods, R.E.: Digital Image Processing, 3rd edn. Addison-Wesley, Reading (1992)
5. Stark, J.A., Fitzgerald, W.J.: An Alternative Algorithm for Adaptive Histogram Equalization. Graphical Models and Image Processing 56, 180–185 (1996)
6. Paranjape, R.B., Morrow, W.M., Rangayyan: Adaptive neighborhood histogram equalization for image enhancement. CVGI 54, 259–267 (1992)
7. Pizer, S.M., Amburn, E.O.P., Austin, J.D., et al.: Adaptive histogram equalization and its variations. CVGIP 39, 355–368 (1987)
8. Kim, Y.T.: Enhancement using brightness preserving bi-histogram equalization. IEEE Trans. Consum. Electron. 43, 1–8 (1997)
9. Sundaram, M., Ramar, K., Arumugam, N., Prabin, G.: Histogram Modified Local Contrast Enhancement for Mammogram Images. Applied Soft Computing 11(8), 5809–5816 (2011)
10. Sundaram, M., Ramar, K., Arumugam, N., Prabin, G.: Histogram based contrast enhancement for mammogram images. In: Proceedings of 2011 International Conference on Signal Processing, Communication, Computing and Networking Technologies, ICSCCN 2011 (2011)
11. Papadopoulos, D.I., Fotiadis, L., Costarido: Improvement of microcalcification cluster detection in mammography utilizing image enhancement techniques. Computers in Biology and Medicine 38, 1045–1055 (2008)

Zone Based Effective Location Aided Routing Protocol for MANET

G.T. Chavan and Vemuru Srikanth

[1] K.L.University, Vijayawada, Andhra Pradesh, India
gt.chavan@gmail.com
[2] Dept. of IST, K.L.University, Vijayawada, Andhra Pradesh, India
vsrikanth@kluniversity.in

Abstract. As the nodes are mobile in MANET, it is very difficult to find the location of destination. Several routing protocols have already been proposed for ad hoc networks. This paper suggests an approach to utilize location information to improve performance of routing protocols for ad hoc networks. By using location information, the proposed method limit the search for a new route to a smaller zone using the concept of request zone and expected zone of the ad hoc n*etwork.*.

Keywords: MANET, Location Zones, LAR1, LAR2.

1 Introduction

Nodes within MANET are organized in different manners, they can be hierarchical or flat, they can move in any direction and speed, and communicate to each other by means of routing protocols. Recent availability of small inexpensive low power GPS receivers and techniques for finding relative coordinates based on signal strengths, and the need for the design of power efficient and scalable networks, provided justification for applying position based routing methods in ad hoc networks. To establish a data transmission between two nodes, typically multiple hops are required due to the limited transmission range. Although many routing protocols have been proposed, still needs some improvement.

2 Routing Protocols in MANET

Routing protocols are divided into three categories: Proactive, Reactive and Hybrid. In Proactive protocols each mobile node maintains a routing table where control packets are broadcast periodically across the whole network. In Reactive routing protocol, the routes are discovered only when needed. Hybrid routing protocols uses the advantages of proactive and reactive approaches. The hybrid protocols separate the network into two layers.

There are two methods of transmitting data in position based protocols: greedy forwarding and directional flooding. GPSR (Greedy Perimeter Stateless Routing) is

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 404–407, 2013.

an example of greedy forwarding where the next hop is the closest node in distance to the destination. LAR (Location Aided Routing) is an example of a directional flooding protocol [1]. To calculate the nodes position, there are two ways to do this. The nodes may use a GPS system. They estimate their (x, y) position using the distance formula

$$d = \sqrt{(x_2 - x_1)^2 - (y_2 - y_1)^2} \quad \& \quad \theta = \tan^{-1} \frac{(y_2 - y_1)}{(x_2 - x_1)} \tag{1}$$

Where d= distance & θ= inverse tangent function to calculate angle.

The second way is to use a directional antenna approach based on directional arrival and signal strength.

2.1 Location-Aided Routing (LAR)

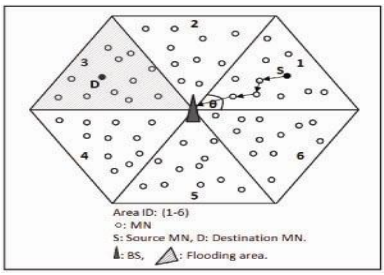
This protocol is an approach that decreases overhead of route discovery by utilizing location information of mobile hosts. It uses location information and concept of request zone and expected zone. In LAR's limited flooding, a node forwards a route request only if it belongs to the request zone and the request zone should include expected zone and other regions around the expected zone. Two LAR algorithms have been presented: LAR scheme 1 and LAR scheme 2.

LAR scheme 1 uses expected location of the destination (so-called expected zone) at the time of route discovery in order to determine the request zone. The request zone used in LAR scheme 1 is the smallest rectangle including current location of the source and the expected zone for the destination. LAR scheme 2 uses distance from the previous location of the destination as a parameter for defining the request zone. Thus, any intermediate node J receiving the route request forwards it if J is closer to or not much farther from the destination's previous location than node I transmitting the request packet to J. Therefore, the implicit request zone of LAR scheme 2 becomes adapted as the route request packet is propagated to various nodes.

3 Proposed Methodology

In order to reduce the area of searching destination node ELAR protocol uses a wireless base station (BS) that covers all MNs in the network. BS efficiently route packets among MNs, it keeps a Position Table (PT) that stores locations of all MNs. PT is built by BS through Broadcasting Small BEACON Packets to all MNs in the Network. Based on the AoA MN, BS can determine the network area in which each MN is located. Protocol first builds and updates Position Table in BS. If any node enter in the network area buildUpdate position table procedure is called to report position of the node to the BS. When S wants to send data to D, protocol will execute Data transmission procedure.

BS divides the network into six areas as shown in Fig.1



Area ID	Range of angle θ
1	$0 \leq \theta < \pi/3$
2	$\pi/3 \leq \theta < 2\pi/3$
3	$2\pi/3 \leq \theta < \pi$
4	$\Pi \leq \theta < 4\pi/3$
5	$4\pi/3 \leq \theta < 5\pi/3$
6	$5\pi/3 \leq \theta < 2\pi$

Fig. 1. Area ID (1-6)

Algorithm ELAR () {

```
Thread (BuidUpdatePositionTable);
While (1) {
        If ( a mobile node enters network area of the base station)
                Thread (BuildUpdatePositionTable)
        If (source mobile node wants to send data to a destination mobile node)
                Thread (data Transmission);
              }// end while
  }// end ELAR
```

BuildUpdate Procedure

- Procedure starts when A enter in the Network range of BS.
- A send PosReq message with its X,Y to BS.
- After receiving message BS will update PT
- Depending upon AoA BS send area ID, distance to node A
- Periodically procedure is repeated
- When MN stop sending packet then it is marked unreachable

Data Transmission Procedure

- Node S request BS to initiates route discovery to D.
- BS check PT is out of date or not if yes then send small message to D for its new position.
- Depending upon AOA BS find area ID and send msg to S
- If BS determines that S and D are not in the same area the data transmission done through BS.
- Else data transmission done without BS

4 Conclusion

This protocol is an optimization to the LAR. Protocol makes significant reduction in the routing message by limiting the area of discovering a new route to a smaller zone. Control packet overhead is significantly reduced and the mobile nodes life time is increased. Position based routing protocols provide better scalability and performance by using geographical position to improve routing decisions & efficiency

References

1. Sarkar, S.K.: AUERBACH "Ad Hoc Mobile Wireless Networks Principles Protocols and Applications"
2. Perkins, C.E.: AD HOC Networking. Pearson, Copyright by Addison-Wisley
3. Niculescu, D., Nath, B.: Ad hoc Positioning System (APS) using AOA. In: Proc. IEEE INFOCOM (2003)
4. Susinder Rajan, G., Sundar Rajan, B.: MSE Optical Algebraicspace time Code. IEEE Transaction on Wireless Comminication 7(7) (July 2008)
5. Okada, H., Takano, A., Mase, K.: Analysis and proposal of position based routing protocols for vehicular ad hoc networks. IEICE Transactions 91-A(7), 1634–1641 (2008)
6. Blazevic, L., Le Boudec, J.-Y., Giordano, S.: A Location-Based Routing Method for Mobile Ad Hoc Networks. IEEE Transactions on Mobile Computing 3(4) (October-December 2004)
7. Ennaji, R., Boulmalf, M.: Routing in Wireless Sensor Networks. IEEE Transaction (2009), 978-1-4244-3757-3
8. Lee, S.B., Hur, K.K., Park, J., Eom, D.-S.: A Packet Forwarding Controller for Mobile IP-Based Networks with Packet buffering. IEEE Transaction on Consumer Electronics 55(3) (August 2009)
9. Thipchaksurat, S., Kirdpipat, P.: Position-based Routing Protocol by Reducing Routing Overhead with Adaptive Request Zone for Mobile Ad Hoc Networks. In: The 8th Electrical Engineering! Electronics, Computer, Telecomm. and Information Technology (ECTI), Paper ID 1420 Association of Thailand - Conference 2011 (2011)
10. Govindaswamy, V., Blackstone, W.L., Balasekaran, G.: Survey of Recent Position Based Routing Mobile Ad-hoc Network Protocols. In: 2011 UKSim 13th International Conference on Modelling and Simulation, 978-0-7695-4376-5/11 $26.00 © 2011 (2011) IEEE, doi:10.1109/UKSIM.2011.95
11. Navidi, W., Camp, T.: Stationary distributions for the random waypoint mobility model. IEEE Transactions on Mobile Computing 3(1) (2004)

Investigation of Modulus Maxima of Contourlet Transforms for Enhancing Mammographic Features

Rekha Lakshmanan and Vinu Thomas

Department of Electronics and Communication Engineering
Govt. Model Engineering College, Thrikkakara, Kochi – 682021, Kerala, India
{rekha,vt}@mec.ac.in

Abstract. Enhancing features of mammographic image assists radiologists in the early detection of breast cancer. In this paper, an enhancement technique using selected modulus maxima of the Contourlet transform is employed to enhance the microcalcification features in mammographic image while simultaneously reducing image noise. Strong edge information corresponding to relevant features was retained based on a parent child relationship among contourlet coefficients at various levels. Simulations were carried out to examine the utility of the proposed technique in mammographic image enhancement. The mini – MIAS database was employed to test the accuracy of proposed method. Contrast improvement index, Peak Signal to Noise Ratio, Target to Background Contrast ratio and Tenengrad Criterion were considered for a evaluating the performance of the proposed methods.

Keywords: Mammogram, Breast Cancer, Modulus-Maxima, Enhancement, Contourlet Transform.

1 Introduction

Breast cancer is one of the leading causes of cancer death among women. According to the statistics in 2010, the expected casualty due to breast cancer is about 39,840 [1], even though there is decrease in death rates since 1990. Mammographic image analysis is a complicated and difficult task which requires opinion of highly trained radiologists. Detection of microcalcification, a possible symptom of breast cancer is a complex task because of the inhomogeneous background and the high noise level in the images due to emulsion artifacts. Microcalcifications are tiny, granular, linear, or irregular deposits of Calcium Phosphate Hydroxide which appear as bright white spots with size ranging from 0.1–1.0 mm and an average diameter about 0.3 mm [2]. The proposed algorithm makes use of the multiscale and directionality properties of the Contourlet Transform(CT) [3] to enhance clinically important microcalcification features in mammograms while suppressing artifacts such as those due to emulsion discontinuity. The performance of the proposed enhancement method is seen to be much better than the approach which employs the local modulus maxima of the Discrete Wavelet Transform (DWT), which is attributed to the directionality and anisotropy properties of contourlet transform.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 408–411, 2013.

2 Proposed Method

With the multiscale and directional analysis of the contourlet transform, the proposed approach can effectively enhance sharp variation points indicating the possible occurrence of microcalcifications. Here we employ local modulus maxima of the directional filter bank outputs in the filter bank implementation of a multilevel contourlet transform. The impulse response of a contourlet coefficient in the directional subband after LP decomposition can be modeled as a smooth function which is oriented in the direction of the impulse response of the corresponding directional filter [4]. Edges can be defined as local sharp variation points of image smoothened by a smooth function [5]. These sharp variation points in the image can be characterized by the contourlet maxima. Clinically important microcalcification structures form edge features which are present at more than one level of resolutions in the contourlet decomposition whereas edges due to image artifacts are very fine structures that are significant at only the finest resolutions. Accordingly, a tree is constructed based on the parent-child-cousin relationship [6] among contourlet coefficients at various levels. The modulus maxima in each directional subband are found out in the direction of orientation of the directional filter impulse response. The tree retains only those modulus maxima that propagate to the coarser levels of the contourlet decompositions, while discarding those modulus maxima that exist only at the finer resolutions. The contourlet coefficients on either side of the selected modulus maxima are boosted by a scale factor. Thus only the coefficients that correspond to significant edge features containing microcalcification features are enhanced. The inverse contourlet transform is performed with the modified directional subband coefficients to obtain an image with enhanced microcalcification features.

3 Algorithm

Step1) Computation of contourlet coefficients of image using Pyramidal directional filter bank (PDFB) decomposition in CT
Step2) Discontinuities in each subbands in different levels are determined using corresponding modulus maxima.
Step3) Coefficients that hold parent–child relationships from finer to coarser levels are retained and others are pruned out.
Step4) Resultant coefficient values obtained from step 3 can be boosted by scaling and is added with coefficients obtained from step1.
Step5) Assign coarse coefficients to zero. Apply PDFB reconstruction on boosted coefficients to get reconstructed image.
Step6) Apply global threshold on coarse suppressed image to get segmented microcalcifications in digitized mammograms.
Step7) Summing up of the result with the original image yields an enhanced image.

4 Results and Discussions

The mini-MIAS mammographic database [7] was employed for benchmarking the proposed algorithm. Three levels of the Laplacian Pyramid with four levels of the Directional Filter Bank at each level of the LP decomposition were employed for the contourlet decomposition of the mammogram images. The PKVA filters [8] were employed for the Laplacian Pyramid and BIOR 9, 7 filters [9] were employed at the DFB stage. In order to evaluate the enhancement results of the proposed method, contrast improvement index (CII), the peak signal-to noise ratio (PSNR), Target to Background Contrast (TBC) and Tenengrad criterion(TEN) [4], [10] were employed.

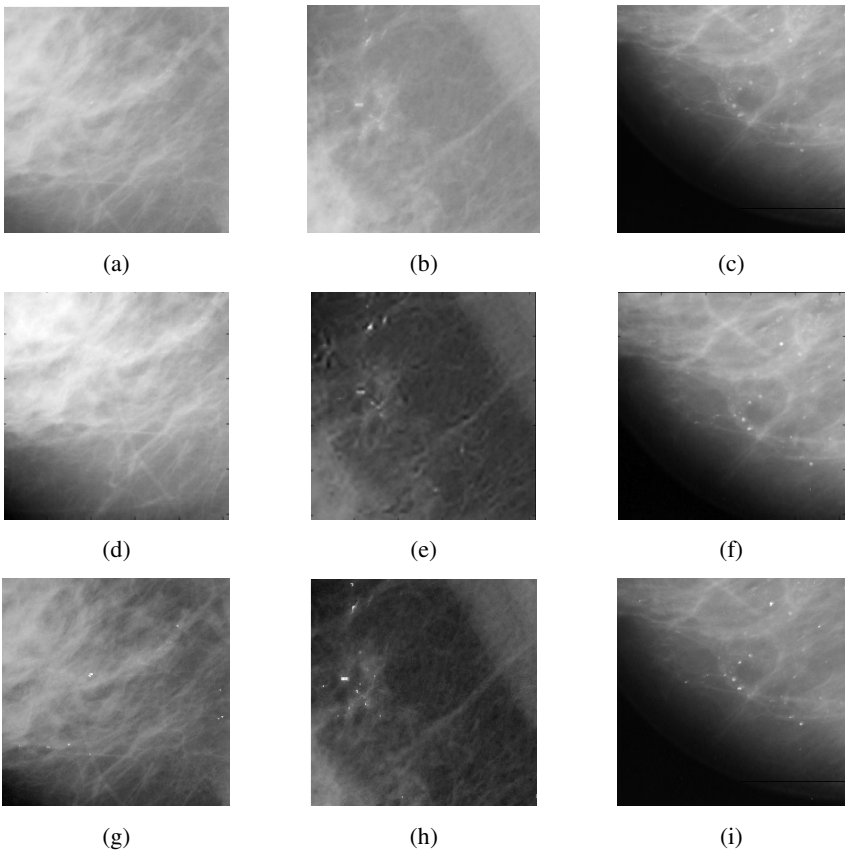


Fig. 1. The enhancement and segmented results of microcalcifications on selected ROI's: (a) (b) (c) original ROI, (d) (e) (f) enhancement by the wavelet approach, (g) (h) (i) enhancement by the contourlet approach using modulus-maxima

Figure 1 shows the resultant reconstructed images obtained after the inverse contourlet transform using modulus maxima technique. The results are shown for the Mdb211, MDb245 and Mdb249 images from the mini MIAS database. Mdb249 is a

mammographic image having dense glandular background tissue with well-defined malignant microcalcifications. Mdb211 is a fatty glandular mammogram with a difficult to detect microcalcification cluster. Mdb245 represents a mammographic image with widely distributed calcifications. The result in figure 2 reveals that the proposed technique enhances the appearance of the microcalcifications against the surrounding dense tissue which may otherwise obscure these structures. On analysis of quality measures, the average TBc and CII value of CT found to be two and four times better than that of DWT respectively. CT has got an average improvement of 64% with the original compared to 29% of DWT for PSNR. Average TEN value of CT is 0.27% more than original while that of DWT is 0.26%. The average execution time for the proposed method is around 3 seconds.

5 Conclusion

In this paper, selected modulus maxima of the contourlet coefficients were employed for fast mammographic image enhancement. Directional subband coefficients on either side of the selected modulus maxima in the corresponding tree were boosted while the modulus maxima that exist at only the finest scales were suppressed. Relevant edge features including microcalcification structures were enhanced while image artifacts that exist at only the finest scales were suppressed. Quality measures including CII, PSNR, TBc and TEN were calculated to determine the efficiency of the proposed method.

References

1. Cancer facts and figures, American Cancer Society (2010)
2. Li, L., Qian, W., Clark, L.P.: X-Ray medical image processing using directional wavelet transform. In: Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing, Atlanta, vol. 4, pp. 2251–2254 (1996)
3. Do, M.N., Vetterli, M.: Contourlet transform: An efficient directional multiresolution image representation. IEEE Trans. on Image Processing 14(12), 2091–2096 (2005)
4. Rekha, L., Rahila, R., Vinu, T.: Enhancement of Microcalcification Features in Mammograms using Zero-Crossings of the Contourlet Transform. JETEAS 2(6) (December 2011)
5. Mallat, S., Zhong, S.: Characterization of signals from multiscale edges. IEEE Transactions on Pattern Analysis and Machine Intelligence 14(7), 710–732 (1992)
6. Po, D.D.Y., Do, M.N.: Directional multiscale modeling of images using the contourlet transform. IEEE Trans. on Image Processing 15(6) (June 2006)
7. Suckling, J.: Mammographic Image Analysis Society Database (1994), `http://peipa.essex.ac.uk/ipa/pix/mias`
8. Vetterli, M., Herley, C.: Wavelets and filter banks: Theory and design. IEEE Trans. Signal Proc. 40(9), 2207–2232 (1992)
9. Phoong, S.M., Kim, C.W., Vaidyanathan, P.P., Ansari, R.: A new class of two-channel biorthogonal filter banks and wavelet bases. IEEE Trans. Signal Proc. 43(3), 649–665 (1995)
10. Chen, Z.Y., Abidi, B.R., Page, D.L., Abidi, M.A.: Gray-Level Grouping (GLG): An Automatic Method for Optimized Image Contrast Enhancement—Part I: The Basic Method. IEEE Transactions on Image Processing 15(8) (August 2006)

Software Configuration Development Process for Banking and Financial Services IT Industry

Poyyamozhi Kuttalam[1], N. Keerthika[2], K. Alagarsamy[3] and K. Iyakutti[4]

[1] Larsen and Toubro Infotech
poyyamozhi@gmail.com
[2] Department of Information Technology, Sethu Institute of Technology, India
ns.keerthika@gmail.com
[3] Computer Center, Madurai Kamaraj University, India
alagarsamymku@gmail.com
[4] School of Physics, Madurai Kamaraj University, India
iyakutti@gmail.com

Abstract. The purpose of this effort is to outline the Software Configuration Management process for Banking and Financial Services Industry which enables the software project managers to manage the configuration items (CI's) that are appropriate for the project level as defined by the high level configuration risk assessment. It will provide for the creation and ongoing maintenance of the project plan component defined for the configuration management plan. This effort describes the processes and procedures used for these tasks:

- Configuration level determination for the project
- Identification of existing configuration items
- Creation of configuration item's package
- Tracking and Monitoring of configuration activities with recorded metrics
- Closing down the project and archiving configuration metrics

Keywords: Software Configuration Management, configuration items, process diagram for SCM process.

1 Introduction

This paper provides support for three levels of users:

- First time users should be familiar with all of the information in this paper
- Users with some experience in this process should follow the steps outlined in the Process Steps section
- Experienced users should follow the information illustrated in the Process Flow section

1.1 Benefits

This process will allow the project manager, and supporting IT participants in the BFS Industry, to properly define and maintain configuration items (CI's) during the

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 412–423, 2013.

project life cycle. The process will also provide the ability to gather, document, analyze, and report on configuration item metrics. It will provide for visible updates to CI's by all parties and reporting on the progress and current status of CI's being utilized by the project team and customers. [1]

1.2 Interdependencies

1.2.1 Project Management Policies

This process should fully satisfy the section on Configuration Management in the BFS organization's IT Management Plan statement. It should meet all of the stated objectives. [3] [5] [20] [4]

SEI Capability Maturity Model Key Process Areas/Goals

This process fully satisfies all of the key practices of the Configuration Management (SCM) which is a Level 2 (Repeatable Level) Key Process Area (KPA) of the Capability Maturity Model (CMM). [4]

1.2.2 PMI PMBOK

This process is in full compliance with the guidelines set down in the Project Management Institute's Project Management Book of Knowledge. Areas satisfied include Project Integration Management, Project Cost Management, and Project Communications Management [3] [5]

1.2.3 Process Interdependencies

This process is fully interdependent with the following processes:

- Estimating Process
- Project Planning Process
- Software Quality Assurance Process
- Requirements Development / Management Process
- This process is fully dependent on the following processes:

Requirements Development portion of the Requirements Development / Management Process:

2 Process Description

This process will define the roles and responsibilities of the project manager as well as the configuration control board (CCB) during the life cycle of the project. It will define key activities to be performed, measurements/metrics to be gathered, documented, analyzed, and reviewed, reviews or inspections to be performed, and verifications/validations of the tasks being performed by the configuration control board. [2] [6]

3 Activity Diagram

3.1 Configuration Level Determination for the Project

The first activity in the configuration management process is to define the level of configuration control to be maintained by the project team. This task is extremely important because it begins the creation of a critical piece of the Project Plan, which is the Configuration Plan. [11] [13] [17] [18]

3.2 Configuration Level Determination Procedure Notes

- Determine the level of configuration risk for the project using the Risk Level Determination Spreadsheet Template
- Define when the work is to begin, how long it is estimated to last, and when the expected completion date is
- Define any known assumptions to be taken into consideration in the creation and/or maintenance of configuration items
- Define any known constraints that the Project Team must adhere to in the ongoing maintenance of the configuration items (CI's)
- Define any dependencies which already exist or will exist by the introduction of new CI's into the overall configuration environment
- Define any applicable standards that the project team will have to adhere to
- Define the environment in which the CI's will exist (current infrastructure and architecture)
- Define the current communications network and the expected reporting structure
- Define the overall context of the work to be performed by the project team
- Define the skill set required in order to perform the work requested
- Define the tools to be used, if any, in the ongoing maintenance of the CI's

3.3 Creation of Configuration Items Package

The third step is to prepare the configuration package that will be maintained by the project team. This package should contain all of the information necessary for the project team to keep up with the current status of all new and existing CI's. The package will follow the project development through deployment. [11] [14] [16] [8]

3.4 Creation of Configuration Items' Package Procedure Notes

- Gather all information together which is part of the CI package
- Review the CI package with the project team, the customer, and other interested parties (such as CCB and PMQA group)
- Obtain the go ahead to submit the package for initial baseline

- Document and report on the status of the package created with the date, numbers and names of CI's, whether the CI is new or existing, the type of CI's, and whether the CI package is part of a light, medium, or heavy configuration
- Update the Configuration Plan component in the Project Plan

3.5 Tracking and Monitoring of Configuration Activities with Recorded Metrics

The next step is to track the CI's as the project proceeds through the life cycle. This procedure should involve the project manager, the development team (whether there are internal or external resources being used), management, the customer, and the Configuration Control Board (CCB). [11] [10] [12] [7]

3.6 Tracking and Monitoring of Configuration Activities Procedure Notes

- Gather the information on every CI from the person(s) responsible for the creation and maintenance of the CI
- Review the results on each CI with the team and the CCB
- Gather metrics on the CI according to a predetermined set of guidelines
- Document the metrics and keep the documentation in the project repository
- Report on the metrics according to a predetermined schedule
- Monitor the CI's as they pass from phase to phase in the Project Life Cycle
- Report any variances in the CI metrics which are outside of organization or management defined thresholds

3.7 Closing down the Project and Archiving Configuration Metrics

Upon closing down the project, it is important to retain the information gathered and documented on the configuration items. This will assist the next

Project team and project manager in determining what the status of the configuration items are when that project begins. [11] [9] [15] [8]

3.8 Closing Down the Project and Archiving Configuration Metrics Procedure Notes

- When the project is complete, archive the metrics on actual numbers and types of CI's in the project repository archive.
- Also archive any appropriate assessment results in the control of the CI's.
- Archive any ratings, contacts, and other pertinent information useful in the selection of CI's for future projects in the project repository.

3.9 Entry Criteria

This process assumes that the following processes or steps of the life cycle are complete before this process begins:

- The SRF for the Project is complete
- The identification of the location of the project in the repository is complete
- The environment for the project has been identified
- The current status of CI's in the environment has been documented and reported

3.10 Inputs

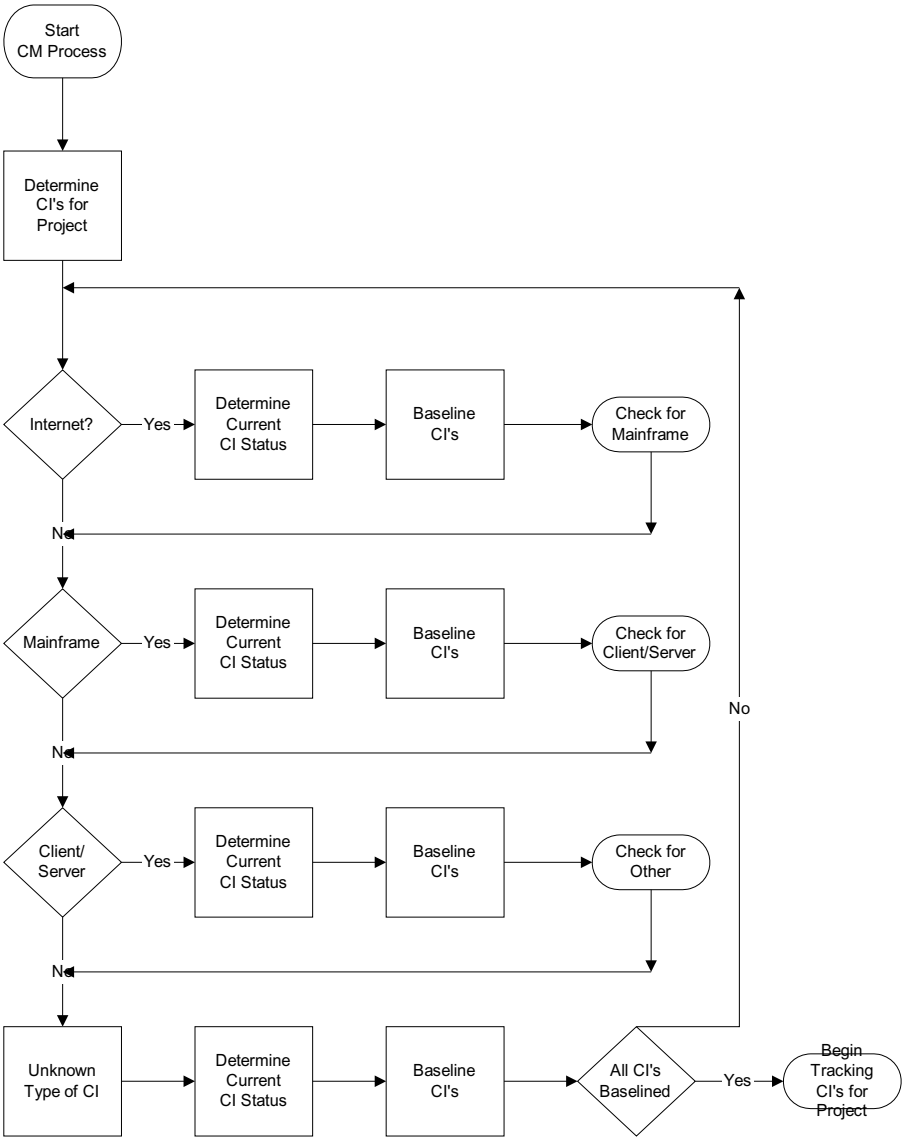
The following inputs are required (or are desirable to have) for the completion of the process: [1]

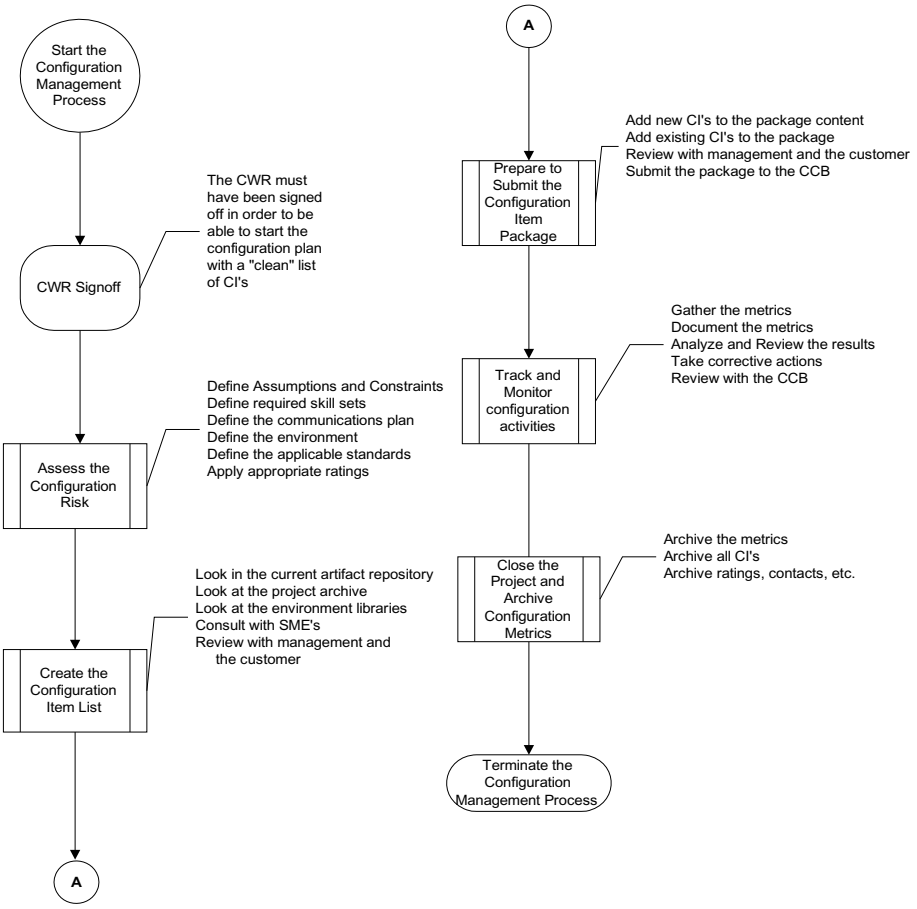
- Project Level Selection Form
- Risk Level Determination Spreadsheet
- Configuration Plan
- Initiation and Draft Scope Document
- Requirements Definition Document
- Initial Forecast and/or Preliminary Estimate
- Assumptions, Constraints, and Dependencies Lists

3.11 Outputs

The following documentation is output from the Configuration Management Process: [1]

- List of appropriate CI's in the Configuration Item Package to be deployed
- Updated Project Schedule
- Updated Configuration Plan
- Updated Communications Plan
- Updated Quality Plan
- Updated Risk Plan
- Updated Estimates
- Updated Project Budget
- Updated Training Plan
- Updated Assumptions, Constraints, and Dependencies Lists
- Updated Test Plan (indicating what CI's will be tested)
- Completed Statement of Work for delivery to the subcontractor
- Completed Request for Proposal


Start CM Process
Determine CI's for Project
Internet?
Yes
Determine Current CI Status
Baseline CI's
Check for Mainframe
No
Mainframe
Yes
Determine Current CI Status
Baseline CI's
Check for Client/Server
No
Client/ Server
Yes
Determine Current CI Status
Baseline CI's
Check for Other
No
Unknown Type of CI
Determine Current CI Status
Baseline CI's
All CI's Baselined
Yes
Begin Tracking CI's for Project
No



3.12 Exit Criteria

The tasks to be performed following the completion of the Configuration Management Process are: [1]

- Update the Project Plan
- Complete a Project Review
- Obtain an Authorization to Proceed
- Make notations of configuration item status in the next status reporting period
- Complete the Integration Process for the project

4 Roles and Responsibilities

Role	Responsibility
Project Manager (PM)	It is the responsibility of the PM to ensure that all configuration items for the project are signed off and ready to be submitted to the CCB. It is also the PM's task to review the selection of the CI's to be used. Finally, it is the PM's task to track the metrics for the CI's and report to management and the customer on the current status of all appropriate CI's and their baseline statistics.
Senior Management	It is the responsibility of Senior Management to review the CI Package that will be submitted for deployment. Senior Management must ensure that when the package is deployed, it is according to standards. It is also the responsibility of Senior Management to review the metrics on the CI's and to take corrective action if there are variances in CI metrics outside the agreed upon boundaries.
Team Members	It is the responsibility of the team members to advise the PM on prior CI utilization and what were the high and low points of the CI's durability and reusability. It is also the task of the team members to assist in keeping the CI's up to date to ensure that the CI's are performing at peak efficiency.
Customer	It is the job of the customer to also advise the PM on CI's and their content. There must be a strong customer relationship with IT. The customer can also advise on any specific constraints that should be placed on the CI's.
Sponsor	It is the responsibility of the Sponsor to provide oversight of the process of configuration administration. The Sponsor should review the performance of the CI's and provide feedback, through the Product Champion or other customers on their rating of the CI's effectiveness.
Product Champion	It is the responsibility of the Product Champion to ensure that CI's are fully documented in terms of the impact to the business that the CI might make. It is also their task to ensure that the CI's are as business-friendly as possible.
CCB	It is the CCB's responsibility to read and understand the CI content, submitting an appropriate response to the project for their services. The CCB must also make the company aware of any conditions that might alter the relationship after the project is underway. The CCB must come up to speed as rapidly as possible on any business or technical disciplines required to perform their assigned tasks.
SQA Assessment Group	It is the responsibility of the SQA Assessment Group to review the projects periodically and have appropriate questioning about process on any interaction between the PM, customers, team, management, and the CCB.

5 Measurement and Analysis

There are several appropriate metrics to be gathered, documented, analyzed, and reported during the configuration management process. The following information details those activities.

5.1 Metrics

- Number of CI's by type estimated
- Actual number of CI's by type
- Number of CI's created, enhanced, deleted
- Number of CI's base lined
- Hours estimated to create CI's
- Duration hours estimated
- Actual hours expended
- Actual duration hours expended
- Estimated staff FTE to create and maintain CI's
- Actual staff FTE
- Estimated costs to maintain CI's
- Actual costs

5.2 Collection

- All actuals will be collected on a monthly basis or as required by the CCB or service provider.
- All indirect costs will be collected on a monthly basis
- All direct costs will be collected on a monthly basis or as required by management or the customer

5.3 Analysis

- Estimates versus actuals analysis and reporting will be performed monthly
- Direct Costs versus budget analysis and reporting will be performed monthly
- Indirect Costs versus budget analysis and reporting will be performed monthly

5.4 Verifying Implementation

The following steps will be taken to verify the successful implementation of the process:

- Project Reviews
- Peer Quality Reviews
- Inspections
- SQA Assessments
- CCB Reviews

5.5 Training

Training for the project manager, the team members, management, and the customer will be based on skills inventory and the requirements of the project.

Training may consist of the following:

- Introduction to CMM for Software
- Project Management Fundamentals
- Project Management Tools and Techniques
- Estimating
- Project Tracking and Oversight
- Project Life Cycle
- Process Based Management
- Requirements Management
- Configuration Management
- Project Integration

5.6 Tools

The tools that are available for assisting in the tracking of the CI's used by the project are:

- MS Project 2000 (for planning and tracking work performed)
- MS Excel (for recording and reporting metrics)
- MS Word (for recording and reporting metrics)
- Mainframe CI librarians (Endevor, CA-Librarian, Panvalet, etc)
- Client/Server librarians (PVCS, Source Safe, etc)

6 Glossary

Configuration Item or CI refers to the fundamental structural unit of a configuration management system. Examples of CIs include individual requirements documents, software, models, plans, and people

Baseline: A baseline may be a single work product, or set of work products that can be used as a logical basis for comparison.

Banking and Financial Services (BFS): Banking and Financial Services (also known as BFS) is an industry name. This term is commonly used by IT/ITES/BPO companies to refer to the services they offer to companies in these domains. Banking may include core banking, retail, private, corporate, investment, cards and the like. Financial Services may include stock-broking, payment gateways, mutual funds etc. A lot of data processing, application testing and software development activities are outsourced to companies that specialize in this domain.

Business: Generic name of the departments that require and use the Technology applications and systems to process information as required for their respective business purposes.

The Capability Maturity Model (CMM) (a registered service mark of Carnegie Mellon University) is a development model that was created after study of data collected from organizations that contracted with the U.S. Department of Defense, who fund the research. This model became the foundation from which CMU created the Software Engineering Institute (SEI). Like any model, it is an abstraction of an existing system.

Metrics: A unit of measurement (such as source lines of code, effort days, project cost vs. estimates or document pages of design).

Process: A sequence of actions, changes, or functions, for example, the software development process, that achieves an end or result.

Project: All activities required to complete an approved undertaking (i.e., Business Case) that are usually engaged in by a team, to produce a finite, one-time result.

A Project Plan, according to the Project Management Body of Knowledge, is a formal, approved document used to guide both project execution and project control. The primary uses of the project plan are to document planning assumptions and decisions, facilitate communication among stakeholders, and document approved scope, cost, and schedule baselines. A project plan may be summarized or detailed.

Risk: Any event that may negatively affect the delivery of work products or completion of a milestone. From a project perspective, a risk is defined in terms of the likelihood of an event occurring that may affect the delivery and/or expected results of the project.

Risk Acceptance Register: A document completed when a Lifecycle deliverable, that had been previously agreed to be produced, will now not be produced.

Risk Assessment is the determination of quantitative or qualitative value of risk related to a concrete situation and a recognized threat (also called hazard)

Software Configuration Management (SCM) is the task of tracking and controlling changes in the software

Software Metrics is a measure of some property of a piece of software or its specifications.

SQA: Software Quality Assurance: An independent, dedicated, planned and systematic means for assuring management that defined software related standards, practices, procedures, and methods are applied to produce a software product that satisfies customer requirements.

7 Acronyms

SCM Software Configuration Management
CI Configuration Items
CWR Configuration Work Request
SQA Software Quality Assurance

Acknowledgement. We are grateful to Larsen and Toubro Infotech –India, Madurai Kamaraj University-India, Citigroup-USA and Nordea AB-Sweden for this effort to get realized. The Second author thanks Dr. M. Jeyanthinath and Mr.V.Ragavendran for their help during the manuscript preparation.

References

1. Kuttalam, P., Iyakutti, K., Alagarsamy, K.: Software Development Life Cycle Standards for Banking and Financial Services IT Industry. International Journal of Wisdom Based Computing, ISSN: 2231-4857
2. Zhang, J.-G.: Method Study of Software Project Configuration Management
3. PMP Exam Prep, 6th edn. Rita Mulcahy PMP (2009)
4. Pressman, R.S.: Software engineering: a practitioner's approach, 5th edn. McGraw-Hill series in computer science (2001)
5. A Guide to the Project Management - Body of Knowledge (PMBOK Guide), 4th edn. An American National Standard ANSI/PMI 99-001-2008 Also available for download by PMI members at `http://www.pmi.org`
6. Ren, Y., Xing, T., Quan, Q., Zhao, Y.: Software Configuration Management of Version Control Study Based on Baseline
7. Kim, D.-Y., Youn, C.: Traceability Enhancement Technique through the Integration of Software Configuration Management and Individual Working Environment
8. Nuraminah, R., Fauzi, S.S.M., Nasir, M.H.N.M.: The Development of Software Configuration Management Repository
9. Ploskas, N., Berger, M., Zhang, J., Wintterle, G.-J.: A Knowledge Management Framework for Software Configuration Management
10. Thao, C., Munson, E.V., Nguyen, T.N.: Software Configuration Management for Product Derivation in Software Product Families
11. Leeten, B.: Software Configuration Management, Fundament for Evolution of Large Existing Code Bases
12. Nguyen, T.N.: Object-Oriented Software Configuration Management
13. Chou, I.-H., Fan, C.-F.: A Regulatory Software Maintenance Environment Using Agent-Based Software Configuration Management
14. Nguyen, T.N., Munson, E.V., Boyland, J.T., Thao, C.: Configuration Management for Designs of Software Systems
15. Shen, H., Sun, C.: A Complete Textual Merging Algorithm for Software Configuration Management Systems
16. Nguyen, T.N., Munson, E.V., Boyland, J.T., Thao, C.: Architectural Software Configuration Management in Molhado
17. Romanski, G.: Requirements, Configuration Management and Traceability for Safety Critical Software
18. van der Hoek, A., Carzaniga, A., Heimbigner, D., Wolf, A.L.: A Testbed for Configuration Management Policy Programming
19. Estublier, J., Leblang, D., Clemm, G., Conradi, R., van der Hock, A., Tichy, W., Wiborg-Weber, D.: Impact of the Research Community for the Field of Software Configuration Management
20. Awad, E.M.: Systems Analysis and Design, 2nd edn.

VHDL Based Analysis of the Channel Estimator Algorithm and Frequency Offset Estimation for OFDM System

Saxena Minal and Khare Kavita

Manit, Bhopal, India
minal_sudhir@rediffmail.com

Abstract. OFDM has gained considerable attention in the last few years. This Research work is the hardware simulation of Van de Beek algorithm in VHDL. It is designed in hardware descriptive language ,synthesized and simulated to obtain the timing and frequency offsets using XILINX ISE 10.1.The cyclic prefix is exploited at the receiver end to obtain the start of the frame using the argmax function.

Keywords: OFDM, Cyclic prefix, timing offset frequency offset, channel estimation.

1 Introduction

Orthogonal Frequency Division Multiplexing (OFDM) is the basis of some WLAN and WiMax air interfaces. It also has been proposed for use in next generation cellular systems such as Super-3G and HSOPA (High Speed OFDM Packet Access) in the 3GPP standards group. OFDM has a number of advantageous features, such as good tolerance to multi-path fading and inter-symbol interference (ISI). By using a number of sub-carriers, the symbol length can be kept long and a guard period (the cyclic prefix) used to mitigate ISI(2). Data is allocated to a number of sub-carriers so that any nulls in the frequency domain do not knock out a whole allocation. This gives forward error correction (FEC) a greater chance to recover the data in the receiver.

2 OFDM System

The complex data symbols *xk* are modulated on *N* subcarriers by an inverse discrete Fourier transform (IDFT) and the last *L* samples are copied and put as a preamble (cyclic prefix) to form the OFDM symbol $[s_0 \ldots\ldots\ldots s_{N+L-1}]$(4). This data vector is serially transmitted over a discrete-time channel, whose impulse response is shorter than *L* samples. At the receiver, the cyclic prefix is removed and the signal r(k) is demodulated with a discrete Fourier transform (DFT).

The insertion of a cyclic prefix avoids ISI and preserves the orthogonality between the tones, resulting in the simple input-output relation.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 424–429, 2013.

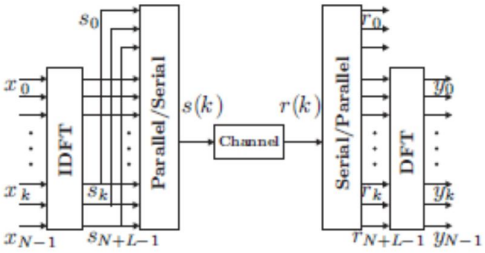


Fig. 1. OFDM System

$$y_k = h_k x_k + n_k \qquad k = 0 \ldots\ldots\ldots\ldots N-1.$$

The uncertainty in the arrival time of the OFDM symbol is modelled as a delay in the channel impulse response, *i.e.*, $\pm(k\ ;\ \mu)$, where μ is the integer-valued unknown arrival time of a symbol.

The uncertainty in carrier frequency due to a difference in the local oscillators in the transmitter and receiver gives rise to a shift in the frequency domain. Such behaviour is modelled as a complex multiplicative distortion of the received data, with a factor,where $e^{j2\pi k\epsilon/N}$ " denotes the difference in the frequency of the transmitter and receiver oscillators as a fraction of the inter-carrier spacing. Hence, the received signal is

$$r(k) = s(k-\theta)e^{j2\pi k\epsilon/N}$$

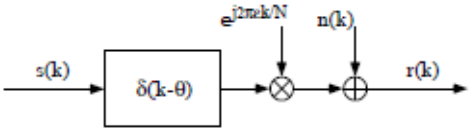


Fig. 2. OFDM System Model

This channel model is shown in Figure 2. This system model is also a special case of the signal model shown in Figure 1. The goal is to estimate θ and ε from the observation signal $r(k)$(8).

3 Cyclic Prefix

Transmitting a cyclic prefix of the data during guard interval transforms the linearly convolutive channel into circularly convolutive channel. Hence, the channel equalizations problem is simplified. Specially, channel equalizations in the frequency domain can be done using one tap filters. This is because cyclic prefixing makes the channel matrix circulant, which is diagnosed by IDFT and DFT operations(11).

The added guard interval and its effect reducing the ISI as shown in the figure The OFDM demodulation must be synchronized with the start and end of the transmitted symbol period. If it is not then ISI will occur (since information will be decoded and combined for two adjacent symbol period) and ICI will also occur because Orthogonality will be lost. The guard interval helps to solve this problem.

With a cyclic extension the symbol period is prolonged but it represents slightly extended frequency spectrum. As long as the correct no of samples are to be taken for decoding, they may be taken from anywhere within the extended symbol. Since a complete period is integrated, Orthogonality is maintained.

Therefore both ISI and ICI are eliminated. Although the cyclic prefix introduces a loss in SNR. This may be considered as a small price to pay to mitigate interference. Note that some bandwidth efficiency is lost with addition of the guard period, but the advantages that we get are many and hence compromising solution is required

4 Algorithm

The WiMAX OFDM preamble is defined differently for the uplink and the downlink. The specification also defines variants for doing initial ranging on the downlink. In this case, the time domain signal has a repeated pattern. The long preamble, used for downlink ranging, consists of a 4x64 pattern symbol, where a 64-sample pattern is repeated 4 times. The uplink uses a short preamble with just a 2x128 pattern symbol(3). These preamble symbols all have the usual cyclic prefix attached.The pattern of the preamble is like this

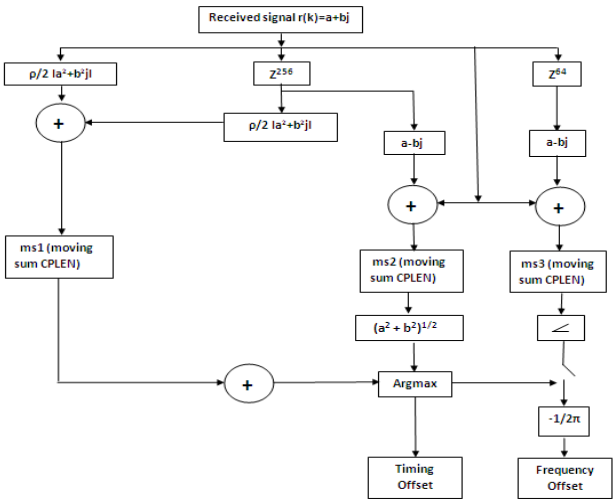


Fig. 3. Block diagram of the algorithm

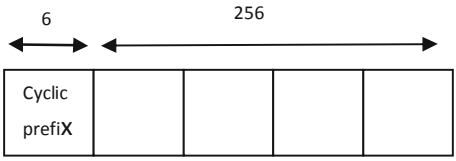


Fig. 4. Preamble

5 Data Flow

The received signal is r(k) =a+bj

This algorithm data flow consists of three parts-

ms1= $\frac{\rho}{2}\sum_{k=m}^{m+L-1} |r(k+N|^2 + |\mathrm{r(k)}\ |^2$

(N =256)

ms2= $\frac{\rho}{2}\sum_{k=m}^{m+L-1} r^*(k)r(k+n)$

ms1 and ms2 are compared to get the maximum value ie; argmax.

Here even the start of the frame is detected when the cyclic prefix is correlated with the data sequences from which it was extracted(3).

ms3= $\frac{\rho}{2}\sum_{k=m}^{m+L-1} r^*(k)r(k+n)$

Here (N=64) and ρ (SNR) is set to 1.

In IEEE802.16-2004 the cyclic prefix can be 64, 32, 16 or 8 samples long(4). The correlation with delay 64 is used to calculate an angle that is used only when the magnitude of the difference between ms1 and ms2 reaches a maximum (argmax). This operation ensures that the frequency offset calculation is done at the best time, i.e., when the correlation over the actual received symbol cyclic prefix is complete(5). Synthesis and simulating the algorithm in VHDL.

The algorithm is simulated in VHDL (VLSI Hardware Description Language).The RTL design structure has been obtained and is as follows

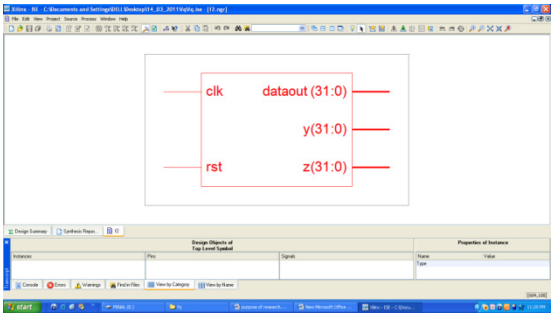


Fig. 5. RTL outputs

6 Estimating the Timing and Frequency Offsets

As per the algorithm the received signal is r(k) which contains both real and imaginary parts. It is divided into three parts .

The first part is multiplied by $\frac{\rho}{2}$ and the mod of the square of both real and imaginary parts.

The second part is delayed by 256 bits and multiplied by $\frac{\rho}{2}$ and the mod of the square of both real and imaginary parts.

Both the first and second part are added to form ms1(moving sum 1).From the second part which is delayed with 256 bits a portion is separated and its complex conjugate is taken and added with r(k) to form ms2.

The third part of r(k) is delayed by 64 bits and one part of this is again taken out to calculate its complex conjugate and added with r(k)to form ms3.

Now ms1 and ms2 are compared to find the start of the frame by detecting the maximum peak of the received signal. Here the signal is correlated with its own subpart so that of the frame is detected. Also shown in the equation that r(k)=r(k)*r(k+n).The max value is the argmax . At this value the frequency offset is calculated. The maximum value and ms3 are compared to get the frequency offset(3).

7 Results

The following outputs for the estimation of frequency offsets are obtained after simulating and synthesis of the algorithm for two type of delays ,128 and 256.The output obtained are as follows.

The values of angle is .0005457 for delay of 128 bits and .00034232 for delay of 256 bits.hence it is inferred that as the delay increases the performance of this algorithm also optimizes to a better valu of frequency offset.

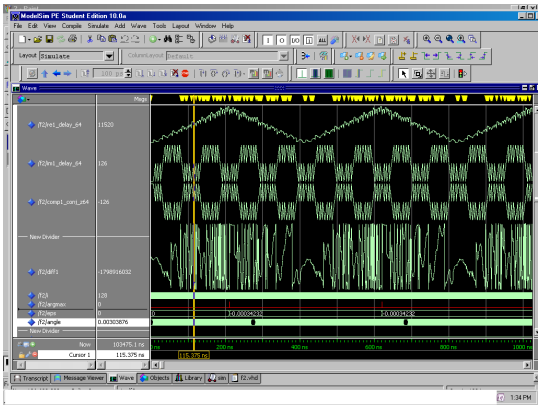


Fig. 6. Output for delay of 256 bits

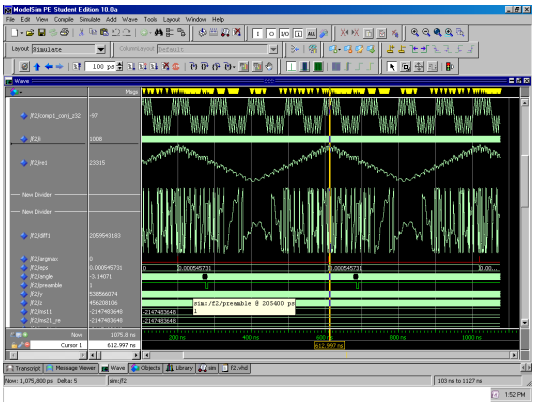


Fig. 7. Output for delay of 128 bits

References

1. Chin, W.-L.: ML Estimation of Timing and Frequency Offsets Using Distinctive Correlation Characteristics of OFDM Signals Over Dispersive Fading Channels. IEEE Transactions on Vehicular Technology 60(2), 444–457 (2011)
2. Morelli, M., Kuo, C.-C.J., Pun, M.-O.: Synchronization techniques for orthogonal frequency division multiple access (OFDMA): A tutorial review. Proc. IEEE 95(7), 1394–1427 (2007)
3. Implementation of an OFDM Wireless Transceiver using IP Cores on an FPGA, Lattice Semiconductor white paper, frequency Offset in OFDM Systems (2005)
4. OFDM Transceiver Reference Design, Lattice Semiconductor OFDM Transceiver design package (2005)
5. Yücek, T., Nezami, M.K.: Joint Channel and Frequency Offset Estimation for OFDM Systems. In: IEEE Military Communications Conference, Monterey, CA, October 31-November 3 (2004)
6. Minn, H., Bhargava, V.K., Letaief, K.B.: A robust timing and frequency synchronization for OFDM systems. IEEE Trans. Wireless Commun. 2(4), 822–839 (2003)
7. Liu, H., Tureli, U.: A high efficiency carrier offset estimator for OFDM communications. IEEE Communications Letters CL-2(4) (1998)
8. Schmidl, T.M., Cox, D.C.: Robust frequency and timing synchronization for OFDM. IEEE Trans. Commun. 45, 1613–1621 (1997)
9. van de Beek, J.-J., Sandell, M., Ola Börjesson, P.: ML Estimation of Time and Fr[3] and Frequency Offset in OFDM Systems. IEEE Transactions on Signal Processing (April 1996)
10. Moose, P.: A Technique for Orthogonal Frequency Division Multiplexing Frequency Offset Correction. IEEE Transactions on Communications 42(10), 2908–2914 (1994)
11. Dalal, U.: Wireless communications. Oxford University Press (2009)

Dot Based Image Analysis Using Local Binary Pattern and Genetic Algorithm

Purshottam Jeevatram Assudani and L.G. Malik

G.H. Raisoni College of Engineering, Nagpur
pjassudani@gmail.com, lgmalik@rediffmail.com

Abstract. Analysis and matching of Dot Patterns is necessary for many of the image analysis and pattern recognition problems. The paper has stated and used local binary pattern for extracting the Dot pattern image features. It has also stated that only the more discriminated features can be retained by discarding the less discriminated features using Genetic Algorithm. The optimized features obtained can be used for matching the dot patterns using Euclidean Distance.

Keywords: Dot pattern, Local Binary Pattern, Genetic Algorithm.

1 Introduction

In processing visual information one often encounters dot patterns instead of gray or color images. A dot pattern is a set of points in 2-D or 3-D space arranged to represent objects in the feature space. For example, objects in an image are often represented by their spots, corners, etc. Pattern recognition procedures like classification and clustering operate on dot patterns. The nighttime sky is a natural dot pattern.

Processing a dot pattern (DP) is useful and important in pattern recognition. Dot patterns encountered in various problems include points in feature space [1], pixels in a digital image [2], physical objects like stars in the galaxy [3] or spatial data [4,5].

Sample Dot Patterns

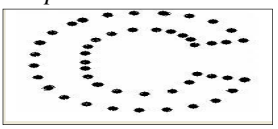


Fig. 1. A Dot Pattern image

Fig. 2. A Dot Pattern image

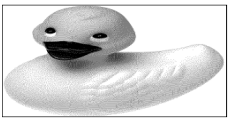


Fig. 3. A Dot-Pattern image

Genetic Algorithm (GA) is one of the optimization algorithms. GAs are used to find an optimal subset of features from a set of features for a particular problem [10].

Dot Pattern Matching (DPM) is important step in many of pattern recognition problems. Some of the application areas concern image registration, object recognition, objects tracking, remote sensing, biomedical imaging etc [6, 7].

A simulated annealing technique was used in dot pattern matching [8]. Zhang et. al. [6] employed genetic algorithm for dot pattern recognition and used the reference

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 430–434, 2013.

triplet points as the chromosome representation that reduced the search space significantly. The effective use of Genetic Algorithm (GA) in dot pattern recognition is also reported in various literatures [6].

The paper is organized as follows. The following section has described about local binary pattern used for image feature extraction which can be optimized by genetic algorithm. Section 3has described in detail about Genetic Algorithm. Section 4has described about Euclidean distance for pattern matching. Section 5 gives partial implementation details. Section 6 concludes the paper.

2 Dot Pattern Feature Extraction

2.1 Local Binary Pattern

LBP [9] is a gray-scale texture operator which characterizes the spatial structure of the local image texture. Features of an image are first extracted from LBP in the form of histogram only. Given a central pixel in the image, a pattern number was computed by comparing its value with those of its neighbourhoods.

$$LBP_{P,R} = \sum_{P=0}^{P-1} s(g_p - g_c)2^P \tag{1}$$

$$s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \tag{2}$$

For example, LBP pattern 00000000 has a U value of 0 and 01000000 of 2. The uniform LBP has limited transition or discontinuities (U<=2)[9].

2.1.1 Rotation Invariant Variance Measures (VAR)

A rotation invariant measure of the local variance can be defined as [9]

$$VAR_{P,R} = {}^1\!/_P \sum_{P=0}^{P-1} (g_P - u)^2 \qquad (3) \qquad where\ u = {}^1\!/_P \sum_{P=0}^{P-1} g_p$$

2.1.2 LBP Variance (LBPV)

$LBP_{P,R}$=$VAR_{P,R}$ is powerful because it exploits the complementary information of local spatial pattern and local contrast [9]. However VAR_P, has continuous values and need to be quantized. The LPBV descriptor proposed here offers a solution to the problems of $LBP_{P,R}$=$VAR_{P,R}$. The LBPV is a joint LBP and contrast distribution method. LBPV gives features of an image in the form of a histogram. We have computed the LBPV histogram by the following equations

$$LBP_{P,R}(k) = \sum_{i=1}^{n} \sum_{j=1}^{m} w(LBP_{P,R}(i,j), k), k \in [0,k] \tag{4}$$

$$W(LBP_{P,R}(i,j), k) = \begin{cases} VAR_{P,R}(i,j), LBP_{P,R}(i,j) = k \\ 0 \qquad\qquad otherwise \end{cases} \tag{5}$$

The ideal feature selection technique removes features that are less discriminative.

3 Dot Pattern Feature Selection (Using Genetic Algorithm)

Genetic Algorithm (GA), first introduced by John Holland, is based on the principles of natural selection. To solve a problem, GA maintains a population of individuals and modifies it by operators such as selection, crossover and mutation, to get an optimal solution.Genetic Algorithm will be used for retaining only the features that have high discriminative power. To describe GA for optimal feature selection, consider the feature vector in Fig. 4.

[30 16 24 161 21 212 …..]

Fig. 4. Sample feature vector

Furthermore, consider the vector shown in Fig. 5 as a candidate real-coded mask.

[0.5 0.1 0.7 0.2 0.8 0……]

Fig. 5. Real-coded feature mask

For optimal feature selection a threshold value of 0.5 is used to create a binary coded candidate feature mask which will be used as condition for masking features. If the random real number generated is less than the threshold (0.5 in this case), then the value corresponding to the real generated number is set to 0 in the candidate feature mask vector or 1 otherwise. Fig. 6 shows the candidate binary coded feature mask..

[1 0 1 0 1 0….]

Fig. 6. Binary coded candidate feature mask

Figure 7 shows the result of the features in Figure 4 when feature masking (Figure 6) is applied to a feature vector of figure 1

[30 0 24 0 21 0.....]

Fig. 7. The Resulting feature vector after feature masking

4 Dot Pattern Matching

The dot pattern images are matched for similarity scores based on their features. We have used Euclidean distance to perform matching between two or more Dot patterns.The Euclidean distance is the distance between two points that one would measure with a ruler, and is given by the Pythagorean formula In Cartesian coordinates, if $p = (p_1, p_2., p_n)$ and $q = (q_1, q_2, . q_n)$then distance d is

$$d(p,q) = d(q,p) = sqrt((q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2) \quad (6)$$

5 Results

Dot pattern image feature extraction using Local Binary Pattern:

Fig. 8. Dot pattern image

Fig. 9. Histogram of extracted features

Extracted Features: 49.04, 56.44, 0, 0, 54.04, 0, 43.03, 0, 0, 0, 0, 55.95, 0, 63.47, 0, 80.33, 0, 0, 0, 0, 0, 0, 46.39, 0, 53.30, 0, 73.59, 0, 12.66, 0, 0, 0, 0, 0, 0, 0, 0, 71.11, 0, 74.80, 0, 9.35, 0, 0, 0, 0,0, 0, 77.88, 0, 13.23, 0, 0, 0, 0, 10.30, 0, 771.40, 512.00

Fig. 10. Dot pattern image

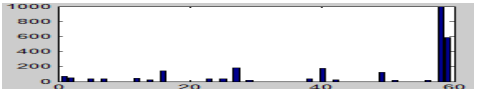


Fig. 11. Histogram of extracted features

Extracted Features: 68.66, 47.10, 0, 0, 36.49, 0, 34.96, 0, 0, 0, 0, 42.62, 0, 26.20, 0, 142.43, 0, 0, 0, 0, 0, 37.73, 0, 39.13, 0, 180.08, 0, 18.92, 0, 0, 0, 0, 0, 0, 0, 0, 33.62, 0, 175.73, 0, 20.56, 0, 0, 0, 0, 0, 0, 124.14, 0, 16.43, 0, 0, 0, 0, 18.56, 0, 998.83, 582.90.

Euclidean distance for first image, with itself=0, with second image=178.7756.
Euclidean distance for second image, with itself=0, with First image=178.7756

6 Conclusion and Future Work

Dot pattern matching is important in image analysis. This paper used LBP for dot pattern feature extraction and Euclidean distance for finding image similarity. The features will be optimized by GA and then the matching will be performed using Euclidean distance. In future character recognition of English alphabets in Dot form will be performed by these methods for which the input sets have been prepared.

References

1. Duda, R.O., Hart, P.E.: Pattern Classification and Scene Analysis. Wiley, New York (1973)
2. Faugeras, O.: Three-Dimensional Computer Vision: A Geometric Viewpoint. The MIT (1993)
3. Ogawa, H.: Labeled pattern matching by Delaunay triangulation and maximal cliques. Pattern Recognition 19, 35–40 (1986)
4. Laurini, R., Thompson, D.: Fundamentals of Spatial Information Systems. The A.P.I.C. Series, vol. 37. Academic Press, London (1992)

5. Taylor, P.J.: Quantitative Methods in Geography: An Introduction to Spatial Analysis. Houghton Mifflin Company, Boston (1977)
6. Sprinzak, J., Werman, M.: Affine point matching. Pattern Recog. Letters 15(4), 337–339 (1994)
7. Zhang, L., Xu, W., Chang, C.: Genetic algorithm for point pattern matching. Pattern Recog. Letters 24, 9–19 (2003)
8. Starink, J.P., Backer, E.: Finding point correspondences using simulated annealing. Pattern Recogn. 28, 231–240 (1995)
9. Ojala, T., Pietikäinen, M., Mäenpää, T.T.: Multiresolution gray-scale and rotation invariant texture classification with local binary pattern. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(7), 971–987 (2002)
10. Adams, J., Woodard, D.L., Dozier, G., Miller, P., Glenn, G., Bryant, K.: GEFE: Genetic & Evolutionary Feature Extraction for Periocular-Based Biometric Recognition. In: Proceedings 2010 ACM Southeast Conference, Oxford, MS, April 15-17 (2010)

Smartphone Terminal Using VNC Protocol

K.P. Chaudhari[1] and Santosh T. Warpe[2]

[1] M.I.T. College of Engineering Aurangabad
Department of Computer Science & Engineering, Aurangabad [M.H], India
Kiran_Chaudhari@rediffmaill.com
[2] P.E.S. College of Engineering Aurangabad
Department of Computer Science & Engineering, Aurangabad [M.H], India
Santoshwarpe@gmail.com

Abstract. This paper describes VNC client-server application for mobile operating system based devices. Virtual network computing (VNC) based architecture for accessing the desktops of remote computers from a cellular phone. A viewer is provided on the cellular phone that enables the user to see and manipulate the desktop of various remote systems such as Windows, Macintosh, and Linux. The system to be accessed must be running a VNC server and it must be attached to a network. VNC protocol was chosen due to multitude of client application on variant platforms and open specification.

Keywords: TCP/IP, Linux, Windows, Macintosh, Symbian/Bada OS, VNC, Smart cell phone.

1 Introduction

A Smartphone viewer/user is provided on the cellular phone that enables him to see and manipulate the desktop/laptop of various remote systems such as Windows, Macintosh and Linux. User can access their system from anywhere in the world anytime with more security and that too with graphical support that is the main objective of our paper. The virtual network computing system is a thin client system.

2 VNC System Overview

VNC system is a simple protocol for remote access to graphical user interfaces. It works at the frame buffer level and therefore applies to all operating systems, windowing systems, and applications—indeed to any device with some form of communications link. The protocol will operate over any reliable transport such as TCP/IP. A VNC system consists of a client, server and communication protocol. Here the mobile is the VNC client and our linux/windows/mac system is the VNC server.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 435–438, 2013.

3 VNC-Based Architecture

To achieve the above-mentioned goals while at the same time considering portability and generality, we propose a VNC based architecture. VNC is an implementation of a remote display system based on a Remote Frame Buffer (RFB) protocol.

3.1 Structure

Figure 1 depicts the VNC architecture. It consists of VNC servers running on one or more remote computers,

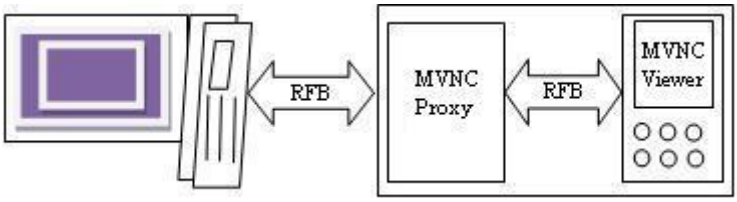


Fig. 1. VNC based architecture

A MobileVNC (MVNC) proxy and a MVNC viewer on a cellular phone. A VNC server sends a remote desktop display as bitmap images in RFB protocol. A MVNC proxy converts (crops, shrinks and resample) the display image and then transfers the converted image to a MVNC viewer in response to a user request that was received from that MVNC viewer. The transfer is performed in our own CRFB (RFB), our simplified RFB protocol. Then, the MVNC viewer displays the transferred images. Key events received by the MVNC viewer are transmitted to a MVNC proxy that coverts them and sends them to the server. When the user first tries to connect to a remote computer, he must specify his user name and password for authentication as well as the host name of the computer that is running a VNC server. If authentication succeeds, the MVNC proxy establishes a *session* with the VNC server and the MVNC viewer starts user services. To suppress network traffic, encoding is changed depending on contexts. Usually, colored display images are transferred from the MVNC proxy to the MVNC viewer. However, while the user is manipulating the remote desktop, such as scrolling and moving the pointing device, the display images are gray-scaled to reduce the number of bytes required to encode the image.

3.2 Maintaining the States of a Session

In order to recover quickly from an unscheduled disconnection, the MVNC proxy maintains its own database containing each session's unique information such as the user name, the password, the target host name, and other internal states.

When it is first connected to a MVNC viewer, the MVNC proxy searches its own database using the user name and the target host name pair as a key. It determines

whether or not there has been any previously established session. If such a session did exist, the proxy restores the stored states. A session's state in the proxy is discarded when the user explicitly terminates it on the viewer.

4 Simulation Result

We have taken a thin client usage session offline by storing the sequence of graphical updates. This session consisted of, starting from the desktop background, opening Open Office writer, typing a text, closing down the office program, opening an internet browser, performing a Google search, followed by visiting the homepage of a local newspaper containing multimedia content (in the form of banners) which was scrolled down to the bottom. Finally the browser was closed, ending the trace with the desktop background again.

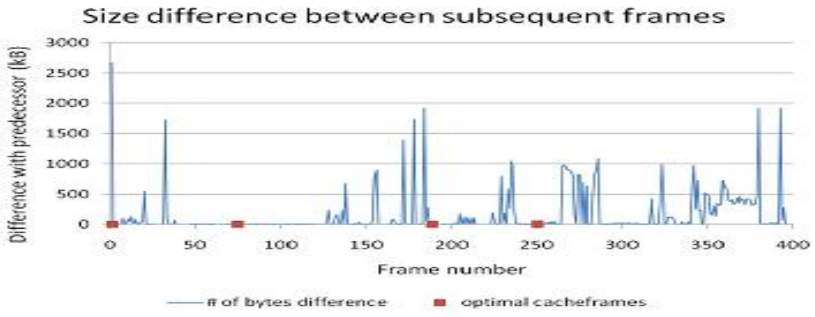


Fig. 2. Number of bytes differing between subsequent frames in a thin client computing session

We have computed the byte-per-byte differences between subsequent full screen frames (in uncompressed format). We have taken the number of different bytes as a measure for the resemblance between frames. Figure 2 presents these difference frames, which show a spiked path. The peaks are interpreted as big differences between two subsequent frames. Generally speaking these peaks are followed by a tail of frames that do not differ much from their predecessor. Through a matrix of mutual distances, we were able to identify the optimal combination of a predefined number of cache frames, indicated by squares in figure 2. Optimal cache frames are those that combined; result in the smallest distance to the complete sequence. After visual inspection we found that the optimal cache frames represented the applications that were executing, i.e. the desktop background, the office program, the Google startup page and the homepage of the local newspaper.

4.1 Compression Factors and Difference Frames

We cannot simply state that the raw byte size of an image is a straight guideline for the compressed byte size. A lot depends on the content of the image. This is why it is difficult to predict the bandwidth gain of encoding difference frames from the cache in function of the bandwidth used by coding the frames themselves.

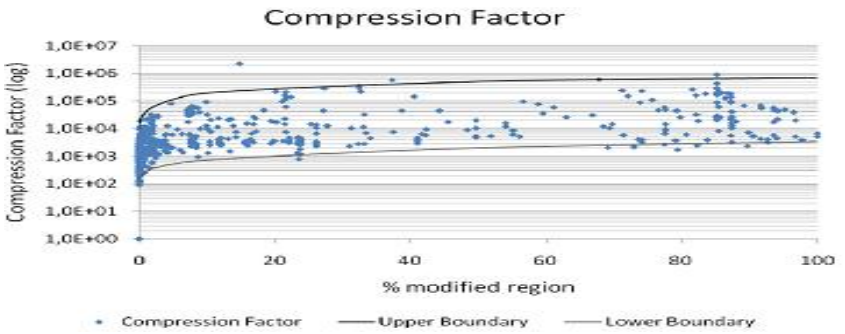


Fig. 3. Compression factor for varying image sizes

5 Conclusion

In this paper we are going to develop VNC client/server application for Smartphone. This application is the first step to develop complete automatic test environment dedicated to test mobile application with graphical interface. The next phase will be to develop scriptable VNC client able to process file with test sets, generates through VNC protocol events and capture screenshots. We have proposed a system to remotely access a computer desktop using a Smartphone, despite the physical & bandwidth limitations of cellular phones. The system has VNC based architecture for accessing a remote desktop from Smartphone. We have to implement a prototype for this system using J2ME and check the operation on Java enabled cellular Smartphone.

Acknowledgments. The author is extremely thankful for the respected guide Prof.K.P.Chaudhari for their encouragement.

References

1. Richardson, T., Stafford-Fraser, Q., Wood, K.R., Hopper, A.: Virtual Network Computing. IEEE Internet Computing
2. VNC Mobile Solution for Automotive- RealVNC
3. Richardson, T.: The RFB Protocol. RealVNC Ltd.
4. Richardson, T., et al.: Teleporting in an X Window System Environment. IEEE
5. Scheifler, R.W.: X Window System Protocol X Consortium Standard X Version 11
6. Microsoft Corporation. Windows 2000 Server Remote Desktop Protocol (RDP)

Improved Performance Analysis of AODV Routing Protocol with Priority and Power Efficiency in Mobile Adhoc Homogeneous and Ad Hoc WiMAX Heterogeneous Networks

Sanjay Sharma and Pushpinder Singh Patheja

MANIT, Bhopal
ssharma66@rediffmail.com,
pspatheja@gmail.com

Abstract. In mobile large heterogeneous network we have to service different type of users working on different platform and different type of hardware at different geographic locations. Network application distributed on a geographical area, may be used for simulations, processing and analyzing data, and visualizing results. We can provide priorities to these applications. Our proposed protocol, called AODV-PP, improved AODV in Priority models and in Power consumption. We also measure performance indicators for some metrics, such as throughput, energy, end-to-end delay, network lifetime, time required for the first node to die and number of times nodes tried to save battery.

1 Introduction

A major drawback of all existing ad hoc routing protocols is that they do not have provisions for conveying the energy and priority and/or quality of a path during route setup. Hence they cannot balance the load on different routes therefore we have taken Priority and Power into consideration [1].

In this paper, we proposed a new variant of AODV routing protocol, called AODV-PP, which performed some improvement based on priority models and energy consumption [5, 12].

2 Literature Review

2.1 Optimized AODV

In [1, 4, 9] provided a very comprehensive and in depth survey about the QoS of routing in MANETs. Their papers offer a recent survey of major contributions to the MANETs routing protocols published in the period of 1997-2010. It discussed about the QoS routing metrics, protocols and the factors that affect the performance of protocols, created the classification of protocols. Next, Thathacar et al. [16] presented

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 439–445, 2013.

variety of views on the learning automata as an alternative optimization algorithm. According to RFC3561 [13] AODV protocol still has many weaknesses [18, 19].

2.2 AODV with Energy Model

Very small amount of routing proposals have focused on the energy constraints of wireless nodes. Such nodes dying adversely affect the operational life time of ad hoc network. First, many applications where the dying nodes are the communication end points will fail. Second, even when the dying nodes are not the communication end points, network connectivity will become sparser and network partition becomes more likely. The goal of our protocol is routing or re-routing around nodes low on battery power as far as possible. This will prolong the network lifetime [10, 17].

Gupta N and Samir R. Das studied "Energy-Aware On-Demand Routing for Mobile Ad Hoc Networks" [6]. They took two-step approach to design the adaptive energy-aware protocol. First, the nodes were classified according to their remaining battery energy. Second, a new cost function was used as routing metric taking into consideration both the hop-wise distance and the battery levels of the nodes. They classified the nodes into three zones according to their energy zones (battery power).

In [15], 2008, introduced in their paper, a mechanism involving the integration of load balancing approach and transmission power control approach to maximize the life-span of MANETs. The mechanism is applied on Ad hoc On-demand Vector (AODV) protocol to make it as energy aware AODV (EA_AODV).

In these algorithms only one parameter of battery is taken care of by any intermediate node. This is taken as challenge and drawbacks are removed in our proposed AODV-PP algorithm which considers the minimum battery of all intermediates node lies between source and destination along-with the remaining battery power of the path.

2.3 Priority

Dola S S R, Bandyopadhyay S, Ueda T and Tanaks S in IEEE Communication Society, 2004, [14] have proposed a scheme for supporting priority-based QoS in mobile ad hoc networks by classifying the traffic flows in the network into different priority classes, and giving different treatment to the flow-rates belonging to different classes. They have adopted a control-theoretic approach to adaptively control the low-priority flows so as to maintain the high priority flow-rates at their desired level.

Alia A, Jamil Y. K and Mahata K [2] in 2011 supported that priority based packet transmission techniques commonly used in communication networks to support multimedia services. Their paper introduces a novel two stage traffic prediction and type of service based priority technique for an infrastructure based Wireless Local Area Network. The developed algorithm alters the priority of transmission queues and services in a radio access network based on the predicted traffic volume and the conventional type of service priority technique.

3 Description of Proposed Protocol

MANETs are power constrained since nodes operate with limited battery energy. To maximize the lifetime of these networks the remaining energy in each mobile node must be very well managed. Assuming that all nodes start with a finite amount of battery capacity, AODV-PP presents a new on demand routing protocol for mobile Ad Hoc networks with an energy management and priority that balances the route inside the network in order to increase the battery lifetime of the nodes and hence the overall useful life of the Ad Hoc network.

The aim is to modify AODV that has capability to determine battery of intermediate node along with the remaining battery on path and to also include the priority of the application. These parameters are adjusted by modifying the RREQ, RREP and RERR packets. In proposed AODV-PP checks the battery of all the intermediate nodes. Here if a new path is available where minimum battery of all intermediate nodes is specifically more than the minimum battery of any one intermediate node of present path. Then the new path is selected and the present path is dropped, hereby increasing the network lifetime as a whole.

Also, if the minimum battery of available and new path is almost same then we check the "Average Remaining Battery on Path" of both the paths. If "average remaining battery on path" of new path is specifically more than the "average remaining battery on path" of present path. Then this new path is selected. This also increases the network lifetime of the whole network.

Once path is established all the intermediate nodes check the priority of the packet sent. If the priority is low or moderate then after depleting specified percentage of battery the intermediate node informs the source that it is leaving the path. Now the source will reselect the path. Therefore, network life time is increased by distributing the route among nodes homogeneously to make sure that some nodes are depleted of their battery more than the others does not arise.

The algorithm route discovery process in AODV-PP is as follows:

1) Find the energy level of the route.
2) Calculate the average route energy of route and the battery power of lowest charge node.
3) Select and reselect the path accordingly.
4) Select the high average energy route for data transmission.
5) Check the priority of applications and take decision of leaving existing path based on battery depletion.

4 Simulation Environment and Parameters

We have used EXata/Cyber 1.1 for our simulations for comparison of AODV and modified AODV protocol (AODV-PP) with Priority and Power Efficient parameters. The study used 50, 100, 150, 200 and 250 nodes in a grid of dimensions 1500x1500m. The study runs each simulation for 2500 sec. The system used a Constant Bit Rate

(CBR) source as the data source for each node. Each source node transmitted packets with a packet size of 512 bytes. The position of the nodes is arranged according to the random waypoint model. The antenna used is omni directional. Linear battery model has been implemented according to Gaussian distribution (from website random.org) with mean charge value of 5.0 and standard deviation of 1.0 and depends upon the number of nodes in the scenario.

Performance Metrics: We use the following performance metric to evaluate the effect of each scheduling algorithm:

- **Average throughput:** It is defined as the ratio of total packets received to the simulation time.
- **Energy consumption** is the energy used for various node density and speed.
- **First node dies:** This battery parameter calculates the time required for the first node to die in the network. It is measured in seconds.
- **Number of times node tries to save battery:** In AODV once route is established the node does not try to save battery. In AODV-PP node tries to save battery by changing the established route.

5 Result and Analysis

In this research paper, we evaluate the performance of AODV-PP routing protocol in WiMAX environment. The first case study is large heterogeneous environment with 45 percent Ad hoc nodes 45 percent infrastructure nodes and 10 percent Wimax nodes. We did the comparative study of AODV-PP along with the standard AODV routing protocol. In the second case study we had all ad hoc nodes.

Based on the simulations we can conclude that using power awareness to find routes is very beneficial because the difference in battery consumption between various nodes is reduced. This typically means longer network life and longer time to node failure. The study also gives priority to the data packets for better and fast communication. The following performance analysis graph shows that AODV-PP generates good results.

AODV-PP performs better in case of first node dies as compare to AODV. The result of this is more network life time of MANET.

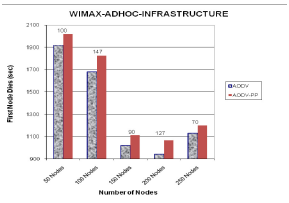


Fig. 1. Analysis of Time required for the node to die Vs number of nodes

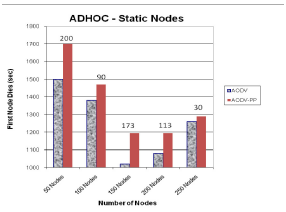


Fig. 2. Analysis of Time required for the node to die Vs number of nodes

Time required for the first node to die is better in both the case studies. These values are represented in Figure 1 and Figure 2. It is evident that the lifetime of the network is improved using AODV-PP.

The throughput is increased in most of the cases. As overall network life time is increased the throughput is also increased in most of the circumstances.

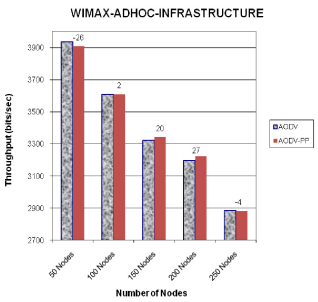


Fig. 3. Anlysis of throughput Vs number of nodes

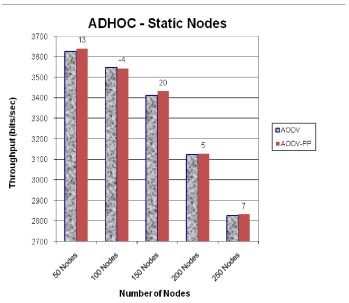


Fig. 4. Anlysis of throughput Vs number of nodes

In Figure 3 and Figure 4 throughput is represented. Even if throughput is decreased, it is observable that it is not compromised in significant way.

As AODV-PP is concentrating on battery constraints of node, it improves the ratio of number of node dies per second. As the frequency of path establishment is reduced the number of node dies per second is less.

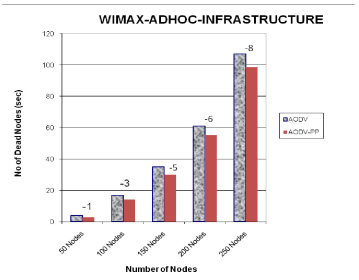


Fig. 5. Anlysis of number of nodes dead in the situation of different number of nodes

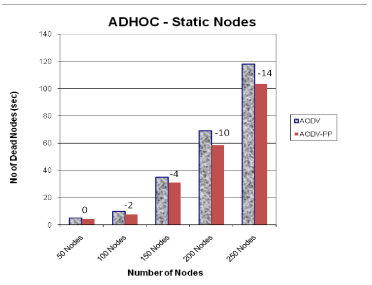


Fig. 6. Anlysis of number of nodes dead in the situation of different number of nodes

In Figure 5 and Figure 6 we can visualize that the number of dead nodes decreases. These also strengthen the fact that the network lifetime of the scenario increases.

In AODV node does not try to save battery once first route is established. This parameter is not present is AODV but is the strength of AODV-PP. It is evident that node is changing route depending upon the battery calculations and priority and hence improving the overall network lifetime. As nodes do not save battery in AODV because in AODV once route is established, the route is not changed. In AODV-PP node inform source after depleting specified percentage of battery and thus overall network lifetime is increased.

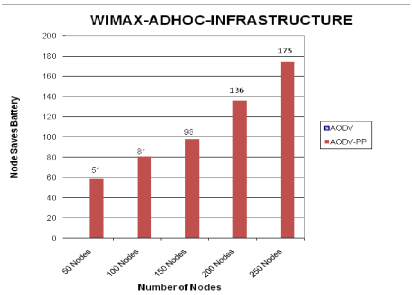


Fig. 7. Anlysis of number of nodes tries to save battery in the situation of different number of nodes

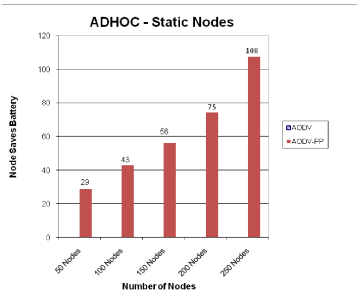


Fig. 8. Anlysis of number of nodes tries to save battery in the situation of different number of nodes

Figure 7 and Figure 8 represents the number of times intermediate node tries to save battery. This parameter is introduced in AODV-PP therefore this parameter can not be compared rather we can display the value of number of node tries to save battery in AODV-PP.

6 Conclusion

Successful delivery of message is very important in a MANET as a lot of route discovery effort is wasted if a new route discovery process has to be reinitiated. Our simulation results show that AODV-PP protocol has better network lifetime with minor change in end to end delay.

In this study the on-demand routing protocols AODV & AODV-PP are analyzed and their performances have been evaluated. This paper can be enhanced by analyzing other MANET routing protocols with different network sources. The results obtained from implementing these techniques are favorable and encouraging.

Summarizing, simulation results show clearly that significant improvement is observed regarding the battery life of nodes and hence the overall useful life of ad-hoc networks. In detail, in AODV-PP the energy consumption is decreased without affecting the other performances of the network.

The results of this research study are also deemed suitable for many other distributed network centric applications. It is further concluded that the present work results will motivate the design of routing solutions for large-scale heterogeneous applications. Protocol designed for this research work will certainly help to improve communication in the large dynamic networks and distributed computing environments.

References

1. Syarif, A., Sari, R.F.: Performance Analysis of AODV-UI Routing Protocol With Energy Consumption Improvement Under Mobility Models in Hybrid Ad hoc Network. International Journal on Computer Science and Engineering (IJCSE) 3(7), 2904–2918 (2011) ISSN: 0975-3397

2. Asheralieva, A., Khan, J.Y., Mahata, K.: Traffic Prediction Based Packet Transmission Priority Technique in an Infrastructure Wireless Network. In: Published in Wireless Communications and Networking Conference (WCNC), pp. 404–409. IEEE (2011)
3. Di Caro, G., Ducatelle, F., Gambardella, L.M.: Special Issue on Self-Organization in Mobile Networking. AnthocNet: An Adaptive Natured-Inspired Algorithm for Routing in Mobile Ad hoc Networks. European Transactions and Communications 16 (2005)
4. Chen, L., Heinzelman, W.B.: A Survey of Routing Protocols that Support QoS in Mobile Ad hoc Networks. IEEE Communication Society Networks 21 (2007)
5. Ding, L., Wan, L.: Improvement Suggestions to the AODV Routing Protocol. In: International Conference on Wireless Networks and Information Systems, WNIS, December 28-29, p. 370 (2009)
6. Nishant, G., Das, S.R.: Energy-Aware On-Demand Routing for Mobile Ad Hoc Networks. In: Proceedings of the 4th International Workshop on Distributed Computing, Mobile and Wireless Computing, pp. 164–173
7. Feeney, L.M.: An Energy Consumption Model for Performance Analysis of Routing Protocols for Mobile Ad hoc Networks. In: Mobile Networks and Applications, pp. 239–249 (2001)
8. Hamidian, A.: Performance of Internet Access Solutions in Mobile Ad hoc Networks, Master Thesis, Lund University (2001)
9. Hanzo, L., Tafazolli, R.: A Survey of QoS Routing Solution for Mobile Ad hoc Networks. IEEE Communications Surveys & Tutorials 9(2) (2007)
10. Jing, X., Lee, M.J.: Energy-Aware Algortihm for AODV in Ad hoc Networks. In: Proceeding of International Conference on Mobile Computing and Ubiquitous, ICMU, Japan, January 8-9 (2004)
11. Kim, C., Talipov, E., Ahn, B.: A Reverse AODV Routing Protocol in Ad Hoc Mobile Networks. In: Zhou, X., Sokolsky, O., Yan, L., Jung, E.-S., Shao, Z., Mu, Y., Lee, D.C., Kim, D.Y., Jeong, Y.-S., Xu, C.-Z. (eds.) EUC Workshops 2006. LNCS, vol. 4097, pp. 522–531. Springer, Heidelberg (2006)
12. Pirzada, A.A., Portmann, M.: High Performance AODV Routing Protocol for Hybrid Wireless Mesh Networks. In: 4th Annual International Conference on Mobile and Ubiquitous System: Networking & Services, Mobi Quitous, Philadelphia, USA, August 6-10 (2007)
13. Request For Comment (RFC) AODV, `http://www.ietf.org/rfc/rfc3561.txt`
14. Dola, S., Roy, S., Bandyopadhyay, S., Ueda, T., Tanaka, S.: IEEE Communication Society, 4172–4176 (2004)
15. Tamilarasi, M., Palanivelu, T.G.: Integrated Energy-Aware Mechanism for MANETs using On-demand Routing. International Journal of Computer and Information Engineering, 212–216 (2008)
16. Thathacar, M.A.L., Sastry, P.S.: Varietes of Learning Autmata: An Overview. IEEE Transactions on Systems, Man, and Cybernetics Part: B, Cybernetics 32(6) (December 2002)
17. Tolba, F.D., Magoni, D., Lorenz, P.: Energy Saving and Connectivity Tradeoff bys Adaptative Transmission Range in 802.11g MANETs. In: International Conference on Wireless and Mobile Communication, ICWMC, IEEE Conference, July 29-31, p. 45 (2006)
18. Rizvi, M.A., Sharma, S.: Enhancing Routing Efficiency Using Agents, a research thesis, MANIT (2011)
19. Kapoor, R.K., Rizvi, M.A., Sharma, S., Malik, M.M.: Exploring Multipath Routing Protocols in Mobile Ad hok Network. Journal of Computer and Mathematical Sciences 2(5), 725–740 (2011)

An Area Efficient and Low Power Multiplier Using Modified Carry Save Adder for Parallel Multipliers

S. Murugeswari[1] and S. Kaja Mohideen[2]

[1] Professor, ECE Department, Sri Ramanujar Engineering College,
Chennai-48, India
eswarimurthy2k7@yahoo.co.in
[2] Professor & Head of the Department, ECE,
B.S. Abdur Rahman University, Chennai-48, India
hodece@bsauniv.ac.in

Abstract. There are different factors that one would like to optimize when designing a VLSI circuit. Often they cannot be optimized simultaneously, only improve one factor at the expense of one or more others. The design of an efficient integrated circuit in terms of Power, Area and Speed has become a very challenging problem. Low power design of VLSI circuits has been identified as a critical technological need in recent years due to the high demand for portable consumer electronic products with more backup and less weight. Adders and multipliers are the most important arithmetic units in a general purpose processors and the major source of power dissipation. The objective of a good multiplier is to provide a physically compact, high speed and low power consuming chip. Fast multipliers are the key components of many high performance systems such as FIR filters, Microprocessors, Digital signal processors etc. A system performance is generally determined by the performance of the multiplier because the multiplier is generally the slowest element in the system. Furthermore, it is generally the most area consuming. Hence, optimizing the speed and area of the multiplier is the major design issue. However, area and speed are usually conflicting constraints so that improving speed results in larger areas. In this paper we are using the Modified Carry Save adder which is used to speed up the final addition in many parallel multipliers.

Keywords: Fast Multiplier, High Speed Adder, Power-Delay Product, Field Programmable Gate Array, Parallel Multiplier.

1 Introduction

A binary multiplier is an electronic circuit used in digital electronics, such as computer, to multiply two binary numbers. It is built using binary adders. Most techniques involve computing a set of partial products, and then summing the partial products together. For high speed multiplications, a huge number of adders or compressors are to be used to perform the partial product addition.

Fast multipliers are essential parts of digital signal processing systems. The speed of multiply operation is of great importance in digital signal processing as well as in

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 446–450, 2013.

the general purpose processors today. Multiplication can be considered as a series of repeated additions.

It is possible to decompose multipliers into two parts. The first part is dedicated to the generation of partial products, and the second one collects and adds them. The major speed limitation in any adder is in the production of carries. Basically, carry save adder is used to compute sum of three or more n-bit binary numbers.

Carry save adder is same as a full adder. When adding together two numbers, using a half adder followed by a ripple carry adder is faster than using two ripple carry adders. This is because a ripple carry adder cannot compute a sum bit without waiting for the previous carry bit to be produced, and thus has a delay equal to that of n full adders. A carry-save adder, however, produces all of its output values in parallel, thus the total computation time for a carry-save adder is less than ripple carry adders .

2 Modified Booth Multiplier

The block diagram of proposed multiplier is shown in Fig. 1.A multiplier has two stages. In the first stage, the partial products are generated by the booth encoder and the Partial Product Generator (PPG), and are summed by compressors [4]. We have used the Modified Booth Encoding (MBE) scheme proposed in [1]. It is known as the most efficient Booth encoding and decoding scheme.

In the second stage, the two intermediate products are added to form the final product through the Modified Carry Save Adder.

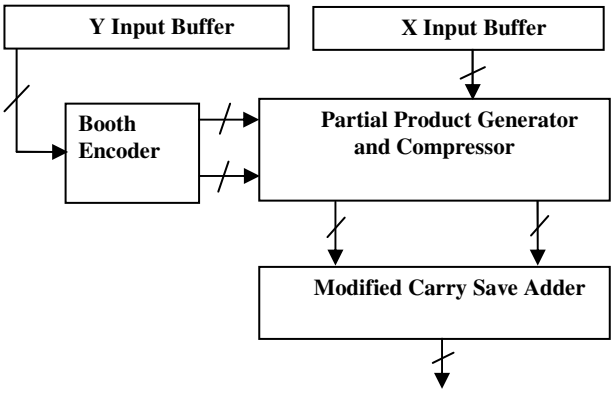


Fig. 1. Block Diagram of the Proposed Multiplier

X and Y are the input buffers. Y is the Multiplier which is recoded by the booth encoder and X is the multiplicand. PPG module and compressor form the major part of the multiplier. Modified Carry Save Adder is the final adder used to merge the sum and carry vector from the compressor module.

To multiply X by Y using the modified Booth algorithm starts from grouping Y by three bits and encoding into one of {-2, -1, 0, 1, 2}[2]. Modified Booth Encoder Scheme[6] is shown in Table 1.

Table 1. Truth Table of Modified Booth Encoder Scheme

Y_{i+1}	Y_i	Y_{i-1}	Value	X1_b	X2_b	Neg	Z
0	0	0	0	1	0	0	1
0	0	1	1	0	1	0	1
0	1	0	1	0	1	0	0
0	1	1	2	1	0	0	0
1	0	0	-2	1	0	1	0
1	0	1	-1	0	1	1	0
1	1	0	-1	0	1	1	1
1	1	1	0	1	0	1	1

3 Modified Carry Save Adder

The 16-bit conventional CSA [3] has 17-half adders (H) and 15-full adders(F). Since the Ripple Carry Adder (RCA) is used in the final stage, this structure yields large carry propagation delay.

To reduce the delay, the final stage of CSA is divided into 5 groups as shown in Fig 2.The first group includes n $1+\log_2$ n-bit value and other groups include $\log_2$n-bit value, where n is the bit size of the adder. The divided groups are listed as follows:

1. {c4, s [4:0]}, output s[4:0] is directly assigned as the final output.
2. {c7, x [7:5]} manipulates the partial result by considering c4 is 0.
3. {c10, x [10:8]} manipulates the partial result by considering c7 is 0.
4. {c13, x [13:11]} manipulates the partial result by considering c10 is 0
5. {X [17:14]} manipulates the partial result by considering c13 is0.

Depending on c4 of the first group, the second group mux gives the final result without the carry propagation delay from c4 to c7; depending on c7 of the second group final result, the third group mux gives the final result without the carry propagation delay from c7 to c10; depending on c10 of the third group final result, the fourth group mux gives the final result without the carry propagation delay from c10 to c13 and depending on c13 of the fourth group final result, the fifth group mux gives the final result without the carry propagation delay from c13 to s17.

The main advantage of this logic is that each group computes the partial results in parallel and the multiplexers are ready to give the final result “immediately” with the minimum delay of the mux. When the Cin of each group arrives, the final result will be determined “immediately”. Thus the maximum delay is reduced in the carry propagation path. This same logic has been used for 32 and 64-bit adder structures to achieve higher speeds. The area indicates the total cell area of the design and the total

power is sum of leakage power, internal power, net power and dynamic power. The proposed result shows that the Modified Carry Save Adder (MCSA) [3] has reduced area, delay and consumes lesser power than CSA.

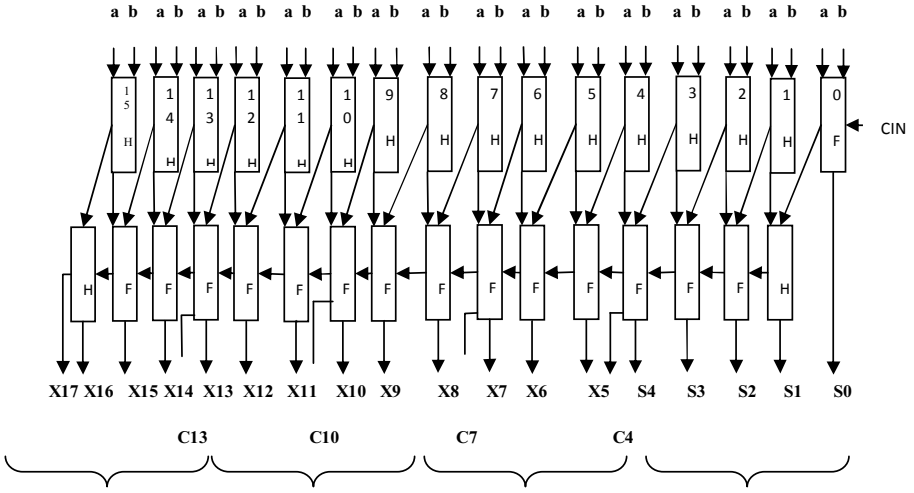


Fig. 2. Modified Carry Save Adder (MCSA)

4 Results and Simulation

The comparison of conventional Carry Save Adder with Modified Carry Save Adder is shown in the Table 2. We have designed all adders using Hardware Description Language (HDL) for 8-bit unsigned data. To get power, delay and area report, we use XILINX ISE 10.1 as synthesis tool and Model-Sim 6.3c for simulation. FPGA-Spartan III is used for implementation. The modified Carry Save adder gives better result than conventional adders in terms of Speed, area and power consumption as shown in the Table-2. The Simulation Result of the Proposed Booth Multiplier is shown in Fig 3. The performance comparison of proposed Booth multiplier is shown in Fig. 4.

Table 2. Comparison of Delay, Power and Area of Carry save Adder with Modified Carry Save Adder

Properties	Carry save adder	Modified carry save adder
LUT	134	106
SLICES	72	56
DELAY(ns)	34.643	19.969
POWER(mw)	65	53

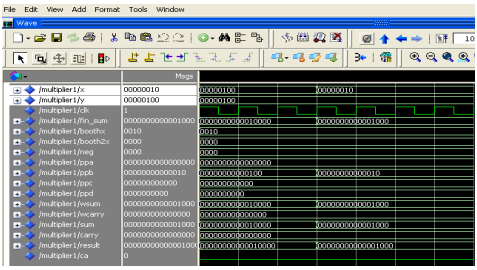


Fig. 3. Booth Multiplier with Modified Carry Save Adder

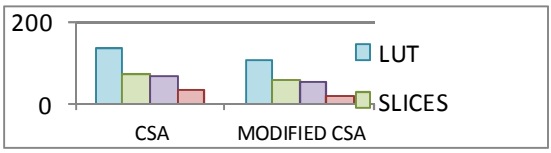


Fig. 4. Performance comparison of Proposed Booth multiplier

5 Conclusions

This paper presents a detailed study of Modified Booth Multiplier with specialized Modified Carry Save Adder, focusing on Xilinx device. From the obtained results, we can conclude that our proposed architecture gives better area utilization and less power consumption than the conventional architecture.

References

1. Rashidi, B., Rashidi, B., Pourormazd, M.: Design and implementation of low power digital FIR filter based on low power multipliers and adders on Xilinx FPGA. In: 2011 3rd International Conference on Electronics Computer Technology, ICECT, April 8-10 (2011)
2. Kim, S., Cho, K.: Design of High-speed Modified Booth Multipliers Operating at GHz Ranges. World Academy of Science, Engineering and Technology 61 (2010)
3. Ramkumar, B., Kittur, H.M., Mahesh Kannan, P.: ASIC Implementation of Modified Faster Carry Save Adder. European Journal of Scientific Research 42(1), 53–58 (2010) ISSN 1450-216X
4. Bagherizadeha, M., Eshghi, M.: A Low Power and High Speed Carbon Nanotube 5-to-3 Compressor. In: 2011 Faible Tension Faible Consommation, FTFC (2011)
5. Wang, J.-P., Kuang, S.-R., Liang, S.-C.: High-Accuracy Fixed-Width Modified Booth Multipliers for Lossy applications. IEEE Transactions on Very Large Scale Integration (VLSI) Systems 19(1) (January 2011)
6. Macpherson, N., Stewart, R.W.: Area Efficient FIR filters for high speed FPGA Implementation. IEEE Proc.-Vis. Image Signal Process. 153(6) (December 2006)

Biosignal Acquisition System for Stress Monitoring

Joydeep Sengupta, Nupur Baviskar, and Surbhi Shukla

Department of Electronics Engineering,
Visvesvaraya National Institute of Technology, Nagpur
joydeep44@rediffmail.com, {nupurbaviskar,author2}@gmail.com

Abstract. In this paper, it focus on the acquisition of biosignals like electroencephalogram and heart rate and study their relation with stress levels of a person. A common day-to-day life situation of car driving and a few stress causing elements to see this relation have been picked up. We have used a car racing simulator for this purpose, with varying traffic levels, different tracks (straight and with turns) as well as varying number of opponents to vary the stress developed on the driver. These signals were procured using our hardware, then interfaced to the LabView using Atmega 16 development board. 2 Virtual Instruments were built to help in analyzing the signals acquired-(i) used to filter the ocular artifacts from EEG using RLS-wavelet method. (ii) was used to separate out the various components of the EEG signals. We further try to correlate frequency of eye-blink rate, components of EEG and their presence with the stress level developed. We found that the stress levels of a person measured in terms of eyeblink rate, heart rate showed an increase in stressful situations (increase in traffic and number of turns, crashes, introducing obstacles, etc). The various components of EEG were also separately studied to find if a person is relaxed. Finally, a statistical model was produced to combine these results and make a decision on whether a person is stressed. This line of research may be very useful to society.

Keywords: Biofeedback, EEG, Eye-blink rate, Heart rate, Stress.

1 Motivation and Focus

Stress is a topic which seems to surface in more and more conversations. It can be considered a topic which is woven deeply into our social and economic fabric. Fortunately, in recent years, we have improved our understanding of how the brain and body work in response to stress. Hans Selye, M.D., a pioneer in stress research, defines stress as *"the non-specific response of the body to any demand made on it (when external demands exceed resources"*. Years ago, many health professionals and researchers felt that the mind and body behaved completely independently of one another. Today, it is clear that the mind and body are interrelated. Stress, therefore can be considered a complex physical and emotional reaction. This reaction is largely based on perception. In other words, if one perceives a situation as stressful then they will respond accordingly. What is stressful to one person may not be to another. But one must realize that the brain is amazingly adaptable, and capable of learning. It can also learn to improve its own performance, if only it is given cues about what to

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 451–458, 2013.

change. By making information available to the brain about how it is functioning, and asking it to make adjustments, it can do so. When the mature brain is doing a good job of regulating itself, and the person is alert and attentive, the brain waves (EEG) show a particular pattern. This is the concept behind Biofeedback. Biofeedback therapies, in general, have been around for about 40 years and have primarily been used to assist in relaxation and for conditions related to stress. EEG biofeedback is an effective way to train the brain to be more efficient in its patterns and control of physiological functions. The demand for biofeedback treatment is undergoing rapid growth, and there is also an increasingly greater scope of issues it is being used to address.

Our focus is on detecting eye blinks of test subjects from EEG and correlate eye blink frequency with the experienced level of stress. For our experiment we chose EEG as the technique to capture brain activity. Its detection technique is based on the electrochemical brain activity. It has excellent temporal resolution in contrast to blood flow techniques such as fMRI and PET. The high temporal resolution together with the low cost make EEG a great solution for our research. Along with EEG we also monitor the heart rate of a person and try to correlate the increased heart rate with the amount of stress experienced.

In our experiment we acquire brain activity using EEG equipment, convert and remove artifacts using software, extract and select features characteristic for eye blinks and finally classify the signal, using the selected features as eye blinks and heart rate of the subject. This line of research may be very useful to society. Human activities like driving vehicles could be made safer when being able to sense that the driver has irregular or fast eye blinks, indicating drowsiness or stress. There are numerous other applications where eye blink detection may be used to enhance stress monitoring.

2 Introduction to Stress Monitoring Hardware for Biosignal Acquisition Software for Biosignal Processing Experimentation

The participant was asked to drive on a track, either a 10 km long straight road, or three laps on a curved race track. The variable which is expected to influence the participant's performance and the resulting EEG recordings and heart rate is the type of track, straight or curved.

Table 1. Details of Tracks

Number	Track
1	Straight road Max speed: 98 mph
2	Race track Max speed: 240 mph

The test administrator would give simple commands to change lanes or to alternate between left and right lanes in order to simulate a real life driving situation and the EEG recordings were marked when the commands were given. The heart rate of the subject was also observed during this time.

3 Observations

The test administrator observed the EEG and the heart rate signals of the subject. He made manual recordings with the help of markers of the locations in the EEG signal where an ocular artifact occurred. The following graphs were observed under different experimental set-ups.

3.1 EEG of the Subject under No Stress While Driving on a Straight Road

Fig. 1. Car on a straight road and eeg of the person driving the car

3.2 EEG of a Subject under Stress When Car Crashes

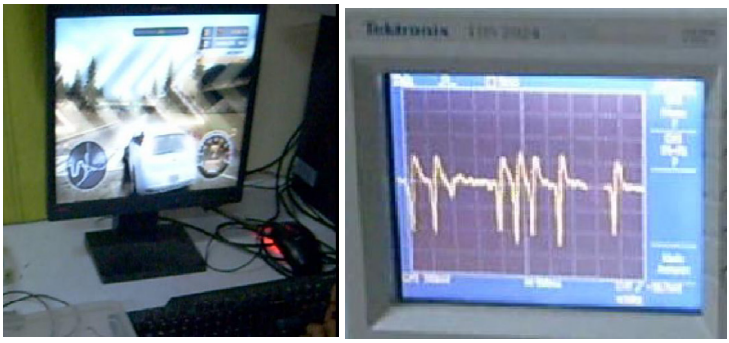
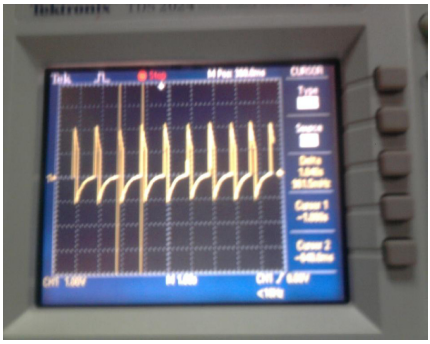


Fig. 2. Car when it crashes and eeg of the person driving the car

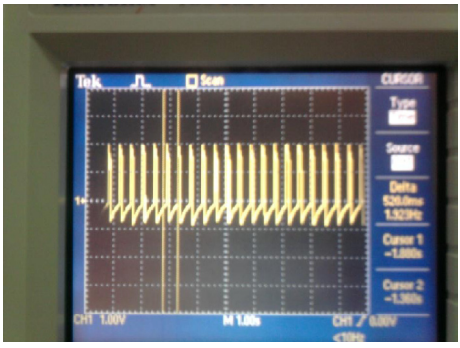
3.3 EEG of a Subject While Maneuvering the Car

Fig. 3. Car while maneuvering and EEG of the person driving the car

3.4 Heart Rate of the Subject While Driving on a Straight Road



3.5 Heart Rate of the Subject When the Car Crashes



4 Analysis

Different features appear in the data that have been recorded during the experiment. Important features of this experiment are the eyeblinks and heart rate, as their frequency is used to measure the perceived level of stress the participant is under. The potential difference between sensors Fp1 and Fp2 and their neighbors is increased during the duration of the eye blink, which is typically between 200 and 400 ms. The longitudinal differences say between Fp2-F4 of Fp1-F3 allow easy visual identification of the eye blinks as peaks on the differences between Fp1 and Fp2 and their surrounding electrodes. The transverse differences measures the potential differences between Fp1 and Fp2. But since both sensors Fp1 and Fp2 will be triggered simultaneously during an eye blink, the eye blinks are barely noticeable in the difference between Fp1 and Fp2. Hence the EEG signals obtained between Fp1-F3 and Fp2-F4 sensors were observed and analyzed. The ocular artefacts were separated from the EEG signal using the software described above and the following graphs were obtained. These artefacts were indicative of the eye-blink frequency of the subject.

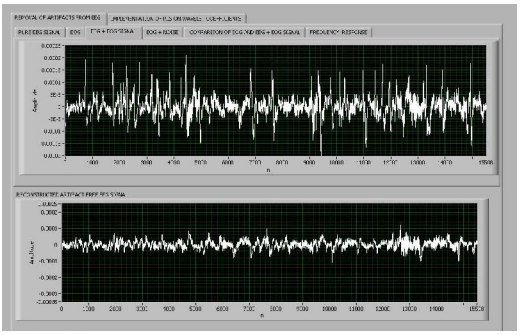


Fig. 4. Contaminated EEG (EEG+ ocular artefacts) and reconstructed artifact free EEG signal

The filtered EEG signal was then separated into its various components using the second software described above. The various bandpass filters used in the software were able to filter out the alpha, beta, gamma and delta components from the EEG signal. They are as shown in the figure. The alpha and the beta components were then analyzed to see if there was any change in them during stressful situations like car crash.

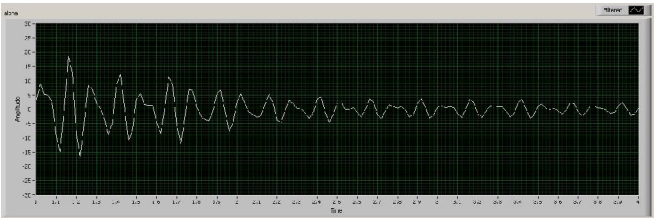


Fig. 5. Alpha Waves in the EEG signal

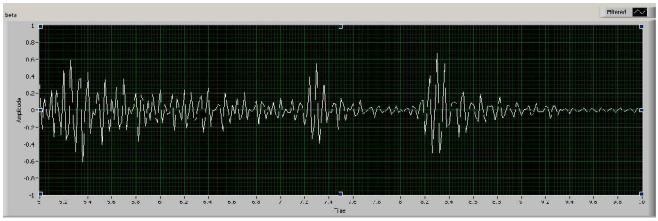


Fig. 6. Beta Waves in the EEG signal

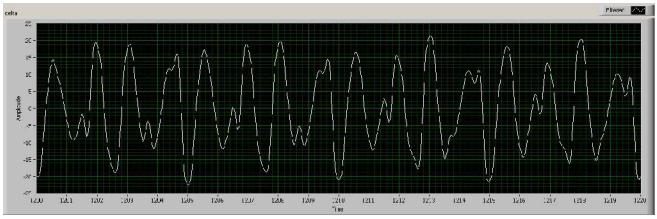


Fig. 7. Delta Waves in the EEG signal

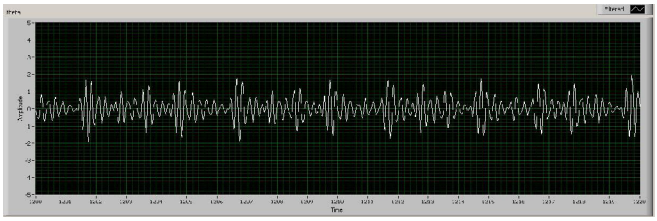


Fig. 8. Theta Waves in the EEG signal

The number of heart beats occurring per minute was also evaluated to determine their significance as a parameter for stress monitoring.

5 Results

Our results show that there is a strong correlation between eye blink frequency and emotional stress. This was even more apparent during more confronting situations, such as the simulated car crashes. Although the temporal increase in eye blink frequency can already be used as a measure of stress, our method is not accurate enough to be used in commercial applications. Further isolating the eye blink from the EEG signal using other features and possible enhanced stress detection using additional sensors, for example attached to the eye lids are being tried out. The detection of eye blinks in an EEG signal using the differences of sensors Fp1 and F4 or Fp2 and F3 done by our software can easily be verified by comparing it with the eye-blinks located visually by the test administrator.

EEG of someone under stress also displays decrease in both alpha (11–12 Hz) and sensorimotor rhythm (SMR, 12–15 Hz) activity, and increases in EEG amplitude in the 19–22 Hz and high beta (23–35 Hz) ranges. Emotional intensity, particularly

relating to anxiety, correlates with 19–22 Hz band activity, while activity within the 23–36 Hz band reflects an active brain state. The Alpha waves which reflect a calm, open, and balanced psychological state showed a decrease in activity during stress. The beta waves were found to be dominating frontally when the subject was stressed and showed increased amplitude.

Stress induced a change in autonomic functioning. Blood pressure and heart rate increased during stress, reflecting a predominance of sympathetic nervous system activity. It was found that when the car crashed the heart rate of the subject showed a rise as compared to his/her heart rate while driving on a straight road.

6 Future Scope

The hardware system designed by us provides the capability of procuring EEG signals as well as Heart rate signals. The software provides the facility to separate the ocular artifacts which occlude the underlying EEG signal. It also allows the user to filter out the various components in the EEG signal. The 2 parameters monitored by us in this project can be used as a part of a bigger system for Stress Monitoring. Apart from these 2 parameters the following parameter can be used in a commercial Stress Monitoring System:

6.1 SRI(Stress Response Inventory) and SAM(Self Assessment Manikin) with Cortisol Level Monitoring

Stress can be assessed with the SRI, which participants complete prior to physiological measurements being obtained. The SAM has been used widely to assess the emotional response of subjects to experimental stimuli. Visual stimuli from the International Affective Picture System (IAPS) are used to evoke emotional responses. The stimuli are presented as a set of pleasant images and a set of unpleasant images. Each set lasts for 5 min, with individual images presented for 15 s. Participants make a rating with the SAM after each image set. Salivary cortisol is collected after the SRI had been completed. Cortisol levels quantify the endocrine response to stress.

Our results suggest that stress may be assessed by increase in the frequency of eye-blinks which can be detected using EEG as well as increase in the heart rate. Indeed, our variables could be combined with the above stated parameters in order to develop a commercial chronic stress monitoring system based on the proposed statistical model.

A set of healthy male and female subjects aged between 18-21 years can be considered for experimentation. They can be asked to describe their level of stress using SRI and SAM mentioned above. 10 samples each can be taken from 5 males and 5 females.

7 Proposed Statistical Model

We can use Bayes classifier (optimal statistical classifier) to classify a person as either stressed or non-stressed. This method is similar to one used for object recognition in

image processing. The Bayes classifier minimizes the probability of mis-classification.

We have 2 classes: W1- Unstressed W2- Stressed

Feature vector X used for event description is 2-D, average eye-blink time interval (seconds) and heart beat rate (beats/minute). m1(X), m2(X) are mean vectors and C1(X), C2(X) are covariances obtained from experimental data. The decision function dj(X) is given by p(X/Wj)*P(Wj). Initially probabilities of occurance of both classes can be initialised to 0.5 each and suitably modified after results of a first few trials. p(X/Wj) is as usual assumed to follow Gaussian distribution. Once the classes are obtained from the experimental data, any unknown test subject whose stress level is to be monitored can be easily classified into one of the 2 classes- stressed and unstressed using this Bayes classifier.

References

1. Haak, M., Bos, S., Panic, S., Rothkrantz, L.J.M.: Detecting stress using eye blinks and brain activity from EEG signals
2. Senthil Kumar, Arumuganathan, Sivakumar, Vimal: An Adaptive method to remove ocular artifacts from EEG signals using Wavelet Transform. Journal of Applied Sciences Research 5(7), 741–745 (2009)
3. Shah, J.D., Panse, M.S.: EEG purging using labview based wavelet analysis. In: NCCI 2010-National Conference on Computational Instrumentation (2010)
4. Teplan, M.: Fundamentals of EEG measurement. Measurement Science Review 2, Section 2 (2002)
5. Shaker, M.M.: EEG Waves Classifier using Wavelet Transform and Fourier Transform. International Journal of Biological and Life Sciences 1, 2 (2005)
6. Prutchi, D., Norris, M.: Design and development of medical electronic instrumentation-a practical perspective of the design, construction, and test of medical device. John Wiley & Sons Publication
7. Seo, S.-H., Lee, J.-T.: Stress and EEG, Department of Computer Science & Engineering, Pusan National University 30, Jangjeon-Dong, Geumjeong-Gu Busan, Korea
8. Garcés Correa, A., Laciar, E., Patiño, H.D., Valentinuzzi, M.E.: Artifact removal from EEG signals using adaptive filters in cascade. In: 16th Argentine Bioengineering Congress and the 5th Conference of Clinical Engineering
9. Kaban, A.: Introduction to Bayesian Learning. School of Computer Science, University of Birmingham

Design of a Fuel Free Electric Vehicle

Chellaswamy Chellaiah[1], Kaliraja Thangamani[3], P. Glaret Subin[2], P. Rathinakumar[1], and P. Muthukrishnan[1]

[1] SRM University, Vadapalani Campus, Chennai, 600122 India
[2] Jawaharlal Nehru Technological University, Kakinada
[3] SKR Engineering College, Anna University, Chennai, India

Abstract. Nowadays electric vehicles are appearing in different shapes and sizes. The fuel vehicles are emitting toxic gases and polluting the environment. The aim of this paper is to design a Fuel free Electric Vehicle (FEV). An advanced battery system and drive is integrated through Controller Area Network (CAN) for signaling and controlling operations. The design of automatic charging system meets different design parameters and automatically starts and stops the charging system, eliminating the need for driver involvement. The proposed system contains two turbines with generator, one is for charging the auxiliary battery and the other one is for two purposes: for charging the HV battery system and drive the motor directly if the speed exceeds twenty kilometers. The performance of the vehicle is studied and we believe that it will eliminate total diesel exhaust and help for healthy generation

Keywords: electric vehicles, power generator, battery charging, inverter.

1 Introduction

The world will never actually run without oil. It will just be more and more expensive and environmentally destructive to extract the remaining supply. As the world population grows, the effects of the pollution they produce become more and more destructive. The cost of controlling these pollutants is increasing every day. Mercury has poisoned our fisheries and acid rain from sulfur has killed forests and lakes. Acidified oceans have damaged coral reefs all over the world. Oil sands, shale mining and mountaintop removal pollute our water and leave a wasteland in their wake. Fossil fuels were once so cheap that we quickly developed wasteful ways that made us addicted to them. Now that the easy deposits have been depleted, it becomes more and more expensive and destructive to extract them from the earth. Electric vehicles typically have less noise pollution than an internal combustion engine vehicle, whether it is at rest or in motion. An electric vehicle emits no tailpipe CO2 or pollutants such as NOx, NMHC, and CO.

Hybrid vehicle is the combination of linear engines and linear alternator that are explained in [1]. The operation of each component of the system has been examined and the design procedure for the alternator has been outlined. The design and analysis of the proof of concept model has indicated areas in which the alternator can be

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 459–464, 2013.

improved, such as decreasing the winding resistance for improved efficiency and lowering the weight of the translator for increasing the speed of operation. The permanent magnet generator output is rectified and passed through a buck converter before connection to the main DC bus bar for getting maximum engine efficiency [2]. The choice of power electronic interface dictates design details in the generator and exact level of output voltage. With the buck converter a high output voltage is required and this necessitates the use of a double layer winding. A self reconfigurable electric motor controller for hybrid vehicle is proposed in [4]. The failure of current sensors are easily detected and estimate the phase current using Luenberger type observer The developed observer is implemented on a 3-phase four pole permanent magnet motor with digital signal processor. The bi-directional ac-dc converter for plug-n hybrid vehicle is presented in [5]. The converter topologies are analyzed in detail and the combined topology is more advantages with respect to cost and weight.

The fuel consumption and emission of hybrid vehicle and the energy management strategy is studied in [6]. From the starting general optimal control problem a new cost functional is defined and account for electrical energy supplied from the grid. Design and implementation of an electric drive system for in-wheel motor using Matlab SIMULINK model is developed and the performance values are calculated in [7]. The hub type drive has more advantages and it can be used in vehicle technology.

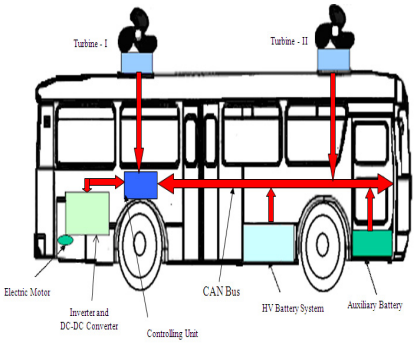


Fig. 1. Different components present in the vehicle

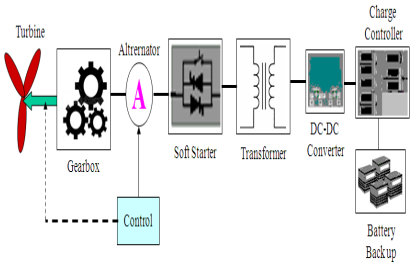


Fig. 2. Block diagram of Turbine-I control system for charging

We believe that the implemented system has many advantages and can be adopted by developing countries at all levels. The rest of this paper is organized as follows: section 2 describes the design of proposed system; section 3 explains the performance evaluation of FEV; and finally, conclusion is discussed in section 4.

2 Design of FEV

2.1 Physical Model of Proposed System

The advanced induction motor and the driver system are developed specifically for vehicle use. The micro controller based power electronic system with sensors

communicates with the Controller Area Network (CAN). The design also contains high power Switched Mode Rectifier (SMR) and inverters for fast charging of the batteries. Various components present in the bus are shown in fig.1 In this system we are incorporating two different capacity turbines with gear arrangement and variable speed Lundell alternators, one is for charging the auxiliary batteries automatically if the vehicle is under running condition as shown in figure 2, and the turbine-II is for two purposes: charging the HV battery and directly run the motor if it exceeds 20 kilometer speed. The motor is taking power from the battery at the starting time and it will automatically switch over to direct supply after attaining the speed of 20-kilometers is shown in figure 3. The controlling operations are maintained by the CAN controller, which is integrated with the system. The alternator-I is a 15KW capacity to recharge the battery system. The wind turbine rotor is connected to the induction generator through a gearbox. The induction generator has many advantages like robustness, mechanical simplicity and low cost. The control system is used to ensure the proper operation of turbine under all conditions. By using a soft starter the output of generator is connected to the load. We are including a variable speed wind turbine in turbine-II for reducing mechanical stress and improve the power quality. The capacity is around 25KW and the output power is maintained with the help of a transformer. Frequency converters become more significant control for wind turbines. The most important properties of frequency converter are 1.The rotor behaves as energy storage 2. Loads on the gear and drive train can be reduced. 3. The acoustic emission noise can be reduced. 4. The power absorbed at low wind speed can be improved. Frequency converter can control the reactive power such that the power quality can be improved. The voltage stability is improved and flicker level is reduced.

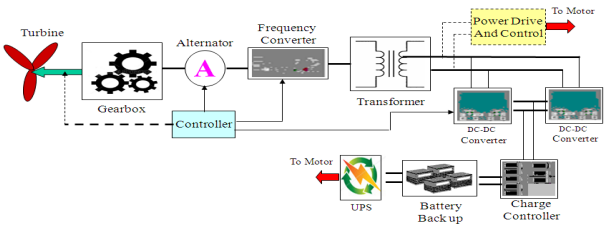


Fig. 3. Block diagram of Turbine-II control system for charging and direct drive

Different battery types and technologies have emerged and become more economical. Lead acid, CNG, Nickel Metal Hydride, Zebra are extensively used in electric vehicles [8], [9]. We are incorporating Zebra batteries because of the storage capacity, good performance and less cost. The battery is completely sealed and include terminal conductors that are connected through a battery management system via CAN bus communication.

2.2 Electrical Design Equation

The current density due to the magnet, J_m is given by

$$J_m = \frac{4}{\tau}\frac{B_r}{\mu_o \mu_r}\sin\frac{W_n}{2}M_d \ \text{A/m}^2 \tag{1}$$

Where B_r is the remenence of magnet, τ is the pole pitch, M_d is the magnet diameter and $W_n = \Pi n/\tau$ described by [10]. The distribution of flux density is almost sinusoidal and is given by

$$B_x = \left[\frac{J_m \mu_o}{W_n}\frac{\sinh W_n t_m}{\sinh W_n g/2}\right]\cos W_n x \tag{2}$$

The magnetic flux density distribution around each magnet with radius r from the center is described by

$$\phi_A = 2\Pi B\left[\left(\frac{1}{u_1}\right)^2(\cos u_1 - 1) + \frac{1}{u_1}(r \sin u_1 r)\right] \tag{3}$$

The armature coil is concentrated at its axial position then the total flux linkage is given by

$$\phi_f = \frac{N}{3}(\phi_A) \text{ and the coil emf is } E_c = \frac{2\Pi}{\sqrt{2}} f\phi_f$$

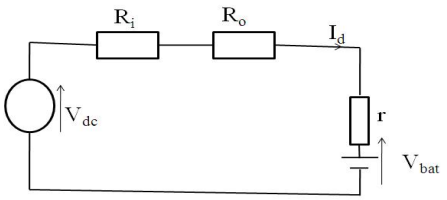


Fig. 4. Equivalent dc circuit for battery charging

The three phase generator output is rectified and directly used for charging batteries or connect to the inverter. The equivalent circuit of battery charger is shown in fig. 4

3 Performance Evaluation

A DC power supply from the battery is used to energize the field winding with 5A. The output voltage of the alternator is varying based on the speed of the vehicle

because the alternator is mounted on the vehicle. A special type of blade is connected to the alternator shaft. The output voltage of alternator at different speed is shown in figure 5. The SMR plays an important role in maximizing the power at low speed by varying the output voltage, and also at high speed to limit the output voltage.

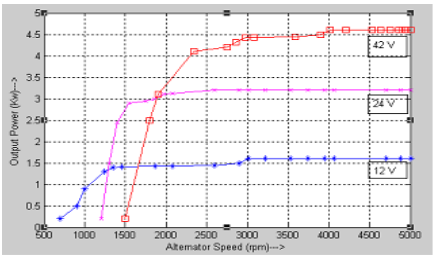


Fig. 5. Output power versus different speed

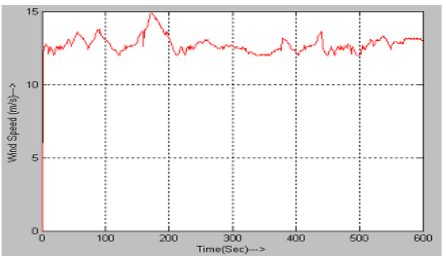


Fig. 6. Simultaneous changes in wind speed

The wind direction can vary between day and night according to the location. Here we are incorporating the turbine on the bus so it will automatically rotate if the vehicle is in motion. The speed of the turbine and the vehicle speed are proportional and it is not always constant. The simultaneous changes in the wind (ie: change in speed of vehicle) and the extreme speed are shown in figure 6.

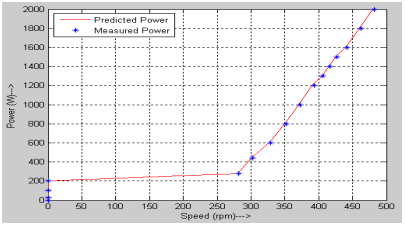


Fig. 7. Load Characteristics of inverter

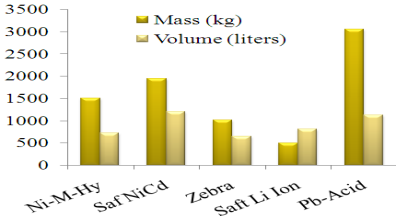


Fig. 8. Comparison of mass and volume of different batteries

The inverter is connected to all the coils and the measured and predicted power is shown in figure 7. The test points are plotted for 400 rpm of generator. Several manufacturers use advanced technology for producing batteries and these batteries remain relatively expensive. From all the batteries zebra provides lowest weight with sufficient energy storage. Figure 8 shows comparison between mass and volume of various battery technologies. Mass and volume of the zebra battery is less and it is higher than saft Li Ion. But other parameters of this technology are much higher so we can prefer zebra batteries for our proposed system.

4 Conclusion

Electric vehicles have a tremendous potential to improve energy efficiency and to provide pollution free world for our future generation. We are introducing a set of

new design and control techniques for automotive alternators that yield dramatic improvements in performance and functionality as compared to conventional systems. In this paper we are using two alternators for reducing the number of batteries as a result the weight. The system automatically recharges the battery at the running time of the vehicle. So charging stations are not required for charging the batteries present in the vehicle. The CAN controller can handle the system efficiently if it is operating under battery mode or in direct mode of operation. Our system reduces the total weight of the system around 27 percentage compared with hybrid vehicles. We believe that the developing countries can adopt the system for pollution free environment.

References

1. Cawthorne, W.R., Famouri, P., Chen, J., Clark, N.N., McDaniel, T.I., Atkinson, R.J., Nandkumar, S., Atkinson, C.M., Petreanu, S.: Development of a Linear Alternator–Engine for Hybrid Electric Vehicle Applications. IEEE Transactions on Vehicular Technology 48(6) (November 1999)
2. Brooking, P., Bumby, J.R.: An Integrated Engine-Generator Set With Power Electronic Interface For Hybrid Electric Vehicle Applications. In: Power Electronics, Machines and Drives Conference, April 16-18, vol. (487) (2022)
3. Merkisz, J., Pielecha, J.: Emissions and Fuel Consumption during Road Test from Diesel and Hybrid Buses under Real Road Conditions. In: IEEE Conference on Power and Propulsion Conference (VPPC), pp. 1–5 (2010)
4. Parsa, L., Toliyat, H.A.: A Self Reconfigurable Electric Motor Controller for Hybrid Electric Vehicle Applications. In: IEEE Conference on Industrial Electronics Society, pp. 919–924 (2003)
5. Shi, L., Meintz, A., Ferdowsi, M.: Single-Phase Bidirectional AC-DC Converters for Plug-in Hybrid Electric Vehicle Applications. In: IEEE Vehicle Power and Propulsion Conference (VPPC), Harbin, China, September 3-5 (2008)
6. Stockar, S., Marano, V., Rizzoni, G., Guzzella, L.: Optimal Control for Plug-in Hybrid Electric Vehicle Applications. In: American Control Conference, Marriott Waterfront, Baltimore, MD, USA, June 30-July 02, pp. 5024–5030 (2010)
7. Nejat Tuncay, R., Ustun, O., Yilmaz, M., Gokce, C., Karakaya, U.: Design and Implementation of an Electric Drive System for In-Wheel Motor Electric Vehicle Applications. In: Vehicle Power and Propulsion Conference (VPPC), pp. 1–6 (2011)
8. MES-DEA, ZEBRA Batteries for Electric Cars, http://www.cebi.com/content/index_html?a=8&b=151&c=208
9. MES-DEA, ZEBRA Technical Information for Z36-371-ML3X-76 Type, http://cebinew.kicms.de/cebi/easyCMS/FileManager/Files/MES-DEA/batteries/Zebra_Z36.pdf
10. Bumby, J.R., Martin, R., Spooner, E., Brown, N.L., Chalmers, B.J.: Electromagnetic design of axial flux permanent magnet machines. Proc. IEE – Electrical Power Applications 151(2), 151–160 (2004)
11. Mohan, N., Underland, T.M., Robbins, W.P.: Power Electronics – Converters, Applications and Design, 3rd edn. John Wiley and Sons (2003) ISBN 0-471-429078-2

Power System Generator and Voltage Stability Enhancement by the Hardware Circuit Implementation of 3-Ph Static Var Compensator (SVC)

Venu Yarlagadda[1], B.V. Sankar Ram[2] and K.R.M. Rao[3]

[1] VNR VJIET, India
venuyar@gmail.com
[2] JNTUH, India
bvsram4321@yahoo.com
[3] MJCET, India
drkrmr@yahoo.com

Abstract. The design, Fabrication of 4-Kvar 3-phase fixed capacitor-thyristor controlled reactor (FC-TCR) type SVC has been developed in the laboratory for a Single Machine Single Bus (SMSB) Test System. The test system is setup in the laboratory to evaluate the FC-TCR type SVC in stabilizing bus voltage. It comprises of a 3-phase Synchronous Machine of 5kva capacity and a 3-phase squirrel cage Induction Motor of 5HP Rating. 3-Ph SVC have been designed and fabricated and tested by connecting it to a SMSB Test System and experimental results have been presented in this paper. The Power – Voltage (P-V) and Power – Load angle (P-δ) Curves of the SMSB Test system with and without SVC have been plotted which shows the effectiveness of SVC on Generator and Voltage Stability enhancement.

Keywords: FC-TCR type SVC, SVC control, open loop control of SVC, Single Machine Single bus System, Voltage Stability Enhancement and P-V Curves, P-δ Curves and Power System Stability.

Nomenclature: α is the Firing Angle, TCR : Thyristor Controlled Reactor, FC: Fixed Capacitor-V: Power – voltage, B_{SVC} : Net susceptance of SVC, B_{TCR} : susceptance of TCR, B_c : Capacitive susceptance, X_L: Inductive reactance, Q_{SVC} : Reactive Power supplied or absorbed by SVC, I_C : Current Through the Capacitor.

1 Introduction

This paper deals with the Design, Fabrication and Testing of three phase 4KVAR SVC by connecting it to a SMSB Test System. Thyristor Controlled Reactors (TCRs) are one of the most important building blocks of SVC. Although TCR can be used by itself, it is most often employed in conjunction with fixed or thyristor switched capacitors to provide rapid, continuous control of reactive power over the entire selected lagging-to-leading range. SMSB test system consisting of a 3-phase 5-kVA synchronous machine and a three phase 5-horse power squirrel cage induction motor has been set up in the laboratory to evaluate the performance of the SVC. Static Var

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 465–469, 2013.

Compensators (SVCs) have been used as FACTS controllers and have a proven track record in the voltage stability and Generator Stability Improvement. An open loop control has used to adjust the equivalent susceptance of the SVC.

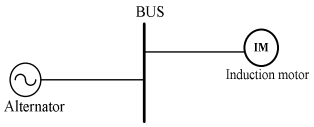


Fig. 1. SMSB Test System

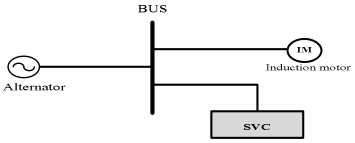


Fig. 2. SMSB Test System with SVC

Fig.1 shows the single line diagram of SMSB test system and Fig.2 shows the single line diagram of SMSB test system with SVC. The laboratory setup of the SMSB test system which has been developed with and without SVC has been presented in the paper. The SVC consists of a 2-kVar thyristor controlled reactor in parallel with 4-kVar capacitor bank comprising of two banks of each rating 2-Kvar along with mechanical switches. The results corresponding to the injection of Vars to maintain constant voltage profile are presented in this paper. The P-V *curves* of the test system with and without the SVC have been obtained experimentally. The results demonstrate the effectiveness of the SVC in enhancing the Generator and voltage stability. Therefore, the firing angle can be in the range of $90° \leq \alpha \leq 180°$ to vary the TCR current from zero to its maximum

2 Power System Stability

Successful operation of a power system depends largely on the engineer's ability to provide reliable and uninterrupted service to the loads which depends on Power System Stability.

2.1 Classification of Stability

Power System Stability is broadly classified into two categories one is Generator Stability or Rotor angle Stability and the other is Voltage Stability. Rotor angle stability is mainly interlinked to real power transfer whereas voltage stability is related to reactive power variations.

2.2 Stability Indices

2.2.1 P-V Curve

As the power transfer increases, the voltage at the receiving end decreases. Finally, the critical or nose point is reached. The curve between the variation of bus voltages with output power (P) is called as P-V curve or 'Nose' curve. PV curves are used to determine the loading margin of the power system. The margin between the voltage collapse point and the current operating point is used as voltage stability criterion.

2.2.2 P-δ Curve

The relation between input power and the load angle is called power angle characteristics. The equation is given by, P=EVsinδ/X. The steady state stability limit is EV/X and it occurs at 90^{o}.

3 Operation of SVC

The Fig.3 shows the SVC which behaves like a shunt-connected variable reactance, which either generates or absorbs reactive power in order to regulate the PCC voltage magnitude. In its simplest form, the SVC consists of a TCR in parallel with a bank of capacitors.

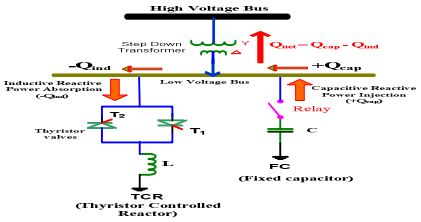


Fig. 3. Single line Diagram of FC-TCR type SVC

The SVC regulates voltage at its terminals by controlling the amount of reactive power injected into or absorbed from the power system. When system voltage is low, the SVC generates reactive power (SVC capacitive). When system voltage is high, it absorbs reactive power (SVC inductive).

4 Design of SVC

The design of SVC is based on test system Requirements. For any system find the variation of reactive power with load variations i.e $Q_d = (Q_1, Q_2, .., Q_n)$.Whare Q_d is the Reactive Power Demand and $Q_1, Q_2, ..., Q_n$ are the variations in demand.

Set the Reference Voltage it may be set to 1.0 p.u or as closer as possible to it and find corresponding Reactive Power demand. Set reference reactive power Q_{ref} as 1.0 p.u (corresponding to voltage value of 1.0 p.u).

The fixed Capacitor (FC) value of the SVC can be obtained as $Q_c = Q_{max} - Q_{ref,}$ Where Q_c is the Reactive Power of the FC and Q_{max} is the maximum Reactive Power Demand.

$Q_c = V^2 / X_C$ (1), $X_C = 1/2\ fC$ (2), In SVC circuit X_C is in parallel with X_L

$Q_{tcr}(\alpha) = Q_{ref} - Q_{min}$ (3), $X_L = = V^2 / Q_{tcr}(\alpha)$ (4), $Q_{svc} = Q_c - Q_{tcr}(\alpha)$ (5)

hence SVC can supply dynamic Reactive power support for maintaining the constant Voltage which will enhance the Generator and voltage stability margins tremendously. Based on design criterion we have chosen FC is equal to 4 Kvar and a Reactor of 2 Kvar Rating which suits to the test system requirements.

5 Experimental Results

Fig.3 Shows the Single Line diagram of SMSB Test System with SVC Table.1 shows the results without SVC and Fig.4 Shows the P-V Curve without SVC. Table.2 shows the results With SVC and Fig.5 Shows the P-V Curve with SVC. Fig.7 shows the P-V Curves without and with SVC and Fig.8 shows the P-Delta Curves without and with SVC

Table 1. Results without SVC

S.No	Bus Voltage (V)	P (W)	I_L (A)	δ (°)
1	425	240	4.6	2.35
2	415	560	5	3.9
3	410	760	5.5	5.4
4	409	880	6	7.35
5	405	1200	7	10.7
6	400	1280	8	11.4

Table 2. Results with SVC

Sno	Bus Voltage (V)	I_L (A)	P (W)	δ (°)	α (°)
1	415	1.5	240	2.1	145
2	415	1.7	400	5.1	136
3	415	2	560	7	132
4	415	2.5	760	8	112

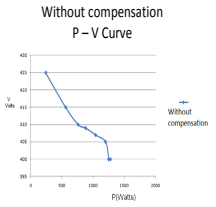


Fig. 4. P-V Curve Without SVC

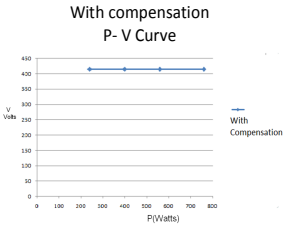


Fig. 5. P-V Curve with SVC

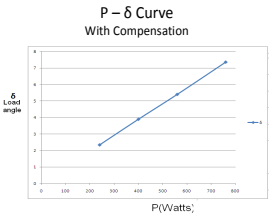


Fig. 6. P-δ Curve with SVC

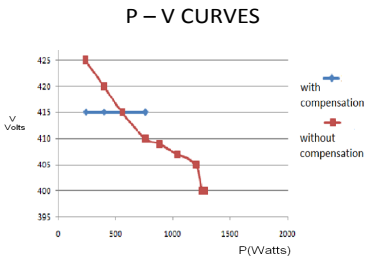


Fig. 7. P-V Curve with and Without SVC

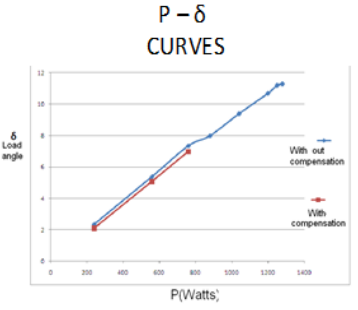


Fig. 8. P-δ Curves with and without SVC

6 Conclusion

In this paper, a 3-Phase SVC is designed, fabricated and tested manually to get the constant voltage by supplying variable reactive power to the system.. The power circuit of a three phase 4KVAr FC-TCR type SVC have been tested experimentally using a SMSB test system.. P-V and P-δ curves have been drawn for the Open loop control of SVC and it is found that the voltage has been maintained constant against load variations.

The proposed model is experimentally verified and is found to give very fast and precise compensation characteristics. The test results demonstrate the effectiveness of the 3-ph SVC. Both the Generator and voltage stability margins have been increased tremendously. Finally constant bus voltage demonstrates the effectiveness of the compensator.

7 Future Scope

The variable shunt compensation using SVC can be extended to the large rating machines as well as for the large Interconnected Power Systems. The SVC can be fabricated by using IGBT's and the control strategy can be implemented by using DSP. TSC-TCR based SVC can also be implemented for SMSB Test System

Acknowledgements. We are immensely thankful to Management and Principal of VNR VJIET for providing R&D lab and other facilities to carry out this work.

References

1. Hingorani, N.G., Gyugyi, L.: Understanding FACTS. IEEE Press, New York (1999)
2. Mathur, R.M., Varma, R.K.: Thyristor-Based FACTS Controllers for Electrical Transmission Systems. IEEE Press, Wiley Interscience, New York, USA (2002)
3. Tagare, D.M.: Reactive Power Management. McGraw-Hill Publication, India (1982)
4. IEEE Power Engg. Society/CIGRE, "FACTS Overview", Publication 95 TP 108. IEEE Press, New York (1995)
5. Kundur, P.: Power System Stability and Control. McGraw-Hill, Inc., New York (1994)
6. IEEE Power Engineering Society, "FACTS Applications", Publication 96 TP 116-0. IEEE Press, New York (1996)
7. Paserba, J.J.: Fellow, IEEE, has proposed, Introduction to FACTS Controllers (2005)
8. Chaghi, A., Louchene, A., Bensaadi, R.: A microcontroller-based static var compensator for unbalanced loads. IEEE Press, India (2008)
9. Taufik, Paet, B.: A small scale Static VAR Compensator (SVC) for Laboratory Experiment. IEEE Press, India (2009)
10. Ahmed, T., Nishida, K., Soushin, K., Nakaoka, M.: Static VAR compensator-based voltage control implementation of single-phase self-excited induction generator. IEEE Press, India (2008)

STATCOM for Improvement of Active and Reactive Power at the Wind Based Renewable Energy Sources

S. Narisimha Rao, J. Sunil Kumar, and G. Muni Reddy

Department of EEE, SKIT, Srikalahasti, Andhra Pradesh, India
{snrinskit,sunilkumarjelledi}@gmail.com,
gmunireddy@rediffmail.com}

Abstract. The Renewable energy sources like wind energy which has been predictable to be a promising alternative energy supply, can bring new challenge when it is connected to the power grid. The major problems occurring are, when the wind flow rate is low the out coming of power is low compared to normal condition and problems like voltage stability, Reactive power, active power variations may occur. To overcome these problems a new emerging devices are implementing at the wind plants. So Flexible ac Transmission Systems (FACTS) device, STATCOM is implementing at the wind generating plants to improve the voltage stability problems and to improve the power profile of the system.

Keywords: Wind energy, Voltage Stability, Active and Reactive power, FACTs, STATCOM.

1 Introduction

The location of generating power form wind energy is determined by wind energy resource availability like wind flow, regularly for high voltage (HV) power transmission network and major consumption centers [1]. To have sustainable growth and social progress, we have to save the conventional sources and we have to utilize the nonconventional energy sources like wind, solar, biomass etc. It is necessary to meet the energy need by utilizing these renewable energy resources. The power of wind has been utilized from thousands of years. It may be of different forms. The first wind turbines for electricity generation had already been developed at the beginning of the twentieth century. The technology was improved step by step from the early 1970.By the end of 1990s wind energy has reemerged as one of the most important sustainable energy resources.

Wind energy is gaining increasing importance throughout the world. These fast developments of wind energy technology and of the market have large impact on a number of people and institutions who are working for instance, for scientists who research and teach future wind power and electrical engineers at universities. With the rapid improvement of wind technology, the cost of testing equipment is more. So far analyzing the whole system we have to develop a prototype, but it is also becoming high costs, so in order to analyze most of the people are recommended to the simulation studies. [2]

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 470–476, 2013.

Power electronic devices have a revolutionary impact on the electric power systems around the world. Flexible AC Transmission System (FACTS) devices are new comings of power electronic devices, which have found a wide spread application in the power industry for active and reactive power control. The StaticSynchronous Compensators (STATCOM) can be used to provide reactive power to stabilize a wind farm also.. In this paper, we will in particular look at the application of STATACOM for the purpose of transfer capability of power with high penetration of wind power. The paper investigates and demonstrates how, at effective FACTS based power flow control can be applied to relieve the transmission congestion and improve the power capability in the system with high penetration of wind power, while improving the voltage profile.

2 Overview of Wind Energy

The working principle of the wind turbine includes the following conversion processes: the rotor extracts the kinetic energy from the wind creating generates torque and the generator converts the torque into electricity and feeds it into the grid. Presently there are three main turbine types available, like Squirrel-cage induction generator, Doubly fed induction generator, Direct-drive synchronous generator.

India is drastically improving the usage of wind energy in the early mid 1990s. Especially in India the Tamil Nadu state is leading with this development, but the state Gujarat is also playing a key role in developing the wind energy. There has been an extensive growth and quick development in the exploitation of wind energy in recent years. The individual units can be of large capacity up to 2 MW, feeding into distribution network, particularly with customers connected in close proximity [3].The main scenery of generators contributes much to the transient stability of wind power generation. Wind turbine generators are generally classified into two main types, fixed speed generators and variable speed generators. Fixed speed generators mostly refer to squirrel cage induction generators.. Specially, during a network fault, the rotor speed increases and absorbs a large amount of reactive power upon fault clearance which may cause voltage instability to the electrical power system.

Asynchronous generator block operates in either generator or motor mode. The mode of operation is dictated by sign of the mechanical torque, If the mechanical torque is positive the machine acts as a motor, if it is negative the machine acts as a generator.

$$V_{qs}=R_s i_{qs}+\frac{d}{dt}\varphi_{qs}+\omega\varphi_{ds} \tag{1}$$

$$V_{ds}=R_s i_{ds}+\frac{d}{dt}\varphi_{ds}+\omega\varphi_{qs} \tag{2}$$

$$V^1_{qr}=R^1_r i_{qr}+\frac{d}{dt}\varphi^1_{qs}+(\omega-\omega_r)\varphi^1_{dr} \tag{3}$$

$$V^1_{dr}=R^1_r i_{dr}+\frac{d}{dt}\varphi^1_{ds}+(\omega-\omega_r)\varphi^1_{qr} \tag{4}$$

$$T_e=1.5\ P\ (\varphi_{ds} i_{qs}-\varphi^1_{qr} i_{ds}) \tag{5}$$

With power rating from a few kilo watts to a few megawatts, synchronous generators provide great flexibility to meet special technical requirements in realistic wind energy systems [5].Some modeling equations of the synchronous generator when they are transformed in to abc to dq reference

a. Armature equations:

$$V_d=-R_S i_d-\omega\lambda_q-(L_{ls}+L_{md})\frac{di_d}{dt}+L_{md}\frac{di_{fd}}{dt}+Lmd\ \frac{di_k}{dt} \quad (6)$$

$$V_q=-R_S i_q-\omega\lambda_d-(L_{ls}+L_{mq})\frac{di_q}{dt}+L_{mq}\frac{di_{kq}}{dt} \quad (7)$$

b. Field equations

$$V_{fd}=-R_{fd}i_{fd}-L_{md}\frac{di_d}{dt}+(L_{lfd}+L_{mq})\frac{di_{fd}}{dt}+L_{md}\frac{di_{kd}}{dt} \quad (8)$$

c. Damper winding

$$0=-R_{kd}i_{kd}-L_{md}\frac{di_d}{dt}+L_{md})\frac{di_{fd}}{dt}+(L_{md}+L_{lkd})\frac{di_{kd}}{dt} \quad (9)$$

$$0=-R_{kq}i_{kq}-L_{mq}\frac{di_q}{dt}+(L_{mq}+L_{lkq})\frac{di_{kq}}{dt} \quad (10)$$

The Asynchronous generator can be constructed with a large number of poles and operate at a speed that directly matches the turbine blade speed. Such a direct-drive system does not need a gearbox. This results a reduction in installation and maintenance costs and provides an advantage over induction generator (IG) based turbines where use of a gearbox is a must. The SG wind energy system is normally controlled by full capacity power converters for variable-speed operation [2], ensuring maximum wind energy conversion efficiency throughout its operating range. With full-capacity converters, the system is able to meet various grid codes [3], including leading/lagging reactive power control and fault ride through operation, without the need for additional equipment. The main parts of the wind energy system are blade or rotor, which converts the energy in the wind to rotational shaft energy, a drive train, usually including a gearbox and a generator, a tower that supports the rotor and drive train, brake, yaw motor, Nacelle, and other equipment, including controls, electrical cables, ground support.etc..

3 Implemented STATCOM Model

The concept of FACTS was first defined by Hingorani. N. G. It was in 1988 that the concept of FACTS was first introduced as “the combination of the application using high power electronic components to enhance the controllability of the power systems”. Flexible AC Transmission Systems (FACTS) are devices are capable of altering voltage, phase angle and/or impedances at particular points in power system. FACTS device provide new control facilities, both in steady state power flow control and dynamic stability control.. The increased interest in these devices is essentially due to two reasons. Firstly, the recent development in high power electronics has made these devices cost effective and secondly, increased loading of power systems, combined with deregulation of power industry, motivations the use of power flow control as a very cost effective means of dispatching specified power transactions.

FACTS devices are of many types based on connection series, shunt, series-series, and series-shunt. STATCOM is one of the most versatile FACTS device which is connected in shunt, searching for a more effective control strategy in implementing to the renewable energy sources, is the objective of the researcher. Until now most of the investigations is done, on STATCOM implementation in the grid systems to solve the voltage flickers, power quality problems [4].

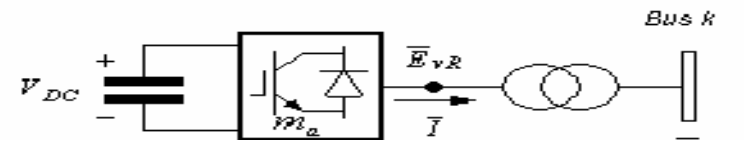


Fig. 1. STATCOM's schematic representation

Based on the equivalent circuit shown in Fig. 01, the active and reactive power equations are

At node k

$$\begin{aligned} P_k = V_k^2 G_{kk} + V_k V_m (&G_{km}\cos(\theta_k - \theta_m) \\ &+ B_{km}\sin(\theta_k - \theta_m)) \\ &+ V_k V_{cR}(G_{km}\cos(\theta_k - \theta_{cR}) \\ &+ B_{km}\sin(\theta_k - \theta_{cR})) \\ &+ V_k V_{vR}(G_{vR}\cos(\theta_k - \theta_{vR}) \\ &+ B_{vR}\sin(\theta_k - \theta_{vR})) \end{aligned} \tag{11}$$

$$\begin{aligned} Q_k = -V_k^2 B_{kk} + V_k V_m (&G_{km}\sin(\theta_k - \theta_m) \\ &- B_{km}\cos(\theta_k - \theta_m)) \\ &+ V_k V_{cR}(G_{km}\sin(\theta_k - \theta_{cR}) \\ &- B_{km}\cos(\theta_k - \theta_{cR})) \\ &+ V_k V_{vR}(G_{vR}\sin(\theta_k - \theta_{vR}) - B_{vR}\cos(\theta_k - \theta_{vR})) \end{aligned} \tag{12}$$

There are diverse publications regarding to model the STATCOM, for example, steady state studies [7], or transient stability ones [8]. There are other ones applied to voltage control problem using novel technical [9]. The focus in the present work is the analysis of the STATCOM interconnected with the wind energy system. The proposed STATCOM at the wind energy generation is majorly used to mitigate the voltage flickers and to improve the power quality and below given are the main objectives. The proposed system is composed of three wind plants and interconnected turbine and generator

Fig. 2. Representation of single wind energy plant with capacitive bank

4 Simulation Models

The proposed model has been simulated using the Matlab/Simulink. For simulating the model we had considered the speed to the turbine is constant, variable, and incremental speed and the synchronous generator is rotating. The proposed simulated model with three wind mills and a STATCOM is shown in Fig. 03. Here all the three wind plants are interconnected and these are simulated. The implementation of the proposed model with STATCOM is shown in below fig 03. And the STATCOM model is shown in below fig 04

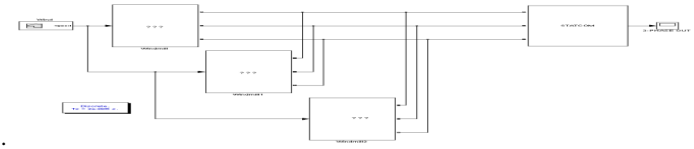


Fig. 3. Simulation model of Three wind turbines with STATCOM

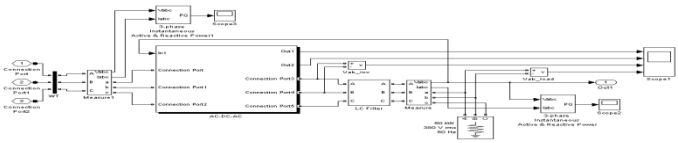


Fig. 4. Simulation model of STATCOM

5 Results and Discussion

The results are shown below for three phase output for the interconnected wind plant without and with STATCOM in the fig 05 and fig 06.On comparing these results the fluctuations occurring at the normal interconnected wind system are eradicate with implementation of STATCOM at the plant.The comparative graphical results of active and reactive powers of the two wind mills without STATCOM and with STATCOM is shown below in fig 07and Fig.08 on observing these results there are many variations in the powers when the wind plants are run without STATCOM. These variations are washed-out when the system is interconnected with the STATCOM. Majorly the reactive power compensation is decreased. Fig.09 shows the graphical results of the STATCOM the voltage profile has been maintained at 1.0pu and by implementing the LC filters the harmonics which are present after the STATCOM are reduced.

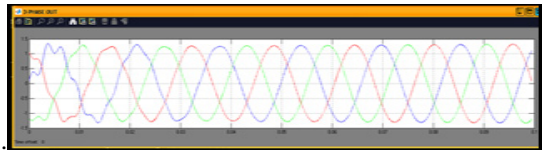


Fig. 5. Graphical results of Three-phase output of the Three wind plants without STATCOM

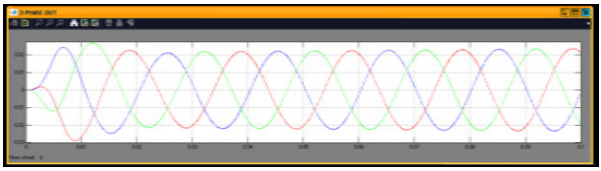


Fig. 6. Graphical results of Three-phase output of the Three wind plants with STATCOM

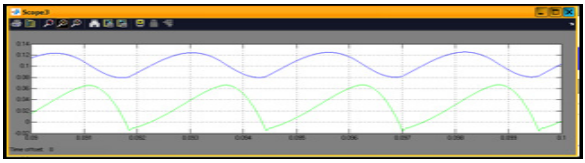


Fig. 7. Graphical results of Total Active and Reactive powers of the three wind plants without STATCOM

Fig. 8. Graphical results of Total Active and Reactive powers of the three wind plants with STATCOM

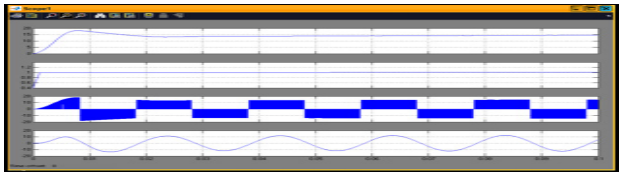


Fig. 9. Graphical results of STATCOM

6 Conclusion

This paper has exploited a detail influence of the voltage stability and the behaviors of wind plants during different wind speed conditions. The wind plants are run without the implementation of any peripheral devices in these circumstances there are variations in the system like voltage profile, active and reactive powers. These problems can be mitigated with the implementation of the STATCOM near to the plants. So the FACTS device at plants like STACOM can provides evidence to the power profiles which have been improved and the voltage fluctuations have been diminished.

References

1. Farias, M.F., Battaiotto, P.E., Cendoya, M.G.: Wind Farm to Weak-Grid ConnectionUsing UPQC Custom Power Device. IEEE (2010) 978-1-4244-5697-0/10
2. Wang, J., Ma, Y., Hu, Z., Yang, X.: Modeling and Real-Time Simulation of Non-Grid Connected Wind Energy Conversion System. IEEE (2009) 978-1-4244-4702-2/09
3. Ackermann, T.: Wind Power in Power Systems. John Wiley & Sons, Ltd., The Atrium, Southern Gate, Chichester (2005)
4. Mohod, S.W., Aware, M.V.: A STATCOM-Control Scheme for Grid Connected Wind Energy System for Power Quality Improvement. IEEE (2010), doi: 10.1109/JSYST, 2052943, 1932-8184
5. Wum, B., Lang, Y., Zargari, N., Kouro, S.: Power converstion and control of wind energy systems. IEEE Press, A John Wiley & Sons, Inc.
6. Devaraj, D., Jeevajyothi, R.: Impact of Wind Turbine Systems on Power System Voltage Stability. In: International Conference on Computer, Communication and Electrical Technology – ICCCET 2011, March 18-19 (2011)
7. Acha, E., Fuerte-Esquivel, C.R., Ambriz-Pérez, H., Ángeles-Camacho, C.: FACTS Modelling and Simulation in Power Network. John Wiley & Sons Ltd. (2004)
8. Cañizares, C.A.: Modelling of TCR and VSI Based FACTS Controllers. Internal Report for ENEL and POLIMI (September 9, 1999)
9. Wang, H.F., Li, H., Chen, H.: Application of Cell Immune Response Modelling to Power System Voltage Control by STATCOM. IEE Proc.-Gener. Transm. Distrib. 49(1) (January 2002)

PI Control of Quasi-resonant Buck Converter

A. Rameshkumar[1] and S. Arumugam[2]

[1] Government College of Technology, Coimbatore, India
arkr1966@yahoo.com
[2] Nandha Engineering College, Pichandapalayam, Erode, India
arumugamdote@yahoo.co.in

Abstract. A linear controller based on PI (Proportional-Integral) is designed and its application to the regulation of 54V, 2.916kW Quasi-Resonant Buck Converter is investigated. The algorithm of PI controller involves two main parameters namely Proportional gain (K_P) and Integral time (K_i) whose manual tuning provides an appropriate control action. Evaluation of output is initially performed by simulation in MATLAB 7.0 software at 200 kHz and the results are depicted for five different operating conditions where peak overshoot and settling time are mainly used to measure the dynamic performance of the converter. It is further implemented using a low cost, quadruple operational amplifier LM 324 and the experimental results are in good agreement with simulation and hence validate the control strategy.

Keywords: Generalised State Space Averaging (GSSA), Quasi-Resonant Buck Converter, PI Controller, Line and Load Regulation.

1 Introduction

Aerospace and industrial applications as of today demand stringently specified high power density converter and it is imperative to design such converter to operate at higher frequency to achieve high power density. Pulse Width Modulated Converter drew prime attention in the last two decades wherein the switch experiences hard switching at switching transitions and it leads to switching loss (P_s) as estimated from equation (1)

$$P_s = 0.5\, V_g I_o\, (t_{on} + t_{off}).f_s \tag{1}$$

where V_g is the input voltage, I_o is the load current, f_s is the switching frequency and t_{on} and t_{off} are the ON-period and OFF-period of converter. With Zero Current Switching (ZCS) property, the switch in Quasi-Resonant Converter works on the state of zero current during switching ON and OFF [1] to offer advantages such as self – commutation, low switching stress and loss, reduced electromagnetic interference and noise, and faster transient response to line and load variations [2]. In modeling and analysis of ac and dc behaviour of switched converter, State Space Averaging (SSA) is used because the different state matrix of each switching interval is replaced by an

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 477–485, 2013.

equivalent matrix which is the weighted average of individual matrix in each switching interval [3] and is mainly based on two assumptions namely (i) the switching frequency is much higher than the highest natural frequency of the converter in each switching mode of operation and (ii) the input to converter in each mode of operation must be either time- independent or slow time-varying variables compared with the switching frequency. In Quasi-Resonant Buck Converter, the natural frequency in the quasi-resonant tank is in the same order as the switching frequency and the state variables associated with the resonant tank are fast time-varying in each switching cycle of the converter compared with those in the low-pass filter formulate the State-Space Averaging method cannot be directly applied to its modeling. Therefore, Generalized State-Space Averaging [4] is needed for the modeling and analysis of Quasi-Resonant Buck Converter to establish more accurate results; here, a time average instead of weighted average is proposed and only the basic assumption (i) of State-Space Averaging method is used while the input variables can be fast time-varying. Losses and non-idealities adversely affect the steady-state output of Quasi-Resonant Buck Converter and in order to achieve desired regulation, control is indispensably required. A PI control algorithm is initially developed and simulation results are depicted for five operating conditions in which the two time-domain specifications are mainly used to measure the performance of Converter. It is verified that the linear PI controller optimized for a small-signal transient at an operating condition offers an excellent dynamic performance at the output however, for large-signal disturbances; the transient response at start-up is associated with overshoot and undershoot resulting in sluggish response. The algorithm is implemented using a low-cost, quadruple LM324 Operational Amplifier and the experimental results agree with simulation to an appreciable degree.

2 Generalized State-Space Averaging (GSSA) Methodology

The state equation of a periodically switched network with k different modes in each switching cycle is represented as

$$\dot{X}(t)=A_i x(t)+B_i(t),\ i=1,2,...k \tag{2}$$

The i^{th} equation of (2) is defined on time interval between the limits (t_{i-1}, t_i) where

$$t_{i-1}=\tau_0+\sum_{j=1}^{i-1}\tau_j \text{ and } t_i=t_{i-1}+\tau_i$$

2.1 Theorem

For the network described by equation (2), if the switching frequency is much higher than the highest natural frequency of the state matrix A_i in each switching mode of operation and if the input control variable function B_i is bounded [5], then, equation (2) is characterized by Generalized State-Space Averaging equation (3) as

$$x=\{\sum_{i=1}^{k} d_i A_i\}x+1/T\sum_{i=1}^{k}\int_{ti-1}^{ti} B_i(\lambda)d\lambda \quad (3)$$

where T is the switching period and is defined as $T=\sum_{j=1}^{k} \tau_j$

$f_s = 1/T$ and is defined as the switching frequency,

f_o is the natural frequency of matrix A_i.

This methodology is used to obtain small-signal model of the Quasi-Resonant Buck Converter and its line or input-to-output voltage transfer function is utilized in order to implement a conventional PI controller to the regulation and dynamic control of the output voltage of Quasi-Resonant Buck Converter.

3 Small Signal Model of Power Converter

A Quasi-Resonant Buck Converter in Fig.1 uses a unidirectional switch S [6] and is operated in discontinuous conduction mode and its input-to-output voltage transfer function [7] is obtained by analyzing the converter in four modes of operation using the Generalized State Space Averaging technique and is depicted in equation (4).

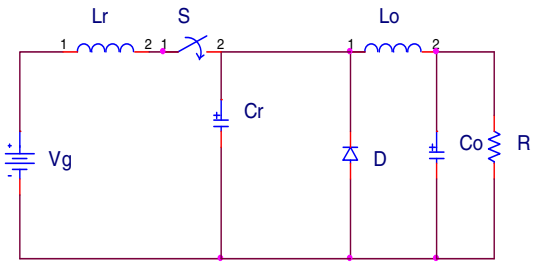


Fig. 1. ZCS – QR Buck Converter

$$\frac{\hat{v}_o}{\hat{v}_g} = \frac{M\left(1-\frac{J_i}{H_i}\right)}{s^2 L_o C_o + s\left(\frac{L_o}{R} - RC_o\frac{J_i}{H_i}\right) + 1 - \frac{J_i}{H_i}} \quad (4)$$

4 PI Controller

A PI Controller fuses the properties of P and I controllers and the algorithm provides a balance of complexity and capability in order to be widely used in process control applications. The equation (5) which describes P controller is

$$u(t) = K_P e(t) \quad (5)$$

where K_p is proportional gain, e (t) is the error and u(t) is the perturbation in output signal of PI controller from the base value corresponding to the normal operating conditions. To remove steady-state offset in controlled variable of a process, an extra intelligence is added to the P controller and this intelligence is t integral action. The controller is a PI controller whose mathematical notation is depicted in equation (6).

$$u(t) = K_p \left[e(t) + \frac{1}{K_i} \int_0^t e(t)\, dt \right] \tag{6}$$

where k_i is the integral or reset time. A PI controller has thus two tuning parameters K_p and K_i; the integral action [8] removes steady-sate offset in the controlled variable.

5 Design Data

It is imperative to design a feedback system such that the output voltage of DC-DC Converter is accurately regulated at desired level in presence of fluctuations either in input voltage or load [9]-[10]. Hence, to study the time-domain performance of PI controller to the regulation of Quasi-Resonant Buck Converter, it is designed with the specifications given in Table 1 with the two constraints [4] as mentioned hereunder.

Table 1. Design Parameters

No.	Parameter	Symbol	Value
1	Input Voltage	V_g	120 –100 V
2	Output Voltage	V_o	54 V
3	Resonant Inductor	L_r	1.6 μH
4	Resonant Capacitor	C_r	0.064 μF
5	Filter Inductor	L_o	0.2 mH
6	Filter Capacitor	C_o	20 μF
7	Load Resistance	R	100 –10 Ω
8	Switching Frequency	f_s	200 kHz
9	Natural Frequency	f_o	2.5165 kHz
10	Load Current	I_o	0.54-5.4 A
11	Output Power	P_O	2.916 kW

(i) The switching frequency (f_s) is much higher than the natural frequency of the low-pass filter (L_0-C_0) and hence the state variables associated with the filter state shall be regarded as constant in each switching cycle of operation and

(iii) The rating of filter components must be much higher than the rating of the resonating components.

6 Simulation of Quasi-resonant Buck Converter

A simulation is an attempt to model a hypothetical situation on a computer and becomes a usual part of modeling many systems in engineering to gain insight into the operation of system. It shortens the design because it is easier to study the influence of a parameter on the behaviour of a system rather than accomplishing it in the laboratory on a breadboard [2], [11]. MATLAB software is used for simulation and two categories of test namely the line and load regulation [12] are performed in order to evaluate the closed-loop performance of PI controller wherein line regulation is defined as the ability of a power-electronic converter to maintain constant voltage at output when there is a fluctuation in input and the load regulation of such power-electronic converter is to maintain the output voltage constant when there is a fluctuation in resistive load. Five operating conditions [13] spanning the entire range of the Converter is selected and the two vital time-domain specifications such as peak overshoot [14] and settling time [15]-[16] are measured. Schematic diagram is designed using the Simulink block of MATLAB software and is depicted in Fig. 2.

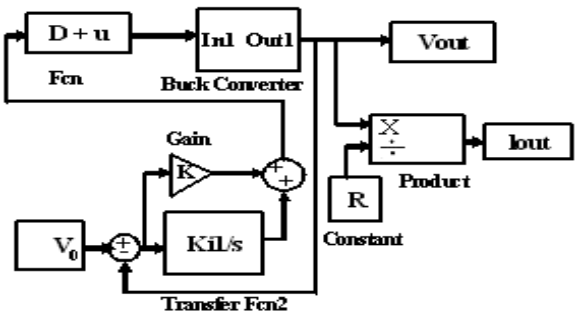


Fig. 2. Schematic diagram of a PI controller with Quasi-Resonant Buck Converter

The results of digital simulation for J_i/H_i parameter equal to -0.4 for two conditions of supply and load variations are shown in Fig.3 and Fig.4 in chronological order.

J_i/H_i parameter = -0.4

Case 5: Maximum line and Light load

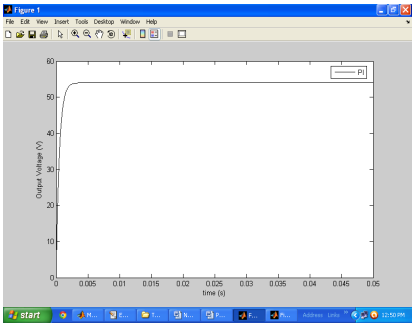


Fig. 3. Output Voltage for Case 5

Case 1: Minimum line and Maximum load

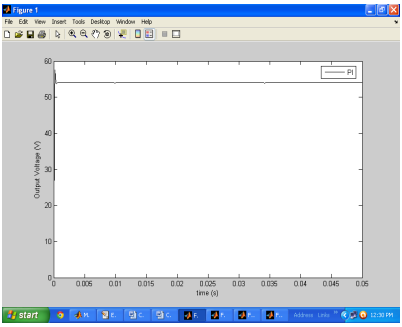


Fig. 4. Output Voltage for Case 1

Case 5: Maximum line (Vg=120V) and Light load (R=100Ω):

PI controller optimized for a small-signal transient performance at this operating point offers an excellent small-signal dynamic performance measured in terms of settling time (1.19949mS) and peak overshoot (54.005V).

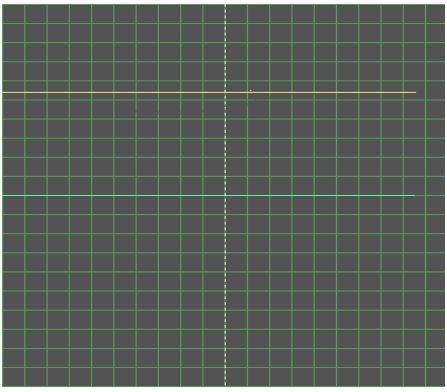
Case 1: Minimum line (Vg=100V) and Maximum load (R=10Ω):

However, for large-signal disturbances, such as for variation in input voltage from 120V to 100V and for variation in load resistance from 100Ω to 10Ω, the transient response of converter at start-up is associated with high overshoot (57.511V) and undershoot whereas settling time is reduced to 0.18278mS; the initial overshoot is due to the rapid building of current in the resonant inductor and the initial charging of the resonant capacitor. However, in PI controller the selection of gains is crucial and a gain which is ideal for one type of disturbance is not ideal for another type of disturbance and consequently the two gains must be selected in order to provide a satisfactory performance for both line and load disturbances.

7 Experimental Results

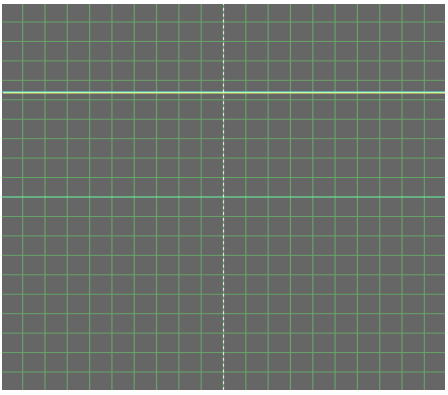
Experimental result is mandatory in-order to validate the control strategy and hence five test cases are performed in order to access the closed-loop performance of Quasi-Resonant Buck Converter. The switch S is replaced by Power MOSFET IRF940 whose main advantages are high commutation speed and good efficiency at low voltages. [17]-[18]. A PI controller is implemented using the National Semiconductor Corporation's LM324 [19]-[20] which has distinct advantages such as computability with all forms of logic and low power drain with an input common-mode voltage range from 0V to V ± 1.5V appropriate for battery operation.

Fig. 5. Snapshot of the Experimental system of PI controller



Vertical 10V/div, Horizontal 1ms/div

Fig. 6. Maximum line and Light load



Vertical 10V/div, Horizontal 1ms/div

Fig. 7. Minimum line and Maximum load

A snapshot of the experimental arrangement is depicted in Fig.5 and experimental results for J_i/H_i parameter = -0.4 is depicted hereunder in Fig.6 and Fig.7 and such results illustrate the effectiveness of the control strategy. Input and Output voltage is measured using Digital Multimeter and the Output waveform is clearly displayed in Scientific Instruments' 30MHz oscilloscope SM410.

8 Conclusion

Quasi-Resonant Buck Converter in its open-loop configuration is insensitive to variations in temperature or other factors such as changes in the input or resistive load and the output voltage of such converter cannot be maintained constant without either a linear or a non-linear controller. Among the linear control methodologies available

at present, a PI control algorithm was initially verified by simulation performed in time-domain in MATLAB environment and the results for line and load regulation for five different operating conditions spanning the entire range of Quasi- Resonant Buck Converter were depicted and the two vital time-domain specifications namely peak overshoot and settling time were primarily used to measure the dynamic performance. It was further implemented using an operational amplifier LM 324 and the experimental results were in good agreement with simulation and hence validate the control strategy. A control algorithm based on Artificial Intelligence may be propounded to overcome the inherent disadvantages of PI controller and enhance the performance of Quasi-Resonant Buck Converter.

References

1. Liu, K.H., Oruganti, R., Lee, F.C.Y.: Quasi-Resonant Converters-Topologies and Converters. IEEE Trans. Power Electron. 2(01), 62–71 (1987)
2. Mohan, N., Undeland, T.M., Robbins, W.P.: Power Electronics- Converters, Applications and Design, ch. 9. John Wiley and Sons, Inc., Singapore (2003)
3. Witulski, A.F., Erickson, R.W.: Extension of State-Space Averaging to Resonant Switches and Beyond. IEEE Trans. Power Electron. 5(1) (1990)
4. Xu, J., Lee, C.Q.: A Unified Averaging Technique for Modelling of Quasi-Resonant Converters. IEEE Trans. Power Electro. 13(03), 556–563 (1998)
5. Xu, J., Lee, C.Q.: Generalised State-Space Averaging Approach for a Class of Periodically Switched Networks. IEEE Trans. on Circuits and Systems—I: Fundamental Theory and Applications 44(11), 1046–1052 (1997)
6. Hart, D.W.: Introduction to Power Electronics, pp. 338–344. Prentice Hall International, Inc., New Jersey (1997)
7. Batarseh, I., Siri, K.: Generalised Approach to the Small Signal Modelling of DC-DC Resonant Converters. IEEE Trans. Aerospace and Electronic Systems 29(03), 894–909 (1993)
8. Gopal, M.: Digital control and State variable methods, pp. 133–137. Tata Mc Graw Hill Publishing Company Ltd., New Delhi (2009)
9. Erickson, R.W., Maksimovic, D.: Fundamentals of Power Electronics, 2nd edn., pp. 187–192. Springer(India) Pvt. Ltd., New Delhi (1999)
10. Lian, K.Y., Liou, J.J., Huang, C.-Y.: LMI Based Integral Fuzzy Control of DC-DC Converter. IEEE Trans. Fuzzy Syst. 14(01), 71–80 (2006)
11. Maksimovic, D., Stankovic, A.M., Thottuvelil, V.J., Verghese, G.C.: Modeling and Simulation of Power Electronic Converters. Proc. IEEE 89(6), 898–912 (2001)
12. Krein, P.T.: Elements of Power Electronics, pp. 93–97. Oxford University Press, Oxford (2003)
13. Viswanathan, K., Srinivasan, D., Oruganti, R.: A Universal Fuzzy controller for a Non-Linear Power electronic Converter. In: Proc. IEEE Int., Conf., Fuzzy Syst. (IEEE 2002), Honolulu, HI, USA, pp. 46–51 (2002)
14. Bandyopadhay, M.N.: Control Engineering: Theory and Practice, pp. 88–92. Prentice Hall of India, New Delhi (2003)
15. D'Azzo, J.J., Houpis, C.H.: Feedback Control System Analysis and Synthesis, pp. 80–84. Mc Graw Hill, Singapore (1984)

16. Gopal, M.: Control Systems-Principles and Design, pp. 362–365. Tata Mc Graw Hill Publishing Company Ltd., New Delhi (1997)
17. Bell, D.A.: Electronic devices and circuits, pp. 367–375. Oxford University Press, New Delhi (2010)
18. Singh, M.D., Khanchandani, K.B.: Power Electronics, pp. 159–174. Tata Mc Graw Hill Education Pvt. Ltd., New Delhi (2011)
19. LM124, LM124A, LM224, LM224A, LM324, LM324A, LM2902 Quadruple Operational Amplifiers datasheet, Texas Instruments, Dallas, Texas 75265
20. LM124, LM224, LM324, LM2902 Low power Quadruple Operational Amplifiers datasheet, National Semiconductor Corporation, California

Cache Oblivious B-tree and Sorting Using Sequential Accessing

Parmeshwar S. Korde[1] and Praskash B. Khanale[2]

[1] Department of Computer Science, Shri Shivaji College, Parbhani (M.S.) India
korde.parmeshwar@rediffmail.com, korde.parmeshwar@yahoo.com
[2] Department of Computer Science, D.S.M. College, Parbhani (M.S.) India
prakashkhanale@hotmail.com

Abstract. Cache oblivious Algorithm is designed to in inherently form hierarchy of caches. It need not know about exact structure. In this paper, we present a cache oblivious B-tree and Sorting using accessing technique. The algorithm uses sequential access elements that are based on linking elements. We show that these data structure are linearizable appeared to serialized order that save memory.

Keywords: Cache oblivious Algorithm, Sorting, B-tree, and Sequential accessing.

1 Introduction

Cache Oblivious B-tree is designed to achieve good data locality at one level of the memory hierarchy and for one fixed block size. It perform at all level of the memory hierarchy and block sizes [6]. It performs search operation with $O(\log_B N+1)$ cache misses for all possible block sizes simultaneously, when block sizes is unknown.

The CO B-tree perform near optimally in theory, and in recent experiments have shown outperforming traditional B-tree [9].

Recently, Rahman and Raman[9] made an empirical study of the performance version of search tree implementations, with focus on showing the significance of also minimizing translation look aside buffer (TLB) misses. Based on exponential search trees, they implemented Van Emde Boas layout supporting searches and updated in $O(\log_{B(}n)+\log\log N)$ memory transfers.

The concept of cache oblivious algorithm has been introduced by Frigo et. al[5]. In essence, this designates algorithm optimized in the I/O model, except that optimizes to a block size B and a memory size M which are unknown. This seemingly simple change has significant consequences: Since analysis holds any block and memory size, it holds all levels of the memory hierarchy.

Cache Oblivious B-tree [7] present dynamic search tree data structure that performs well in the setting of a hierarchical memory. But do not depend upon the number of memory levels, the block sizes and number of blocks at each level, or relative speed of memory access. Cache oblivious data structure supports insertion and deletion in $\theta(\log_B N+1)$ amortized memory transfers, which matches B-tree worst case bound.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 486–493, 2013.

In this paper, we investigate the probably most basic algorithm for B-tree. In contrast we want to demonstrate that it is possible to construct an algorithm where memory address changes only with link stepwise one, which also eliminate the need for address arithmetic. We will therefore concentrate on the basic idea for the algorithm and present some of the nice properties that result from this approach.

In cache oblivious setting the computational model is a machine with two levels of memory: a cache limited capacity and secondary memory of infinite capacity. The core assumption of the cache oblivious model is that M and B are unknown to the algorithm where as in the related I/O model introduced by [1] the algorithm know M and B and algorithm perform block transfers explicitly. Cache oblivious model and its relation to multilevel memory hierarchies are given in [2].

Many problems can be reduced to cache obvious sorting. In particular Arge et al[3] developed a cache oblivious priority queue based on a reduction to sorting. They furthermore showed how a cache obvious priority queue can be applied to solve a sequence of graph problems.

Bordal and Fagerberg in [4] showed how to modify cache obvious funnel sort algorithm to solve several problems with in computational geometry.

In this paper we investigate the probably most basic nontrivial algorithm, the sorting of array. We want to emphasize that to produce the sorting methods on specific computer architecture. We want to demonstrate that it is possible to construct an algorithm where memory address copies stepwise one, which also eliminate the need for transpositions. We will therefore concentrate on the basic idea of the algorithm and present some of nice properties that result from this approach.

1.1 Bubble Sort

A simple introductory example of the application of the incompressibility method is Bubble Sort. It is found in [5], it is well known that Bubble Sort uses $\theta(n^2)$ comparisons on average. Depending upon the Programming language used, the elements of array required to execute is proportional to n^2 where n is the number of input items.

In this paper, we will present an approach that uses an ordering of min elements array. However our sequential sorting which totally avoid the transposition process.

In Bubble Sort we make passes from left to right over the permutation to be sorted and always move the currently largest right by exchanges between it and right adjacent element – if that one is smaller. We make at most n-1 passes, since after moving all but one element in the correct place the single remaining element must be also in its correct place. The total number of exchanges is obviously at most n^2, so we only need to consider the lower bound. Let B be a Bubble Sort algorithm. For a permutation of elements 1, 2, -------- n. we can describe the total number of exchanges by

$$M = \sum_{i=1}^{n-1}$$

In sequential processing sorting we move minimum number from array and store into array 2. These processes repeat n times. In this method there are eliminating process of swapping and exchanging elements.

It was proved that [1] sorting requires $\theta\left(\frac{N}{B}\log_{M/B}\frac{N}{B}\right)$ block transfers and permuting an array requires $\theta\left(\min\left\{\frac{N}{B}\log_{M/B}\frac{N}{B}\right\}\right)$ block transfers. Lower bounds hold for the cache-oblivious model. The lower bounds from[1] immediately give a lower bound of $\Omega\left(\frac{N}{B}\log_{M/B}\frac{N}{B}\right)$ block transfers for cache-oblivious sorting. The upper bound from [1] cannot be applied to the cache-oblivious setting since these algorithm make explicit use of B and M.

In sequential processing we transfer the value of min to array 2 while processing and reformulate the algorithm as algorithm 1.

```
Algorithm 1
    1. Repeat steps 2,3,4 for i = 1 to n
    2. Repeat for j = 1 to n
       a) If ( a[ j] > = min ) and ( a[ j ] < > 0 ) then
          i)      Set min =a [ j ]
          ii)     Set loc = j
                  [ end of if ]
                  [ end of j loop ]
    3. Set b[ i ] = min
    4. Set min = max value
       [ End of i loop ]
    5. Exit
```

In this process we formulate the Bubble sort method; here we used two arrays as array 1 and array 2. In first array we find the location of min value and place as null value and copy value in array 2. The process may continue up to last elements. Every phase we must store min value as max value. It is compulsory for every external loop. We make at most n passes, since after moving all but one element in the correct place the single remaining element must be also in its correct place. The total number of exchanges is obviously at most n^2, so we only need to consider the lower bound. By using same technique we may design algorithm for Cache oblivious Sequential Sorting for descending.

We want to emphasize that as in fig 3 that is not the aim of paper to produce the fastest algorithm for sorting on specific computer architecture. We want to demonstrate that it is possible to construct an algorithm where memory address copy only with stepwise one and it present nice properties that result from this approach.

1.2 Cache Oblivious B-tree

The four memory layout for trees: DFS, in-order, BFS and Van Emde Boas layout. A complete tree of size n is stored in an array of n node records.

1.3 Memory layout of Trees

DFS Layout: The node of T are stored in the order they are visited by a left to right depth first traversal of T(that is a pre order traversal)

In-Order Layout: The nodes of T are stored in the order that they are visited by left to right in- order of traversal of T.

BFS Layout: The nodes of T are stored in order they are visited by a left to right breath first traversal of T.

Van Emde Boas Layout: The layout is defined recursively; A tree with only one node is a single node record. A Tree with only one node is a single node record. The Van Emde Boas layout of T consist of the Van Emde Boas layout of T_0 followed by the Van Emde Boas layout of T_1, T_2------ T_k.

A compete tree with height four fig 1 gives DFS, In-order, BFS and Van Emde Boas layout.

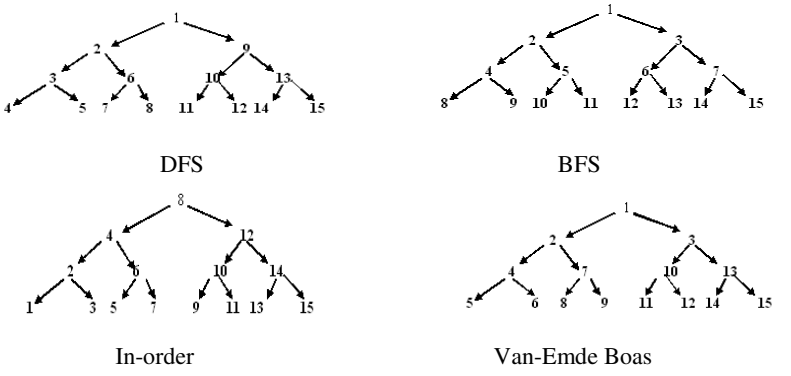


Fig. 1. The DFS, In-order, BFS and Emde Boas layout for a complete tree with height fours

2 Sequential Accessing B-tree

A binary tree T is in memory. The algorithm does a DFS (preorder) traversal of T, applying an operation PROCESS to each of its nodes. An array stack is used to temporarily hold the addresses of nodes as in algorithm 2.

```
Algorithm 2
    1. Set Top=1, Stack [1] = Null and PTR= ROOT
    2. Repeat steps 3 to 5 while PTR< > NULL
    3. Apply PROCESS to INFO[PTR]
    4. [Right child ?]
       If RIGHT[PTR] < > NULL then [ Push on STACK ]
       Set TOP = TOP+1, and STACK[TOP] = RIGHT[PTR]
       [End of if structure]
    5. [Left child]
       If LEFT[PTR] < > NULL then
        Set PTR= LEFT[PTR]
       Else : [Pop from STACK]
        Set PTR=STACK[STACK[TOP] and TOP=TOP-1
       [ End of If structure]
       [ End of 2 loop]
    6. Exit.
```

The worst case number of memory transfers during a top down of a path using the above layout schemes assuming each block stress B nodes. With the BFS layout, the topmost [log(B+1)] levels of the tree will be contained in at most two blocks whereas each of the following blocks read only contains one node form the path. The total number of memory transfers is therefore θ (log (n /B). For DFS and In-order layout we get worst case bound when path to rightmost leaf, since first [log (n+1)] - [log B] node have distance at least B in memory, where as the last log(B+1)] nodes stored in at most tow blocks. Prokop[5] observed, in the van Emde Boas layout there are at most ($\log_B N$) memory transfers. Only van Emde Boas layout has asymptotically optimal bound achieved by B-tree [7].

In Sequential Accessing Process we change the sequence of storing binary trees in memory.

2.1 Linked Representation of Binary Tree Using Sequential Accessing

A binary tree T in memory by using three parallel arrays INFO, LEFT, RIGHT and a pointer variable ROOT as follows. Each node N of T will correspond to a location K such that as in fig 2, fig 3.

1) INFO [K] contains the data at the node N.
2) LEFT [K] contains the location of the left child of node N.
3) RIGHT[K] contains the location of the right child of node N

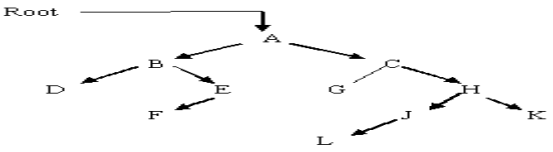
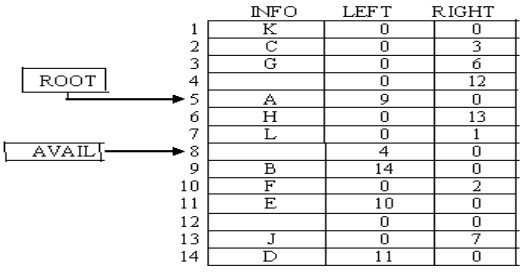


Fig 2. A Binary Tree T



	INFO	LEFT	RIGHT
1	K	0	0
2	C	0	3
3	G	0	6
4		0	12
5	A	9	0
6	H	0	13
7	L	0	1
8		4	0
9	B	14	0
10	F	0	2
11	E	10	0
12		0	0
13	J	0	7
14	D	11	0

Fig 3. Memory Representation of Sequential Processing

In this process we divide binary tree T as s Left-sub-tree and Right sub-tree which linked with each other for sequential processing. The four memory layout DFS, in-order, BFS, van Emde Boas when processed the algorithm 2 will be convert as algorithm 3.

CO B-tree executes in linking order elements as left-sub tree and right sub-tree. In sequential access method it process linking order of data structure.

We have perform two steps

1) It executes in sequential order.
2) The same approach are used for all four layouts

```
Algorithm 3 ( DFS Preorder)
    1. Set PTR=ROOT
    2. While ( PTR < > Null )
        a) Apply Process of INFO[PTR]
        b) Set SAVE=PTR
        c) Set PTR=LEFT[PTR]
        [ End of loop]
    3. Set PTR=RIGHT[SAVE]
    4. While ( PTR < > Null )
        d) Apply Process of INFO[PTR]
        e) Set PTR=RIGHT[PTR]
        [ End of loop]
    5. EXIT
```

3 Implementation

Algorithm 1 is simple scheme we developed in first sections. The algorithm takes n phases as min1, min 2 ----- min n to indicate the sorting process. In second section algorithm 3 is a Cache oblivious B-tree which uses two phase of loop it repeat until PTR goes to null. In programming language like C or C++, Java directly used these schemes.

We tested algorithm 1 with different array size and their execution speed shown in table 1 and Four layout schemes with Sequential B-tree in table 2.

Table 1. Cache Oblivious Sorting time slots

Sr.no.	Array Size	Bubble Sort	Sequential Sort
01	10	0.164835	0.164835
02	100	0.274725	0.274725
03	500	0.659341	0.659341

Table 2. Cache Oblivious B-tree time slots

Sr.no.	Layouts	Length	Time
01	DFS	10	0.10989
		25	0.164835
02	In-order	10	0.10989
		25	0.164835
03	BFS	10	0.10989
		25	0.164835
04	Van Emde Boas layout	10	0.10989
		25	0.164835
05	Sequential Access	10	0.10989
		25	0.164835

The comparative ratio of Bubble sort and Sequential sort as well as Four layout and Cache Oblivious B tree is in graphical format as shown in fig 4, fig 5 respectively.

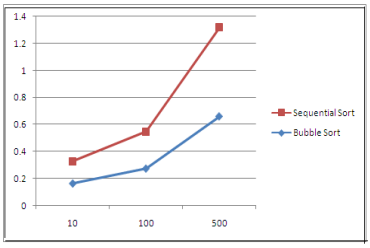


Fig. 4. Cache ratio for Bubble sort Oblivious and Sequential Sort

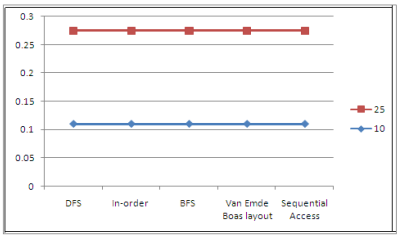


Fig. 5. Cache ratio for Cache B-tree

4 Cache Oblivious Performance

In sorting algorithm mainly consists of two parts: we assume the input Sequence contains n elements and each equal-sized subsequence has w elements. Then, the average time complexity is O (n (log w) 2). Because w is a constant in our implementation, the time complexity is O (n). As for the sort in the following steps, it contains log (n / w) levels. In each level, it consists of t elements. Thus, every element needs O((log t) 2) compare operations in each level, and the time complexity of sort is $O(n \log(n / w) \bullet (\log t)2) = O(n \log n)$. As a result, the time complexity of our sort is $O(n) + O(n \log n) = O(n \log n)$. However, the sort does not work in-place: we use two buffers and switch each other in different levels of the tree. In addition, we need a little more auxiliary space for the patching of misaligned address. Thus the space requirement of our algorithm is $2n+c$. The constant c depends on the number of splitters used.

The combined number of tree nodes in threes sub trees is less than 2k, it remains to be shown that these nodes fit in O (k/B+1) blocks. Consider all sub tree blocks {T_1……..T_2} to update from left to right. Then there exist an i such that the sub trees rooted at parent (T_i) contains the sub trees T_1……. T_i and parent (T_{i+1}) contains the sub trees T_{i+1}……. T_r Thus the tree nodes to update are contained in tow sub trees, with total size at most 4k. All of these sub trees, therefore, are laid out in O (k/B+1) blocks.

Additionally, some nodes, not in sub trees, along with the path to the root must be updated. We note that we only update nodes with a right child that changes. Thus we update only a single path to the root, which requires at most O ($\log_B$ N+1) blocks.

5 Conclusion

The paper focuses on correctness issues more than performance issues in section first we have presented Sequential processing for sorting algorithm that has better locality features. Using ideal cache model cache misses is of order of θ ($\frac{N}{B}$ log $_{M/B}$ $\frac{N}{B}$) block transfers and permuting an array requires θ (min { $\frac{N}{B}$ log $_{M/B}$ $\frac{N}{B}$}) block transfers. This is asymptotically optimal for any algorithm that is based on sorting. It totally avoids the need for swapping for address arithmetic. While this fact is not fully exploited on standard computers, it may be considerable advantage for hardware implementations of sorting techniques.

In second section, the basic idea of data structure is to maintain binary tree of height log $_n$ +$\mathcal{O}$ (1) using existing methods. We also investigate the practicality of cache obliviousness in the area of trees by providing comparison of different methods for laying out a tree in memory. One further observation is that the effects from the space saving and in size caused by implicit layout are notable.

References

1. Aggarwal, A., Vitter, J.S.: The input/output complexity of sorting and related problems. Communications of the ACM 31(9), 1116–1127 (1988)
2. Demaine, E.D.: Cache-oblivious algorithms and data structures. In: Lecture Notes from the EEF Summer School on Massive Data Sets. LNCS, BRICS, University of Aarhus, Denmark (2002)
3. Brodal, G.S., Fagerberg, R.: Cache Oblivious Distribution Sweeping. In: Widmayer, P., Triguero, F., Morales, R., Hennessy, M., Eidenbenz, S., Conejo, R. (eds.) ICALP 2002. LNCS, vol. 2380, pp. 426–438. Springer, Heidelberg (2002)
4. Arge, L., Bender, M.A., Demaine, E.D., Holland-Minkley, B., Munro, J.I.: Cache-oblivious priority queue and graph algorithm applications. In: ACM (ed.) Proceedings of the 34th Annual ACM Symposium on Theory of Computing (STOC 2002), pp. 268–276. ACM Press (2002)
5. Frigo, M., Leiserson, C.E., Prokop, H., Ramachandran, S.: Cache-oblivious algorithms. In: 40th Annual Symposium on Foundations of Computer Science, pp. 285–297. IEEE Computer Society Press (1999)
6. Bender, M.A., Brodal, G.S., Fagerberg, R., Ge, D., He, S., Hu, H., Iacono, J., López-Ortiz, A.: The cost of cache-oblivious searching. In: FOCS 2003, pp. 271–282 (2005)
7. Bender, M.A., Demaine, E., Farach-Colton, M.: Cache-oblivious B-trees. In: Proc. 41st Ann. Symp. on Foundations of Computer Science, pp. 399–409. IEEE Computer Society Press (2000)
8. Rahman, N., Cole, R., Raman, R.: Optimised Predecessor Data Structures for Internal Memory. In: Brodal, G.S., Frigioni, D., Marchetti-Spaccamela, A. (eds.) WAE 2001. LNCS, vol. 2141, pp. 67–78. Springer, Heidelberg (2001)
9. Kasheff, Z.: Cache-oblivious dynamic search trees. MS thesis, Massachusetts Institute of Technology, Cambridge, MA (2004)

Locality Based Static Single Assignment Partial Redundancy Elimination

Sivananaintha Perumal and N. Aneesha

Dept of CSE, MES College of Engineering,
Kuttippuram, Kerala, India

Abstract. The Partial Redundancy Elimination (PRE) is an important redundancy elimination transformation technique. Locality based Static Single Assignment Partial Redundancy Elimination method classifies expressions into three categories: Local, Confined and Other. In order to reduce optimization time, LSSAPRE removes unnecessary Φ's during the Φ -Insertion phase itself, which is the first step of SSAPRE. The unnecessary Φ – assignments for the local and confined expressions can be easily identified and removed in Φ – Insertion step. Even though this LSSAPRE system has reasonably good performance, it is not considering the other category expression based data flow problems in SSA form. In this paper, tried to propose a new Locality based SSAPRE method to remove unnecessary Φ assignments for the other category expressions during Φ - Insertion step itself and thereby reducing the analysis time and optimization time required for SSAPRE algorithm. But it theoretically proved that only based on the concept of locality of expressions and the dominance/post dominance relationship, it is not possible to remove the unnecessary Φ assignments for other category expressions during Φ - Insertion step and it is not possible to reduce the analysis time required for the SSAPRE algorithm.

Keywords: Partial Redundancy Elimination, Static Single Assignment form, Locality of expression, Dominance relationship, Φ-assignment.

1 Introduction

Partial redundancy elimination (PRE) is an important redundancy elimination transformation first developed by Morel E and Renvoise C [1] which eliminates the partially redundant expressions in a program. Dhamdhere D.M [2] did a modification to the algorithm proposed by Morel E and Renvoise C [1] by suppression of partial redundancies. Knoop J et al. [3] presented an optimal bit vector algorithm called Lazy Code Motion (LCM) for the elimination of partial redundancies. This algorithm decomposed the bidirectional structure of the PRE into a sequence of unidirectional analyses.

All these PRE systems have modeled the problem as systems of data flow equations. Regardless of how efficiently the systems of data flow equations can be solved, a large amount of time needs to be spent in scanning the contents of each

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 494–497, 2013.

basic block in the program to initialize the local data flow attributes that serves as input to the data flow equations. So another approach to PRE is needed which does not require the dense initialization of the data flow information. Robert Kennedy et al. [4] presented a new algorithm Static Single Assignment Partial Redundancy Elimination (SSAPRE) for performing PRE based on SSA form.

Jongsoo Park and Jaejin Lee [5] presented a locality based SSAPRE algorithm which gives a practical improvement to the analysis time of the original SSAPRE algorithm proposed by Robert Kennedy et al. [4]. It is based on locality of expressions and dominance relationship. Locality based SSAPRE identifies and filtered out most of the unnecessary Φ- assignments during the Φ – insertion step itself.

The rest of this paper is organized as follows. Section 2 describes the working of original Locality based SSAPRE. Section 3 discusses whether the other category expressions are necessary to perform SSAPRE and whether it is possible to reduce the analysis time required for SSAPRE algorithm or not. Section IV concludes the paper.

2 Working of Locality Based SSAPRE

Locality based SSAPRE algorithm proposed by Jongsoo Park and Jaejin Lee [5] provides a practical improvement in the analysis time of the original SSAPRE algorithm. The analysis time of SSAPRE is determined by the number of Φ – assignments inserted in the SSA form. Locality based SSAPRE identifies and removes most of the unnecessary Φ assignments during the Φ – insertion step itself and thus it reduces the analysis time in the following stages. Locality based SSAPRE identifies that Φ assignments for confined expressions are unnecessary to perform SSAPRE. Φ-assignment for local expression is necessary only if the expression post dominates the Φ –assignment. In order to reduce analysis time and optimization time, LSSAPRE removes unnecessary Φ's during the Φ-Insertion phase itself, which is the first step of SSAPRE. The remaining expressions are treated in the same way as the original SSAPRE algorithm.

This locality based SSAPRE algorithm is not considering the unnecessary Φ assignments for the other category expressions in the Φ - Insertion step. Even though this LSSAPRE system has reasonably good performance, it still contains unnecessary Φ 's. If it is possible to remove the unnecessary Φ 's for other category expressions in the Φ - Insertion step, then the proposed locality based SSAPRE algorithm will become more faster than the current locality based SSAPRE algorithm.

3 Proposed Work

In the Locality based SSAPRE, based on only locality and dominance/post dominance relationship, SSAPRE considers only the local and confined expressions. We tried to find Φ –assignments for other expressions are necessary or not and whether unnecessary Φ –assignments for other expressions can be eliminated during Φ – insertion step itself or not. Theoretically it is proved that only based on the locality of expressions and dominance/post dominance relationship it is not possible to conclude that Φ –assignments for other expressions are unnecessary to perform SSAPRE and hence not possible to remove all of them during Φ –insertion step itself.

Lemma 1 [5]. If v ∈ DF+ (u) then u does not strictly dominates v. i.e. ¬ (u sdom v) means ¬ (u dom v) or u=v.

Lemma 2. If E is an other expression at some basic block set U, ¬ (u dom v) does not implies that E is fully available at any point in v.

Proof of Lemma 2. By the definition of dominance relation, ¬ (u dom v) means there is another path p from entry to v that does not goes through u. Since E is an other expression, sometimes P contains an occurrence of E. Also it is possible to occur a definition to any of the operands of E on any node on the path p between the last occurrence of E and v. If such a definition is there, E is not fully available at v. E is not occurred on the path p, it is not fully available at v. The locality of expression and dominance relationship is not giving any information about the occurrence of E and the definition of E's operands. Hence ¬(u dom v) does not imply that E is fully available at any point in v.

Lemma 3. If E is an other expression at some basic block set U, ¬(u pdom v) does not implies that E is fully anticipable at any point in v.

Proof of Lemma 3. By the definition of post dominance relation, ¬ (u pdom v) means there is another path p from v to program exit that does not goes through u. Since E is an other expression, sometimes P contains an occurrence of E. Also it is possible to occur a definition to any of the operands of Eon any node on the path p between the first occurrence of E and v. If such a definition is there, E is not fully anticipable at v. E is not occurred on the path p, it is not fully anticipable at v. The locality of expression and dominance relationship is not giving any information about the occurrence of E and the definition of E's operands. Hence ¬ (u pdom v) does not imply that E is fully anticipable at any point in v.

Theorem 1. If E is an other expression at some basic block set U, and ¬(u pdom v) such that v ∈ DF+(u) then based on locality and post dominance relationship it is not possible to conclude that Φ –assignments for other expressions are unnecessary to perform SSAPRE and hence not possible to remove all of them during the Φ – insertion step itself.

Proof of Theorem 1. Assume that Φ –insertion at v is necessary. i.e either fully available or fully anticipable. By Lemma 1 v ∈ DF+ (u) means ¬ (u dom v) or u = v. By Lemma 2, ¬(u dom v) not implies that E is fully available at v. Otherwise u=v, there exist at least one path p from entry to v. Since E is other expression it is possible to include a definition to any of the operands of E at any node on the path p. If such a definition occurred between the last occurrence of E and v, E is not fully available at v. But dominance relation and locality of expressions are not giving any information about the definition of operands. So it is not fully available at v. By Lemma 3, it is not possible to conclude that E is fully anticipable at v.

That is only based on the locality of expressions and dominance/post dominance relationships, it is not possible to conclude that Φ – insertion at v is necessary and hence not possible to remove the unnecessary Φ assignments for other expressions during Φ-insertion step, first step of SSAPRE.

The number of Φ assignments inserted in the SSA form is the main factor that determines the analysis time of SSAPRE steps. If it is possible to remove unnecessary Φ assignments during Φ - Insertion step itself, then analysis time required for the SSAPRE can be reduced. The new Locality based SSAPRE method theoretically proved that only based on the locality of expressions and dominance/post dominance relationships, unnecessary Φ assignments for other category expressions are not possible to remove during Φ -Insertion step itself. Hence in the new Locality based SSAPRE method, it is not possible to reduce the analysis time required for the SSAPRE method.

4 Conclusion

The Locality based SSAPRE algorithm performs the partial redundancy elimination by taking the advantages of SSA forms. In this work, Φ assignments for other category expressions are considered: whether they are necessary to perform SSAPRE and whether it is possible to remove unnecessary Φ assignments during Φ Insertion step itself. It theoretically proved that only based on the concepts of locality of expressions and the dominance/post dominance relationship, it is not possible to remove the unnecessary Φ assignments for other category expressions during Φ Insertion step itself. Hence in the new Locality based SSAPRE method, it is not possible to reduce the analysis time required for the SSAPRE method.

References

1. Morel, E., Renvoise, C.: Global Optimization by Suppression of Partial Redundancies. Communication of the ACM 22(2), 96–103 (1978)
2. Dhamdhere, D.M.: Practical adaptation of the global optimization algorithm of Morel and Renvoise. ACM Transactions on Programming Languages and Systems 13(2), 291–294 (1991)
3. Knoop, J., Ruthing, O., Steffen, B.: Lazy code motion. In: Proceedings of the ACM SIGPLAN 1992 Conference on Programming Language Design and Implementation (ACM SIGPLAN), vol. 27(7), pp. 224–234 (1992)
4. Kennedy, R., Chow, F., Chan, S., Liu, S.-M., Lo, R., Tu, P.: A new Algorithm for Partial Redundancy Elimination based on SSA Form. In: Proceedings of ACM SIGPLAN 1997 Conference on PLDI, pp. 273–286 (1997)
5. Park, J., Lee, J.: Practical Improvement to the Partial Redundancy Elimination in SSA Form. Journal of Computing Science and Engineering 2(3), 301–320 (2008)

The FIRMER-Square Contour Algorithm

Rangeet Mitra[1], Sumit Kumar[2], and Santosh Kumar[2]

[1] Ozyegin Universitesi
Istanbul, Turkey
[2] Department of Computer Science and Engineering,
Indian Institute of Technology Guwahati,
Guwahati – 781039, India
{mitra.rangeet,sumit.itech,mrsonuk}@gmail.com

Abstract. The recently proposed Square Contour Algorithm (SCA) provides a good solution to the blind equalization problem as it does not suffer from any type of phase rotation. However, its suitability to Code Division Multiple Access (CDMA) system has been explored in this paper, and it is found that its convergence speed has reduced in multiuser scenarios. An algorithm for mitigating this effect is proposed, and simulations reinforce the suitability of this proposal to multi-user environments.

Keywords: FIRMER, Square Contour Algorithm, MAI, FIRMER-Square Contour Algorithm, Digital signal processing.

1 Introduction

Equalization is basically a signal processing operation that is the inverse of the corruption introduced into the symbol by the wired/wireless channel. Blind equalizer is a class of equalizer which don't require training and hence remove this training overhead and conserve bandwidth. It basically restores some observable and positive definite property [1] of the transmitted signal in order to achieve equalization.

One of the earliest contributions in this field was the Constant Modulus Algorithm (CMA) [1] which suffered from an arbitrary phase rotation problem. The Multi-Modulus-Algorithm (MMA) [2] solved this problem to a certain extent, but still had a tendency to converge to diagonal solutions. The recently pro -posed Square Contour Algorithm (SCA) [3] fully solved this problem of phase rotation, but showed a slow convergence characteristic. Many variants of SCA [4] have been proposed to fasten its convergence.

These schemes basically try to remove the Inter-Symbol-Interference (ISI).But there exist some schemes that remove the ISI and Multiple Access Interference (MAI) simultaneously in a CDMA multi-user environment. These include FIRMER-CMA [5, 6] and FIRMER-Modified-CMA (MCMA) [7] based schemes that use the CMA and MCMA as their "*parent* "algorithms. However, when MCMA/MMA based schemes are used in an equalizer structure in [2] in a multi-user environment as in [5] we may get a disadvantage of phase rotation. Even if we use a Decision-Directed

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 498–501, 2013.

(DD) based scheme, we are at a disadvantage as such rotations would increase the time for safe DD to take place. Hence, we should go for a solution that does not provide any phase rotation like the SCA.

The main aim of this paper is to explore the suitability of the SCA to CDMA multi-user environment. We simulate the SCA in a multi-user environment, and we find that it suffers from a slow convergence characteristic. To improve its convergence and remove MAI, a variant of SCA is proposed that is found to give better results in a multi-user scenario. The SCA in a multi-user scenario has been unexplored ever since its proposal in [3].

Our terminology throughout the paper will be as follows: y would denote the equalizer output. y_R would denote the real part of the output and y_I would denote the imaginary part of the output. $\boldsymbol{w}(n)$ would denote the equalizer weights at time instant n. ∂ would denote the gradient operator. In this paper we will analyze a class of stochastic gradient algorithms. $\mathbf{H}$ would denote the convolutional matrix of the chip dispreading sequence at the receiver for a user.

2 Overview of SCA

The cost function of SCA is given by [3]

$$J_{SCA} = (max(|y_R|,|y_I|)^2 - R_{SCA})^2 \tag{1}$$

$$= ((|y_R + y_I| + |y_R - y_I|)^2 - R_{SCA})^2$$

Where, n can be choosing as 2 that is called the Generalized–SCA(GSCA). The expression for *RSCA* has also been given in [3], which is nothing but the modulus for the SCA. Please note that a constant of proportionality has been absorbed into R_{SCA} in the latter expression.

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) - \mu \partial_{\boldsymbol{w}} J_{SCA} \tag{2}$$

Where $\boldsymbol{w}$, is the weight vector for the linear equalizer. μ is the step size. Due to the interdependency introduced between the real and imaginary weights in the equalizer structure in [2], we get no diagonal solutions.

3 Review of FIRMER – CMA

To suppress MAI and improve the convergence in a multiuser environment, FIRMER-CMA was proposed [5] that optimizes a cost function,

$$J_{FCMA} = (|y''|^2 - R_{FCMA}^2) \tag{3}$$

where,

$$y'' = \boldsymbol{w}^T \boldsymbol{H} \boldsymbol{x}$$

wis a chip-spaced equalizer, and ***H*** is a convolutional matrix consisting of a particular user's code and **x** is the received block of data. Such a code- filtering helps to remove MAI in a multi-user CDMA environment.

4 Proposed Scheme

The SCA was simulated in a CDMA-multiuser environment and due to MAI; the results were not very impressive as we describe in future sections. To improve up on this, we propose a variant of the SCA named FIRMER-SCA for CDMA multiuser environment. Our main purpose is to optimize a cost function,

$$J_{SCAF} = (max(|y_R''|,|y_I''|)^2 - R_{SCA})^2$$
$$= ((|y_R''+y_I''|+|y_R''- y_I''|)^2 - R_{SCA})^2 \quad (5)$$

where y_R'' and y_I'' are real and imaginary parts of the equalizer output y following the same terminology as in previous section. For each user, the chip-spaced equalizer weights are adapted by,

$$\boldsymbol{w}(n+1)=\boldsymbol{w}(n)-\mu\,\partial_{\boldsymbol{w}}J_{SCAF} \quad (6)$$

Where

$$\partial_{\boldsymbol{w}}\, J_{SCAF} = ((|y_R''+y_I''|+|y_R''-y_I''|)^2 - R_{SCA})(sign(y_R''+y_I'')(1+j)+sign(y_R'-y_I'')(1-j))\mathbf{Hx} \quad (7)$$

5 Results and Discussion

We considered a channel having an impulse response $\left[0.5exp\left(-j\frac{3\pi}{4}\right)\right]\ 1\ \left[0.5exp\left(-j\frac{3\pi}{4}\right)\right]$ and considered two scenarios, one with low MAI and another with a higher MAI and considered 4 users for our case.

At high MAI the interference power is three times the signal power for a particular user. We plot the ISI metric [8] curves for a particular user (in Fig1 (a)) in this scenario and find that the FIRMER-SCA converges faster.

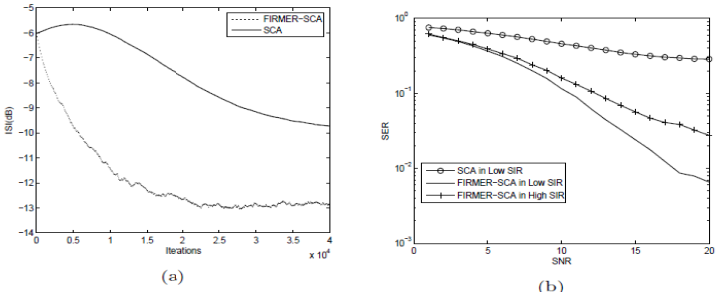


Fig. 1. (a) ISI Comparison, (b) SER Comparison between SCA and FIRMER-SCA

We consider low MAI where interference power is 0.1 times the signal power. We show Symbol Error Rate (SER) results in Fig 1(b). We find that the SER for FIRMER-SCA is lower than SER of SCA at lower Signal-to-Interference Ratio (SIR). Also, even at higher SIR the FIRMER-SCA outperforms the SCA (which has been simulated at lower SIR). This confirms the better suitability of this algorithm in a multi-user scenario.

6 Conclusion

A variant of SCA which is better suited to multiuser scenarios is introduced and validated through simulations, which may find use in equalizer structures in high speed communication systems where CDMA is used as the multiple access technique. Our present work analyzed the SCA using stochastic gradient algorithms and proposed a new variant of SCA for the multiuser scenario. As far as we know, this is the first extension of SCA to multiuser cases. Our future goal aims at finding an analog of the least squares for the SCA and using it for multiuser scenarios.

References

1. Treichler, J., Agee, B.: A new approach to multipath correction of constant modulus signals. IEEE Trans. on Acoust., Speech, and Signal Proc. 31(2), 459–471 (1983)
2. Yang, J., Werner, J.J., Dumont, G.A.: The multimodulus blind equalization and its generalized algorithms. IEEE Journal on Selected Areas in Communications 20(5), 997–1015 (2002)
3. Thaiupathump, T., Kassam, S.A.: Square Contour Algorithm: A New Algorithm for Blind Equalization and Carrier Phase Recovery. In: Proc. of the 37th Asil. Conf. Sig. Sys. Comp., pp. 647–651 (2003)
4. Sheikh, S.A., Fan, P.: New Blind Equalization techniques based on improved square contour algorithm. Signal Proc. 18, 680–693 (2008)
5. Hadef, M., et al.: Blind Chip Rate Equalization for DS-CDMA Downlink Receiver. In: Asil. Conf. Sig. Sys. Comp., vol. 2, pp. 1283–1287 (2003)
6. Hadef, M., Weiss, S.: Concurrent constant modulus algorithm and decision directed scheme for synchronous DS-CDMA equalization. In: Proc. IEEE 13th Workshopon Stat. Signal Processing, Bordeaux, France, pp. 203–208 (July 2005)
7. Mitra, A.: Modified CMA based blind equalizer with decision directed scheme. IEE Electronics Letters 44(6) (March 2008)
8. Shalvi, O., Weinstein, E.: New criteria for blind DE convolution of no minimum phase systems (channels). IEEE Trans. Inform. Theory 36(2) (March 1990)

Performance Evaluation of DSDV, DSR, OLSR, TORA Routing Protocols – A Review

Davinder Singh Sandhu and Sukesha Sharma

University Institute of Engineering and Technology,
Panjab University, Chandigarh, India
davinder2572@gmail.com,
er_sukesha@yahoo.com

Abstract. Mobile ad Hoc networks (MANETs) can make dynamic changes in topology. The most realistic characteristic which differentiates MANETs from other networks is that it is capable of changing its location. Different routing algorithms are used while maximizing packet delivery ratio (PDR), throughput and minimizing end to end delay, routing load to improve the performance. Maximum PDR and throughput gives best performance. This review paper describes the performance of the four routing protocols in case of change in traffic, no. of nodes and node mobility. The routing protocols include Destination Sequenced Distance Vector (DSDV), Dynamic Source Routing (DSR), Optimized Link State Routing Algorithm (OLSR) and Temporally Ordered Routing Algorithm (TORA).

Keywords: MANET, Routing protocols, DSDV, DSR, OLSR, TORA, Proactive, Reactive.

1 Introduction (MANETs)

In general, Mobile ad Hoc networks can change its location and configure its network. Mobile nodes act as hosts (running applications) and Routers (forwarding for others). [1] It makes the series of mobile nodes, which are dynamic in nature. MANET is temporary (no fixed infrastructure) and self configuring network [2]. The nodes make the wireless network to each other by making non-centralized administration. We assume that these nodes are laptops, phones, PDA's (Personal data assistant), wireless handsets etc. Packets are sending from source to destination with the help of intermediate node. Due to node mobility, links are established or broken. MANET [3] has the several applications in military network environment, disaster operation, rescue operation and conference rooms. Routing protocols determines the route, which is required. To find the result, the above routing protocols are comparing to each other. These protocols are selected based on previous result, characteristic and features.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 502–507, 2013.

2 Literature Review

The main objective is to review the routing protocols to check out the performance. Two types of routing protocols are:

2.1 On Demand Routing Protocol (Reactive)

The protocol finds the route as it is required. First, protocol sends the route request packet to all the forwarding nodes. This is done till the destination is not found. Destination sends the route reply packet to the source node. Only response route is maintained, rest is deleted.

- **Dynamic Source Routing (DSR):** [4], [5], [6],[7], [8] The Dynamic source routing first finds the route discovery. This route discovery is done by sending a packet frame as the source to next node. That node sends the route reply message back. The other, term this is used as the route maintenance. After that,which node is out of range so there is a break in connection. Then, route error packet is assigned.
- **Temporally Ordered Routing Algorithm (TORA):**[9], [17]TORA provides more than one routes for source to destination. It is more efficient when mobility is more, due to concept of loop-free and adaptive distributing routing.

2.2 Table Driven Routing Protocol (Proactive)

This type of routing protocol makes the fresh list of destinations. The route is maintained by routing table. When the source sends the packet to the next node, there is the change in the route table. Two proactive routing protocols are:

- **Destination Sequenced Distance Vector (DSDV):**[10], [18] DSDV routing protocol is based on Sequence number. Route table is maintained for checking the information between nodes. Exchanging of information is done with Network Protocol Data Unit (NPDU). Bellman-Ford routing algorithm is used in this protocol.
- **Optimized Link State Routing Algorithm (OLSR):**[11], [12], [13] OLSR sends the HELLO message to check its neighbor. The routing table is managed by the information of Topology control (TC) Packets. Control packet is sent to the network, by the nodes called multipoint relays. The path is selected by using the shortest path algorithm.

3 Performance Evaluation

According to the literature review [3], [14], [15], and [16] these parameters are considered as to calculate the performance of routing protocols.

- **Packet Delivery Ratio (PDR):** It is the ratio of the packet sent from source to the number of packet received at the destination. PDR is determined as:

$$PDR = (P_r/P_s) \text{ X } 100 \quad (1)$$

Where P_r is the total packets received and P_s is the total packets sent.

- **End to End Delay (D_{avg}):** This is the delay packet send from source to the destination. The average delay is computed as:

$$(D_{avg}) = \sum (t_r - t_s) / P_r \quad (2)$$

Where t_s is the packet sent time, t_r is the packet receive time and P_r is the total packets received for the same packet at destination.

- **Routing Load:** Routing load is the number of routing control packet transmitted for each data packet delivered at the destination. Routing load is determined as:

$$RL = P_c / P_d \quad (3)$$

Where P_c is the total control packets sent and P_d is the total packets sent.

- **Throughput:** The total successfully received packet to the destination.

4 Survey Analysis

It is very difficult to design a routing protocol, which satisfies all the parameters and gives best performance. Here Table: 1, 2, 3 and Table: 4 gives the summary in which "1" denotes the best performance and "4" denotes the worst performance.

4.1 Apply Effect of Node Density and Packet Length

Table 1. Nodes = 50, Packet length = 50,000 bytes

Parameter / Protocols	Packet Delivery Ratio	End to End Delay	Routing Load	Throughput
DSDV	4	1	3	3
DSR	2	3	1	2
OLSR	1	2	2	1
TORA	3	4	4	4

Table 2. Nodes = 100, Packet length = 50,000 bytes

Parameter / Protocols	Packet Delivery Ratio	End to End Delay	Routing Load	Throughput
DSDV	2	1	2	3
DSR	3	3	4	1
OLSR	1	2	1	2
TORA	4	4	3	4

4.2 Apply Effect of Node Density and Mobility

Table 3. Nodes = 50, Mobility = 30 m/s

Parameter / Protocols	Packet Delivery Ratio (PDR)	End to End Delay	Routing Load	Throughput
DSDV	3	2	2	2
DSR	4	3	3	4
OLSR	1	1	1	1
TORA	2	4	4	3

Table 4. Nodes = 100, Mobility = 30 m/s

Parameter / Protocols	Packet Delivery Ratio (PDR)	End to End Delay	Routing Load	Throughput
DSDV	4	3	3	4
DSR	3	2	2	1
OLSR	1	1	1	2
TORA	2	4	4	3

5 Results and Comparative Discussion

In this paper, we use two types of routing protocols. These two types of protocol are reactive and proactive protocol. Different type of variations is done with nodes is number of nodes, packet length and mobility as shown in Table: 5.

Table 5. Overall performances of protocols

Minimum nodes, Maximum packet length (traffic load).	OLSR> DSR> DSDV> TORA
Maximum nodes, Maximum packet length (traffic load).	OLSR> DSDV> DSR> TORA
Minimum nodes, Maximum mobility.	OLSR> DSDV> DSR> TORA
Maximum nodes, Maximum mobility.	OLSR> DSR> DSDV> TORA

5.1 Effect of Minimum Nodes and Maximum Packet Length

In Table: 1 OLSR gives best performance than other protocols. Because OLSR has maximum PDR and lesser delay as compared to other protocols. OLSR is a proactive protocol due to it maintains a fresh list of nodes. Either it forwards or drops the packet

when it arrives. TORA is worst performer among the given protocols, because TORA has more delay and routing load.

5.2 Effect of Maximum Nodes and Maximum Packet Length

In Table: 2 as the network size increases the performance of OLSR remains the same. Network size has more effect on reactive protocols because proactive protocols make the routing table. Therefore, performance of DSDV is more than DSR.

5.3 Effect of Minimum Nodes and Maximum Mobility

In Table: 3. OLSR has slightly more throughput and PDR. This is because OLSR makes the connection with intermediate node easily, as the mobility is high. DSR has difficulty in choosing routes due to which PDR and throughput decreases. Then DSDV gives better result than DSR. Here, TORA has high PDR than other protocols but due to high delay and routing load gives overall worst performance.

5.4 Effect of Maximum Nodes and Maximum Mobility

In Table: 4. DSDV decreases the performance than DSR because node mobility gives link breakage between nodes. Then the throughput of DSR is more than DSDV. Here DSR has good performance than DSDV, but less than OLSR.

6 Conclusion

In this review paper, we have to analyze the performance by comparing the routing protocols. Overall, when the number of nodes increases then there is decrease in PDR and throughput and increase in delay and routing load. DSR show better result when maximum number of nodes and mobility is high. DSDV gives good result in case of maximum number of nodes and maximum packet length. In all the scenarios, TORA is worst performer and OLSR show best result among four routing protocols.

References

1. Shrivastava, L., Bhadauria, S.S., Tomar, G.S.: Performance Evaluation of Routing Protocols in MANET with different traffic loads. In: 2011 International Conference on Communication Systems and Network Technologies, pp. 13–16 (2011)
2. Blum, I.J., Eskandarian, A., Hoffman, J.L.: Challenges of inter- vehicle Ad hoc Networks. IEEE Transactions on Intelligent Transportation Systems 5(4) (December 2004)
3. Adam, N., Ismail, M.Y., Abdullah, J.: Effect of Node Density on Performances of Three MANET Routing Protocols. In: IEEE 2010 International Conference on Electronic Devices, Systems and Applications (ICEDSA 2010), pp. 321–325 (2010)
4. Maan, F., Mazhar, N.: MANET Routing Protocols vs Mobility Models: A Performance Evaluation, pp. 179–184. IEEE (2011)

5. Barakovi, S., Barakovi, J.: Comparative Performance Evaluation of Mobile Ad Hoc Routing Protocols, pp. 518–523. IEEE (2010)
6. Johnson, D.B., Maltz, D.A.: Dynamic source routing in adhoc wireless networks. In: Imielinski, Korth (eds.) Mobile Computing, pp. 153–181. Kluwer Academic Publishers (1996)
7. Agrawal, R., Tripathi, R., Tiwari, S.: Performance evaluation and comparison of AODV and DSR under Adversarial environment. In: IEEE 2011 International Conference on Computational Intelligence and Communication Systems, pp. 596–600 (2011)
8. Jain, R., Khairnar, N.B., Shrivastava, L.: Comparative study of three mobile Ad-Hoc network routing protocols under different traffic sources. In: IEEE 2011 International Conference on Communication Systems and Network Technologies, pp. 104–107 (2011)
9. Naeemv, M., Ahmed, Z., Mahmood, R., Ajmal Azad, M.: QOS Based Performance Evaluation of Secure On-Demand Routing Protocols for MANET's. IEEE (2010)
10. Barakovi, S., Barakovi, J.: Comparative Performance Evaluation of Mobile Ad Hoc Routing Protocols, pp. 518–523. IEEE (2010)
11. Kulla, E., Ikeda, M., Barolli, L., Miho, R., Kolici, V.: Effect of source and destination movement on MANET performance considering OLSR and AODV protocols, pp. 510–515. IEEE (2010)
12. Kulla, E., Ikeda, M., Hiyama, M., Barolli, L.: Evaluation of MANET testbed in indoor environment considering OLSR protocol. In: IEEE 2011 International Conference on Complex, Intelligent, and Software Intensive Systems, pp. 160–167 (2011)
13. Sarkar, N.I., Lol, W.G.: A Study of MANET Routing Protocols: Joint Node Density, Packet Length and Mobility, pp. 515–520. IEEE (2010)
14. Broach, J., Maltz, D.A., Johson, D.B., Hu, Y., Jetcheva, J.: A Performance Comparison of Multi-hop Wireless Ad Hoc Network Routing Protocols. In: IEEE Proc. of MobiCom, Texas, USA (1998)
15. Hussein Mamon, M.: Important Characteristic of Differences between DSR and AODV Routing Protocol. In: IEEE MCN 2007 Conference (2007)
16. Tanenbaum, A.S.: Computer Networks, 4th edn. Prentice-Hall PTR (2003)
17. Kiwior, D., Lam, L.: Routing protocol performance over intermittent links, pp. 1–8. IEEE (2007)
18. Zhang, J., Peng, H., Jing Shao, F.: Energy consumption analysis of MANET routing protocols based on mobility models. In: IEEE 2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), pp. 2275–2280 (2011)

MJ_2-RSA Signature Scheme in E-Commerce for Fair-Exchange Protocols

E. Madhusudhana Reddy[1] and M. Padmavathamma[2]

[1] Dept of Computer Science & Engineering,
Madanapalle Institute of Technology & Science, Madanapalle-517325, India
[2] Dept of Computer Science,
S.V. University, Tirupati-517502, India
{e_mreddy,prof.padma}@yahoo.com

Abstract. This paper presents the design of our proposed protocols using MJ_2-RSA signature scheme for invoking communication between Customer, Trader, Multi Level Trader and Bank for registration and generation of secret key, publishing of public keys and distribution of private keys with ZTTP.

Keywords: Key-Generation, protocols, MJ_2-RSA signature scheme.

1 Introduction

In this paper we presents the significance of Fair Exchange protocols and propose protocols for invoking communication among different parties such as Customer, Trader, Multi Level Trader and Banker using our developed cryptosystem using MJ_2-RSA cryptosystem and signature scheme. Section 2 describes about J_k-RSA Cryptosystem and MJ_2-RSA Digital Signature. Section 2 presents the importance of Fair Exchange protocols for using MJ_2-RSA cryptosystem. Design of protocols for symmetric key generation is presented in section 3. In section 4 introduces Key-Generation phase which include generation of One Public keys, One Private Key for Customer and One Public Key and Two Private Keys for Multi Trader, Bank (Hierarchy level of authentication) using our proposed protocol MJ_2-RSA crypto scheme.

2 J_K-RSA Cryptosystem

The algorithm for key generation, encryption and decryption of J_k-RSA Cryptosystem is described as follows.

Key Generation: Choose two large primes P and Q such that N = PQ.
Let K be an integer such that $1 \le K \le N$.

Compute $J_k(N) = N^k \prod_{P/N} (1 - 1/p^k)$ and consider $\left(Z_{J_K}(N), +J_k(N), xJ_k(N)\right)$ a commutative ring with unity of order J_k (N) as a message space. Assign the

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 508–511, 2013.

numerical equivalents to the alphabets taken from $Z_{J_K}(N)$. M is the message belongs to $Z_{J_K}(N)$. Select a random integer e such that gcd (e, J_k (N)) = 1, where 1 < e < J_k (N), E≡ M mod J_k (N) ∈ message space $Z_{J_K}(N)$. Select integer d such that ED ≡ 1 (mod J_k (N)) i.e., D = E^{-1} mod J_k (N) where 1 ≤ D ≤ J_k (N).

Public-Key PK = (J_k (N), E)
Private Key SK = (J_k (N), D)

Encryption: Given a public-key (J_k (N), E) and a message $M \in Z_{J_K}(N)$, compute the ciphertext C = M^E mod J_k (N) = M^E mod J_k (N)

Decryption: Given a public-key (J_k (N), D) and ciphertext C, compute the message

$$M = C^D \bmod J_k(N) \quad = C^D \bmod J_k(N)$$

The correctness of J_k-RSA decryption is verified as follows

$C^D \bmod J_k(N) = (M^E)^D \bmod J_k(N) = (M^{ED} \bmod J_k(N) = (ED).M \bmod J_k(N)$
$= 1.M \bmod J_k(N) = M$

2.1 MJ$_2$-RSA Digital Signature

The algorithm for key generation, signature generation and verification of MJ$_2$-RSA Digital Signature is described as follows.

Key generation: 1. The Signer Choose sufficiently large distinct primes, $p_1,p_2,\ldots\ldots,p_r$ at random.

2. Compute N= $\prod_{i=1}^{r} p_i = p_1.p_2.\ldots.p_r$ and $J_2(N) = n^k \prod_{p/k}(1-p^{-k})$

$= (p_1^k - 1)(p_2^k - 1) - - - - - (p_r^k - 1) = \prod_{i=1}^{r} (p_i^k - 1)$

3. Choose a random integer E< J_2(N) such that gcd (E, J_2(N)) = 1.
4. Compute the integer D Which is the inverse of E i.e, ED≡1(mod J_k(N).
5. for 1≤i≤r, compute $D_i \equiv D(\bmod p_i^k - 1)$

Public Key = (2, N, E), Private Key = (2, ($p_1,p_2,\ldots\ldots,p_r$, $D_1,D_2,\ldots\ldots,D_r$))

Signature generation: Using the Private Key (N, E) creates a signature 'σ' on the message M by computing. $\sigma \equiv M^D (\bmod N)$

Signature Verification: After obtaining the signature 'σ' and the message M, Check whether $M \equiv \sigma^e (\bmod N)$ .If the above equation holds then "Accept: the message otherwise.

3 Fair-Exchange

Fair-exchange is an important property that must be ensured in all electronic commerce environments where the merchants and the customers are reluctant to trust each other. This property guarantees that none of the transacting parties suffers because of the fraudulent behavior of the other party in the transaction. In this paper, we propose the design of our MJ_2-RSA protocol for registration phase of Customer, Trader, ML-Trader, and Bank to overcome the problem of fair-exchange for digital products. A fair-exchange protocol allows two parties to exchange items in a fair way so that either each party gets the other's item, or neither party does. We describe a novel method of constructing very efficient fair-exchange protocols by distributing the computation of MJ_2-RSA signatures and MJ_2-RSA dual-signatures. Today, the vast majority of fair-exchange protocols require the use of zero-knowledge proofs, which is the most computationally intensive part of the exchange protocol. Using the intrinsic features of our dual-signature model, we construct protocols that require no zero-knowledge proofs in the exchange protocol.

4 Dual Key Generation for ML-Trader, Bank

The Private key is generated by finding the unique integer D, $1 < D < J_2(N)$, such that $ED = 1 \pmod{J_2(N)}$. The corresponding (primary signer's) partial private key is generated by finding the unique integer D_1, such that $ED_1=1 \pmod{J_2(N)}$. Next, the (cosigner's) partial private key D_2 is computed as $D_2 = D / D_1$. Dual private keys for ML-Trader and Bank are (E,D_1,D_2,N) Server first generates the parameters multiple-prime $P_1,P_2,P_3,\ldots P_r$ and the keys E,D,D_1 and D_2 (D_1,D_2 for ML-Trader/Bank).

4.1 Our MJ_2-RSA Based Dual-Signature Scheme for ML-Trader / Bank

MJ_2-RSA based dual-signature scheme that allows two signers to compute a dual-signature efficiently. The core idea behind his scheme is to *multiplicatively* split the private key d into two partial keys D_1 and D_2, each associated with a different signer. That is, $D \equiv D_1D_2 \pmod{J_2(N)}$. The dual-signature computation is then based on the equation $H(M) \equiv H(M)^{D1D2} \pmod{N}$. In our fair-exchange protocol, we employ an MJ_2RSA-based dual-signature scheme that also splits D into two partial keys, but the splitting is done *additively* instead of multiplicatively.

4.2 Dual-Signature Generation for ML-Trader, Bank

The primary signer and cosigner create their respective partial signatures, $Sig_i = H(M)^{Di} \bmod N$ i=1,2, $Sig_1 = H(M)^{D1} \bmod N$ *for signer,* $Sig_2 = H(M)^{D2} \bmod N$ for cosigner. Using their respective partial private keys D_1 and D_2, These are multiplied modulo N to form the dual-signature Sign. That is, $Sign = H(M)^{D1*D2} \bmod N$. The partial signature al is considered valid if and only if $Sign^E \bmod N = H(M)$. The dual-signature Sign is verified using the public key E.

5 Conclusion

The protocols provide strong fairness. At the end of each exchange, a party obtains his counterpart's item(s) if and only if the other party obtains his item(s). If fairness is not achieved during a normal exchange process, a ZTTP can be invoked for assistance. Our research hypothesis was that fair exchange in these cases could be achieved by encrypting the exchanged signature so that the receiver could verify its correctness, while both its originator and an agreed ZTTP could decrypt the signature; in the case when the ZTTP is involved, the signature's confidentiality is preserved from the ZTTP.

References

1. Abadi, M., Glew, N., Horne, B., Pinkas, B.: Certified Email with a Light On-line Trusted Third Party: Design and Implementation. In: Proceedings of the International World Wide Web Conference, WWW 2002, pp. 387–395. ACM Press, New York (2002)
2. Alur, R., Henzinger, T.A., Mang, F.Y.C., Qadeer, S., Rajamani, S.K., Tapcsiran, S.: MOCHA: Modularity in Model Checking. In: Vardi, M.Y. (ed.) CAV 1998. LNCS, vol. 1427, pp. 521–525. Springer, Heidelberg (1998)
3. Asokan, N., Schunter, M., Waidner, M.: Optimistic Protocols for Fair Exchange. In: Proceedings of the ACM Conference on Computer and Communications Security, pp. 6–17 (1997)
4. Madhusudhana Reddy, E., Padmavathamma, M.: Need for strong public key cryptography in electronic commerce. IJ-CA-ETS, 58–68 (2009) ISSN: 0974-3596
5. Madhusudhana Reddy, E., Padmavathamma, M.: An information security model for e-commerce. The Technology World Quarterly Journal V(I), 203–206 (2009) ISSN: 2180-1614
6. Madhusudhana Reddy, E., Padmavathamma, M.: A middleware e-commerce security using $RMPJ_K$-RSA cryptosystem and $RMPJ_K$-RSA digital signature. International Journal of knowledge Management & e-Learning 1(2), 71–76 (2009)
7. Madhusudhana Reddy, E., Padmavathamma, M.: The impossibility of secure public key encryption. In: Proceedings of International Conference on Information Processing, ICIP 2007, August 10-12, Organized by the Society of Information Processing, Bangalore, UVCE, pp. 253–258 (2007)

Robust Image Watermarking by Establishing Adaptive Attributes

Seema Hiremath and Anita Sahoo

Department of Computer Science & Engineering, JSS Academy of Technical Education, Noida, India
{anushab.chinmaya,Anitasahu1}@gmail.com

Abstract. In this paper a new digital image watermarking technique have been proposed to protect the copyright of multimedia data. The goal of this paper is to develop a watermarking algorithm based on image segmentation and adaptive Harris corner detector. The Image segmentation is done using Expectation Maximization algorithm based on intensity values of the pixels. Adaptive Harris corner detector is used to identify the rotation and scaling invariant feature points. Circular region is defined around each selected feature points, and selected regions are normalized .Then a pseudorandom sequence of watermark is generated and embedded in the circular region. Watermark embedding strength is adjusted adaptively using the noise visibility function. The effectiveness and accuracy of the proposed scheme is established through experimental results. The results show that our proposed watermark scheme is robust to rotation, scaling and common signal distortions.

Keywords: Image segmentation, Noise Visibility Function, Adaptive Harris Corner Detector, Watermarking.

1 Introduction

Digital Watermarking is a technique used to hide data or identifying information within digital multimedia. The study goal of this paper was to design a simple & speed watermarking scheme that was resistant to RST, JPEG compression, and median filtering.

Say Wei Foo [1] proposes the scheme, in which feature points are first extracted from host image and several circular patches centred on these points are generated. The patches are used as carriers of watermark information. Zernike transform is then applied to the patches to calculate local Zernike moments. Dither modulation is adopted to quantize the magnitudes of the Zernike moments followed by false alarm analysis. Wei Wang, [2] proposes the scheme which selects the phase information in DFT domain as feature of image, and uses generalized Radon transformations to identify the geometric transformations. A secret key related to the original image is generated using the phase feature, maxima and the watermark sequence. The key is used later in image authentication and watermark extracting. Sirvan Khalighi [3] proposes scheme based on the Contour let Transform in which watermark is embedded into the highest frequency sub band of the host image in its Contour let domain, this method enables to embed more amounts of data into the directional sub

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 512–515, 2013.

bands of the host image without degrading its perceptibility. Our proposed method is very fast & efficient and highly robust. Histogram based image segmentation is used to make segmentation fast & simpler. Adaptive Harris corner detector is used to identify the features points, once the features are identified circular regions are defined around each feature points and each circular region is normalized and watermark is embedded in that circular region.

2 Proposed Watermarking Scheme

Proposed architecture of the watermarking system consists of Image segmentation, feature extraction, circular region alignment, normalization, watermark generation and embedding. Similar steps are followed for watermark extraction.

Segmentation: Histogram based image segmentation proposed by Sateesh & Raja [4] is applied on spatial domain using patchwork algorithm. Histogram based algorithm is used to obtain the number of regions and the initial parameters like mean, variance and mixing factor. The final parameters are obtained by using the Expectation and Maximization algorithm. The segmentation of the approximation coefficients is determined by Maximum Likelihood function. It is observed that the proposed method is computationally efficient allowing the segmentation of large images and performs much superior to the earlier image segmentation methods.

Feature Extraction: Harris corner detector is based on the local auto-correlation function of a signal. Given a shift (x, y) and a point the auto-correlation function is defined as

$$c(x,y) = \sum_{W} \left[I(x_i y_i) - I(x_i + \Delta x, y_i + \Delta y) \right]^2 \tag{1}$$

Where I (xi, yi) represent the image function and (xi, yi) are the points in the window W centered on (x, y). The shifted image is approximated by a Taylor expansion truncated to the first order terms.

$$I(x_i + \Delta x, y_i + \Delta y) \approx \left[I(x_i y_i) + \left[I_x(x_i + y_i) I_y(x_i + y_i) \right] \right] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \tag{2}$$

where Ix (xi, yi) and Iy (xi, yi) indicate the partial derivatives in x and y respectively. With a filter like [-1, 0, 1] and [-1, 0, 1] T, the partial derivates can be calculated from the image by substituting.

$$c(x,y) - \begin{bmatrix} \Delta x & \Delta y \end{bmatrix} \begin{bmatrix} \sum_{W} (I_x(x_i, y_i))^2 & \sum_{W} (I_x(x_i, y_i) I_y(x_i, y_i)) \\ \sum_{W} (I_x(x_i, y_i) I_y(x_i, y_i)) & \sum_{W} (I_y(x_i, y_i))^2 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} - \begin{bmatrix} \Delta x & \Delta y \end{bmatrix} c(x,y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \tag{3}$$

C(x, y) the auto-correlation matrix which captures the intensity structure of the local neighborhood. Let α1 and α2 be the eigenvalues of C(x, y), then we have 3 cases to consider:

1. Both eigenvalues are small means uniform region (constant intensity).
2. Both eigenvalues are high means Interest point (corner)
3. One eigen value is high means contour(edge)

To find out the interest points, Characterize corner response H(x, y) by Eigen values of C(x,y).

(i) C(x, y) is symmetric and positive definite that is α1 and α2 are >0
(ii) α1 α2 = det (C(x, y)) = AC –B2
α1 + α2 = trace(C(x, y)) = A + C
(iii) Harris suggested: that the HcornerResponse = α1 α2 – 0.04(α1 + α2)2

Finally, it is needed to find out corner points as local maxima of the corner response.

Watermark Embedding: Circular regions are defined around these selected features. Each circular region is normalized. Watermark W is generated using Pseudo Random Sequence generator and a secret key. Watermark embedding strength is adjusted using the Noise Visibility Function. Below formula calculates the watermark embedding strength using NVF .

$$Ys=Xs+ (1-NVF)*W$$

Xs is original image data in circular region , Ys is watermarked image data ,W is the generated Watermark sequence.
Each whole circular region is filled with Ys data which is an original image data embedded with watermark sequence.

Watermark Extraction: Image segmentation, feature extraction and circular region alignment are similar to the method used for embedding the watermark. Once the circular regions are located, watermark is extracted from that circular region and watermark sequence is generated using PN generator and the same secret key that was used to generate the watermark during watermark embedding. Correlation Coefficient is used to find the distortion of the extracted & generated watermark.

3 Experiments and Results

PSNR and Correlation Coefficient without distortion and after RST & general Image processing are listed in Table 1 below. Results are compared with N.R. Nantha Priya's work [5]. We can observe, that we are able to retrieve the same watermark which was embedded , comparing this result to Nantha's, we can see that out method

Table 1. PSNR & CC of test images without distortions and after RST and common signal processing attacks

Attacks /Images	PSNR	CC[ours] / CC[Nantha's]	Rotate - 90°/ -45°	Scaling 0.5	Gaussian noise of 0.1	3x3Median+ JPEG 85%
Leena	37.969	1/0.6931	1/ 0.9409	1	0.96711	0.928
Baboon	35.9361	1/1	1/0.9068	1	0.92558	0.9288
Barbara	40.4365	1/0.6184	1/0.8423	1	0.93772	0.9633
Pepper	36.5905	1/0.9007	0.9585/0.9011	1	1	1

has performed extremely well in retrieving watermark when image is not distorted. We can also observer from the below table that proposed method is highly robust to scaling, rotation and general image processing operations.

4 Conclusion

Our proposed watermarking scheme is highly robust against geometrical & common signal processing attacks and demonstrated its superior performance using different experiments. This approach outperforms other results has limitation for resizing.

References

1. Foo, S.W., Dong, Q.: A Feature-based Invariant Watermarking Scheme Using Zernike Moments. International Journal of Electrical and Computer Engineering 5(6), 375–381 (2010)
2. Wang, W., Men, A., Chen, X.: Robust Image Watermarking Scheme Based on Phase Features in DFT Domain and Generalized Radon Transformations. IEEE Transactions on Image and Signal Processing 9, 213–218 (2009)
3. Khalighi, S., Tirdad, P., Rabiee, H.R.: A Contourlet-Based Image Watermarking Scheme with High Resistance to Removal and Geometrical Attacks. EURASIP Journal on Advances in Signal Processing, Article ID 540723, 1155–1168 (2010)
4. Sateesh Kumar, H.C., Raja, K.B.: Automatic Image Segmentation using Wavelets. IJCSNS International Journal of Computer Science and Network Security 9(2), 365–370 (2009)
5. Nantha Priya, N.R.: Robust Feature Based Image Watermarking Process. International Journal of Computer Application 4(5), 11–13 (2010)
6. Gao, X., Deng, C., Li, X., Tao, D.: Geometric Distortion Insensitive Image watermarking In Affine Covariant Regions. Transactions on Systems, Man, and Cybernetics 40(3), 278–286 (2010)
7. Pereira, S., Pun, T.: Robust template matching for affine resistant image watermarks. IEEE Trans. on Image Process. 9(6), 1123–1129 (2000)
8. Cox, I.J., Millar, M.L., Bloom, J.A.: Digital Watermarking (2002) ISBN:1-55860-714-5
9. O'Ruanaidh, J., Pun, T.: Rotation, scale, and translation invariant digital image watermarking. Signal Process. 66(3), 303–317 (1998)
10. Falkowski, B.J., Lim, L.S.: Image Watermarking Using Hadamard Transforms. IEEE Electronics Letters 36(3), 211–213 (2000)
11. Cox, I.J., Millar, M.L., Bloom, J.A.: Watermarking Applications and Their Properties, pp. 6–10 (2000) ISBN 0-7695-0540-6
12. Katzenbeisser, S., Petitcolas, F.A.P.: Information hiding techniques for steganography and digital watermarking, pp. 121–132 (2002) ISBN 1-56053-035-4

A Survey of Reversible Watermarking Techniques for Data Hiding with ROI-Tamper Detection in Medical Images

Amrinder Singh Brar[1] and Mandeep Kaur[2]

[1] University Institute of Engineering and Technology
Panjab University, Chandigarh
amrindersinghb@yahoo.com
[2] IT, University Institute of Engineering and Technology
Panjab University, Chandigarh
mandeep@pu.ac.in

Abstract. Digital watermarking in the field of medical images is becoming a new research focus. Modern medical equipments produce mass digital images and data everyday [1]. The integrity of such records needs be protected from unauthorized modification. One of the security measures that can be used is watermarking as traditional technologies for data security such as cryptography, digital authentication etc cannot fully meet the requirements [2]. With digital watermarking ,confidentiality and integrity of a medical image can be achieved by hiding the Electronic Patient Record (EPR data) in corresponding medical images and can have wide applications in the field of Tele- medicine. But at the same time, diagnostic quality of the medical images should not be compromised. The objective of this paper is to discuss some of the most efficient techniques suitable for watermarking of medical images. These methods are based on modified difference expansion, CDCS, Secondary LSB replacement and are reversible in nature. These techniques aim at increasing the data hiding capacity, without distortion of diagnostically important information and also provides ROI authentication and temper detection.

Keywords: CDCS, Medical image, ROI-based, Tamper Detection, Reversible Watermarking, Histogram modification, Near Reversible, Secondary LSB.

1 Introduction

Healthcare institution that handles a number of patients, opinions is often sought from different experts. It demands the exchange of the medical history of the patient among the experts which includes the clinical images, prescriptions, initial diagnosis etc. But with the widespread and increasing use of Internet, these digital images can be easily accessed and manipulated. Considering patient's privacy and diagnostic accuracy, the prevention of medical images from tampering tends to be an urgent task [1].

Another fact is that in recent years, health care systems involve a large amount of data storage and transmission such as patient information, medical images, electronic patient records (EPR) and graphs [3]. Transmission of such a large amount of data

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 516–522, 2013.

when done separately using ordinary commercial information channels like internet, it results in excessive memory utilization, an increase in transmission cost and time and also make that data accessible to unauthorized personnel [4]. In order to reduce storage and transmission cost, data hiding techniques are used to embed patient information with medical images. These data hiding techniques can also be used for authentication and tamper detection so that integrity of region of interest (ROI) can be judged [5].

In medical imaging applications, there are stringent constraints on image fidelity that strictly prohibit any permanent image distortion by the watermarking [6]. For instance, artefacts in a patient's diagnostic image due to image watermarking may cause errors in diagnosis and treatment, which may lead to possible life-threatening consequences [7]. Thus, to overcome the problem of occurrence of artefacts and to produce zero distorted or noise free watermarked medical images, various reversible, fragile watermarking schemes have been introduced by various researchers out of which some are surveyed in this paper.The rest of the paper is organized as follows:

In Section 2 the reversible watermarking technique with high capacity EPR data hiding using CDCS with ROI tamper detection is reviewed. In section 3 reversible watermarking using two-dimensional difference expansion scheme is discussed. Section 4 gives an overview of Near Reversible Watermarking algorithm based on secondary LSB replacement. The key features and capabilities of all three techniques is discussed in Section 5.

2 Class Dependent Coding Scheme

This scheme works in two phases: the text processing phase and the image processing phase. In text processing phase stream of EPR encoded bits along with hash code bits are prepared and in image processing phase these bits are embedded into the RONI region of corresponding medical image.

2.1 Preparing Payload

In this technique firstly Class Dependent Coding Scheme (CDCS) is applied on EPR data bits to reduce the size of EPR data bits and enhance the embedding capacity. The CDCS technique assigns fixed codes to each character according to their occurrence probability [11, 12]. In this scheme EPR characters are categorized into three different non-overlapping classes. Class A is considered to be most frequently appearing character set, Class B as an average frequency appearing character set and Class C as a less frequently appearing character set. The number of bits needed to represent each character in the respective class is calculated by assuming only capital letters, alphanumeric and few special characters. Then variable length code is designed to represent each class based on Huffman encoding. Using this scheme any character can be represented by only 4-bits prefixed by a class code (1-bit or 2-bit). CDCS combines the advantage of both fixed length and variable length coding to get less number of bits to represent same information.

To enhance the robustness of the technique against various attacks such as image compression redundancy bits are added. Tampering along with bit error correction can be achieved by adding redundancy for each bit prior to embedding [11, 12]. Interleaving of bits is done to spread subsequent bits from each other throughout the image, so that even if any block of stego image undergoes with an attack, EPR bits can be recovered from another blocks [10]. These two techniques along with CDCS provide more robustness.

The 8x8 blocks of the ROI part of image are not considered for embedding so a secure hash function using SHA-256 is calculated on these ROI blocks [11]. These hash bits are appended to EPR data bits to get the Final Bit Stream (FBS). This FBS is embedded in NROI 8x8 blocks.

2.2 Finding Robust Difference Parameter 'α'

We split each 8×8 RONI image block into two subsets A and B. For each of this RONI block, we calculate the difference value α which is defined as the arithmetic average of difference of pixel pairs within the block [13].

$$\alpha = (1/n) * \sum_{i=1}^{n} (ai - bi)$$

The difference value α is expected to be very close to zero due to correlation and spatial redundancy in the pixel values of local block [13]. We select this α as a robust quantity for embedding information bit as the value 'α' is based on the statistics of all pixels in the block, it has certain robustness against attacks.

2.3 Embedding Bits

According to the distribution of α in the original medical image the threshold value't' is chosen. Two cases arise:

In case 1, difference value α is located within defined threshold. In this case if 1 is to be embedded, we shift the difference value α to the right side beyond a threshold, this is done by adding or subtracting a fixed number'd' called offset from each pixel value within one subset. If 0 is to be embedded, the block is intact [13].

In case 2, if the difference value α is located outside the threshold, no matter whether we have to embed bit 1 or 0, we always embed bit 1, thus shifting the value α farther away beyond the threshold. The bit error introduced in this case is corrected by using error correction [13]. While adding or subtracting'd', the pixel values should not lead to overflow/underflow problem.

2.4 Retrieval of EPR Data

To extract the EPR data from stego image firstly it is divided into 8×8 non-overlapping blocks excluding ROI blocks. The difference value α is calculated for each block using above equation [13]. If the calculated difference value lies outside

the threshold, then bit 1 is extracted and pixel value of one subset is restored to its original value [13]. If the difference value α lies within threshold, then bit 0 is extracted without changing the pixel value of that block. So following this we can extract EPR data and embedded hash bits. This hash value is then compared with the new hash calculated over stego image's ROI to check the authenticity of ROI of image and to check whether it is tampered or not. Original image can be restored without any distortion after EPR data is extracted, after extracting all bits of EPR information it can be reconstructed using CDCS.

3 Two-Dimensional Difference Expansion Based Scheme

In this scheme 2-Dimensional Difference Expansion (2D-DE) is used which is modified form of previous difference expansion (DE) techniques. This scheme allows embedding the recovery information of ROI without lossy compression which was used previously. This technique can be used for authenticating ROI, hiding patient's data, finding tampered areas inside ROI, and recovering those tampered areas. Also, the original image is recovered exactly after watermark extraction at the receiver end [3].

In this scheme, the image is divided into non-overlapping blocks of 4×4 pixels, then using Haar wavelet transform in horizontal and vertical direction (which represent two dimensions of the block) each block is transformed into frequency domain. Then in the blocks which do not cause overflow or underflow 16 bits are embedded using difference expansion [3].

3.1 Embedding the Payload

In this technique first of all ROI region of medical image is selected, RONI and the border regions are defined. ROI is divided into 16×16 pixel blocks. Then MD5 algorithm is used to calculate hash message (H) for ROI. The bits of ROI pixels are collected as P. The LSB's of border's pixels are collected as L. The patient data, D, is concatenated with P, H and L and is then compressed using Huffman coding to generate the payload. This payload is embedded into RONI using 2-Dimensional Difference Expansion (2D-DE) technique and embedding map is generated. This embedding map and ROI coordinates are concatenated, compressed and then embedded into LSB's of border's pixels. The watermarked image is now ready and can be stored or transferred safely [3].

3.2 Extraction Process

In the extraction process LSB's of the border's pixels are recovered. This recovered bit stream is depressed; the embedding map and ROI coordinates are extracted separately. ROI and RONI regions are defined using ROI coordinates. Then the payload is extracted from RONI and is decompressed and then decomposed into; H, P, L, D.

Then the hash message of ROI is calculated and compared to extracted one. If they are equal, the image is authentic and then using the bits of *L,* the LSB's of the border's pixels are recovered. Else if they are not equal image is unauthentic and it means some tampering is detected. Then to localize the tampered area and to recover the original ROI bits, ROI is divided into blocks of 16×16 pixels. The average value of each block is calculated and matched with value of the average of corresponding pixel in extracted ROI; *P*. If both are not equal, the block is marked as tampered and is replaced by corresponding pixel *P* [3]. So in the end, the original image is extracted exactly after watermark extraction.

4 Near Reversible Watermarking Algorithm Based on Secondary Lsb Replacement

This paper proposes a near reversible image watermarking algorithm based on the secondary LSB replacement [14]. Simulation results prove that the watermarked image not only can well hide the watermark information by storing the secondary LSB data of host image and the watermark data in the same bit, but also can be recovered to the original host image to a high extent. This technique have strong robustness and low calculating complexity[14].

4.1 Embedding Process

Firstly we read image files, get the data matrixes of host and watermark image *I(i,j)* and watermark image *W(i,j)*, write the size of watermark into the head of the host image.

I(i,j),i=1:*M, j*=1: *N,* M and N are the row and column of host image respectively.

W(i,j),i=1:*m, j*=1:*n,* m and n are rows and column of watermark respectively.

Now by using formula(1) calculate the ratio R of host image and watermarked image

$$R=(M\times N)/(m\times n) \tag{1}$$

According to the value Rdivide the host image into small blocks.

If *R*<8, the size of watermark image extends the capacity of host image, and algorithm stop;

If *R*≥8, the watermark image can be embedded into the host image. The larger R is, the less the host image data will be lost.

Now embed the watermark data into the divided small blocks in the following manner. The 8^{th} row secondary LSB data of the host image block

replace the LSB data of the same row, and the 8^{th} row original LSB data is discarded. Then, the 7^{th} row secondary LSB data replace the 8^{th} row secondary LSB data, the 6^{th} row secondary LSB data replace the 7^{th} row secondary LSB data……, the 1 row secondary LSB data replace the 2^{nd} row secondary LSB data, and one byte of watermark data replace the 1 row secondary LSB data of the host image block [14].

4.2 Extraction Process

In the extraction algorithm firstly the size of watermark is extracted from the head of the host image data. Now calculate the value of *R* by (1), and according to these values of *R* divide the watermarked image into small blocks. Then extract the 1st row secondary LSB data of every block and recompose the watermark image. Now recover the original host image. This process is executed as:

The 2nd row secondary LSB data replaces the 1 row secondary LSB data, the 3rd row secondary LSB data replaces the 2nd row secondary LSB data, the 4th row secondary LSB data replaces the 3rd row secondary LSB data……, the 8 row secondary LSB data replaces the 7th row secondary LSB data, and the 8th row LSB data replaces secondary LSB data of the same row [14].

So this technique presents near reversible watermarking algorithm based on the secondary LSB replacement, which allows near lossless recovery of the original host image [14].

5 Discussion on Features and Capabilities

EPR data hiding using Class Dependent Coding Scheme can be used as effective coding scheme for hiding information in medical images which provides better perceptual quality of stego image along with increase in embedding capacity. This technique enhances the robustness against various attacks like JPEG compression, image tampering and image manipulation [10]. This technique also provides flexibility to doctors in selecting the critical area of the medical image as ROI. Also any tamper in the ROI can be easily detected [10]. If lossless compression is applied on stego image and the stego image is not altered before authentication, then hidden data, original image as well as hash bits can be extracted correctly. Else if compression is not so severe then hidden data can be extracted correctly but original image cannot be recovered [10]. So this technique although provide high data hiding capacity but robustness against intentional attacks is reasonable.

Reversible data hiding using 2-dimenstional difference expansion technique can be used for hiding patient's data in large volume and provide authenticity to medical images. This technique not only detects the locations of tampered areas inside ROI of watermarked medical images but also recover the contents of tampered areas. Very good performance is achieved in the terms of hiding capacity and visual quality using this technique [3].

Near reversible watermarking algorithm based on secondary LSB replacement allows near lossless recovery of original host image. This algorithm not only can recover the original host image to a high extent, but also have good performance in robustness, hiding ability and computing complexity. The embedding capacity of this algorithm is mainly decided by the ratio between the size of the host image and watermark, so this algorithm provides low embedding capacity, which is main limitation of this algorithm. So when *R* is less than 8, the algorithm may fail [14].

References

1. Xu, W., Huang, H., Yu, W.: A Study of Medical Image Tamper Detection. In: International Conference of Medical Image Analysis and Clinical Applications (2010)
2. Sun, X., Bo, S.: A Blind Digital Watermarking for color Medical Images Based on PCA. IEEE Transactions (2010)
3. Al-Qershi, O.M., Khoo, B.E.: ROI-Based Tamper Detection and Recovery for Medical Images Using Reversible Watermarking Technique. IEEE Transaction (2010)
4. Anand, D., Niranjan, U.C.: Watermarking medical images with patient information. In: Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 703–706 (1998)
5. Al-Qershi, O.M., Khoo, B.E.: Authentication and Data Hiding Using a Reversible ROI-based Watermarking Scheme for DICOM Images. In: Proceedings of International Conference on Medical Systems Engineering (ICMSE), pp. 829–834 (2009)
6. Zain, J.M., Fauzi, A.R.M., Aziz, A.A.: Clinical Evaluation of Watermarked Medical Images. In: Proceedings of the 28th IEEE EMBS, Annual International Conference, pp. 5459–5462 (2006)
7. Dharwadkar, N.V., Amberker, B.B., Supriya, Panchannavar, P.B.: Reversible Fragile Medical Image Watermarking with Zero Distortion. International Conference on Computer Science and Communication Technology (2010)
8. Lestriandoko, N.H., Wirahman, T.: Reversible Watermarking Using Difference of Virtual Border for Digital image Protection. In: International Conference on Distributed Frameworks for Multimedia Applications (DFmA) (2010)
9. Krishna, S.L.V., Abdul Rahim, B., Shaik, F., Soundara Rajan, K.: Lossless Embedding Using Pixel Differences and Histogram Shifting Technique. IEEE Transaction (2010)
10. Dhavale, S.V., Patnaik, L.M.: High Capacity, Robust Lossless EPR Data Hiding Using CDCS with ROI Tamper Detection. In: Int'l Conf. on Computer and Communication Technology (2010)
11. Al-Qershi, O.M., Ee, K.B.: Authentication and Data Hiding Using a Reversible ROI-based Watermarking Scheme for DICOM Images. World Academy of Science, Engineering and Technology 50 (2009)
12. Mali, S.N., Dhavale, S.V.: High Capacity Secured Adaptive EPR Data Hiding with Integrity Checking using CDCS. International Journal of Computational Sciences 3(6), 657–668 (2009)
13. Ni, Z., Shi, Y.Q., Ansari, N., Su, W.: Robust Lossless Image Data Hiding Designed for Semi-Fragile Image Authentication. IEEE Transactions on Circuits and Systems for Video Technology 18(4), 497–509 (2008)
14. Zhang, B., Xin, Y., Niu, X.-X., Yuan, K.-G., Jiang, H.-B.: A Near Reversible Image watermarking Algorithm. IEEE Transactions (2010)

Effective Site Finding Using Qualified Link Information

K.M. Annammal[1,*], A. Siva Sundari[1], N. Jaisankar[1], and J. Sugunthan[2]

[1] Anna University, MNM Jain Engineering College
[2] Anna University, Infant Jesus College of Engineering and Technology

Abstract. This paper describes a link-based technique for automating the detection of Web spam, that is, pages using deceptive techniques for obtaining an undeservedly high score in search engines. The problem of Web spam is widespread and difficult to solve, mostly due to the large size of the Web that makes many algorithms infeasible in practice. We propose spam detection techniques that only consider the link structure of Web, regardless of page contents. In particular, we compute statistics of the links in the vicinity of every Web page applying rank propagation and probabilistic counting over the Web graph and C5.0 classification algorithm. These statistical features are used to build a classifier that is tested over a large collection of Web link spam.

Keywords: Content analysis, information retrieval, link, spam, Web spam detection.

1 Introduction

The Web is nowadays both an excellent medium for sharing information, as well as an attractive platform for delivering products and services. This platform is, to some extent, mediated by search engines in order to meet the needs of users seeking information. Given the vast amount of information available on the Web, it is customary to answer queries with only a small set of results (typically 10 or 20 pages at most). Search engines must then rank Web pages, in order to create a short list of high-quality results for users. The Web contains numerous profit-seeking ventures, so there is an economic incentive from Web site owners to rank high in the result lists of search engines. All deceptive actions that try to increase the ranking of a page in search engines are generally referred to as Web spam or spamdexing.

A Web search engine must consider that "any evaluation strategy which counts replicable features of Web pages is prone to manipulation" [20]. In practice, such manipulation is widespread, and in many cases, successful. The authors of [9] report that : "among the top 20 URLs in our 100 million page PageRank calculation were pornographic, and these high positions appear to have all been achieved using the same form of link manipulation". In general, we want to explore the neighbourhood of a page and see if the link structure around it appears to be artificially generated with the purpose of increasing its rank. We also want to verify if this link structure is

* Corresponding author.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 523–529, 2013.

the result of a bounded amount of work, restricted to a particular zone of the Web graph, under the control of a single agent. This imposes two algorithmic challenges: the first one is how to simultaneously compute statistics about the neighbourhood of every page in a huge Web graph, and the second is what to do with this information once it is computed, and how to use it to detect Web spam and demote spam pages. In this paper we adapt two link-based algorithms, and apply them for detecting malicious link structures on large Web graphs. The main contributions of this paper are:

• We introduce a damping function for rank propagation [1] that provides a metric that helps in separating spam from non-spam pages.
• We propose a new technique for link spam detection that exploits the distribution of the number of Web page supporters with respect to distance. To this purpose, we present an improved approximate neighbour hood counting algorithm [21].
• We describe an automatic classifier that only uses link attributes, without looking at Web page content, still achieving a precision that is equivalent to that of the best spam classifiers that use content analysis. This is an important point, since in many cases spam pages exhibit pretty "normal" contents.

Our algorithms are tested on a large sample of the .uk domain where thousands of domains have been inspected and manually classified as spam or non-spam domains. Concerning the algorithmic contribution of our work, we propose an approximate counting technique that can be easily "embedded" within a rank computation. This sheds new light and simplifies the method proposed in [21], suggesting that the base rank propagation algorithm provides a general frame work for computing several relevant statistics in large networks. All our algorithms also work in a streaming model: Every computation involves a limited number of sequential scans of data stored in secondary memory [17]. We also provide a framework in which the algorithms use an amount of main memory in the order of the number of nodes, whereas an amount of memory in the order of the number of edges may not be feasible. In conclusion, our methods are scalable to deal with Web graphs of any size. In [4] we study several combinations of the algorithms presented in this paper with other link-based metrics, showing a detailed breakdown of the different techniques and the classification accuracy achieved in Web spam detection. Here we focus on the description and analysis of the two main algorithms (Truncated PageRank and Probabilistic Counting). For the sake of completeness we also include results on their general performance on the spam detection task.

2 Characterizing Spam Pages

In [15], spamming is defined as "any deliberate action that is meant to trigger an unjustifiably favorable relevance or importance for some Web page, considering the page's true value". A spam page is a page that is used for spamming or receives a substantial amount of its score from other spam pages. Another definition of spam, given in is "any attempt to deceive a search engine's relevancy algorithm" or simply "anything that would not be done if search engines did not exist". There are several techniques for spamming the index of a search engine, or spamdexing. A spam page

may contain an abnormally high number of keywords, or have other text features that make content-based spam detection [21,8] possible. A link farm is a densely connected set of pages, created explicitly with the purpose of deceiving a link-based ranking algorithm [2]. In this is called collusion, and is defined as the "manipulation of the link structure by a group of users with the intent of improving the rating of one or more users in the group". This is the kind of page we are interested in. The targets of our spam-detection algorithms are the pages that receive most of their ranking by participating in link farms. A link farm is a densely connected sub-graph, with little relationship with the rest of the Web, but not necessarily disconnected. From non-spam sites by buying advertising, or by buying expired domains used previously for legitimate purposes. It is important to note that there are some types of Web spam that are not completely link-based, and it is very likely that there are some hybrid structures combining link farms (for achieving a high link-based score) and content-based spam, having a few links, to avoid detection. In our opinion, an approach mixing content features, link-based features and user interaction (e.g.: data collected via a toolbar or by observing clicks in search engine results) should work better in practice than a pure link-based method. In this paper, we focus on detecting link farms, since they seem to be an important ingredient of current spam activity. We view our set of Web pages as a Web graph, that is, a graph G = (V,E) in which the set V corresponds to Web pages in a subset of the Web, and every link(x, y) ∈ E corresponds to a hyperlink from page x to page y in the collection. For concreteness, the total number of nodes N = |V | in the full Web is in the order of 1010 [13], and the typical number of links per Web page is between 20 and 30. Link analysis algorithms assume that every link represents an endorsement, in the sense that if there is a link from page x to page y, then the author of page x is recommending page.

3 Qualified Link Analysis

We propose a deep analysis of Web links from the standpoint of quality as defined in [15]. This qualitative analysis has been designed to study neither the network topology, nor link characteristics in a graph.

3.1 Analyzing Web Links

The information retrieval system analyzes the links in a page and extracts several features from that page. The system not only offers information about the number of links whose pointed page can be recovered using information from the link and the page that contains it, but also data about every link. This system is based on classical information retrieval techniques and natural language processing, and it mainly consists of two stages.

3.1.1 Extraction of Relevant Information on a Link

There are many works which have analyzed the importance of the anchor text as a source of information , thus we use the anchor text as the main source of information to

recover a link. However, there are many cases in which the anchor text does not contain enough information. For this reason, the system performs a terminology extraction from other sources of information such as the URL, the page that contains the link, the context of the anchor text, and a cached page version of the analyzed link that can be stored in a search engine (Yahoo!) or digital library The horizontal axis corresponds to the number of recovered links minus the number of non recovered. The vertical axis corresponds to density. The system uses several terminology extraction approaches based on frequency and statistical language modeling (specifically KL divergence), depending on the source of information considered. Specifically, TF-IDF is used to extract terminology from the page that contains the link, which could be unrelated to the searched page, and KL divergence to extract terminology from the cache page, which, if it exists, contains relevant terminology for sure.

3.1.2 Construction of Complex Queries and Request to a Search Engine

The original query is composed of the terms extracted from the anchor text, and this query is expanded using the terms extracted from the other sources of information considered. The different expanded queries are submitted to the selected search engine (Yahoo!), and the top ten ranked documents are retrieved. In this paper, we consider that a link has been recovered if the page pointed by the link is in the set of pages retrieved with some of the queries. Six measures are extracted in this work and described below. We have considered different features for each measure.

Recovery Degree: The most important feature that is extracted thanks to the recovery system is precisely the degree of recovered links. For every page the system tries to retrieve all their links and as result, three values are obtained: 1) the number of recovered links (retrieved within the top ten results of the search), 2) the number of not recovered links, and 3) the difference between both previous values. we can observe that the spam pages concentrate on a separate area of the distribution, which allows us to distinguish them. We can also observe than the rate of recovered links with respect to non recovered is clearly higher in the non spam pages, thus providing a very useful feature for the classifier. The degree of recovered links can be understood as a coherence measure between the analyzed page, one of its links, and the page pointed by this link

4 Experiments

4.1 Classification Algorithms

For the classification tasks, we have used theWeka software because it contains a whole collection of machine-learning algorithms for data-mining tasks. The first step to obtain the best results in the classification task is to select the most appropriate classifier. We selected different classification algorithms to evaluate the introduced features. In particular, we have chosen the following classification algorithms: Metacost, a cost-sensitive wrapper algorithm that takes the base classifier decision tree with bagging C5.0; Naive Bayes, a statistical classifier based on the Bayes'

theorem using the joint probabilities of sample observations to estimate the conditional probabilities of classes given an observation; Logistic Regression, a generalized linear model to apply regression to categorical variables; and finally, SVMs which aim at searching for a hyperplane that separates two classes of data with the largest margin.

We performed an extensive evaluation with these classifiers that are implemented in theWeka toolkit.We used theWeka J48 implementation of a decision tree, the Naïve Bayes and Logistic Regression algorithms, and the sequential minimal optimization (SMO) implementation of an SVM Polynomial kernel .We used the default options of these algorithms, except in the case of the decision tree for which we set a reduced error pruning method. The algorithmic details of these classifiers are beyond the scope of this paper. Additional information about the classifiers is available in most standard machine-learning texts. Optimizing the algorithm parameters could slightly improve our results, which can, therefore, be considered a lower bound of the performance we could obtain with our approach. The evaluation of the learning schemes used in all the predictions of this paper was performed by a ten-fold cross-validation. For each evaluation, the dataset is split into ten equal partitions and is trained ten times. Every time, the classifier trains with nine out of the ten partitions and uses the tenth partition as test data. We have adopted a set of well-known [7] performance measures in Web spam research: true positive (TP or recall), false positive (FP) rate, and F-measure.

5 Discussion and Conclusion

In this paper, we proposed a new methodology to detect spam in the Web, based on an analysis of QLs and a study of the divergence between linked pages. To use QLs and the LM features effectively, we proposed a C5.0 classifier algorithm. The approach of identifying qualified links by computing a number of similarity measures of their source and target pages. The following limitations can be addressed in future work.

- The classifier and similarity measures being used are quite simple. It is expected that the use of a better classification algorithm and an advanced set of similarity measures would produce a better result. For example, examining the similarity of text in and around a link to its target might fare better, especially for multi-topic hubs.
- The computational complexity of “qualified link analysis” is an issue that requires careful consideration. Although the index of the corpus could be made available before hand, computing the similarity scores is still expensive considering the size of the web. Potential solutions include using fewer features, using features that are easy to compute, and utilizing simple classification algorithms. We tested one possible extension, which builds a thresholding classifier based on anchor text similarity. The classifier simply categorizes the links within the first eight buckets as “qualified links”, and the rest as “unqualified”.

References

1. Araujo, L., Martinez-Romo, J.: Web Spam Detection: New Classification Features Based on Qualified Link Analysis and Language Models 5(3) (2010)
2. Baeza-Yates, R., Castillo, C., López, V.: Pagerank increase under different collusion topologies. In: First Workshop on Adversarial Information Retrieval on the Web (2005)
3. Baeza-Yates, R., Ribeiro-Neto, B.: Modern Information Retrieval. Addison Wesley (May 1999)
4. Becchetti, L., Castillo, C., Donato, D., Leonardi, S., Baeza-Yates, R.: Link-based characterization and detection of web spam. Technical report, AEOLUS/Universitá La Sapienza (2006)
5. Benczúr, A.A., Csalogány, K., Sarlós, T., Uher, M.: Spamrank: fully automatic link spam detection. In: First Workshop on Adversarial Information Retrieval on the Web, Chiba, Japan (May 2005)
6. Cohen, E.: Size-estimation framework with applications to transitive closure and reachability. Journal of Computer and System Sciences 55(3), 441–453 (1997)
7. Davison, B.D.: Recognizing nepotistic links on the web. In: AAAI 2000 Workshop on Artificial Intelligence for Web Search, Austin, Texas, pp. 23–28. AAAI Press (July 2000)
8. Drost, I., Scheffer, T.: Thwarting the Nigritude Ultramarine: Learning to Identify Link Spam. In: Gama, J., Camacho, R., Brazdil, P.B., Jorge, A.M., Torgo, L. (eds.) ECML 2005. LNCS (LNAI), vol. 3720, pp. 96–107. Springer, Heidelberg (2005)
9. Eiron, N., Curley, K.S., Tomlin, J.A.: Ranking the web frontier. In: Proceedings of the 13th International Conference on World Wide Web, pp. 309–318. ACM Press, New York (2004)
10. Baeza-Yates, R., Boldi, P., Castillo, C.: Generalizing PageRank: Damping functions for link-based ranking algorithms. In: Proceedings of SIGIR, Seattle, Washington, USA. ACM Press (August 2006)
11. Castillo, C., Donato, D., Gionis, A., Murdock, V., Ilvestri, F.: Know your neighbors: Web spam detection using the web topology. In: Proc. 30th Annu. Int. ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR 2007), pp. 423–430. ACM, New York (2007)
12. Fetterly, D., Manasse, M., Najork, M.: Spam, damn spam, and statistics: Using statistical analysis to locate spam web pages. In: Proceedings of the Seventh Workshop on the Web and Databases (WebDB), Paris, France, pp. 1–6 (June 2004); Flajolet, P., Martin, N.G.: Probabilistic counting algorithms for data base applications. Journal of Computer and System Sciences 31(2), 182–209 (1985)
13. Gibson, D., Kumar, R., Tomkins, A.: Discovering large dense subgraphs in massive graphs. In: VLDB 2005: Proceedings of the 31st International Conference on Very Large Data Bases, pp. 721–732. VLDB Endowment (2005)
14. Gulli, A., Signorini, A.: The indexable Web is more than 11.5 billion pages. Poster Proceedings of the 14th International Conference on World Wide Web, Chiba, Japan, pp. 902–903. ACM Press (2005)
15. Gyöngyi, Z., Berkhin, P., Molina, H.G., Pedersen, J.: Link spam detection based on mass estimation. Technical report, Stanford University, California (2005)
16. Gyöngyi, Z., Garcia-Molina, H.: Web spam taxonomy. In: First Workshop on Adversarial Information Retrieval on the Web (2005)
17. Gyöngyi, Z., Molina, H.G., Pedersen, J.: Combating web spam with trustrank. In: Proceedings of the Thirtieth International Conference on Very Large Data Bases (VLDB), Toronto, Canada, pp. 576–587. Morgan Kaufmann (August 2004)

18. Haveliwala, T.: Efficient computation of pagerank. Technical report, Stanford University (1999)
19. Lipton, R.J., Naughton, J.F.: Estimating the size of generalized transitive closures. In: VLDB 1989: Proceedings of the 15th International Conference on Very Large Data Bases, pp. 165–171. Morgan Kaufmann Publishers Inc., San Francisco (1989)
20. Mitzenmacher, M., Upfal, E.: Probability and Computing: Randomized Algorithms and Probabilistic Analysis. Cambridge University Press (January 2005)
21. Morris, R.: Counting large numbers of events in small registers. Commun. ACM 21(10), 840–842 (1978)
22. Ntoulas, A., Najork, M., Manasse, M., Fetterly, D.: Detecting spam web pages through content analysis. In: Proceedings of the World Wide Web Conference, Edinburgh, Scotland, pp. 83–92 (May 2006)
23. Page, L., Brin, S., Motwani, R., Winograd, T.: The PageRank citation ranking: bringing order to the Web. Technical report, Stanford Digital Library Technologies Project (1998)

Performance Comparison of On-Demand Routing Protocols for Varying Traffic

Shalini Singh

Graphic Era University, Dehradun, India
shalinisingh111@rediffmail.com

Abstract. Mobile Ad hoc network is an infrastructure-less network comprising of mobile nodes which dynamically form a network without the help of any centralized administration. Since nodes are highly mobile, efficient routing protocols are needed for better performance in such networks. This paper compares the performance of two on-demand routing protocols for mobile ad hoc networks – Dynamic Source Routing (DSR) and Ad Hoc On-Demand Distance Vector (AODV). The performance differentials are analyzed using varying traffic. The performance is evaluated by using Network Simulator (NS-2.1), from viewpoint of packet delivery fraction and average end- to-end delay.

Keywords: Ad hoc networks, wireless networks, mobile networks, routing protocols.

1 Introduction

Mobile ad hoc network (MANET) [2] consists of a set of wireless mobile nodes communicating with each other without any centralized control or fixed network infrastructure. MANET nodes are equipped with wireless transmitters and receivers. In MANET, nodes have to announce their presence periodically and listen for their neighbors announcements broadcast to discover and learn how to reach each other. Therefore, the new node is not familiar with its network topology at the beginning. Mobility and scalability are main challenge in infrastructure- less networks. Where mobility means that often changes the network topology drastically and unpredictably while scalability means more traffic and overhead control packets. In addition, wireless links have been significantly affected by constraint resources such as bandwidth and power. Hence, there is a need of efficient and effective MANET routing protocols to allow the nodes to communicate over multi- hop paths, and should perform acceptably in a dynamic, low bandwidth environment [17]. Many routing protocols for MANET have been proposed [1], to provide routes in such dynamic environment. These protocols can be broadly classified onto three main categories, namely, proactive reactive and hybrid. The objective of this paper is to carry out the systematic performance study of two reactive routing protocols for ad hoc networks- *Dynamic Source Routing* (DSR) [7] and *Ad hoc On – Demand Distance Vector* (AODV) [8, 9]. The key motivation behind the design of on- demand protocols is the reduction of the routing load. High routing load usually has a

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 530–536, 2013.

significant performance impact on low bandwidth wireless links. On the other hand, the on – demand approach do not compute until there is a data to be sent. It may cause longer delay to construct a route, and the connection may not be set up due to the long latency and congested channel [4, 5].

This paper presents an extensive simulation study that compares the performance of two on demand routing protocols for ad hoc networks: AODV [3] and DSR [6, 16]. The performance comparison metrics are Packet delivery fraction and average End – to – End delay for the varying traffic sources. The rest of the paper is organized as: 2) Protocol description of DSR and AODV. 3) Performance Metrics and Simulation Scenario 4) Result and Discussion. 5) Conclusion.

2 Protocol Description

2.1 DSR

The main feature of the DSR is to use source routing. That is, the sender of a packet determines complete route from itself to the destination, including all intermediate hops. These routes are stored in route cache. The data packet carries the source route in the packet header [6, 9]. All intermediate hosts forward the packet based on this predetermined route (called source route). No routing decision is made at intermediate hosts. When a node in the ad hoc network attempts to send a data packet to a destination for which it does not already know the route, it uses the route discovery process to dynamically determine such a route. Route discovery process is straight forward; the initiator flooded the route request (RREQ) packet across the network. A RREQ packet contains the address of destination host as well as a route record which records the host that the request has passed. Each node receiving a RREQ, rebroadcast it, unless it is the destination or it has a route to the destination in its route cache. In both cases, the complete route from the initiator to the destination is found. This route is then replies to the initiator by sending back the route reply (RREP) packet. The route request builds up path traversed so far. The path carried back by the RREP packet is cached at the source for further use. If any link on a source route is broken, the source node is notified using a route error (RERR) packet. The source removes any route using this link from its cache. A new route discovery process must be initiated by the source, if this route is still needed.

2.2 AODV

Ad Hoc On- Demand Distance Vector (AODV) [3] combines the features of Destination Sequenced Distance Vector (DSDV) and Dynamic Source Routing (DSR) protocols [10, 15]. AODV uses destination sequence number like DSDV to determine the freshness of routing information. However, AODV maintains route in a distribution fashion, as routing table entries. The AODV keeps routing table entries in the form of <destination, next hop, distance>. An important feature of the AODV is maintenance of timer – based states in each node, regarding utilization of individual routing table entries [14, 18]. A routing table entry expires if not used recently.

In AODV routing, when a source has data to transmit to a new destination, it broadcast a route request (RREQ) for that destination to its neighbors. RREQ packet contains the address of destination node, sequence number of destination node, broadcasting sequence number, sequence number of source node, address of previous hop count. When an intermediate node receives RREQ information, it would forward route reply (RREP) packet by the reverse routing. RREP containing address of source node, address of destination node, hop count and life time. The RREP is unicasted in hop – by – hop fashion to the source [11]. When the source receives the RREP, it record route to the destination and begins sending data. If multiple route replies are received by the source, the route with the shortest hop count and highest destination number is chosen. In case link break is detected the node at the upstream of route broken would broadcast RERR, which contains the address and sequence number of unreachable nodes to the neighbor nodes. As the route error propagates towards the source, each intermediate node invalidates route to unreachable destinations. When the source node receives the RERR, it invalidates the route and reinitiate route discovery. Sequence number in AODV plays a key role in ensuring loop freedom and freshness of the route [12]. A higher sequence number signifies a fresher route. AODV can handle low, moderate and relatively high mobility rates, as well as variety of data traffic [10].

3 Performance Metrics and Simulation Scenario

The performance comparison of the routing protocols is carried out based on Packet Delivery Fraction (PDF) and Average End-To-End Delay (EED) metrics.

Packet delivery fraction (PDF) is the ratio of data packets delivered to the destination to those generated by CBR sources. This metric actually tells us how much reliable the protocol is. It describes the loss rate that will be seen by the transport protocol, which in turn affects, the maximum throughput of the ad hoc network can support.

$$\text{PDF}=\frac{\sum \text{CBR Received Packets by CBR destination}}{\sum \text{CBR Sent Packets by CBR Sources}} x100$$

Average end- to- end delay (EED) includes all possible delays caused by buffering during route discovery latency, queuing at the interface queue, retransmission delays at the MAC propagation and transfer time. The Average EED is computed as follows.

$$EED = \frac{1}{N}\sum_{n=1}^{N}(r_n - s_n)$$

s_n=Time that data packet n was sent
r_n=Time that data packet n was received
N = Total number of data packets received

A simulated network scenario for variety of traffic loads [10, 13] were implemented on NS-2.1. The simulation parameters and their values are given in table 1.

Table 1. Simulation Parameters

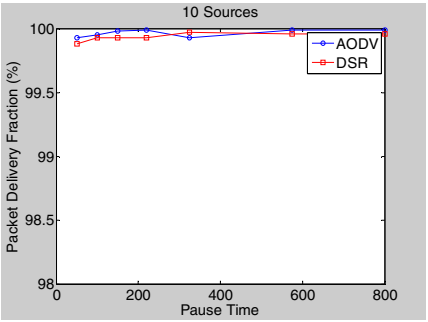
Parameter	Value
Node density	50
Area	500mx500m
Simulation Time	900 sec.
Mobility model	Random Waypoint
Pause Time	50,100,150,220,325,575,800
Node speed	0-5m/s
Wireless MAC	802.11
Protocol used	DSR, AODV
Traffic Source	CBR, 4packet/sec, 512 byte
Number of Sources	10, 18, 32, 45

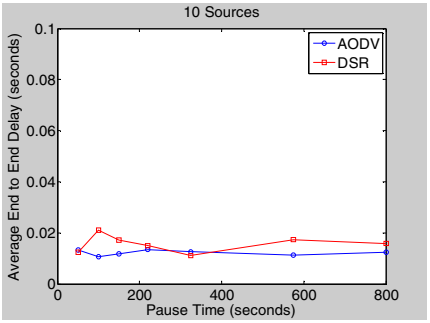
The 50 node density is taken in the square area of 500 m x 500 m. The simulation was carried out by the use of varying number of sources (10, 18, 32 and 45) with the moderate packet rate and changing pause time. The simulation is run for 900 seconds.

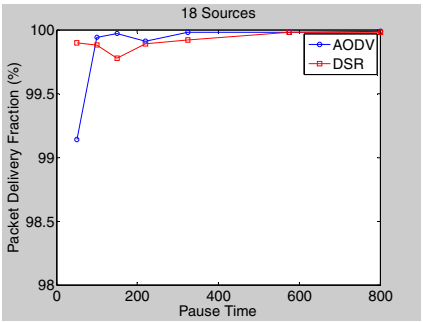
4 Result and Discussion

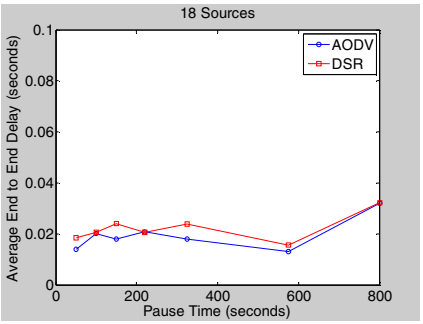
The Figure 1 shows that the packet delivery fractions for DSR and AODV are very similar for 10 and 18 sources. The DSR outperforms AODV for high pause time at low traffic load as 10, 18. Since number of connections and mobility of nodes are low, and the entire route to the destination is available in DSR cache. Hence the packets were delivered to the destination faster than AODV. In contrast AODV does not have entire route to the destination and utilizes hop-by-hop route. But when the sources become more, DSR with small pause time which means high mobility began to perform worse than AODV. Since AODV have more routing control packets, but may always choose the fresh route, while DSR with smaller number of routing packets under stressful situation with the network topology continue to change from time to time, will be inclined to choose wrong routes, thus lower the packet delivery rate.

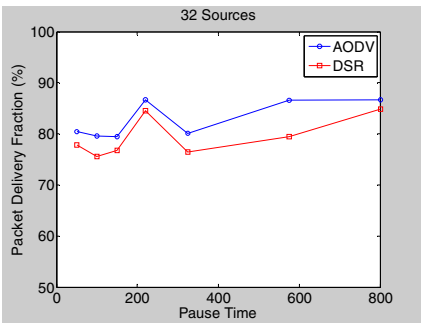
Figure 2 shows that with the low traffic of 18 sources AODV have more delay than DSR, because AODV has much more routing packets than DSR, and those routing packets will consume more bandwidth. When load become heavy which is 32 sources, DSR with small pause time have more delay time since stale routes often be choose, which lost many delivery time. The delay for both the protocols increases with 45 sources at low mobility. This is due to the high level of network congestion and multiple access interference at certain regions of the ad hoc network. This phenomenon is less visible with higher mobility where traffic automatically gets more evenly distributed due to source movements.

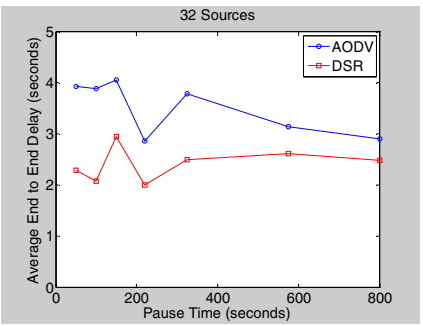


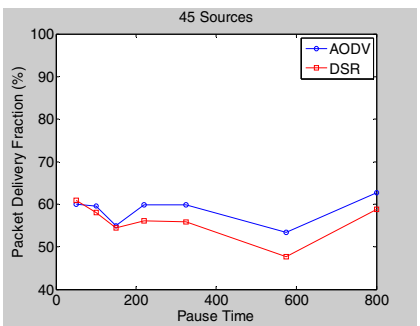












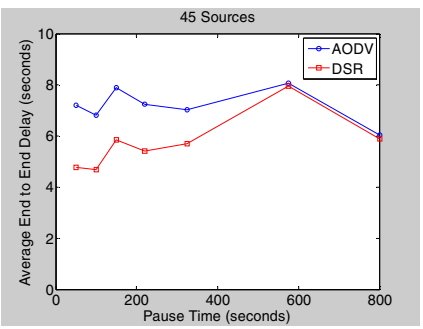


Fig. 1. Packet delivery fraction for the 50 nodes model for varying number of sources

Fig. 2. Average end-to end delay for the 50 nodes model for varying number of sources

5 Conclusions

The general observation from the simulation is that for application oriented metrics such as delay and packet delivery fraction, DSR outperforms AODV in less 'stressful' situations, i.e., smaller number of nodes and lower load. AODV however outperforms DSR in more stressful situations.

The poor delay and PDF performances of DSR are mainly attributed to aggressive use of caching and lack of any mechanism to expire stale routes. Aggressive caching, however seems to help DSR at low loads and also keeps its routing load down. The mechanism to expire routes and/or determine freshness of routes will benefit DSR's performance significantly. On the other hand AODV keeps track of actively used routes, destination also can be searched using a single route discovery flood to control routing load, which is timer –based.

References

1. Royer, E., Toh, C.: A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Network. IEEE Personal Comm., 46–55 (April 1999)
2. Corson, S., Macker, J.: Mobile Ad Hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations. IETF RFC 2501 (January 1999)
3. Cagalj, M.: Performance Evaluation of AODV Routing Protocol; Real Life Measurement. LCA, EPFL. Alexander Zurkinden, SCC (June 2003), http://Icawww.epfl.ch/Publications/Cagalj/ZurkindenCH03.pdf
4. Wu, H., Qiao, C., De, S., Tonguz, Q.: Integrated Cellular and Ad Hoc Relaying Systems: iCAR. IEEE Journal on Selected Areas in Communications 19(10), 2105–2115 (2001)
5. Das, S.R., Castaneda, R., Yan, J.: Simulation- based Performance Evaluation of Routing Protocols for Mobile Ad Hoc Networks. Department of Electrical and Computer Engineering and Computer Science, University of Cincinnati, OH 45221- 0030 USA
6. Borch, J., Maltz, D.A., Johnson, D.B., Hu, Y.-C., Jetcheva, J.: Performance Comparison of Multihop Wireless Ad Hoc Network Routing Protocols. In: Proc. Mobicom, pp. 85–97 (1998)
7. Johnson, D.B., Maltz, D.A.: Dynamic Source Routing in Ad Hoc Wireless Networks. In: Mobile Computing, vol. 353. Kluwer Academic (1996)
8. Perkins, C.E., Royer, E.M.: Ad Hoc On Demand Distance Vector Routing. In: Proc. Workshop Mobile Computing Systems and Applications, CWMCSA 1999 (February 1999)
9. Perkins, C.E., Royer, E.M., Das, S.R., Marina, M.K.: Performance Comparison of two On – Demand Routing Protocols for Ad Hoc Networks. IEEE Personal Communications, 16–28 (February 2001)
10. Network working Group RFC 3561 (July 2003)
11. Yu, D., Li, H.: A Model for Performance Analysis of a Mobile Ad Hoc Networks. Siemens AG, ICM N PG SP RC FR, Gustav-Heinemann–Ring 115, 81730 Munchen, Germany
12. Sirivianos, M., Leontaris, A.: Comparative Evaluation of Ad Hoc Routing Protocols in Highly Dynamic Environments. University of California, Irvine
13. The network simulator- ns-2, http://www.isi.edu/nsnam/ns
14. Das, S.R., Perkins, C.E., Royer, E.M.: Performance Comparison of two On – Demand Routing Protocols for Ad Hoc Networks. In: Proc. INFOCOM, pp. 3–12 (2000)

15. Bai, R., Singhal, M.: DOA: DSR over AODV Routing for Mobile Ad Hoc Networks. IEEE Transaction on Mobile Computing 5(10) (October 2006)
16. Lee, F., Swanson, C.T., Liu, J.: Efficient On – Demand Cache Routing for Mobile Ad Hoc Networks. In: 2nd IEEE International Conference on Computer Science and Information Technology (2009)
17. Conti, M., Giordano, S.: Multihop Ad Hoc Networking: The Theory. IEEE Communications Magazine (April 2007)
18. Wu, L., Qian, X., Dou, W.: Routing protocols for prolonging network's lifetime based on AODV. Computer Engineering and Applications 43(19) (2007)

Analyzing Buffer Occupancy of the Nodes under Acknowledged Delay Tolerant Network's Routing Protocols

Harminder Singh Bindra[1] and A.L. Sangal[2]

[1] Department of Computer Science and Engineering, DR B R Ambedkar National Institute of Technology, Jalandhar
bindra.harminder@gmail.com
[2] DAV Institute of Engineering & Technology, Jalandhar

Abstract. The routing algorithms of DTN have the inbuilt storage management scheme such as Hop based TTL (Spray and Wait) or passive cure (Potential-based Entropy Adaptive Routing PEAR). There has been a significant amount of work in the past regarding buffer management policies. In this paper we have proposed a new message deletion policy for multi-copy routing schemes. This scheme uses the acknowledgement method to remove the useless bundles from the network, preventing the nodes from the buffer overflow problem and avoid transfer of useless message replicas thus relaxing the resources of the nodes. In this paper we have analyzed the buffer occupancy of the nodes under the extended DTN Routing Protocols. Simulative results clearly show that the extended routing protocols proposed in this work greatly relax the buffers of the nodes enabling them to handle more and more messages which in turn will improve the efficiency of the routing protocols.

Keywords: Delay Tolerant Networks, Opportunistic Network Environment (ONE), Epidemic, Prophet, MaxProp, Spray and Wait.

1 Introduction

Connected graphs are used to model the data networks as in these networks there is always an existence of at least one end to end path between any source destination pair. These networks are assumed to bear certain characteristics like low error probability, bidirectional links connecting the nodes, low delivery latency, infrequent power outages and almost 100% network uptime[1]. However certain emerging wireless networks which are deployed in the challenging conditions do not hold these assumptions. These networks are especially deployed in extreme environment like battlefield, volcanic regions, deep oceans, deep space, developing regions etc. where they suffer challenging conditions resulting in excessive delays, severe bandwidth restrictions, and frequent power outages[2]. Under such extreme conditions, the network connectivity is intermittent and thus the end to end path between source destination pair cannot be guaranteed.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 537–544, 2013.

Networks working under such extreme conditions are referred to in the open literature as a Delay or Disruption-Tolerant Networking (DTN).

Number of routing strategies has been proposed in the past for the delivery of data bundles in the DTN[3, 4]. The significant work in the field of Routing in Delay Tolerant Networks started in year 2000 when Amin Vahdat gave Epidemic Routing Protocol[5]. In year 2003 Anders Lindgren proposed the Prophet Routing Protocol[6]. The operation of PROPHET is similar to that of Epidemic Routing. In this protocol when two nodes meet each other, they exchange summary vectors which in this case also contain the delivery predictability information stored at the nodes. In year 2005 Thrasyvoulus Spyropoulus proposed the Spray and Wait routing protocol[7]. This protocol has two phases: Spray phase that sprays the number (fixed number/user defined number) of copies of bundle in the network and Wait phase in which it waits for one of these nodes meets the destination. In 2006 John Burgess gave MaxProp Routing Protocol[8] for effective delivery of DTN messages. MaxProp is based on prioritizing both the schedule of packets transmitted to other peers and the schedule of packets to be dropped so that the packets or bundles should be delivered to the destination with the optimized number of replicas. In 2007 Ram Ramanathan further improved the Epidemic routing protocol and proposed the Prioritized Epidemic Routing[9] for DTN. In 2008 Padma Munder proposed Immunity based Epidemic routing protocol[10]. In this author proposed to include immunity based information disseminated in the reverse direction once messages get delivered to their destination. In 2011 YU Hai-zheng emphasized on the removal of message redundancy for multi-copy routing in delay tolerant networks and proposed a counter based scheme to control the message redundancy for multi copy routing in Delay Tolerant Networks[11]. The routing algorithms of DTN have the inbuilt storage management scheme such as Hop based TTL (Spray and Wait), passive cure (Potential-based Entropy Adaptive Routing PEAR). Till date there are various buffer management policies which work in these networks for the prevention of buffer overflow problem and excessive resource utilization problem. In our scheme we propose to add the acknowledgement technique above the routing scheme of DTN. So our technique frees up the unnecessary resource usage by deleting the replicas of the delivered messages. We analyzed buffer occupancy of the nodes over our proposed method by simulation on two major DTN's routing algorithms: Spray and Wait[7], and MaxProp[8]. The structure of the paper is as follows. Section 2 presents the message replication in DTN and presence of useless replicas in the network after their delivery to the final destination. In section 3 the methodology for the removal of the useless message replicas is given along with the data structure and algorithm of the acknowledgement process. Section 4 discusses the results and finally section 5 concludes the paper.

2 Routing and Message Replication in DTN

2.1 Problem Definition

It is evident from the literature that the multi-copy routing schemes achieve higher delivery probability as compared to the single copy routing scheme [5-8]. This

improvement is achieved at the cost of higher resource utilization i.e. multi-copy routing protocols requires more buffer space, more bandwidth, incur more overheads and consume other vital network resources[12]. In this section we have tried to analyze the impact of message replicas (which exists in the network after the delivery of message in the multi-copy routing scheme) on the performance of the routing protocols.

In the figure 1-4 the existence of message replicas in the network even after the delivery of the message to the final destination is shown. At time t_1 source 'S' generates a message for the Destination 'D' (fig 1 (a)). At this point of time there exists different disjoint regions of networks and S is connected to only 1,2 and 3 nodes. At time t_2, the source S forwards the message to the node 3 retaining the copy of the forwarded message (fig 1 (b)).

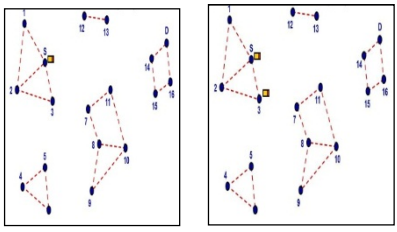


Fig. 1. Network and data bundle status at time instance t_1 and t_2

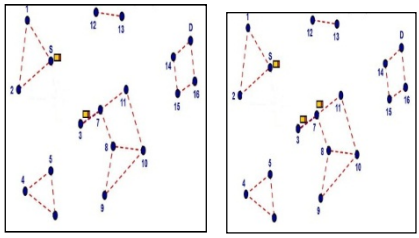


Fig. 2. Network and data bundle status at time instance t_3 and t_4

At time instance t_3, the node 3 moves to the new location and thus the node 3 come in contact of region with nodes 7, 8,9,10 and 11 and hence node 3 forwards the message to node 7 which in turn is forwarded to node 11 at time t_4. In this process the node S, 3 and 7 retains a copy of the forwarded bundle (fig 2(a),(b) & 3(a)).

Then node 11 at time t_6 moves to new point in the network and thus comes in the communication range of node 14. A link is established between 11 and 14 and thus node becomes the member of region with nodes 14, 15, 16 and D.

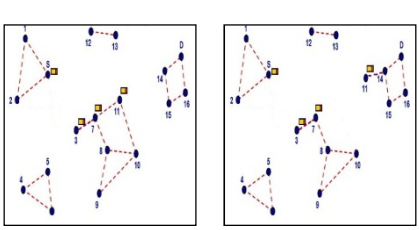


Fig. 3. Network and data bundle status at time instance t_5 and t_6

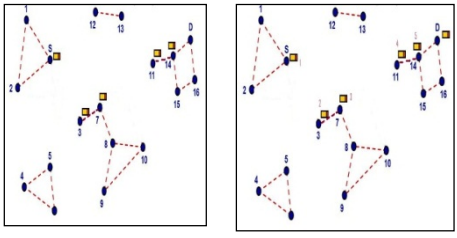


Fig. 4. Network and data bundle status at time instance t_7 and t_8

The message is forwarded to node 14 by node 11 keeping the copy of the original message. This message is forwarded to node D (Destination) by node 14 again retaining the copy of it. At time t_8 the message is delivered to the intended destination. But if we see the figure and the state of the network we can identify that in this process there exists 5 replicas of this delivered message which are now useless and are utilizing the resources. These messages will keep on utilizing the resources and will participate in further replications as these nodes which are holding these replicas are not aware of the delivery of the message to the final destination. ***So there is great need of removing these replicas from the network so as to free up the held resources***. And this can be achieved if the delivery information is communicated back to these nodes.

3 Methodology

3.1 Acknowledgement Mechanism Overview

In case of the multi copy routing scheme the countless replicas of the bundle keep on utilizing the resources held by them even after the delivery of the data bundles to the final destination[13]. After the delivery of the data bundles these replicas become useless and hence there further replication or forwarding will not help in improving the efficiency of the network. Rather they had negative impact on the performance of the network as they continue to utilize the valuable resources of the network. In this paper we have proposed the scheme for the deletion of these useless replicas so as to improve the DTN routing protocols performance.

We have incorporated the acknowledgment from the destination node to communicate the delivery information of the message to the other nodes which are carrying the replicas of the delivered messages. This information is used to remove the useless replicas from the networks thus relaxing the resources of the nodes and preventing the buffer overflow problem.

The section 3.2 details about the data structures required for the implementation of the acknowledgement algorithm

3.2 Data Structures Required

The following are the data structures required to implement the acknowledgement mechanism:

3.2.1 deliveredmessages

This is the hash map maintained by each router and contains the messages received by this router as a final recipient.

3.2.2 del_msgs

This is the list maintained by each router and contains the collection of all the known delivered messages in the network. Whenever two nodes comes in contact of each other they exchange there del_msgs list with each other updates its own del_msgs.

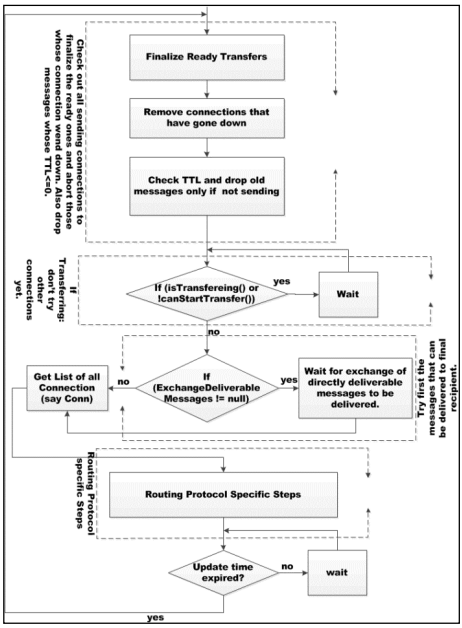


Fig. 5. Flowchart for the unacknowledged communication in DTN

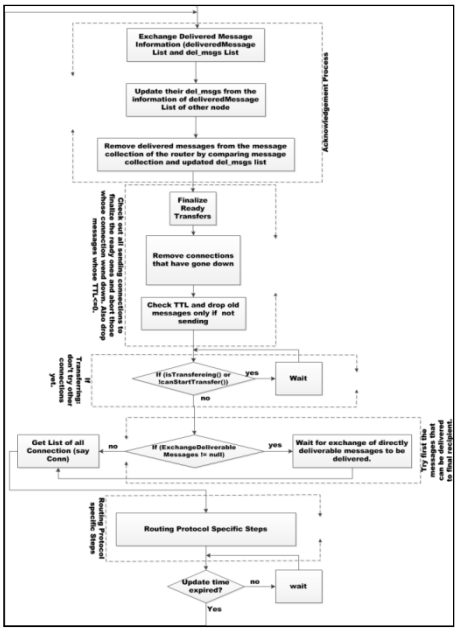


Fig. 6. Flowchart for the acknowledged communication in DTN

The flow chart of the normal communication for the DTN routing protocols is shown in Fig 5.

This communication process basically consists of four major sub processes i.e. (1) Check out all sending connections to finalize the ready ones and abort those whose connection went down. Also drop messages whose TTL<=0. (2) If transferring: don't try other connections yet. (3) Try first the messages that can be delivered to final recipient. (4) Routing Protocol specific Steps.

Fig 6 depicts the flowchart with the acknowledgement process added above these four basic communication processes.

4 Results and Discussion (Analysis of Buffer Occupancy with Respect to Simulation Time)

The analysis of buffer occupancy of the nodes is done by using ONE Simulator. ONE is an agent-based discrete event simulation engine. At each simulation step the engine updates a number of modules that implement the main simulation functions. The main functions of the ONE simulator are the modeling of node movement, inter-node contacts, routing and message handling. The detailed simulation parameters used to carry out this analysis are summarized in table 1.

Table 1. Simulation Parameters

Parameter	Value
Simulation Time	43200 sec (12Hours)
Interface Transmission Range	30m
No of Host Groups	1
No of Hosts Per Group	50
Group Router	MaxProp; Spray And Wait
Buffer Size (Varied)	25M
Message TTL	120 Min
Movement Model	ShortestPathMapBasedMovement
Message generation Event Interval	one new message every 15 to 30 seconds
Message Size	250kB - 2MB
Wait Time	300-900 sec

The base routing protocols (Spray and Wait, MaxProp) are extended in ONE by adding acknowledgement process defined in flowchart in fig 6.

From Fig 7 we observe that the buffer get fully occupied in the early time of the simulation and for the rest of the simulation period it remained fully occupied.

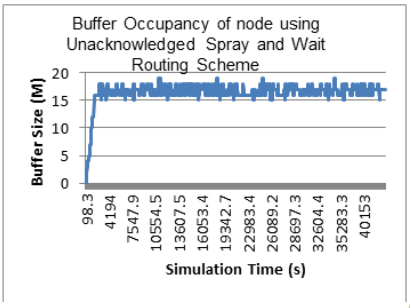


Fig. 7. Buffer occupancy of node using Unacknowledged Spray and Wait Routing Protocol

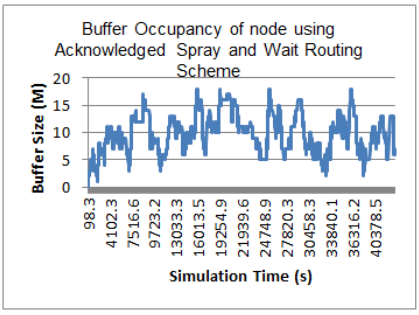


Fig. 8. Buffer occupancy of node using Acknowledged Spray and Wait Routing Protocol

Thus for the rest of the period the new incoming only gets the buffer space when some of the packets are either removed from the buffer due to TTL expiry or due to some of the queue management technique. In this period there are quite high chances that the packets which are already delivered to the destination are holding the buffer space and denying the buffer space to the new incoming undelivered messages.

But from the Fig 8, we see that the buffer gets relaxed due to the removal of the delivered messages because of acknowledgement process. Thus the buffer is available for the most of the time for the new incoming messages and hence they are handled by the router module.

From the Fig 9 and 10 we observe that the buffer is available in both the cases i.e. in unacknowledged and acknowledged MaxProp routing protocol. This is because of the fact the base MaxProp routing protocol has acknowledgment process thus adding the acknowledgement process above it does not have any impact.

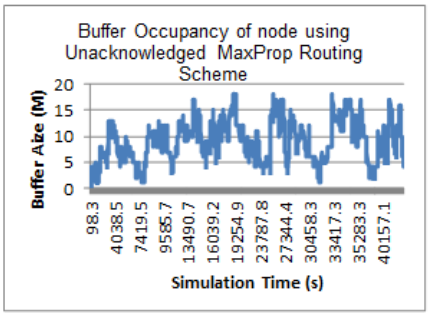


Fig. 9. Buffer occupancy of node using Unacknowledged MaxProp Routing Protocol

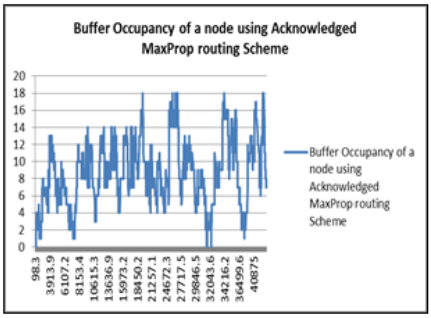


Fig. 10. Buffer occupancy of node using Acknowledged MaxProp Routing Protocol

5 Conclusion

In this paper we have proposed the extension to the base spray and wait routing protocol and MaxProp routing protocol and then analyzed the buffer occupancy of the nodes under the extended and base routing protocols. The simulative results show that the extended Spray And Wait has a much relaxed buffer occupancy as compared to the base Spray and Wait routing protocol. Thus this extended spray and wait routing protocol appears to behave much more efficiently as compared the base spray and wait routing protocol. In future work we will analyze the performance of extended spray and wait routing protocol with respect to delivery probability, Average Latency and Overhead Ratio. Also we will apply this extension to the other multi-copy routing protocols such as Prophet and Epidemic.

References

1. Fall, K.: A delay-tolerant network architecture for challenged internets. In: Proceedings of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, pp. 27–34. ACM, Karlsruhe (2003)
2. Burleigh, S., et al.: Delay-tolerant networking: an approach to interplanetary Internet. IEEE Communications Magazine 41(6), 128–136 (2003)
3. Pelusi, L., Passarella, A., Conti, M.: Opportunistic Networking: Data Forwarding in Disconnected Mobile Ad Hoc Networks. IEEE Communications Magazine 44(11), 134–141 (2006)
4. Zhang, Z.: Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: overview and challenges. IEEE Communications Surveys & Tutorials 8(1), 24–37 (2006)
5. Vahdat, A., Becker, D.: Epidemic Routing for Partially Connected Ad Hoc Networks (2000)
6. Lindgren, A., Doria, A., Schel, O.: Probabilistic routing in intermittently connected networks. SIGMOBILE Mob. Comput. Commun. Rev. 7(3), 19–20 (2003)

7. Spyropoulos, T., Psounis, K., Raghavendra, C.S.: Spray and wait: an efficient routing scheme for intermittently connected mobile networks. In: Proceedings of the 2005 ACM SIGCOMM Workshop on Delay-tolerant Networking, pp. 252–259. ACM, Philadelphia (2005)
8. Burgess, J., et al.: MaxProp: Routing for Vehicle-Based Disruption-Tolerant Networks. In: Proceedings of the 25th IEEE International Conference on Computer Communications, INFOCOM 2006 (2006)
9. Ramanathan, R., et al.: Prioritized epidemic routing for opportunistic networks. In: Proceedings of the 1st International MobiSys Workshop on Mobile Opportunistic Networking, pp. 62–66. ACM, San Juan (2007)
10. Mundur, P., Seligman, M., Jin Na, L.: Immunity-Based Epidemic Routing in Intermittent Networks. In: 5th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks, SECON 2008 (2008)
11. Yu, H.-Z., Ma, J.-F., Bian, H.: Message redundancy removal of multi-copy routing in delay tolerant MANET. The Journal of China Universities of Posts and Telecommunications 18(1), 42–48 (2011)
12. Jain, S., Fall, K., Patra, R.: Routing in a delay tolerant network. SIGCOMM Comput. Commun. Rev. 34(4), 145–158 (2004)
13. Zhensheng, Z.: Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: overview and challenges. IEEE Communications Surveys & Tutorials 8(1), 24–37 (2006)
14. Kernen, A., Ott, J., Kinen, T.: The ONE simulator for DTN protocol evaluation. In: Proceedings of the 2nd International Conference on Simulation Tools and Techniques, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), Rome, Italy, pp. 1–10 (2009)

Multi Infrared (IR) Flame Detector for Tangential Fired Boiler

Himanshu Shekhar[1], Sam Jeba Kumar[2], and Parivesh Singh Rajawat[1]

Essar Project (I) Ltd., 2.SRM University, Chennai
mail2himanshu2009@gmail.com

Abstract . Conventional flame detector sometimes unable to detect the flame flicker due to saturation effect which is called washout. In this state the detector cells maintain high electrical resistance. The continuous background radiation received from the furnace focused on the detector drastically reduces its electrical resistance. This leaves almost no room for the cell to respond to flame flicker modulation. Thus it vote for "flame off" condition to the controller. To limit false alarms caused by hot refractory radiation sources, we propose multi-infrared sensors, which will select specific spectral bands using band-pass filter and being analyzed by advanced mathematical algorithms such as correlations and autocorrelations.

Keywords: Flicker modulation, Flame off, Multi-Infrared, Correlation, Autocorrelations.

1 Introduction

All flames produce electromagnetic radiation. Hence it is important to study the characteristics of flame. It produces radiation in three part of spectrum namely Ultraviolet (UV), Infrared (IR) & Visible region. Only 10% of flame emits radiation in Ultraviolet region rest of the 90% emits radiation in Infrared and visible range. The visible region that we can see is mostly red-yellow caused by the Carbon in a fire. The invisible IR part of the fire we experience as heat. Hydrogen, burns light blue-transparent. It also does not have the CO2 peak at 4.4μ and can therefore is detected in a different way.

The flame produced as a result of combustion is constantly flickering because combustion is a chain of small explosion which change the shape and size of flame. Usually this flicker frequency having the range of 300-600 Hertz .Hence this is very important parameter for designing of Infrared flame scanner. It also discriminate the actual flame From hot refractory radiation and other background Infrared sources. Sometimes the Radiations received from the hot refractory are flicker in nature when it paths is bent, reflected or blocked by ash, carbon or unburned fuels. This happens because the ignition zone of the targeted flame "masks" the bright background low-frequency radiation while the targeted flame is on. When the targeted flame disappears, the background radiation comes into full view.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 545–548, 2013.

2 Conventional Method

When fuel ignites with oxygen during the burning process it initiates many small explosions. Each explosion emits light and IR radiation, giving the flame an appearance of comparatively steady shape and glow. The flame constantly moves – changing shape and brightness. The function of the photo detector is to monitor flame flicker to distinguish between flame and other sources of radiation. The photo detector most commonly used is the PbS (lead sulfide) photo resistor. The PbS cell lowers its electrical resistance relative to amplitude of radiation greater than 400 nm.

3 Multi Infra-red Flame Detector

It comprises more than two Infra-red sensors, sensitive to radiation emitted at various wavelengths in different spectral bands. To limit false alarms caused by hot refractory radiation sources, we use multi-infrared sensors, filtering specific spectral bands along with advanced mathematical algorithms.

3.1 Principle

Every object that has a temperature higher than 0º Kelvin (or -273 ºC) radiates energy and, at room temperature, the energy is already detectable by the most sensitive Infrared sensors. Sometimes, a moving hand close to the sensor is enough to generate an alarm. At 700 K, a hot object already starts to send out visible energy (glowing).

3.2 Methodology

Flickering frequency analysis :Flicker frequency is defined as change in slope Per second .It amplitudes and intensity will vary according to the combustion process.

Fig. 3. Flicker Frequency Graph

$$\mathbf{F_{sen}} = \mathbf{Ff_{on}} - \mathbf{Ff_{off}} \quad (1)$$

Where, $\mathbf{F_{sen}}$= Frequency sensitivity, $\mathbf{Ff_{on}}$= Flicker-frequency of burner on, $\mathbf{Ff_{off}}$ = Flicker frequency of burner off

3.3 Detection Algorithm (Correlations and Autocorrelations)

The spectral emissivity of any radiation sources can be assumed to vary inversely with wavelength irrespective of its chemical composition .Using this assumption ,the apparent temperature of any source can be determined from the measured spectral radiation intensities(I_λ) at two wavelength(900nm & 1000 nm) as

$$\mathbf{T} = \frac{hc}{k} - \left(\frac{1}{\lambda 1} - \frac{1}{\lambda 2}\right) / \ln \left\{\left(\frac{\lambda_2^6}{\lambda_1^6}\right)\left(\frac{I_{\lambda 1}}{I_{\lambda 2}}\right)\right. \quad (2)$$

Where, h=Planck's constant, k=Boltzmann constant, c=speed of the light

The advantage of using two wavelengths close to one another is that assumption of $1/\lambda$ dependence for the emissivity of the source does not seriously affect the apparent source temperature obtained from the equation. From the time series of spectral radiation intensities, measured at 900 & 1000nm,

4 Result and Conclusion

The proposed Multi-IR flame detector will have highest immunity to false alarms which is ideal for multi-burner environment. It will also have highest sensitivity and gain adjustment option unlike conventional detector. Since it uses more than one sensor hence longest detection range.The flame condition can be monitored using mathematical correlation technique.

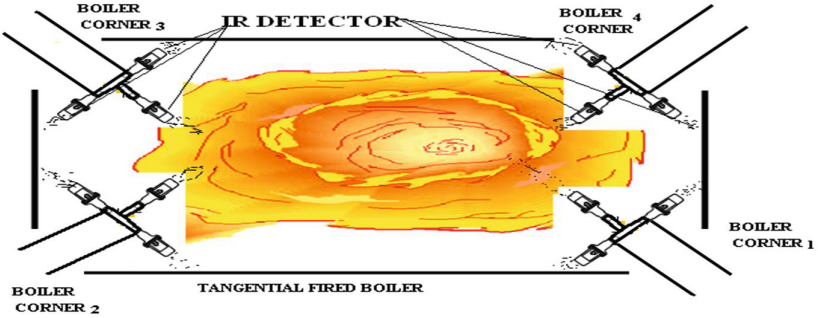


Fig. 1. Arrangement of sensor at boiler corner

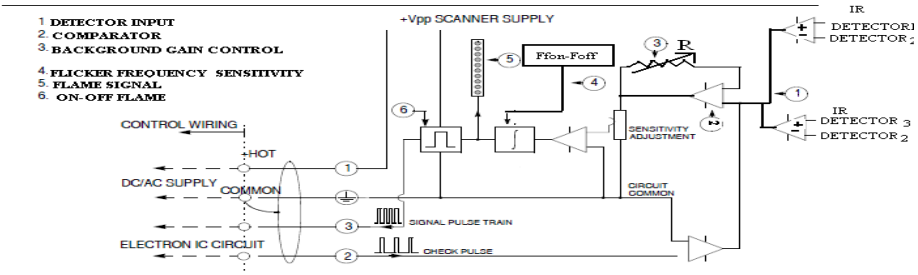


Fig. 2. Multi-IR detector internal arrangement

Acknowledgement. The authors greatly acknowledge the support from the Management of the ESSAR. The work has been carried out as a part of our thermal power project. We would also like to thank our seniors and our colleague for their help and kind support rendered to us.

References

1. Sayre, et al.: Scaling Characteristics of the Aerodynamics and Low NOx Properties of Industrial Natural Gas Burners. Scaling 400 Study–Part IV: The 300 kW BERL Test Results, GRI-94/0186 (November 1994)
2. Sivanthanu, et al.: Miniature Infrared Emission Based Temperature Sensor and Light-Off Detector. In: Conf. Proceedings, Advanced Turbine Systems Annual Program Review Meeting, Poster 5 (October 1997)
3. Savtavicca, et al.: Miniature Infrared Emission Based Temperature Sensor and Light-Off Detector. In: Conf. Proceedings, Advanced Turbine Systems Annual Program Review Meeting, Poster 7 (October 1997)
4. Savtavicca: Miniature Infrared Emission Based Temperature Sensor and Light-Off Detector. In: Conf. Proceedings, Advanced Turbine Systems Annual Program Review Meeting, Poster 9 (October 1997)
5. Khesin, et al.: Continuous On-line Monitoring of Unburned Carbon Case Study on a 650 MW Coal-Fired Unit. In: 1998 USDOE Conf. on Unburned Carbon on Utility Fly Ash. FETC Publications (May 1998)

Management Challenges and Solutions in Next-Generation Networks (NGN)

Maryam Barshan[1] and Maryam Shojaei[2]

[1] Islamic Azad University, Baft branch, Baft, Kerman, Iran
[2] Department of Computer Engineering, University of Isfahan Hezarjarib, Isfahan, Iran
ma_barshan@comp.iust.ac.ir, m_shojaei@eng.ui.ac.ir

Abstract. The next-generation networks should provide users with highly intelligent and integrated services and applications. From the other hand, the increasing numbers of users and services have made the integrated management of the networks complicated. Still traditional management systems are not efficient enough and the heterogeneity and complexity of next-generation networks causes many challenges in network management. Therefore network operators and service providers have to provide users with high quality services. This paper reviews the challenges of management in next-generation networks and the proposed solutions.

Keywords: Management of Next- Generation Network, Next-Generation Network.

1 Introduction

Next-Generation Networks (NGNs) have a low cost and simple architecture with a highly executive power. This architecture that is fairly intelligent and efficient has the ability to provide users with new services and centrally control and manage them. The NGN management is quite different from traditional methods. It requires monitoring and controlling the network services and transfer/ service components. It maintains the network using the management information that obtains through the transmission interfaces between NGNs and the management system and also between the NGN backup management systems and the network components, personnel, service providers and network operators. The current management systems are heterogeneous and complicated and the advent of NGS which has a distributed architecture will increase the complexity. Given the rapid growth of networks, it is needed to have an improved and advanced management service that could overcome the network problems.

In this paper, first in section 2 the next-generation networks are reviewed and compared with the PSTN / IN networks and the functional architecture of next-generation networks is introduced. In Section 3 the current approaches of the next generation network are presented and compared. Then in section 4 the next-generation network management requirements are discussed. In section 5 the standardization in next generation network management and in section 6 the management challenges are discussed. Section 7 describes some proposed approaches for solving management problems and at the end we have concluded the paper.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 549–555, 2013.

2 Review of the Next-Generation Network Features

2.1 Functional Architecture of the Next-Generation Network

Figure 1 shows the functional architecture of the NGNs. According to this figure, the network consists of two layers of service and transmission and the management functions of the layers are shown separately. The NGN interact users, applications and networks (next-generation networks and traditional networks) through UNI, NNI and ANI logical interfaces respectively. In this regard ITU-T has formed some research groups in order to study and do some standardization. There is no general rule and fixed solution for implementing the next-generation network. In order to create this structure the operators evaluate moving towards next-generation networks based on network conditions and the possibilities. For each type of network according to its structure, it is needed to implement a method or combine several methods.

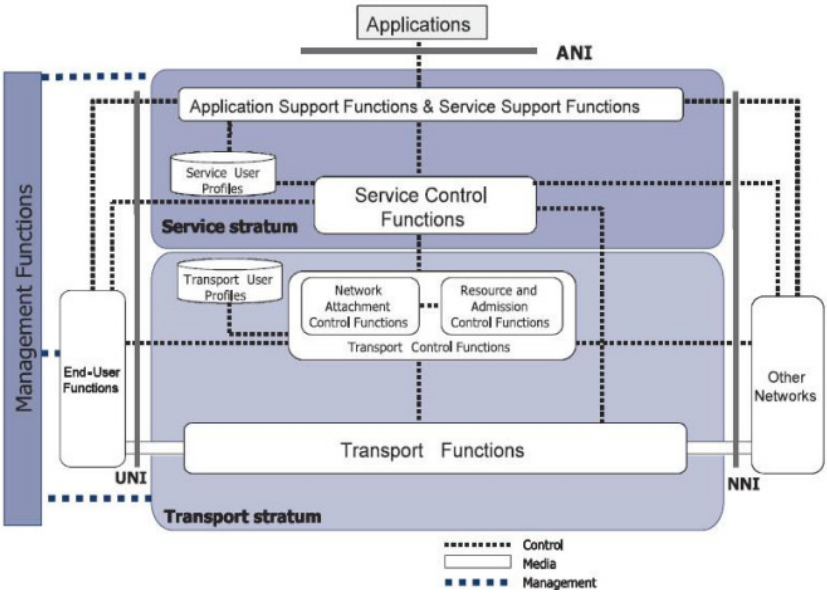


Fig. 1. Functional architecture of next generation network

2.2 Comparison of Next-Generation Internet Network and the PSTN / IN [2]

Next-generation networks and Internet are both packet-based. The structure consists of four separate parts:

- Transfer layer: transmits the digital data and information
- Service layer: create, develop, monitor and manage the services
- Management section: manages the entire next-generation network.
- End-user functions: functions that are implemented in the user's terminal network in order to make it possible to connect to the network.

3 Current Approaches of Network Management

Generally there are two significant technologies for managing the networks. 1) TMN, which was proposed by the ITU-T and is used for telecommunication networks and 2) SNMP that belongs to the IETF and is used in IP networks. In the past decade these two solutions could comply well with the network management requirements. However, as the network technology moves towards next-generation network this scenario will be complicated in the future and the integrated network that will replace the current structure of separate networks will require its own management systems.

3.1 TMN vs. SNMP

TMN is broadly consistent with the telecommunications network and covers a set of standards, including CMIP, GDMO and ANS.1. TMN helps to manage the network in two different ways. 1) Operations and internal communication between different vendors and 2) hierarchical modeling and defining management operations.

SNMP, which is used in current IP based network, uses application layer and UDP protocol in order to exchange management information. SNMP-based applications include two components: the manager and the agent. Each agent has an MIB to store information of managed devices and it is placed beside the management equipments. The manager refers to an entity that interacts with the agents under his management. The manager is located in the vicinity of management applications. Recently RMON and SNMPv3 standards have been added to this protocol. SNMP is a simple and low-cost protocol that has some open Standards. In contrast, TMN based views such as CORBA and CMIP focus on stability and reliability of networks.

4 Processing Models of Next-Generation Network

According to Y.2401, the next-generation network management (NGN-M) has provided management functions for services and NGN resources.

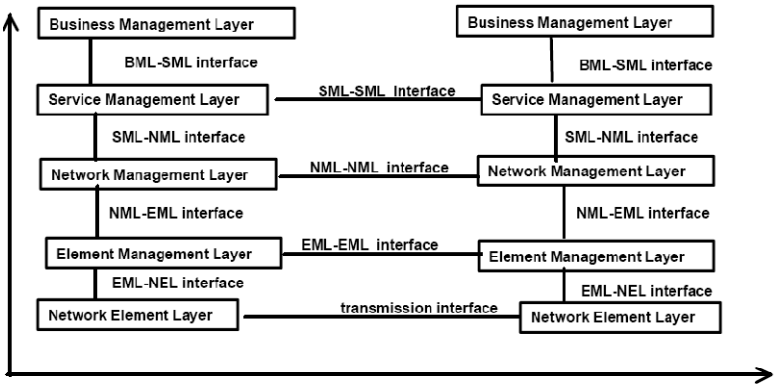


Fig. 2. Management interface FCAPS [10]

Based on NGN-M, the requirements include: fault management, configuration management, accounting management, performance management, security management, mobile management, customer management and the end-node management.

In FCAPS most of the following management parts are related to the wired next-generation networks. These parts includes: fault management, configuration management, audit management, efficiency management, security management and the rest are related to wireless and mobile networks and have to be added.

5 Management Challenges in Next-Generation Network

Although the next-generation networks consist of telecommunication and IP-based networks, its architecture is quite different. Irregularities, competitions and the rapid development of technologies have created challenges in managing and maintaining the network and the services. These challenges include:

1. Dynamic topology: since the next-generation network equipments (such as routers and switches) are programmable, it is possible to control them directly through software and therefore their topology can change easily. The dynamic topology and configuration of NGNs caused many challenges in the traditional network management systems and made them inefficient. In the next-generation networks to manage the dynamic topology of a network that may include hundreds of distributed nodes, a capable and high speed network configuration is needed [2].
2. Multiple services and traffic considerations: given that the NGN provides the quality of service, it should also be able to control and monitor the traffic and services. In TMN until the network communication does not have a fixed rate, the traffic management is not done properly and since the next- generation network is packet based and provides multiple IP-based services, controlling and monitoring the network traffic is an important issue for the service providers [2].
3. Interoperability: the best choice to have interoperability between systems is to support the open network interfaces. An open interface lets the organizations which are outside of the network to access the network without knowing the technical details and how to implement the functions. Internal communications between different operators of next-generation networks and the other networks should be provided through open interfaces [1].
4. Standardization: In the NGN, the most important factor that encourages the service providers to accept and adopt themselves to the new systems and services is the profit increments. In this regard, an agreed set of general requirements, in order to support the next generation core network is needed. Management framework, architecture, information model and management protocols should be standardized and agreed between all the companies [2].
5. Finding the available networks: a terminal should be able to access multiple applications simultaneously. The terminal should be able to find an available network among the heterogeneous networks. A proposed solution is using radio-based software devices that can scan the available networks. After downloading the software from one medium (e.g. a server or a smart card or ...), it has to be configured for the selected network [1].

6. Heterogeneity: in the next-generation network, providing that all the services needed by users are impractical for a service provider, interoperability between heterogeneous entities is important. In addition, for flexible negotiations between individual domains in the next-generation networks, pervasive computing environments is needed and different service providers require to work together to support NGN services. Dealing with the heterogeneous resources is one of the most important challenges of the next-generation network [2].
7. Mobile management: to support vertical and horizontal handover in the NGN, it is required to support mobile management. The challenge occurs when a session with a good quality of service in a network performs handover and goes to another network which has a lower quality of service. In such cases the main question is whether to continue with the low quality of service or put an end to it. In this case a series of agreements between mobile service providers is needed to provide the same services for users who perform handover to other networks [1].
8. QoS Support: recently some researches have been done around the quality of service and NGN protocols, and many solutions have been proposed. Due to the low cost of bandwidth in wired networks, service providers prefer not to get involved in the complicated QoS mechanisms. Although the next-generation networks use higher bandwidth and low-cost channels, because of the high cost of bandwidth in next-generation wireless networks, using this method is not practical and it is certainly needed to support quality of service in the next-generation network so it will remain a major challenge [1].
9. Charging and Billing: Internet access is widely available from ISPs by using a number of access technologies including xDSL, cable modem, FTTH, satellite, and leased lines. The charging models for these traditional Internet access networks are already established and are in use around the world. These established charging models are now being applied with limited success to the new and evolving Internet access networks including WLAN and mobile telecommunication networks offering 2G and 3G services [1].
10. Over the past few years, the Internet and enterprise networks have been plagued by denial of service attacks (DoS), worms and viruses, which have caused millions of computer systems to be shutdown or infected and the stored data to be lost, ultimately causing billions of dollars in loss. The introduction of wireless LANs (e.g., IEEE 802.11) into enterprises has made network security more vulnerable since rogue base stations (i.e., unauthorized private base stations) can easily connect to existing wired networks, potentially becoming the source of security attacks inside firewalls and intrusion detection systems. Furthermore, connecting malicious PCs via a base station that is not well managed is critical as well [1].

6 The Proposed Solutions

- Web service management: In [2] the use of web services to support a large number of distributed services and transparently hide technical implementation details and thereby reduce integration costs is proposed. Because SOAP messages can work with standard Web protocols (such as XML, HTTP, TCP / IP), they can function well across heterogeneous network components. Availability of HTTP and simplicity of SOAP-based XML makes web services

ideal for internal system communications. XML-based approaches have been created for a few years and there are several vendors who have been working on it. Even though they are not still accepted as a standard by major management committee, it has been seen in various products.

- Policy-based Management: Policy-Based Network Management (PBNM) provides a means by which the administration process can be simplified and largely automated. In the traditional network management approaches the uses of network management systems have been limited just to monitor the network status. In such systems, the information model and policy expressions can be made independent of network management protocols by which they are carried policy director for command and management functions as an interface between the acts and policies of the predefined policies.
- In [5] four management model problems are discussed. 1) Models cannot express the same concepts in different implementations. 2) The model should be content-independent. 3) These models express only the current network states. 4) It is difficult to develop these models to adapt with new technologies. Then it has provided a Novell based architecture for end to-end policy management.
- Autonomic network management: [3] believes that autonomic computing seems to be the answer to deal with both the distributed architecture of next-generation services and the economical aspects of reducing costs management. In this paper an autonomic system with four independent features are described: autonomic configuration, autonomic recovery and retrieval, autonomic optimization and autonomic protection to anticipate and resolve intrusions. In their proposed method instead of defining a management layer on the top of service layer, they have considered each service or service component as an autonomic element that is interacting according to an autonomic behavior.
- Another approach is to allow users to manage a part of their service by themselves. This method is called CSM (customer service management). In [8], CSM architecture is presented for next-generation networks. This method reduces management costs and increases the users' conception of services.
- In [7] the service management concepts and challenges to analyze NGS are presented and an architecture for NGS and its management is proposed. The NGS is an advanced telecommunications service. Service management challenges include: complexity and heterogeneity of management in next-generation network, inadequate information model to manage services and lack of standard management interface for different services.
- Seamless Network Management in presence of heterogeneous management protocols: the authors in [4] believe that although web-based service is a suitable solution for network management, it is not efficient enough and it quite suffers some limitations including security management. In the proposed approach, this issue is solved by using the "JCA Container". In this paper a Novel architecture is presented to overcome the complexity of interactions between different management protocols. According to this approach, if the network management changes from one protocol to another one, it is not needed to make any changes in managing the applications.

- A model of performance management in the distributed networks: in [6] a flexible and scalable model is presented to manage the large-scale networks. This model has used web services framework to build the software architecture and XML to build the data exchange interfaces. The policy-based hierarchical event processing mechanism presented in this paper can balance loads and improve system flexibility as well.

7 Conclusion

Most part of this paper is devoted to the challenges of the NGN management. The discussed challenges are: dynamic connectivity, multiple services and traffic considerations, interoperability, standardization, finding available networks, heterogeneity and non-uniformity, mobile management, support and service quality, cost and billing and at the end security challenges. The proposed guidelines for the management problems in NGN such as web services management, policy-based management, automated network management, CSM method and seamless network management in presence of heterogeneous management protocols were provided.

References

1. Choi, M.-J., Hong, J.W.-K.: Towards Management of Next Generation Networks. IEICE Trans. Commun. E90–B(11), 3004–3014 (2007)
2. Li, M., Sandrasegaran, K.: Network Management Challenges for Next Generation Networks. IEEE (2005)
3. Grida Ben, Yahia, I., Bertin, E.: Towards autonomic management for Next Generation Services. IEEE (2006)
4. Sri Karthik, B., Jaiswal, M., Menon, V., Kannan, V., Venkobarao, S., Pande, M., Talukder, A., Das, D.: Seamless Network Management in Presence of Heterogeneous Management Protocols for Next Generation Networks. IEEE (2006)
5. Raymer, D., Strassner, J., Lehtihet, E., van der Meer, S.: End-to-End Model Driven Policy Based Network Management. IEEE (2006)
6. Bin, Z., Tao, H., Wei, W., Zi Tan, L.: A Model of Scalable Distributed Network Performance Management. IEEE (2004)
7. Yahia, I.G., Bertin, E.: Next/new generation networks services and management. In: International Conference on Networking and Services (ICNS 2006), pp. 15–19 (July 2006)
8. Jiang, X., Yang, F., Zou, H.: A novel architecture to customer service management for the NGN. In: International Conference on Communication Technology Proceedings (ICCT 2003), pp. 123–126 (April 2003)
9. Anantharangachar, R., Anantharamaiah, P., Thalapathy, K.: An SOA Compass for Next Generation Networks. IEEE (2006)
10. Sidor, D.: ITU-T Overview of NGN Management, ITU-T Workshop on NGN (jointly organized with IETF) Geneva (May 2005)

Modeling and Dynamic Simulation of Permanent Magnet Brushless DC Motor (PMBLDCM) Drives

P. Ramesh Babu[1], S. Ramprasath[2], and B. Paranthagan[3]

Department of EEE, Saranathan College of Engineering, Trichy, Tamilnadu, India
{rameshbabu-eee,ramprasath-eee,
paranthagan-eee}@saranathan.ac.in

Abstract. This paper deals with modeling and dynamic simulation of PMBLDCM drives using MATLAB and its toolbox Simulink. Mathematical model of whole system which incorporates BLDC Motor, inverter, current controller and speed/torque controller is used for the proposed model of PMBLDCM drives. The current, back emf and torque equations are used to build the mathematical model of BLDC Motor. Inverter is modeled with the help of switching function concept instead of using actual switching device. PI controller is utilized for speed/torque control of PMBLDCM drives. Hysteresis controller is employed for current control of PMBLDCM drives. With the aid of the developed model, the Steady and Transient-state characteristics of speed and torque in addition to voltages and currents of inverter components can be effectively examined and analyzed. This proposed model can be projected to trouble-free design tool for the development of PMBLDCM drives.

Keywords*:* PMBLDCM Drives, Modeling, Dynamic Simulation.

1 Introduction

Brushless DC (BLDC) Motors are able to replace conventional DC Motors in many adjustable speed applications [1]. BLDC Motors are driven by dc voltage but current commutation is done by solid state switches. The commutation instants are decided by the rotor position and the position of the rotor is detected either by position sensors or by sensor less techniques. In order to design BLDC motor drive system, it is necessary to have motor model that gives accurate value of torque which is related to current and back-EMF. Different simulation models [3], [4] have been presented to analyse performance of PMBLDCM drives. Even though all the previous works made a great involvement to modeling of PMBLDCM drives system, the model developed in this paper will be effective for monitoring and analysing the performance of PMBLDCM drives. Hence here a 3 phase, 4 pole, star connected, trapezoidal back-EMF type of BLDC motor has been taken for modeling a complete PMBLDCM drives system subsequently simulated in MATLAB /Simulink.

V.V. Das and Y. Chaba (Eds.): AIM/CCPE 2012, CCIS 296, pp. 556–564, 2013.

2 Modeling of PMBLDC Motor

In PMBLDC Motor, the distribution of flux is trapezoidal in nature [1]. Given the trapezoidal flux distribution, it is sensible to build up a model of the PMBLDC Motor in phase variables. The motor is assumed to be three phase star connected R-L –e (back emf) stator circuit. Due to smooth cylindrical rotor, there is no rotor inductance with angle and assuming three identical phases, therefore self and mutual inductance of three phases are equal. The stator currents are restricted to be balanced and it can be described by the following voltage equations.

$$V_a = i_a R_s + (L - M)\frac{d}{dt} i_a + e_a \tag{1}$$

$$V_b = i_b R_s + (L - M)\frac{d}{dt} i_b + e_b \tag{2}$$

$$V_c = i_c R_s + (L - M)\frac{d}{dt} i_c + e_c \tag{3}$$

Where V_a, V_b, V_c – Stator phase voltages, R_s – Stator resistance, e_a, e_b, e_c - Stator induced emfs, L ,M– Self and Mutual inductances of phase a-b-c, i_a, i_b, i_c – Stator phase currents.

The equations (1), (2) and (3) lead to simplification of PMBLDC Motor model as

$$\begin{bmatrix} V_a \\ V_b \\ V_c \end{bmatrix} = R_s \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} + (L - M) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \frac{d}{dt} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} + \begin{bmatrix} e_a \\ e_b \\ e_c \end{bmatrix} \tag{4}$$

2.1 Mathematical Model of Electromagnetic Torque

In PMBLDC Motor, the induced emfs are trapezoidal, whose peak value, E_p is given as

$$E_p = \lambda_p \omega_r \tag{5}$$

The total electromagnetic torque is given by

$$T_e = (e_a i_a + e_b i_b + e_c i_c) / \omega_r \tag{6}$$

The instantaneous induced emf of phase, a can be written as

$$e_a = f_a(\theta_r)\lambda_p \omega_r \tag{7}$$

Where f_{as} (θ_r) – Rotor position which is dependable for inducing emf in phase, 'a'. λ_p – peak mutual flux linkage of stator winding .Hence the total electromagnetic torque in terms of rotor position is given as

$$T_e = \lambda_p[f_a(\theta_r)i_a + f_b(\theta_r)i_b + f_c(\theta_r)i_c] \text{ N-m} \tag{8}$$

2.2 Electromechanical Modeling

Speed equation in terms of moment of inertia, J in Kg-m^2, friction coefficient, B in N-m/ (rad/sec), load torque, T_L and electromagnetic torque, T_e is given as

$$\omega_r = \frac{1}{J}\int (T_e - T_L - B\omega_r)dt \tag{9}$$

Rotor speed, ω_r, poles, P and rotor position, θ_r are related as

$$\frac{d\theta_r}{dt} = \frac{P}{2}\omega_r \tag{10}$$

3 Block Diagram of the Proposed Model for PMBLDCM Drive

Figure 1 shows the overall control scheme of the PMBLDC Motor drives. The inverter is a six switch voltage source which incorporates constant dc link voltage. The rotor speed is compared with reference speed and the error is amplified by the PI controller then it provides the reference torque, T_{ref} .The current magnitude command is obtained from the torque expression and is compared with stator phase currents. The current errors are amplified and used with hysteresis current controller to produce the switching pulses for inverter.

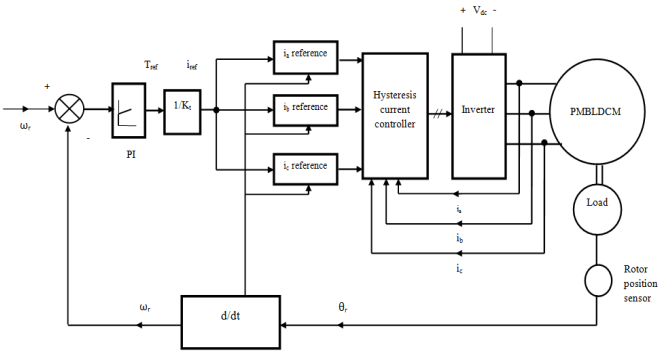


Fig. 1. Control scheme of BLDC Motor drive. With reference to this figure, the proposed Simulink model has been developed and shown in figure 2 which consists of five functional blocks namely back emf block, current generating block, inverter block, hysteresis current controller block and speed-torque controller block.

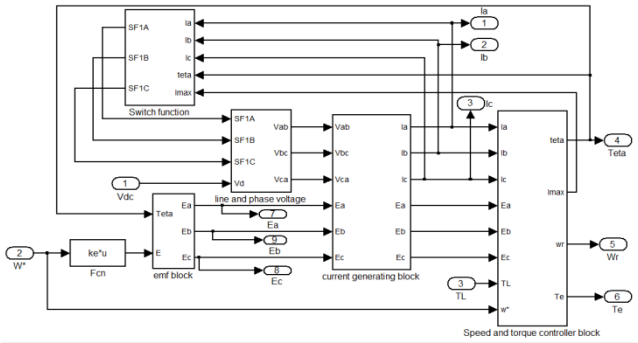


Fig. 2. Block diagram of the proposed model for PMBLDC Motor drive system

3.1 MATLAB/Simulink Model of Back emf Block

The MATLAB function block is shown in figure 3 is implemented to perform the job for producing the trapezoidal back emf wave form of phase, 'a' which depends upon the rotor position. The rotor position ,θ_r peak emf , E_p and $T = 2 * pi$ are the inputs to the MATLAB function block which in turn outputs the back emf voltage, e_a of phase, 'a'. The mathematical expression for back emf, e_a is given by

$$e_a = \begin{cases} (12E_P/T)\theta_r & (0 < \theta_r < T/12) \\ E_P & (T/12 < \theta_r < 5T/12) \\ -(12E_P/T)\theta_r + 6E_P & (5T/12 < \theta_r < 7\,T/12) \\ -E_P & (7T/12 < \theta_r < 11\,T/12) \\ (12E_P/T)\theta_r - 12E_P & (11T/12 < \theta_r < 6T) \end{cases} \quad (11)$$

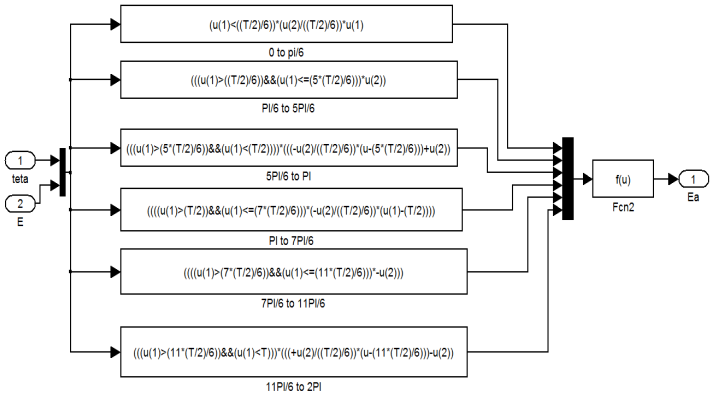


Fig. 3. Back emf, e_a generating block for phase, 'a'. Similarly Back emf generating block for phase, b and phase, c are modeled.

3.2 MATLAB/Simulink Model of Speed and Torque Control Block

The speed and torque control block in MATLAB/Simulink is designed by using the mathematical expression given in equations (9) and (10) .The MATLAB function block is shown in figure 2 is implemented to control speed and torque of the drive system.

3.2.1 Design of PI Controller

PI control is a proportional plus integral controller whose transfer function [5] is:

$$G_{pi}(s) = K_p\left(1 + \frac{1}{T_i s}\right) \tag{12}$$

Where K_p – Proportional gain and T_i – integral time. The selection of the Proportional and Integral, PI controller parameters can be obtained using the Ziegler-Nichols methods which is shown in figure 4

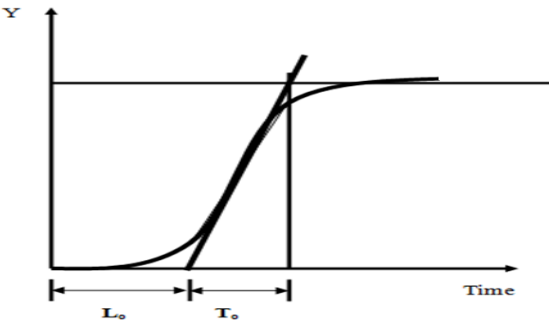


Fig. 4. Ziegler-Nichols PI Parameters. The selection of the Proportional gain and Integral time of PI controller parameters can be obtained using the Ziegler-Nichols method which is shown in figure 4.

With aid of following equations, the PI parameters can be obtained

$$K_p = 1.2\frac{T_o}{L_o} \tag{13}$$

$$T_i = 2L_o \tag{14}$$

3.3 Simulink Model of Hysteresis Current Control and 3Phase VSI Inverter

The MATLAB function block shown in figure 5 is developed to perform the job of the mathematical model of the hysteresis current control [6] and three phase VSI inverter [7]. The i_a, I_{max} and θ_r are the inputs to the MATLAB function block which in

turn outputs the three switching pulses SF_a, SF_b, and SF_c. With the help of voltage equations (1), (2), and (3), the VSI inverter model is simply implemented by using the functional block of Matlab/ Simulink.

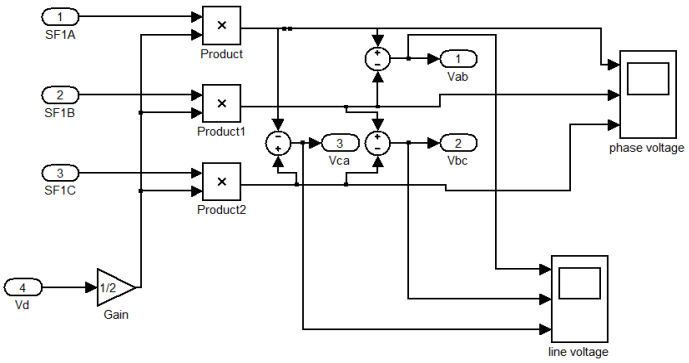
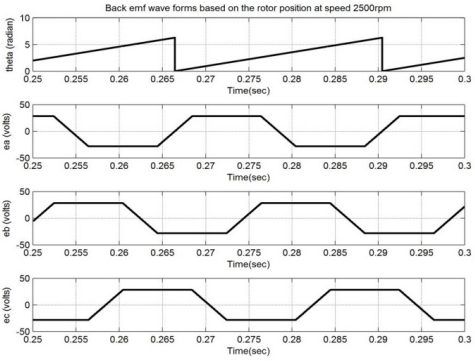
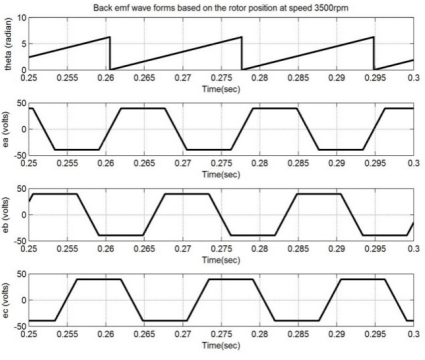


Fig. 5. Simulink Model of Hysteresis control and inverter

4 Analysis and Comparison of Simulation Results



(a) Operating speed: 2500 rpm



(b) Operating speed: 3500rpm

Fig. 6. Simulation results of *(a)* back emf at 2500rpm & *(b)* back emf at 3500 rpm. At 2500 rpm the rotor speed position is varied from 0 to 2 π per electric cycle is 0.024 s, and the back EMF has an amplitude of 28.12 V. On the other hand, at 3500 rpm, the electric cycle is 0.018 s, and the amplitude is 39.36V.

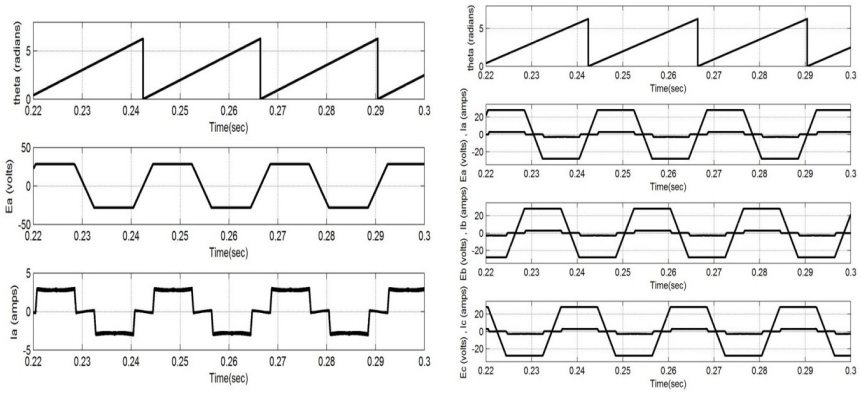


Fig. 7. Synchronization of Phase Currents with back emfs. In this figure the back emfs and phase currents are in well bringing together with one another as per Hysteresis Control algorithm.

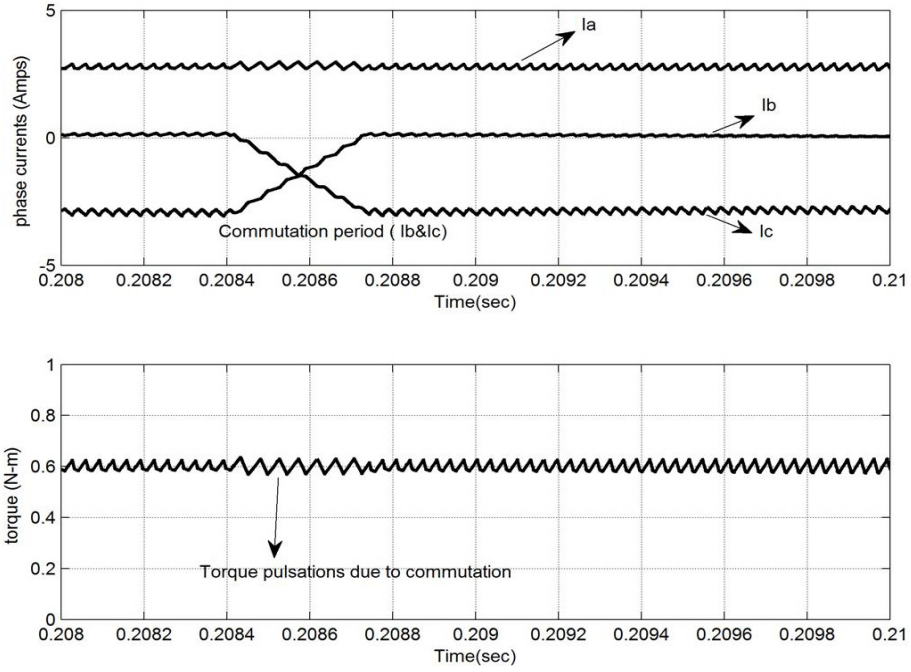


Fig. 8. Phase currents and Torque Pulsations during the commutation period. As shown in fig .9 where the phase current 'b' reaches the I_{max} value before phase current 'c' current departs to zero. As the result, the phase 'a' current decreases, and torque pulsations during the commutation period [8]. Therefore it is noted that in the proposed simulation model, the transient characteristics & steady-state conditions are well observed.

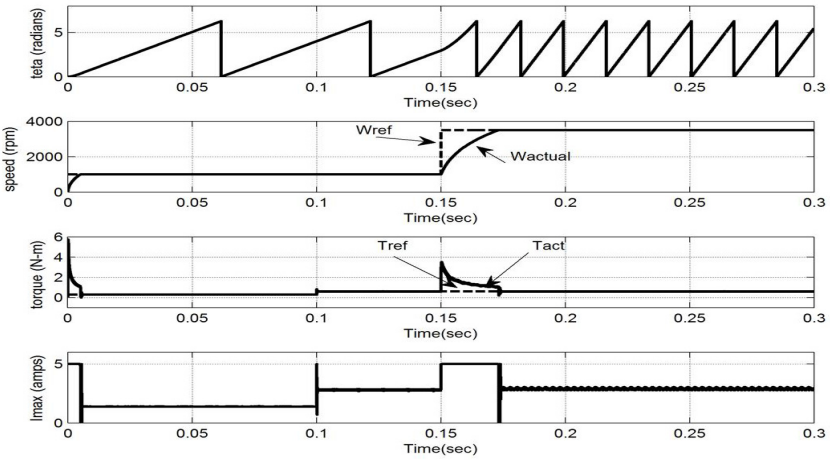


Fig. 9. Shows the dynamic responses of the speed and torque controller. At 0.1s, as per command value the torque changes from 0.3N-m to 0.6N-m but speed remains constant at 1000rpm with corresponding change in I_{max}. At 0.15 s, as per command value the speed changes from 1000 rpm to 3500 rpm within 0.02 s.

5 Motor Data

Rated Power	1 HP	Resistance	0.75Ω /phase
Rated voltage	160 Vdc	Inductance(L+M)	6.1 mH /phase
Rated speed	3500 rpm	Inertia	8.2614e-5 Kg-m^2
Rated Torque	0.6 N-m	Back emf constant	0.1074 V/(rad/sec)
Torque constant	0.21476 N-m	PI controller parameters	K_p=9 ; T_i = 0.01

6 Conclusion

In this paper, modeling and dynamic simulation of PMBLDC Motor drive system has been developed and the performance analysis of drive system has been verified. The dynamic and steady state characteristics are effectively and more accurately analyzed. The proposed model is split into many functional simulation blocks, so that it can be very well applied with a slight modification to other ac machine applications which incorporates the permanent magnet ac motor, the synchronous reluctance motor and the like.

References

1. Krishnan, R.: Electric Motor Drives - Modeling, Analysis and Control, 2nd edn. Prentice-Hall of India Private limited, New Delhi
2. Bimal, K.B.: Modern Power Electronics and Drives. Prentice Hall PTR (2002)

3. Krishnan, R., Pillay, P.: Modeling, Simulation, and analysis of Permanent-magnet motor drives, Part II: the brushless DC motor drive. IEEE Trans. on Industry Applications 25(2), 274–279 (1989)
4. Evans, P.D., Brown, D.: Simulation of brushless DC drives. Pro. of the IEEE 137(5), 299–308 (1990)
5. Al Mashakbeh, A.S.O.: Proportional Integral and Derivative Control of Brushless DC Motor. European Journal of Scientific Research 35(2), 198–203 (2009) ISSN 1450 – 216X
6. Wiechmann, E.P., Ziogas, P.D., Stefanovic, V.R.: Generalized functional model for three phase PWM inverter/rectifier converters. In: Pro. of the IEEE Industry Application Society Annual Meeting, IAS 1985, pp. 984–993 (1985)
7. Ehsani, M., Lee, B.K.: A simplified functional model for 3-phase voltage source inverter using switching function concept. IEEE Trans. on Industrial Electronics
8. Carlson, R., Lajoie-Mazenc, M., Fagundes, D.S.: Analysis of torque ripple due to phase commutation in brushless DC machines. IEEE Trans. on Industry Applications 28(3), 632–638 (1992)

Author Index