Lecture Notes in Electrical Engineering 150

Vinu V. Das
Editor

Proceedings of the Third International Conference on Trends in Information, Telecommunication, and Computing

Springer

Lecture Notes in Electrical Engineering

Volume 150

For further volumes:
http://www.springer.com/series/7818

Vinu V. Das
Editor

Proceedings of the Third International Conference on Trends in Information, Telecommunication, and Computing

Springer

Editor
Vinu V. Das
Saintgits College of Engineering
Kottayam
India

ISSN 1876-1100 ISSN 1876-1119 (electronic)
ISBN 978-1-4614-3362-0 ISBN 978-1-4614-3363-7 (eBook)
DOI 10.1007/978-1-4614-3363-7
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2012943631

Printed on acid-free paper

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

Preface

A warm welcome to the International Joint Conference on Advances in ITC2102 and PEIE2012 on Aug 03–04, 2012, Bangalore, India. The conference intends to provide a forum for worldwide researchers and practitioners to discuss theory, practices, and applications in the fields of Computational Engineering, Computer, Power Electronics, Instrumentation, Control System, and Telecommunication Technology. Our primary objectives are to pursue an in-depth understanding of the fundamental challenges and to formulate a broad vision for the future research direction to support new technologies in practice. To address these objectives, we invited full papers as well as short papers and poster presentations.

The conference received 413 submissions overall. Only 79 papers have accepted and registered to appear in the final proceedings. The Program Committee consists of 34 members from 15 different countries. Every submitted paper received a careful review from the committee and the final accept/reject decisions were made by the co-chairs on the bases of recommendations from the committee members.

As in previous years the conference program includes paper presentations selected, to trigger discussions and exchange ideas. It also provides the possibility for discussions of the papers presented as posters. In addition, the conference organizes two keynote speeches to give the audience a comprehensive overview on emerging technologies.

The conference demonstrates a variety of research that is underway and identifies many interesting challenges and opportunities for further work.

I would like to thank ACEEE, the organizing and program committees of the conference, and all the authors and participants, for their invaluable help in making this conference a successful event.

Dr. Janahanlal Stephen

Contents

Part I
Full Papers

Chapter 1
High Through-Put VLSI Architecture for FFT Computation

S. SreenathKashyap

Abstract Parallel-prefix adders (also known as carry tree adders) are known to have the best Performance in VLSI designs. The Design of the three types of carry-tree adders namely Kogge-Stone, sparse Kogge-Stone, and spanning carry look ahead adder is done and compares them to the simple Ripple Carry Adder (RCA). These designs of varied bit-widths were implemented on a Xilinx Spartan 3E FPGA and power measurements were made with LIBRO. Due to the presence of a fast carry-chain, the RCA designs exhibit better delay performance up to 128 bits. The carry-tree adders are expected to have a speed advantage over the RCA as bit widths approach 256. An Efficient FFT is designed by implementing the adder which consumes low power is replaced in the adder module of FFT.

Keywords FFT · Kogge stone adder · Sparse kogge stone adder · Spanning kogge stone adder · Ripple carry adder · Carry look ahead adder

1.1 Introduction

The saying goes that if you can count, you can control. Addition is a fundamental operation for any digital system, digital signal processing or control system. A fast and accurate operation of a digital system is greatly influenced by the performance of the resident adders. Adders are also very important component in digital systems because of their extensive use in other basic digital operations such as subtraction, multiplication and division. Hence, improving performance of the

S. SreenathKashyap (✉)
M.Tech VLSI Design, SRM University, Chennai, India
e-mail: Kashyap.foru3@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_1,

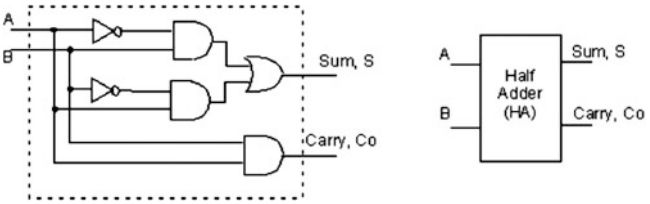


Fig. 1.1 Half adder and Logic Block

digital adder would greatly advance the execution of binary operations inside a circuit compromised of such blocks. The performance of a digital circuit block is gauged by analyzing its power dissipation, layout area and its operating speed.

1.1.1 Types of Adders

The implementation technique of several types of adders and study their characteristics and performance. There are so many types of adders some of them are

- Ripple carry adder
- Carry look-ahead adder
- Carry tree adders

1.2 Basic Adder Unit

The most basic arithmetic operation is the addition of two binary digits, i.e. bits. A combinational circuit that adds two bits, according the scheme outlined below, is called a half adder. A full adder is one that adds three bits, the third produced from a previous addition operation.

1.2.1 Half Adder

A half adder is used to add two binary digits together, **A** and **B**. It produces **S**, the sum of A and B, and the corresponding carry out **Co**. Although by itself, a half adder is not extremely useful, it can be used as a building block for larger adding circuits (FA). One possible implementation is using two AND gates, two inverters, and an OR gate instead of a XOR gate as shown below (Fig. 1.1, Table1.1).

Table 1.1 Half adder truth table

A	B	S	C_0
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1

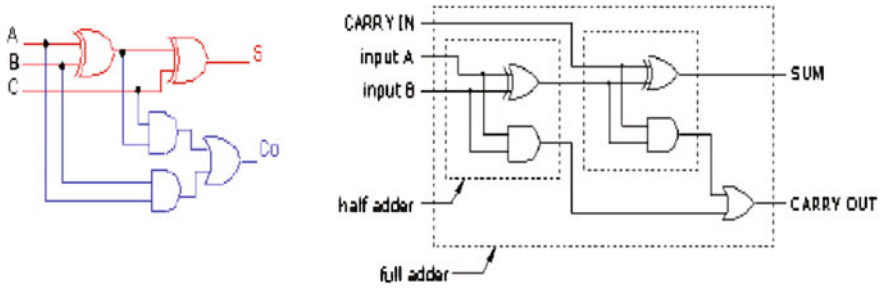


Fig. 1.2 Full adder and Logic Block

1.2.2 Full Adder

A full adder is a combinational circuit that performs the arithmetic sum of three bits: A, B and a carry in, C, from a previous addition. Also, as in the case of the half adder, the full adder produces the corresponding sum, S, and a carry out Co. As mentioned previously a full adder maybe designed by two half adders in series as shown. The sum of A and B are fed to a second half adder, which then adds it to the carry in C (from a previous addition operation) to generate the final sum S. The carry out, Co, is the result of an OR operation taken from the carry outs of both half adders (Fig. 1.2, Table1.2).

1.2.3 Ripple Carry Adder

The ripple carry adder is constructed by cascading full adders (FA) blocks in series. One full adder is responsible for the addition of two binary digits at any stage of the ripple carry. The carryout of one stage is fed directly to the carry-in of the next stage. A number of full adders may be added to the ripple carry adder or ripple carry adders of different sizes may be cascaded in order to accommodate binary vector strings of larger sizes. For an n-bit parallel adder, it requires n computational elements (FA). It is composed of four full adders. The augends' bits of x are added to the addend bits of y respectfully of their binary position. Each bit 6 addition creates a sum and a carry out. The carry out is then transmitted to the

Table 1.2 Full adder truth table

A	B	C	S	C_0
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

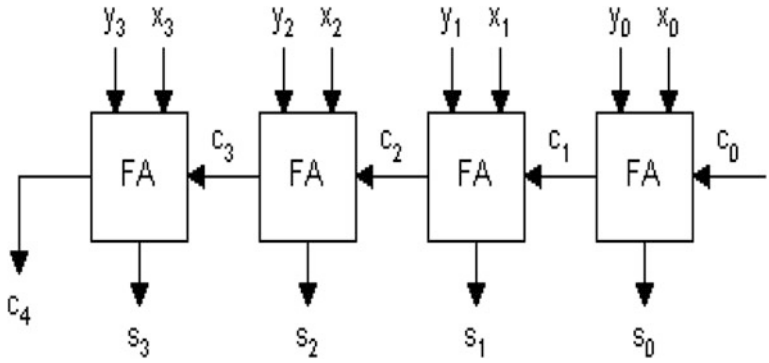


Fig. 1.3 Ripple carry adder

carry in of the next higher-order bit. The final result creates a sum of four bits plus a carry out (c4) (Fig. 1.3).

1.3 Back Ground

Parallel-prefix adders, also known as carry-tree adders, pre-compute the propagate and generate signals. The arrangement of the prefix network gives rise to various families of adders. The different types of carry tree adders are

- Kogge stone adder
- Sparse Kogge stone adder
- Spanning carry look ahead adders

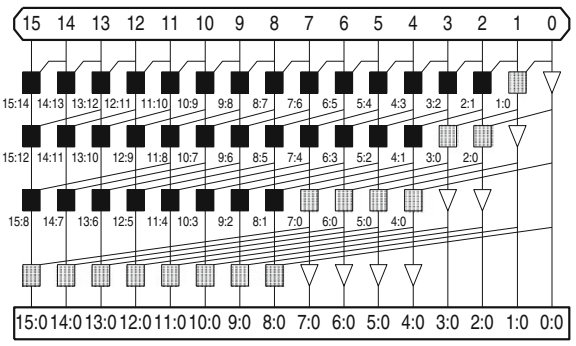


Fig. 1.4 Kogge-stone adder

1.3.1 Kogge Stone Adder

The Kogge-Stone adder is a parallel prefix form carry look-ahead adder.It is widely considered the fastest adder design possible. It is the common design for high-performance adders in industry.It has high speed performance with reduced delay and occupies less area. Each vertical stage produces a "propagate" and a "generate" bit, as shown. The culminating generate bits (the carries) are produced in the last stage (vertically), and these bits are XOR'd with the initial propagate after the input (the red boxes) to produce the sum bits. E.g., the first (least-significant) sum bit is calculated by XOR ing the propagate in the farthest-right red box (a "1") with the carry-in (a "0"), producing a "1" (Fig. 1.4).

1.3.2 Sparse Kogge Stone Adder

Enhancements to the original implementation include increasing the radix and sparsity of the adder. The radix of the adder refers to how many results from the previous level of computation are used to generate the next one. The original implementation uses radix-2, although it's possible to create radix-4 and higher. Doing so increases the power and delay of each stage, but reduces the number of required stages. The sparsity of the adder refers to how many carry bits are generated by the carry-tree. Generating every carry bit is called sparsity-1, whereas generating every other is sparsity-2 and every fourth is sparsity-4 (Fig. 1.5).

Sparse kogge-stone adder is nothing but the enhancement of the koggestone adder. The block in this sparse kogge stone adder are similar to the kogge stone adder. In this sparse kogge stone a reduction of number of stages is being done by reducing the genration and propagate units. The ouputs of the previous GP units are being considered such that the combination of consecutive Gp units produces carry once and that one is being given as inout to the next stage.

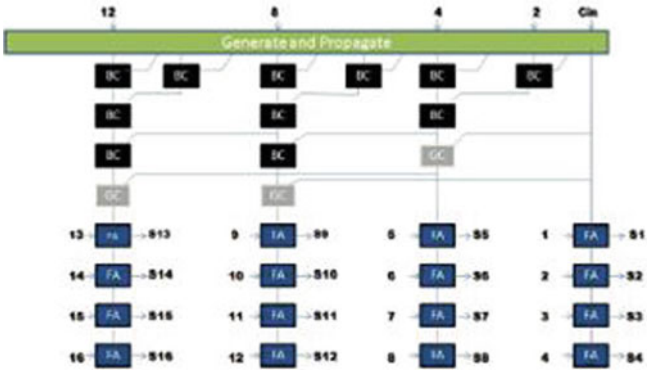
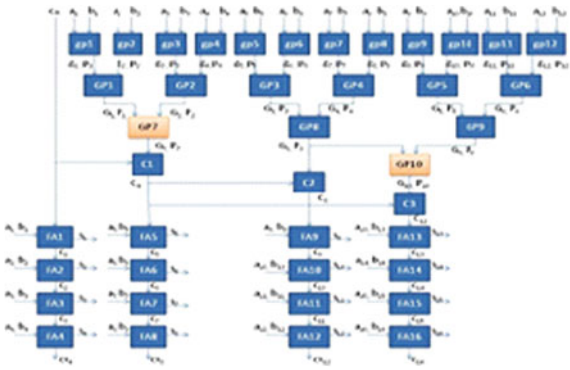


Fig. 1.5 Sparse Kogge-Stone adder

Fig. 1.6 Spanning CLA



1.3.3 Spanning CLA

The spanning CLA is nothing but the enhancement of kogge stone adder and sparse kogge stoone adder. In this spanning carry look ahead adder a reduction of number of stages is being done by reducing the GP units.The combination of consecutive outputsof GP units produces carry once and that one is being given as inout to the next stage. The generation and propagation of carry is being done. The output of previous block will act as inout to the next block and these operations are performed stage by stage this is how reduction of stages is being done and then the final sum is being produced by the operations performed by the combination of GP outs given as inouts to the full adders. The delay, power and area is low and speed is high (Fig. 1.6).

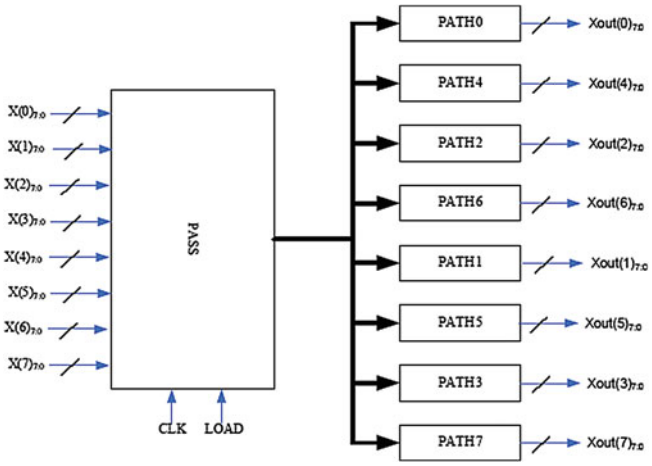


Fig. 1.7 Block diagram of an 8 point FFT processor

1.4 FFT Design

DISCRETE Fourier transform (DFT) is one of the fundamental operations in the field of digital signal processing. This section discusses the approach and method that is chosen to design of FFT. The computational time between DFT and FFT is faster using FFT method because the number of multiplications and additions operation in FFT is less compared to DFT method.

1.4.1 Algorithm of an 8-Point Fast Fourier Transform

The adders namely kogge stone adder is used in the Fast Fourier Transform.The implementation of an 8 point FFT processor involved few modules. All this modules are combined together to produce an 8 point FFT processor.

In the FFT algorithm (generally), the even and odd outputs are computed separately in two main groups. The odd output blocks computation is more complex compared to the even group computation. The odd output computations are represented by Path 1, Path 3, Path 5 and Path 7. The even output computations are Path 0, Path 2, Path 4 and Path 6 (Fig. 1.7).

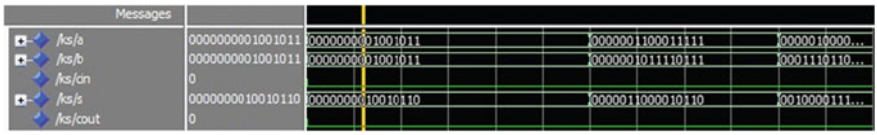


Fig. 1.8 Kogge stone Adder

1.4.2 Pass Module of 8 Point FFT Processor

This module passes the inputs to the sub-modules that do the FFT computations. The Pass module consists of 8 D flip flop registers. The outputs of this block are 8 lines of 8 bit output which are connected to Path 0, Path 1, Path 2, Path 3, Path 4, Path 5, Path 6 and Path 7.

1.4.3 Path 0 and Path 4 Module of 8 Point FFT Processor

The function of Path 0 and Path 4 is to compute and display the result of these computations. The outputs are Xout(0) and Xout(4) respectively.The arithmetic operation for Xout(0) is summation. The Xout(4) arithmetic involves summation, subtraction and division.

$$X(0) = x(0) + x(4) + x(2) + x(6) + x(1) + x(5) + x(3) + x(7)$$
$$X(4) = x(0) + x(4) + x(2) + x(6) - x(1) - x(5) - x(3) - x(7)$$

1.4.4 Path 2 and Path 6 Module of 8 Point FFT Processor

The function of Path 2 and Path 6 is to compute and display the result of these computations. The outputs are Xout(2) and Xout(6) respectively. The arithmetic operation for Xout(2) and Xout(6) involves real and imaginary operation.The arithmetic operation involves summation, subtraction and division. The twiddle factor for this output is either j or –j which contributes to the imaginary component for this path.

$$X(2) = x(0) + x(4) - x(2) - x(6) + jx(1) + jx(5) - jx(3) - jx(7)$$
$$X(6) = x(0) + x(4) - x(2) - x(6) - jx(1) - jx(5) + jx(3) + jx(7)$$

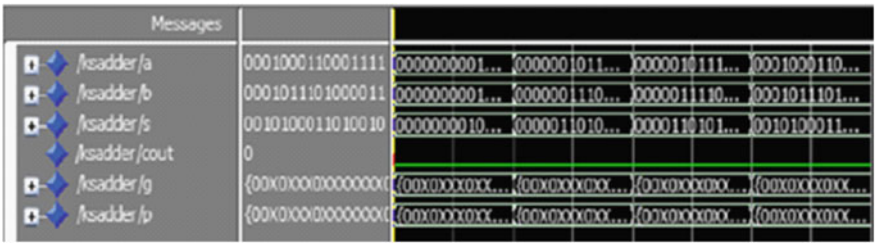


Fig. 1.9 Sparse Kogge Stone Adder

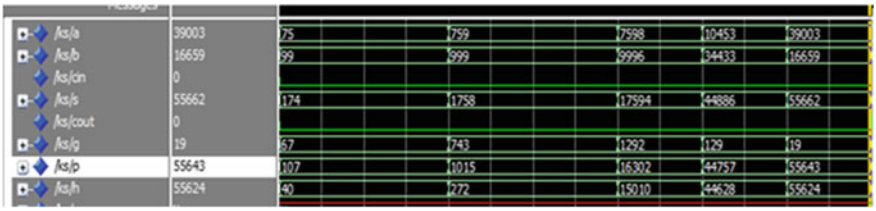


Fig. 1.10 Spanning CLA

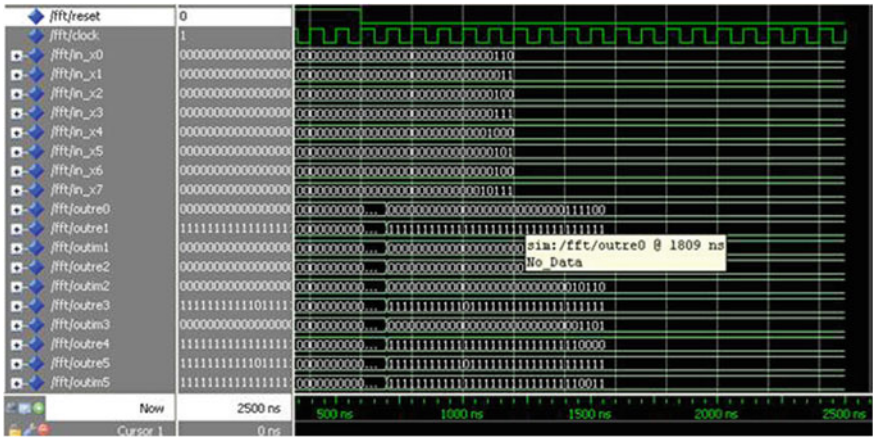


Fig. 1.11 FFT Output

Table 1.3 Delay comparisons in between adders

Adder	Bit widths	Delay (ns)
Kogge stone	16	15.072
Sparse Kogge stone	16	16.396
Span CLA	16	18.247
Koggae stone	128	30.165
Sparse kogge stone	128	54.987
Span CLA	128	84.378
Kogge stone	256	29.742
Sparasekogge stone	256	56.675
Span CLA	256	84.378

Table 1.4 Power Comparison in between adders

Adder	Static power (mW)	Dynamic power (mW)
Kogge stone adder	3	2.28
Sparse kogge stone adder	4.15	3.56
Spanning CLA	7.35	6.23
RCA	12	6.196

1.5 Results and Discussion

The codes for the Adders are written in VHDL. They are designed for different bit widths and simulated using Modelsim. The simulation result of adders for different bits are shown below.The output of the FFT is also shown (Figs. 1.8,1.9,1.10,1.11).

1.5.1 Delay

The synthesis of the above adders is done in Xilinx. The delay variations are being observed in the adders and then the delay variations are being compared in between the above designed adders for different bit widths and then tabulated (Table 1.3).

1.5.2 Power

The synthesis of the above adders is done. The delay variations are being observed in the adders and then the power variations are being compared in between the above designed adders for different bit widths and then tabulated. The static and dynamic powers are shown for different adders. The Power calculations are done using LIBRO tool (Table 1.4).

1.6 Conclusion and Future Work

The Adders namely ripple carry adder, carry look ahead adder, Kogge stone adder, sparse Kogge stone adder, spanning carry look ahead adder are discussed in detail. VHDL code was written for all the modules within the Adder. The adders are designed for the different bit widths namely 16, 128,256 bit widths. Simulation is done in XILINX and the delay is measured. Power is calculated using LIBRO.

This project has resulted in the development of Adders Design with reduced delay and power advantage. The FFT module algorithm can be implemented in the design OFDM transmitter and receivers for the generation of OFDM signal by transforming a spectrum (amplitude and phase of each component) into a time domain signal.

References

1. Westte NHE, Harris D (2011) CMOS VLSI Design, Pearson-Addison-Wesley, 4th edn
2. Brent RP, Kung, HT (1982) A regular layout for parallel adders. IEEE Trans Comput C-31:260–264
3. Harris D (2003) A taxonomy of parallel prefix networks. In: Proceeding 37th Asilomar conference signals systems and computers, pp 2213–2217
4. Kogge PM, Stone HS (1973) A parallel algorithm for the efficient solution of a general class of recurrence equations. IEEE Trans Comput C-22(8):786–793
5. Ndai P, Lu S, Somesekhar D, Roy K (2007) Fine grained redundancy in adders. In: International symposium on quality electronic design pp 317–321
6. Lynch T, Swartzlander EE (1992) A spanning tree carry look ahead adder. IEEE Trans Comput 41(8):931–939
7. Gizopoulos D, Psarakis M, Paschalis A, Zorian Y (2003) Easily testable cellular carry look ahead adders. J Elect Test Theory Appl 19:285–293
8. Xing S, Yu WWH (1998) FPGA adders: performance evaluation and optimal design. IEEE Des Test Comput 15(1):24–29
9. Becvar M, Stukjunger P (2005) Fixed point arithmetic in FPGA. ActaPolytechnicia 45(2):67–72
10. Virtoroulis K, Al-Khalili SJ. Perfromance of parallel prefix adders implemented with FPGA technology
11. Ghosh S, Patrick N, Kaushik R (2008) A novel low overhead fault tolerant Kogge-stone adder using adaptive clocking
12. Rabaey J (1996) Digital integrated circuits: a design perspective. Prentice Hall, India
13. Brent RP, Kung HT (1982) A regular layout for parallel adders. IEEE Tr Comp C-31(3):260–264
14. Gurkaynak FK et al (2000) Higher radix KS parallel prefix adder architectures. ISCAS, May
15. Han T, Carlson DA (1987) Fast area-efficient VLSI adders. In: 8th Symposium on Computer Arithmetic
16. Knowles S (1999) A family of adders. In: Symposium on Computer Arithmetic
17. Kogge P, Stone H (1973) A parallel algorithm for the efficient solution of a general class of recurrence equations. IEEE Trans Comp C-22(8):786–793

Chapter 2
Formal Approach to Reliability Improvement With Model Checker

Kazuhiro Yamada and Shin-ya Nishizaki

Abstract Since the 1960s, Fault Tree Analysis has been extensively used in Safety Engineering and Reliability Engineering, and other methodologies have been proposed. We study reliability analysis with formal methods. Fault tree analysis is one of the most popular methods of reliability analysis. With this, one analyzes the causes of a fault in a top-down manner. Model checking is an automatic verification method and has recently become popular. In this paper, we incorporate model checking into the fault tree analysis and show a case study of a pressure tank control system. Moreover, we propose a formal approach for introducing a fault detection mechanism. We show an example of a fault detection mechanism in the pressure tank control system, in which it is implemented using a set of lights to check electric current. We successfully show that model checking can evaluate the effectiveness of the fault detection mechanism.

Keywords Model checking · Reliability engineering · Fault tree analysis

K. Yamada · S. Nishizaki (✉)
Department of Computer Science, Tokyo Institute of Technology, 2-12-1-W8-69, Ookayama, Meguro-ku, Tokyo 152-8552, Japan
e-mail: nisizaki@cs.titech.ac.jp

K. Yamada
e-mail: kazuhiro.yamada@lambda.cs.titech.ac.jp

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_2,

2.1 Introduction

We start this paper by explaining the background and the motivation of our work

Traditionally, system verification has focused on correctness with respect to requirement specification: a check is made that the system meets the specified requirements. For correctness, it is implicitly assumed that

- Each component of the system operates normally;
- The environment surrounding the system is normal;
- The system outputs the results only through its interface components.

Even if the correctness of the system is verified, faults may make the system unserviceable. Above all, for critical systems such as those in aircraft and atomic reactors, correctness is not sufficient but reliability is required, which ensures that fatal and unusual situations do not occur.

Fault Tree Analysis [1], FTA, is a traditional method of reliability analysis that was first proposed in the 1960 s. FTA is a top-down, deductive analysis in which an undesired state of a system is decomposed using Boolean connectives into a series of lower level events. This method is mainly used in the field of safety engineering and reliability engineering. FTA can be used to understand the logical structure leading to the undesired state and to show compliance with the system reliability requirements.

In software engineering, model checking [2] 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 [3] and UPPAAL [4]. In UPPAAL, models are described in terms of timed automata using a GUI and specifications expressed by temporal logic Computational Tree Logic (CTL).

2.2 Motivation of Our Research

We introduce formal methods into reliability engineering and want to make it possible to analyze system reliability more carefully and in more detail. In this paper, we begin by applying model checking to the fault tree analysis. We propose a method of improving the reliability of a system. We incorporate a fault detection mechanism into a target system, analyze its effect by model checking.

2.3 Formal Reliability Improvement with FTA and Model Checking

In this paper, we propose a method to improve the reliability of a target system. We incorporate model checking into the Fault Tree Analysis. By introducing a fault detection mechanism to the target system, we can partially automate the evaluation of effectiveness and difficulty of fault detection.

Investigation of a system to be analyzed and description of a system model.

We investigate the system's behavior and requirements, and then describe a model formulating the system. We also define the undesired events to study in the reliability analysis. We should notice that the model and the undesired events to study should be formulated as simply as possible: if they are complicated, model checking in the succeeding phase could cause a state explosion, that is, a combinatorial blow-up of the state-space in the model checking.

Applying FTA to the system.

We apply the Fault Tree Analysis to the target system and find basic events (i.e. errors and faults) which cause the undesired events. When we enumerate the basic events, we identify components related to the basic events. The identified components are later used in evaluating the optimal fault detection mechanism.

Describing a model of the target system. By describing a model representing the behavior of the target system, we can verify the adaptability of the system's behavior to the required specification. We should choose a model checker appropriate for the model, considering the structure of the target system. Whichever is chosen, we should simplify the model as far as possible, in order to avoid the state explosion of the model checker.

Introducing a fault detection mechanism and evaluating its effectiveness and difficulty.

We introduce a fault detection mechanism to the target system model and analyze how effective it is in detection of faults and errors in the target system. The fault detection mechanism enables us to evaluate how many undesired events it detects and prevents.

In order to analyze and evaluate the fault detection mechanism, we utilize model checking. Since a model checker can report on the status of components in the target system's model, we can analyze and verify the fault detection mechanism by relating the fault detection to states in the model appropriately. We should notice that the fault detection mechanism itself can develop problems; if the fault detection is comparatively complicated, we should formulate its fault in the target system model.

Reflecting on the result of the analysis and the evaluation.

By introducing the fault detection mechanism, we can estimate the difficulty of fault detection of undesired events. Such a result is helpful for improving the reliability of the target system,

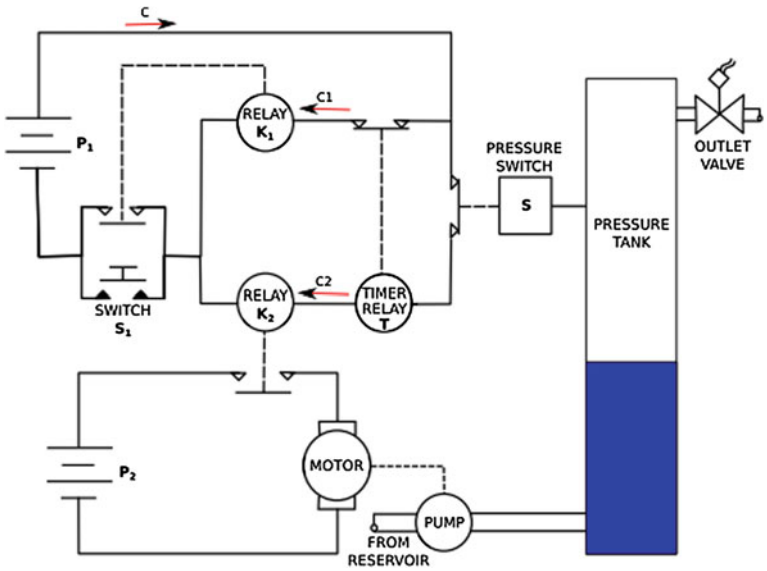


Fig. 2.1 Pressure tank control system

2.4 Example of Formal Reliability Improvement

We consider an example of a pressure tank control system in Sect. 8 of "Fault Tree Handbook" [5].

2.4.1 The Pressure Tank Control System

The pressure tank control system [5], which is depicted in Fig. 2.1, consists of three components:

- An electric circuit, powered by P_1, which controls the pump's motor circuit. The power supply P_1 electrifies the electric circuits c_1 and c_2. The circuit c_1 electrifies the coil of the relay K_1, if the push button S_1 or K_1 and the timer T are closed. The circuit c_2 electrifies the coil of relay K_2 and the timer T and c_2 is closed if S_1 or K_1 and the pressure switch S are closed.
- An electric circuit powered by P_2 and isolated from the above circuit powers the pump motor. The power supply P_2 electrifies c_2 which supplies power to the motor if the relay K_2 is closed.
- A pressure tank, which the pressure switch S contacts, is fitted with an outlet valve and is connected to an infinitely large reservoir through the pump.

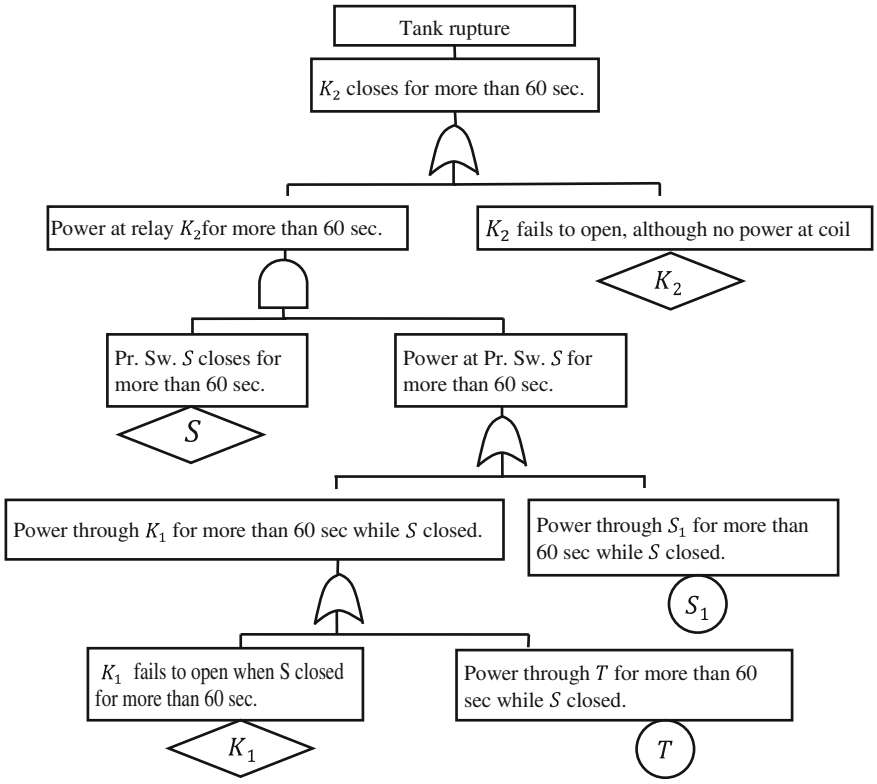


Fig. 2.2 Failure tree of the pressure tank

If pressure switch *S*is pressed, then circuit $c_1$is closed and $K_1$is electrified. If the pressure tank is empty, pressure switch S is closed. Then circuit $c_2$electrifies K_2, motor circuit $c_2$is closed, and consequently the motor starts. If the tank is full, then pressure switch *S*opens, relay K_2 opens, and the motor stops. If the outlet valve drains the tank and the tank is empty, the cycle described above is repeated.

2.4.2 Applying FTA to the Pressure Tank Control System

In this section, we briefly present a process in which we establish a fault tree of the pressure tank control system, as depicted in Fig. 2.2. Since the direct cause of tank rupture is continuous pumping for more than 60 s, we can identify the fault event as being that relay K_2 is closed for more than 60 s. Next, this event is decomposed into more basic events: a fault event of relay K_2 and two fault events due to other causes, which are connected by an OR-gate. The latter two fault events are further

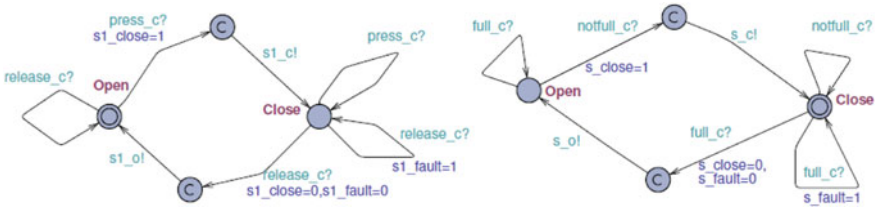


Fig. 2.3 Model of the pressure tank system (1)

decomposed and we finally identify fault events corresponding individually to the components. The following five fault events are identified as the most basic ones.

- S: Power switch S is closed for more than 60 s.
- S_1: The circuit is powered through switch S_1 for more than 60 s while switch S is closed;
- K_1: Relay K_1 fails to open when S is closed for more than 60 s;
- T: Circuit c_1 is powered through timer relay T for more than 60 s while pressure switch S is closed.
- K_2: Relay K_2 fails to open, although no power is supplied at the coil of relay K_2.

We can derive the minimal cut set of the fault tree $\{K_2\}, \{S, S_1\}, \{S, K_1\}, \{S, T\}$ from Figs. 2.2, 2.3, 2.4.

2.4.3 Description of a Model of the Pressure Tank Control System

We adopt the model checker UPPAAL for verifying and analyzing the pressure tank control system, since it is appropriate for modeling real-time systems. The synchronization is represented by channel communication in Promela, the modeling language of UPPAAL.

2.4.4 Introducing a Fault Detection Mechanism and Evaluating Its Effectiveness and Difficulty

We evaluate the effectiveness of a fault detection mechanism that lamps display electric current in circuits. The reasons we adopt lamps for fault detection are as follows:

- The lamps are simple and have a low fault rate. Moreover, we can easily improve this fault rate by redundancy.
- We can implement the fault detection mechanism at low cost.
- We can make correspondence between the model and its implementation.

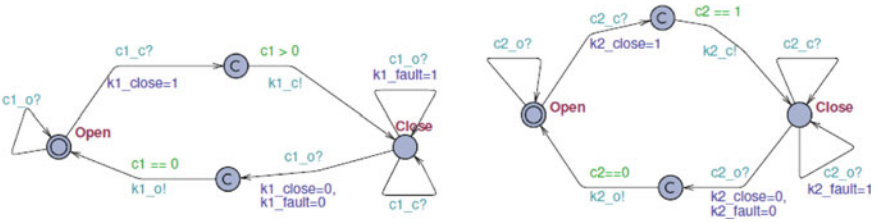


Fig. 2.4 Model of the pressure tank system (2)

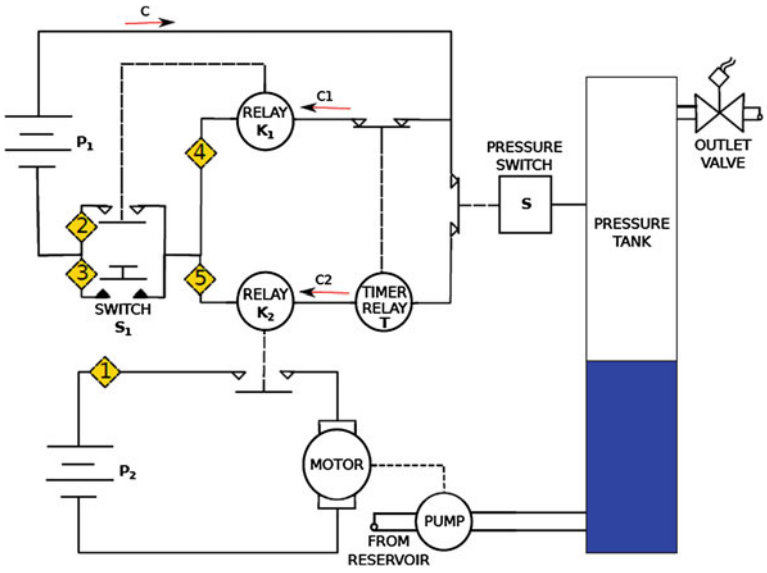


Fig. 2.5 Insertion of fault detecting mechanism into the pressure tank system

- The lighting patterns of the lamps provide a variety of information.

Figure 2.5 shows an improved circuit pressure tank system in which five lamps are inserted for fault detection. We show the correspondence between the five lamps and the sub-components in the pressure tank system.

Identification of the system state by lighting patterns. By observing the lighting patterns of the five lamps, we can identify the system state and the existence of faults in the system. We represent a lighting pattern of the five lamps as a five-bit sequence. For example, a five-bit sequence "01010" means that lamps #1, #3 and #5 are switched off and #2 and #4 are lit. Some of the lighting patterns are impossible with respect to the system's specification. For example, "00011" means that an electric current is flowing through the sub-circuits c_1 and c2 but relay K_1 and pressure switch $S_1$are turned off, which is impossible. The possible patterns are obtained by UPPAAL verification, as follows (Table 2.1):

Table 2.1 Correspondence between lamps and sub-components

Lamp	Sub-components
#1	Relay K_2
#2	Relay K_1
#3	Pressure switch S_1
#4	Sub-circuit c_1
#5	Sub-circuit c_2

Table 2.2 The possible lighting patterns

Pattern No.	Lamp#1	Lamp#2	Lamp#3	Lamp#5	Lamp#6
1	0	0	0	0	0
2	0	1	0	1	0
3	0	1	1	1	0
4	1	0	0	0	0
5	1	0	1	0	1
6	1	1	0	0	0
7	1	1	0	1	0
8	1	1	0	1	1
9	1	1	0	1	1
10	1	1	1	1	0
11	1	1	1	1	1

1. A TCTL formula representing that all lamps are lit is described.
   For example, a lighting pattern "00000" corresponds to a TCTL formula
   E <> relay_k2.Open & relay_k1.Open & switch_s1.Open & current1.NoCurrent & current2.NoCurrent
2. UPPAAL is used to check the TCTL formulas corresponding to all lighting patterns with respect to the model of the pressure tank control system.
3. The possible lighting patterns from the result of model checking with UPPAAL are obtained.

The obtained lighting patterns are shown in Table 2.2:

Correspondence of lighting patterns to top-level events. We next check the relationship between the above lighting patterns and the tank rupture. The tank rupture is a top-level event of the fault tree, which we should avoid as a priority. Thus, we know that lighting patterns #2 and #3 will never occur with a tank rupture. In other words, we can see from lighting patterns #2 and #3 that the tank will not rupture in such a situation.

Correspondence of lighting patterns to faults. In order to analyze the correspondence of lighting patterns to faults, we attach flags to the models representing whether each model is out of order or not. We then analyze the correspondence of the lighting patterns to faults by incorporating the fault flags into verification formulas. We summarize the result of the analysis in Table 2.2. In the table,

Table 2.3 Lighting pattern and faults

Pattern No. No.	Relay (K_1)	Relay (K_2)	Sensor (S)	Sensor (S_1)	Timer (T)
1	0	0	?	0	0
2	0	0	0	0	0
3	0	0	0	?	0
4	0	1	1	0	0
5	0	?	1	?	0
6	1	?	1	0	0
7	0	1	0	0	0
8	0	?	?	0	?
9	1	?	1	?	0
10	0	1	0	?	0
11	0	?	?	?	?

Table 2.4 Correspondence of lighting patterns to faults

	Relay (K_1)	Relay (K_2)	Sensor (S)	Sensor (S_1)	Timer (T)
Specified	11	6	8	6	9
Unspecified		5	3	5	2

- "0" means that the indicated component is operating normally,
- "1" means that the indicated component is out of order
- "?" means that a state of the indicated component is not specified.

From Table 2.3, we know that some of faults can be specified from lighting patterns, and the others cannot. This is summarized in Table 2.4.

According to Table 2.4, the lighting pattern #11 specifies that relay K_1 is out of order and the other patterns specify that it is operating normally. On the other hand, pattern #6 does not specify which relay is out of order; relay K_2 or sensor S_1. From Tables 2.3 and 2.4, we know that the state of sensor S_1 is difficult to specify.

2.5 Conclusion and Related Works

In this paper, we incorporate model checking into the fault tree analysis and show a case study on the pressure tank control system. Moreover, we propose a formal approach to introducing a fault detection mechanism. We show an example of a fault detection mechanism in the pressure tank control system, in which it is implemented with a set of lights for checking electric current. We successfully show that model checking can evaluate the effectiveness of the fault detection mechanism.

The first study on using a formal approach to fault tree analysis was made by Schellhorn et al. [6]. In their study, fault trees were formalized in the interval temporal logic and various properties were formally formulated. Thums et al. [7]

studied the application of model checking to the formal tree analysis by using a real-time model checker RAVEN. Faber [8] developed a formal FTA tool using the model checker UPPAAL.

Acknowledgments This work was supported by Grants-in-Aid for Scientific Research (C) (24500009).

References

1. Ericson CA II (2011) Fault tree analysis premier, create space
2. Clarke E (1997) Model checking. In foundations of software technology and theoreteical computer science. Lecture notes in computer science, vol 1346. Springer-Verlag, Berlin, pp 54–56
3. Holzmann GJ (2003) SPIN Model checker: the primer and reference manual, Addison-Wesley, Reading
4. Pettersson P, Larsen KG (2000) UPPAAL2k. Bull Eur Assoc Theor Comput Sci 70:40–44
5. Vesley WE, Goldberg FF, Roberts NH, Haasl DF (1981) Fault tree handbook, Office of Nuclear Regulatory Research, Rockville
6. Schellhorn G, Thums A, Reif W (2002) Formal fault tree semantics. In: IDPT-2002, society for design and process science, pp 739–757
7. Thums A, Schellhorn G (2003) Model checking FTA. In: FME 2003. Lecture notes in computer science. Springer-Verlag, Berlin, pp 739–757
8. Faber J (2005) Fault tree analysis with Moby/FT. http://iist.unu.edu/sites/iist.unu.edu/files/biblio/ToolPresentationMobyFT.pdf
9. Kumamoto H, Mizuno T, Narita K, Nishizaki S (2010) Destructive testing of software systems by model checking, In: The Proceedings of international symposium on communications and information technology (ISCIT), 2010, IEEE (2010), pp 26–29
10. Nishizaki S, Ohata T (2012) Real-time model checking for regulatory compliance. In: The Proceedings of AIM2012 (in printing)

Chapter 3
DDoS Attacks Defense System Using Information Metrics

P. C. Senthilmahesh, S. Hemalatha, P. Rodrigues and A. Shanthakumari

Abstract A Distributed Denial-of-Service (DDoS) attack is a distributed, coordinated attack on the availability of services of a target system or network that is launched indirectly through many compromised computing systems. A low-rate DDoS attack is an intelligent attack that the attacker can send attack packets to the victim at a sufficiently low rate to elude current anomaly-based detection. An information metric can quantify the differences of network traffic with various probability distributions. In this paper, an anomaly-based approach using two new information metrics such as the generalized entropy metric and the information distance metric, to detect low-rate DDoS attacks by measuring the difference between legitimate traffic and attack traffic is proposed. DDoS attacks detection metric is combined with IP traceback algorithm to form an effective collaborative defense mechanism against DDoS attacks.

Keywords Information metrics · IP traceback · Low-rate DDoS attack

P. C. Senthilmahesh (✉)
Anna University, Chennai, India
e-mail: senthilmagesh@yahoo.com

S. Hemalatha
Anna University, Chennai, India
e-mail: hemalathasenthilmahesh2004@yahoo.co.in

P. Rodrigues
Velammal Engineering College, Chennai, India
e-mail: drpaulprof@gmail.com

A. Shanthakumari
Department of Computer Science and Engineering, Arunai Engineering College, Tiruvannamalai, India
e-mail: shanthakumaritvm@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_3,

3.1 Introduction

Distributed Denial of Service (DDoS) attacks are one of the most serious threats in the Internet. The aim of the attack is to overload the victim and render it incapable of performing normal transactions. A low-rate DDoS attack has significant ability of concealing its traffic and elude anomaly-based detection schemes. It is a serious threat to the security of cyberspace. It typically exhausts bandwidth, processing capacity, or memory of a targeted machine or network.

Today, a large-scale DDoS attack is usually combined with multiple low-rate attacks, which are distributed on the Internet to avoid being detected by current detection schemes. An attacker can use the "BOTNETS" (Collection of compromised computers on which, a software called 'bot' is automatically installed without user intervention and are remotely controlled via command and control server) to launch a low-rate DDoS attack, producing network behavior that appears normal. Therefore, it is difficult to detect and mitigate such attacks.

The rest of the paper is organized as follows: Sects. 3.2 and 3.3 describe the Motivation and Background of DDoS attacks. The system architecture, the DDoS attack detection algorithm and an IP traceback algorithm are proposed in Sects. 3.4, 3.5 concludes the paper.

3.2 Motivation

Existing DDoS attack detection metrics are mainly separated into two categories: the signature-based metric and anomaly-based metric. The Signature-based detection metric depends on technology that deploys a predefined set of attack signatures such as patterns or strings as signatures to match incoming packets. Limitation is, it is more specific and cannot detect a wide range of DDoS attacks. Anomaly-based detection metric typically models the normal network (traffic) behavior and deploys it to compare differences with incoming network behavior. This method also has several limitations.

A low-rate DDoS attack has significant ability of concealing its traffic and elude anomaly-based detection schemes. Early detection and detection accuracy (such as a low false positive rate) of DDoS attacks are the two most important criteria for the success of a defense system. Aim of this paper is to detect Low-rate DDoS attacks earlier with high detection accuracy and to propose an IP traceback scheme by using information metrics.

3.3 Background of DDoS Attacks

There are two categories of DDoS attacks—Typical DDoS attack and DRDoS (Distributed Reflection Denial-of-Service) attacks. In a typical DDoS attack, the army of the attacker consists of master zombies and slave zombies. The hosts of both categories are compromised machines that have arisen during the scanning process and are infected by malicious code. The attacker coordinates and orders master zombies and they, in turn, coordinate and trigger slave zombies. In DRDoS attacks, slave zombies are led by master zombies to send a stream of packets with the victim's IP address as the source IP address to other uninfected machines (known as reflectors), exhorting these machines to connect with the victim.

3.4 Proposed Work

The proposed work can be defined as a two stage procedure for effectively defending against the low-rate DDoS attacks. The stages are—The DDoS attack detection algorithm and The IP Traceback algorithm.

The metrics used are: Generalized Entropy Metric—The generalized information entropy is a family of functions for quantifying either the diversity uncertainty or randomness of a system. It is defined as:

$$H_\alpha(x) = \frac{1}{1-\alpha}\log_2\left(\sum_{i=1}^{n} p_i^\alpha\right) \tag{3.1}$$

where p_i are the probabilities

Information Divergence Metric—The information distance or divergence is a measure of the divergence between P and Q, where P and Q are the discrete complete probability distributions. It is defined as (Fig. 3.1).

$$D_\alpha(P\|Q) = \frac{1}{\alpha-1}\log_2\left(\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right),\ \alpha \geq 0 \tag{3.2}$$

3.4.1 The DDoS Attack Detection Algorithm

The DDoS attack detection algorithm describes the steps involved in detecting the attack traffic. Initially, the detection threshold and the sampling period values must be assigned, then the network traffic from the systems are captured. The number of packets in the traffic with recognizable characteristics is determined in order to

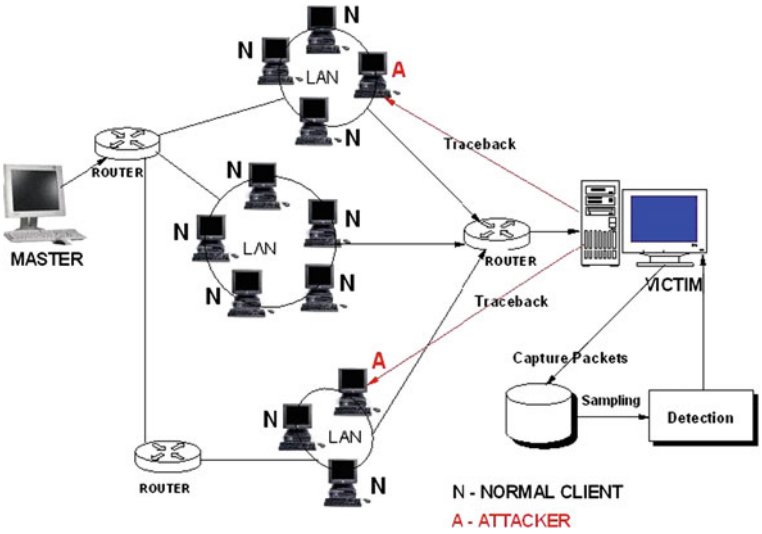


Fig. 3.1 System Architecture

Fig. 3.2 The DDoS attack detection algorithm

Step 1: Set the sampling period as T, and the collaborative detection threshold as σ

Step 2: Capture the network traffic from the systems in the sampling period several times.

Step 3: Calculate in parallel the numbers of packet which have various recognizable characteristics (e.g., the source IP address , protocol, the packet's size, etc.)

Step 4: Calculate the probability distributions of the network traffic from the systems.

Step 5: Calculate their information distances using the formula: $D\alpha(P,Q) = |D\alpha(P\|Q) + D\alpha(Q\|P)|$, where

$$D_\alpha(P\|Q) = \frac{1}{\alpha - 1}\log_2\left(\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right) \text{ and}$$

$$D_\alpha(Q\|P) = \frac{1}{\alpha - 1}\log_2\left(\sum_{i=1}^{n} q_i^\alpha p_i^{1-\alpha}\right)$$

Step 6: Compare the information distance value with the detection threshold.

Step 7: If the information distance is more than the detection threshold σ, then the system detects the DDoS attack, and begins to raise an alarm.

calculate the probability distributions. With the values of probability distribution, the information distance is calculated and compared with the threshold value (Fig. 3.2).

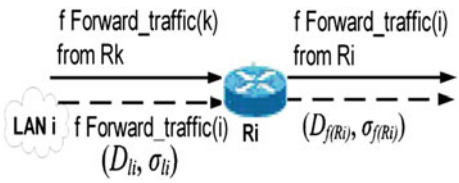


Fig. 3.3 Local traffic, forward traffic, information distance and threshold σ at a router

3.4.2 IP Traceback

IP traceback is the ability to find the source of an IP packet without relying on the source IP field in the packet, which is often spoofed. The proposed DDoS attacks detection metric is combined with IP traceback algorithm [2, 3] and filtering technology together to form an effective collaborative defense mechanism against network security threats in Internet (Fig. 3.3).

When the proposed attacks detection system detects an attack on a victim, the proposed IP traceback algorithm will be launched immediately. The victim initiates the pushback process to identify the locations of zombies: the victim first identifies which of its upstream routers are in the attack tree based on the flow entropy variations it has accumulated, and then submits requests to the related immediate upstream routers. The upstream routers identify where the attack flows came from based on their local entropy variations that they have monitored. Once the immediate upstream routers have identified the attack flows, they will forward the requests to their immediate upstream routers, respectively, to identify the attacker sources further; this procedure is repeated in a parallel and distributed fashion until it reaches the attack source(s) (Fig. 3.4).

3.5 Conclusion

In today's computer-dominated society, the practice of securing and administrating computer systems and enterprise network become critical and challenging. This defense mechanism is the combination of DDoS attack detection and mitigation, thus it provides protection to the internet users against the threatening DDoS attacks in the networks.

In this paper, two new and effective information metrics for low-rate DDoS attacks detection are proposed: generalized entropy and information distance metric. They outperform the traditional Shannon entropy and Kullback–Leibler distance approaches, respectively, in detecting anomaly traffic. As the proposed metrics can increase the information distance (gap) between attack traffic and legitimate traffic, they can effectively detect low-rate DDoS attacks early and reduce the false positive rate clearly.

Fig. 3.4 The IP Traceback algorithm

Step 1: Assign the detection threshold, σ_{fi}, σ_{li}.

Step 2: Check attacks on the traffic by calculating Information Distance, D_{fi}.

Step 3: Information Distance for forward traffic is calculated by using the formula,

$$\left|D\sigma_{fi}(p_f, p_l)\right| = \left|D\sigma_{fi}(p_f \| p_l) + D\sigma_{fi}(p_l \| p_f)\right|$$

Step 4: Compare this information distance value with the detection threshold.

Step 5: If $D_{fi} > \sigma_{fi}$, calculate the information distance for local traffic, D_{li}. Otherwise forward the packet.

Step 6: D_{li} is calculated using the value of σ_{li} in the formula for information distance.

Step 7: If $D_{li} > \sigma_{li}$, stop forwarding the attack traffic to downstream routers. Else forward the packet.

References

1. Ashley C, Jaipal S, Wanlei Z (2009) Chaos theory based detection against network mimicking DDoS attacks. IEEE Commun Lett 13(9):717–719
2. Xiang Y, Li K, Zhou W (2011) Low-rate DDoS attacks detection and traceback by using new information metrics. IEEE Trans Inform Forensics Secur 6(2):426–437
3. Yu S, Zhou W, Doss R, Jia W (2011) Traceback of DDoS attacks using entropy variations. IEEE Trans Parallel Distribd Sys 22(3):412–425
4. Li K, Zhou W, Yu S (2009) Effective metric for detecting distributed denial-of- service attacks based on information divergence. IET Commun 3(12):1859–2860
5. Yu S, Zhou W, Doss R (2008) Information theory based detection against network behavior mimicking DDoS attack. IEEE Commun Lett 12:319−321
6. Sheng Z, Zhang Q, Pan X, Xuhui Z (2010) Detection of low-rate DDoS attack based on self-similarity. In; Proceeding International Workshop on Education Technology and Computer Science pp 333–336
7. Liu Y, Yin J, Cheng J, Zhang B (2010) detecting ddos attacks using conditional entropy. International conference on computer application and system modeling (ICCASM 2010)
8. Giseop N, Ilkyeun R (2009) An efficient and reliable DDoS attack detection using a fast entropy computation method. ISCIT
9. Lee W, Xiang D (2001) Information-Theoretic measures for anomaly detection. In: Proceeding IEEE Symposium Security and Privacy pp 130–143

Chapter 4
CEAR: Cluster based Energy Aware Routing Algorithm to Maximize Lifetime of Wireless Sensor Networks (WSNs)

H. Sivasankari, R. Leelavathi, M. Vallabh, K. R. Venugopal, S. S. Iyengar and L. M. Patnaik

Abstract Technological development in wireless communication enables the development of smart, tiny, low cost and low power sensor nodes to outperform for various applications in Wireless Sensor Networks. In the existing Tabu search algorithm, clusters are formed using initial solution algorithm to conserve energy. We propose a Cluster Based Energy Aware Routing (CEAR) algorithm to maximize energy conservation and lifetime of network with active and sleep nodes. The proposed algorithm, removes duplication of data through aggregation at the cluster heads with active and sleep modes. A comparative study of CEAR algorithm with Tabu search algorithm is obtained. Comparative study shows improvement in the Lifetime and energy conservation by 17 and 22 % respectively over the existing algorithm.

Keywords Clusters · Delay · Energy aware routing · Lifetime and wireless sensor networks

H. Sivasankari (✉) · R. Leelavathi · M. Vallabh · K. R. Venugopal
Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore University, Bangalore 560 001, India
e-mail: sankari@yahoo.com

S. S. Iyengar
Florida International University, Miami, USA

L. M. Patnaik
Indian Institute of Science, Bangalore, India

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_4,

4.1 Introduction

Wireless Sensor Networks (WSNs) are deployed with hundreds or thousands of computable and low cost sensors. These sensor nodes are multi-functional and with computing capabilities. These sensor nodes are integrated on a single board in a few cubic inches with embedded microprocessor, radio receivers, sensing, and computing and communication unit. They are powered by 50 W and can last for 2–3 years with very less duty cycling and these sensor nodes are prone for failures. In WSN there are three types of communications, they are clock driven, event driven and query driven. In the clock driven approach, data collection and transmission takes place at regular periodic intervals. In the event driven and query driven approaches, data collection is triggered by events or queries. Data Aggregation, Clustering, effective routing and data compression, min–max and averaging methods are used to reduce energy consumption in WSNs. Tabu search algorithm [1] form clusters of sensor nodes and routes data from source to the destination through cluster heads. The sensor nodes are active and transmits data without aggregation all the time and this approach consumes more energy as there is no aggregation. We propose Cluster based Energy Aware Routing (CEAR) in order to maximize lifetime and energy conservation of WSNs. In a cluster, sensor node works in active and sleep mode randomly. It routes the data from source to destination through cluster head with aggregation. Thus, it reduces the energy consumption for sending the large set of data. The rest of the paper is organized as follows: Related work is discussed in Sect. 4.2. Problem Definition is formulated in Sect. 4.3. Algorithm is developed in Sect. 4.4. Simulation and Performance Analysis are analyzed in Sect. 4.5. Conclusions are presented in Sect. 4.6.

4.2 Literature Survey

Heinzelman et al. [2] developed a Low Energy Adaptive Clustering Hierarchy (LEACH) algorithm for clustering. LEACH is a distributed approach and it requires no control information from the base station. In this approach data is required to be sent in the allotted time. Madan et al. [3] proposed TAG: Tiny Aggregation Service for Ad-Hoc Sensor networks. Interface for data collection, aggregation, executed aggregation queries in the sensor networks. Aggregation queries are executed in time, efficiently. Each mote is required to transmit only a single message per epoch, regardless of its depth in the routing tree. It is sensitive to resource constraints and lossy communication. Liang et al. [4] have designed a genetic cost model for data gathering in sensor networks, as routing tree used for query evaluation and therefore the network is balanced. The proposed heuristic algorithm showed that online gathering problem is NP complete. It is focused only on the event driven data and data gathering queries are assumed sequential. Ghiasi et al. [5] proposed a balanced K clustering algorithm to minimize energy

consumption using minimum cost network flow. Agarwal et al. [6] have proposed sub exponential algorithms to solve the K-center problem to compute the optimal K-center. They considered the metrics for planar case, i.e., smaller dimensions. Noritaka et al. [7] have proposed centralized and distributed approaches for Clustering. These methods prolong the network lifetime than the conventional methods.

4.3 Problem Definition and Mathematical Model

Given a set of Wireless Sensor Nodes $S_i \in V$ where i = 1, 2,…, n. we consider a single sink with maximum coverage radio and power. The objectives are to reduce the energy consumption and to maximize the lifetime of the network. The assumptions are (i) All source sensor nodes are static. (ii) Links in between the nodes are bidirectional. (iii) Sink has long communication range and energy than the source sensor nodes (iv) The highest energy node becomes the cluster head. (v)Sensor nodes act in either active or sleep mode.

4.3.1 Mathematical Model

All sensor nodes are deployed with equal amount of energy. Clusters are formed based on Euclidean distances. Cluster size is increased as the number of node increases in the deployment. Graph (G) i.e., G = (V, E) where V is a set of vertices and E is a set of edges. Neighbor of a node is selected based on the following constraints. In a graph, v_i is the source node and v_j next neighbor node to receive data as the destination. The distance between source and destination is calculated based on Euclidean distance.

$$Dist(v_i - v_j) < Range \tag{4.1}$$

$$Euclidean\ Distance = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{4.2}$$

$$List\ of\ Neighbors\ L\ (N_i) = \bigcup_{V_j \in V} \{V_j / Dist(v_j, v_i) < r\} \tag{4.3}$$

Theorem Aggregation at the cluster head removes duplication of data thus it increases energy conservation and lifetime.

$$C(e) = \sum_{i=1}^{n} (C_i)(n_i) \tag{4.4}$$

$$A(e) = l(v)(1 - \sigma_{uv} X_e) \tag{4.5}$$

$$P_{total} = T(e) + A(e) + R(e) + C(e) \tag{4.6}$$

Where *T(e)* is Transmission cost, *A(e)* is Aggregation cost, σ_{uv} is reduction ratio and X_e is the Aggregation factor if $X_e = 1$ there is an aggregation else there is no aggregation, *R(e)* Reception cost, *C(e)* Cluster formation cost, load on node *v*, C_i represents the *ith* cluster and n_i is the number of nodes in a cluster.

4.4 CEAR: Cluster Based Energy Aware Routing Algorithm

In a Cluster based Energy Aware Routing (CEAR), clusters are formed as the nodes are deployed in the region. When the first node is deployed, the cluster is formed and it becomes the cluster head. Later, the deployed nodes identify the neighboring nodes by sending hello packets. When it receives successive echo packets from neighboring nodes then it is added to the cluster. If the node is neither in the range nor in neighbor list of any cluster then a new cluster is formed and node becomes the cluster head. As the more and more nodes are deployed, the cluster size is increased. This procedure is repeated until there is no more node to be deployed.

In a cluster, after each iteration it checks for the highest energy node. It elects the highest energy node as the cluster head. In a cluster, a set of source nodes act in active and sleep mode. Before a node goes for sleep mode it ensures that enough number of nodes are in active mode in order to ensure that any event during the transition of active and sleep mode is not missed. Energy Aware Routing helps the data to be routed through cluster heads (Table4.1).

4.5 Simulation and Performance Analysis

In the setup of MATLAB simulation, a 100 × 100 m region is considered with 150 sensor nodes. All sensor nodes have equal amount of energy initially with 50 J. Radio model is lossy in nature and communication range is 150 m, energy consumption per bit is 60 pJ. Control packet size is 500 bytes. Fig 4.1 shows the graph between the energy consumption with number of sensor nodes or clusters. The proposed algorithm CEAR consumes lower energy than the existing algorithm Tabu. The energy savings in CEAR is 22 % higher than the existing Tabu search algorithm. Fig 4.2 depicts that the lifetime is increased by 17 % in CEAR algorithm in comparison with the existing algorithm.

Table 4.1 Cluster based energy aware routing algorithm

The Subgraph G' $\in$ G, $\forall$ nodes N_i.

Begin
  while deployment of nodes is True do N ++;
   if(Id_N_j == 1) then Cluster C_1; N_j as Cluster Head for C_1;
    else neighbor node selection() get node id of N_j i.e. id_N_j;
     for n 1 to N do
       if $hello_{jn}$ = = 1 then Add node N_n & N_j as neighbors;
        Update Neighbor list;
   endif
  endfor
for i 1 to M do for n 1 to N do
  if id_N_j $\in$ neighbor N_n & & in_range(C_i) then Make node $N_j \in C_i$; return;
   else if (n == N && i == M) then M ++; Create new Cluster C_M;
    Elect node N_j as Cluster head for C_M;
  endif return;
endif, endfor, endif, endwhile
  K = ActiveNode(); Cluster Head Election Algorithm
   for i 1 to M do
   for p 1 to K do $\forall$ $k_p \in C_i$;
    if $P_{u_p}(t) \leq P_{v_p}(t)$
   Select active node v_p(t) at time t as Cluster Head to the current Cluster C_i;
  else
  Select active node u_p(t) at time t as Cluster Head to the current Cluster C_i;
  endif endfor endfor
if event then Select $\forall$ $k_p \in C_i$ at time t;
  if(dist(event, k_p) $\leq$ min) then min = dist (event, k_p); $S_i = k_p$;
  endif, endif
Route : for non sink active pair(u, v) do
                for neighbor j to k of S_i do
                 if Near $S_j \neq$ active() then
                   make S_j as active node;
                endif
                Euclidean_distance = Euclidean_distance + dist(i, j); S_i = Sj;
               endfor Compute minimum Euclidean Distance;
               Calculate fusion benefit δ_{active}(u, v) $\forall$ $k_p \in C_i$;
if δ_{active} (u, v) == 0 then
  Select active pair(u, v) as non fusion pairs $\in$ set E_n;
else Select active pair(u, v) as fusion pairs $\in$ set E_f;
endif
if $S_{i+1} \neq$ St then $go to **Route**;
else return;
endif, End

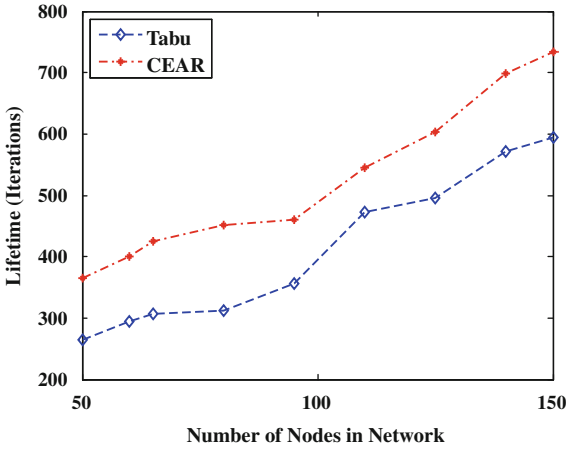


Fig. 4.1 Lifetime Comparison between Tabu Search and CEAR

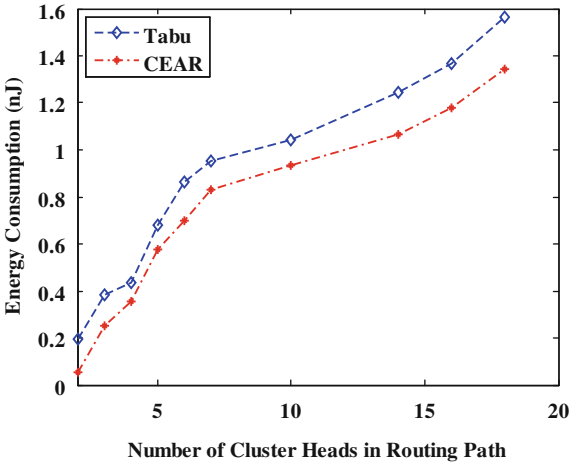


Fig. 4.2 Energy consumption versus number of cluster heads in routing path

4.6 Conclusions

Wireless sensor networks works with limited battery and lifetime. We propose Cluster based Energy Aware Routing (CEAR) to conserve energy and maximize the life time of the WSNs. In the case of Tabu search algorithm, clusters are formed and routed through clusters heads without aggregation and active/sleep modes. Therefore, it consumes more energy. The data aggregation and periodic active/sleep mode in CEAR algorithm improves the life time and energy conservation, This work can be enhanced in future to reduce delay.

References

1. El Rhazi A, Pierre S (2009) A Tabu search algorithm for cluster building in wireless sensor networks. IEEE Trans Wirel Comput 8(4):433–444
2. Heinzelman W, Chandrakasan A, Balakrishnan H (2002) An application specific protocol architecture for wireless microsensor networks. IEEE Trans Wirel Commun 1(4):660–670
3. Madden SR, Franklin MJ, Hellerstein JM, Hong W (2002) TAG: tiny aggregation service for ad-hoc sensor networks. In: Proceedings of the fifth symposium operating systems design and implementation (OSDI 02), pp 131–146
4. Liang W, Liu Y. (2007) Online data gathering for maximizing network lifetime in sensor networks, IEEE Trans Mobile Comput 6(1):2–11
5. Ghiasi S, Srivastava A, Yang X, Sarrafzadeh M (2002) Optimal energy aware clustering in sensor networks. IEEE Trans Mobile Comput 2:258–269
6. Agarwal PK, Procopiuc CM (2002) Exact and approximation algorithms for clustering. Algorithmica 33(2)201–226
7. Shigei N, Miyajima H, Morishita H, Marda M (2009) Centralized and distributed clustering method for energy efficient wireless sensor networks. In: Proceedings of the International Multiconference of Engineers and Computer Scientists(IMECS), Vol.2174, No 1, LNCS, pp 423–427

Chapter 5
TGAR: Trust Dependent Greedy Anti-Void Routing in Wireless Sensor Networks (WSNs)

H. Sivasankari, R. Aparna, K. R. Venugopal, S. S. Iyengar and L. M. Patnaik

Abstract In Wireless Sensor Networks (WSNs), energy and reliable data delivery are two major issues. Sending data from source to destination without void problem is an objective of any routing algorithm. The existing Greedy Anti-void Routing (GAR) uses the Rolling ball Undirected Traversal to guarantee the packet delivery from source to the destination. In the case of sparse network when it encounters an obstacle in the route it fails to deliver the data. To address this issue, we propose Trust dependent Greedy Anti-void Routing (TGAR) to find the reliable path from source to sink. We use Bayesian estimation model to calculate the trust value for the entire path. Simulation results show that TGAR achieves successful data delivery and energy conservation in sparse networks when compared with the existing Greedy Anti-void Routing (GAR) Algorithm.

Keywords Sparse network · Unit distance graph · Void and wireless sensor networks

H. Sivasankari (✉) · R. Aparna · K. R. Venugopal
Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore University, Bangalore 560001, India
e-mail: sankari@yahoo.com

S. S. Iyengar
Florida International University, Miami, USA

L. M. Patnaik
Indian Institute of Science, Bangalore, India

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_5,

5.1 Introduction

Sensor nodes are battery operated and have limited energy in WSNs. It is desirable to design an energy efficient protocol for WSNs. There are three types of routing techniques i.e., flat, hierarchical and location based routing. In flat routing, nodes are collaborated to sense an event. There is no unique id for all nodes since they are deployed in large numbers and it is not feasible to give identification for each node. In hierarchical approach, high energy node can be used to process and send the information and low energy nodes are used only for sensing. In location based routing, sensor nodes are identified by their locations. The Greedy Anti-void routing is used to forward a packet in an alternative path if it encounters a void in the path. The alternative path is inefficient and there is a chance of encounter obstacles. We propose Bayesian estimation model to calculate trust value for each path, to avoid unreliable and inefficient paths. We propose Trust dependent Greedy Anti-void Routing(TGAR) for successful data delivery and energy conservation and better than the existing GAR algorithm. The rest of the paper is organized as follows: Related work is discussed in Sect. 5.2. Problem Definition is formulated in Sect. 5.3. Algorithm is developed in Sect. 5.4. Simulation and Performance Analysis are analyzed in Sect. 5.5. Conclusions are presented in Sect. 5.6.

5.2 Literature Survey

Shigang et al. [1] have proposed"Right Hand Rule" method to route a packet out of a dead end in WSNs to avoid a path with void. These paths are inefficient and long. Sungoh et al. [2] developed a new Geographic routing algorithm to alleviate the effect of location errors in routing in wireless ad-hoc networks. Significant errors were, present in getting estimation of the locations. Kung et al. [3] presented a Greedy Perimeter Stateless Routing (GPSR) protocol for wireless networks. Here, routing is based on the current positions of routers to forward packets. When GPSR is not possible, packet is recovered with help of Perimeter routing. Hannes et al. [4] have introduced a planar graph routing on geographical cluster based on the current location of devices. It is failed to connect a path since it has one hop neighbor information only. It is avoided in [5], by distance routing for wireless Ad-hoc Networks with multihop. Tassos et al. [6] have developed a Geographic routing around void in sensor networks, for efficient routing decisions. It is a cross-layered approach to improve routing decisions and respond immediately for topological changes. Wen et al. [7] addressed an issue of un-reachability problem. They proposed an Greedy routing with Antivoid traversal for wireless sensor networks. Boundary Map(BM) and Unit Disk Graph (UDG) are used to increase the efficiency and solve the void problem. In this approach, the optimal path between source and destination is not discussed.

5.3 Problem Definition and Mathematical Model

Given a set of Wireless Sensor Nodes $S_i \in V$ where $i = 1, 2,\ldots, n$. we consider a single sink with maximum coverage radio and power. The objectives are to reduce the energy consumption and to improve the reliability of data delivery. The assumptions are (i) All source sensor nodes are static. (ii) Links in between the nodes are bidirectional. (iii) Sink has long communication range and higher energy than the source sensor nodes (iv) Neighbor list includes node within the range of communication. (v) The path is selected when the trust values are good, otherwise the path is not selected.

5.3.1 Mathematical Model

5.3.1.1 Determination of Trust Value

All sensor nodes are deployed with equal amount of energy. We assume WSN as a Graph (G) i.e., G = (V, E) where V is a set of vertices and E is a set of edges. Neighbor of a node is selected based on the range of transmission. In a graph, v_i is the source node and v_{j+1} the next neighbor node to receive data as the destination. The distance between source and destination is calculated based on Euclidean distance. A Bayesian network is a belief network model to represent a set of random variables and their conditional dependencies in a graph. Nodes are conditionally dependent when they are connected. Nodes are independent in nature when they are not connected. The optimal path is obtained through a good estimate of the target state ϕ(t) i.e., sink from the measurement history z(t). P(ϕ) denotes Priori Probability Distribution Function(PDF) about the sink state ϕ(t). Priori Probability Distribution Function (PDF) of z given ϕ is given as P(z/ϕ). The posteriori distribution of ϕ given the measurement z, is defined by the likelihood function as P(ϕ/z) and also referred to as the current belief. Bayes theorem gives the relationship between the posteriori distribution P(ϕ/z), the priori distribution P(ϕ) and the likelihood function P(z/ϕ).

$$P\left(\frac{\phi}{z}\right) = \frac{P\left(\frac{z}{\phi}\right)P(\phi)}{\int P\left(\frac{z}{\phi}\right)P(\phi)d(\phi)}. \tag{5.1}$$

$$P(z) = \int \left(\frac{z}{\phi}\right)P(\phi)d(\phi). \tag{5.2}$$

and P(z) is the normalizing constant to make value less than one since the maximum probability value is one. Therefore, Eq. (5.1) can be rewritten as

$$P\left(\frac{\phi}{z}\right) = k\, P\left(\frac{z}{\phi}\right) P(\phi). \tag{5.3}$$

$$i.e., P\left(\frac{\phi}{z}\right) \propto P\left(\frac{z}{\phi}\right) P(\phi). \tag{5.4}$$

$\hat{\phi}(t)$ is an estimation of the optimal path which is close to the true value of $\phi(t)$ according to the measurement. Minimum Mean Squared Estimator (MMSE) is used to estimate the mean value of the distribution

$$P(\phi/z_1, z_2, z_3, z_4, \ldots..., z_N) \tag{5.5}$$

Mean is calculated by the equation given below

$$\overline{\phi} = \int \phi\, P(\phi/z_1, z_2, z_3, z_4, \ldots..., z_N) \tag{5.6}$$

Uncertainty is approximated by covariance Σ as given below

$$= \int \left(\phi - \overline{\phi}\right)\left(\phi - \overline{\phi}\right)^T P\left(\frac{\phi}{z_1}, z_2, z_3, z_4, \ldots..., z_N\right) d(\phi) \tag{5.7}$$

We must pay a cost for communication but to make important decision belief status are used to reduce cost in communication. In a centralized Bayesian estimation model with N sensor nodes in the deployment, at any time t, each sensor n(i = 1, 2…N) informs about the measurement $z_n(t)$. The central unit updates the belief state using Eq. (5.4). If the sensor measurements are mutually independent on the target location then

$$P(\overline{z}^{(t)}/\phi^{(t)}) = \prod\nolimits_{n=1,2,3,\ldots,N} P(z_n^{(t)}/\phi^{(t)}) \tag{5.8}$$

For reliable communication, power consumption is exponentiation (α) of the distance

where $\alpha = 2$ is the pathloss exponent.

5.3.1.2 Computation of Energy Level at a Node

$$D_{ij} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}. \tag{5.9}$$

$$E_i = E_t - (E_i) * P_i * D_{ij}. \tag{5.10}$$

where D_{ij} is the distance between two nodes *i, j* and E_i is the energy at node *i*. P_i is the total number of packets and E_t is the total amount energy present in the node.

Table 5.1 TGAR: Trust dependent greedy anti-void routing algorithm

```
Begin
. Step 1: Deploy the sensor nodes.
. Step 2: Identify the Neighborlist for all nodes by sending hello
   packet.
. Step 3: Select the Nearest Neighbor for routing from the Neighbor
   list.
. Step 4: Obtain the Trust value for each path as derived in
   Equation(5.8).
. step 5: Select the path if the Trust value is good(= 1); otherwise
   reject
when the Trust value is zero.
. Step 6: Calculate Energy using Energy Equation(5.10) for each node in
   the, path.
. Step 7: Calculate residual Energy in each path. If the nodes energy is
   zero, lifetime of the network expires.
End
```

5.4 TGAR: Trust Dependent Greedy Anti-Void Routing Algorithm

In a Trust based Greedy Anti-void Routing (TGAR) each source sensor node finds a neighbor by sending hello packets. Nodes within the range of communication responds to hello packets and are included in the neighbor list. It selects the minimum distance neighbor from the neighbor list. Unit Distance Graph (UDG) gives all possible paths from the source to sink node. The Bayesian estimation expression gives the estimate of reputation value for each path. If the reputation value is *good* then the path is selected for routing. History $Z(t)$ gives whether the path is reliable for communication, then the reliable and efficient path is selected from the history of the paths (Table 5.1).

5.5 Simulation and Result Analysis

In the MATLAB simulation, a 100 m $\times$ 100 m region is considered with 100 sensor nodes. All sensor nodes have equal amount of energy initially with 50 J. Radio model is lossy in nature and communication range is 150 m, energy consumption per bit is 50 pJ. Control packet size is 300 bytes. Figure 5.1 shows the graph between the energy consumption with number of sensor nodes deployed. The proposed algorithm TGAR consumes 12 % lower energy than the existing algorithm GAR. Figure 5.2 shows the delay between TGAR and GAR in sparse network. Delay is 33 % less in proposed algorithm TGAR than the GAR. Since in our approach routing is based on history, directs the data in correct path towards the sink, and avoids unnecessary search thus saving time in data delivery.

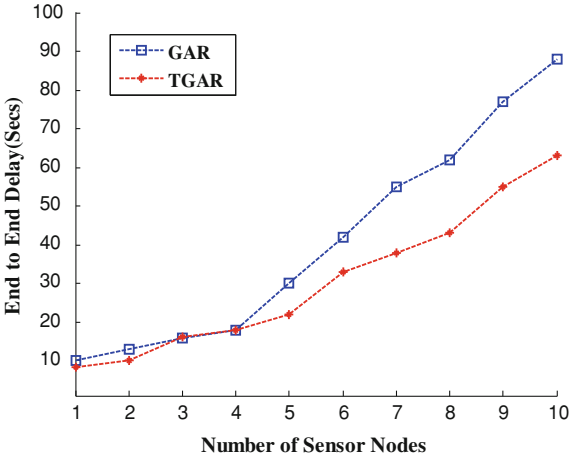


Fig. 5.1 End to end delay versus number of nodes

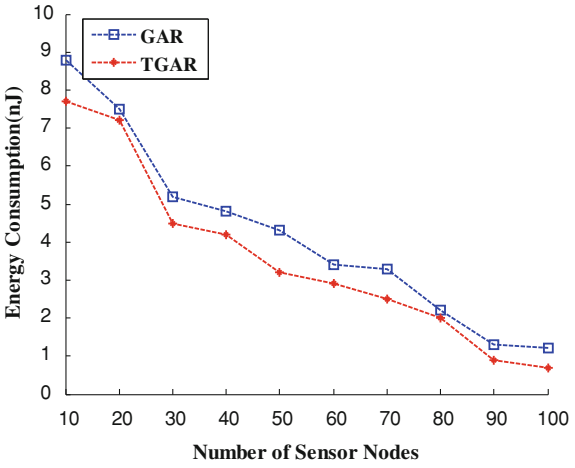


Fig. 5.2 Energy consumption versus number of nodes

5.6 Conclusions

Wireless Sensor Networks are constrained with limited battery and lifetime. We propose Trust dependent Greedy Anti-void Routing (TGAR) to conserve energy and to maximize the reliability of data delivery. The TGAR algorithm overcomes the void and obstacle problem encountered in GAR that results in larger delay and higher consumption of energy. We design a mathematical model to obtain the trust values from the history using Bayesian Estimation model. It identifies the reliability of this path by checking it's trust value from the history of Bayesian estimation model. Simulation result shows that TGAR saves 12 % energy and 33 % faster than GAR.

References

1. Chen S, Fan G, Cui JH (2006) Avoid "Void" in geographic routing for data aggregation in sensor networks. Int J Ad-hoc Ubiquitous Comput 1(4):169–178
2. Kwon S, Shroff NB (2006) Geographic routing in the presence of location errors. Int J Comput Netw 50(15) 2902–2917
3. Karp B, Kung HT (2006) GPSR: greedy perimeter stateless Routing for wireless networks. In: Proceedings of the Mobicom, pp 1158–1168
4. Frey H, Gorgen D (2005) Planar graph routing on geographical clusters. J Adhoc Netw 3:560–574
5. De S, Caruso A, Chaira T, Chessa S (2006) Bounds on Hop Distance in Greedy Routing Approach in Wireless Ad Hoc Networks. Int J wirel Mobile comput 1(2):131–140
6. Dimitriou T, Krontiris I (2007) Gravity: geographic routing around voids in sensor networks. Int J Pervas Comput Commun 2(4):351–361
7. Liu W-J, Feng K-T (2009) Greedy routing with anti-void traversal for wireless sensor networks. IEEE Trans Mobile Comput 8(7):910–922

Chapter 6
DoS Attack-Resistant Framework for Client/Server Intelligent Computing

Shintaro Tabata and Shin-ya Nishizaki

Abstract Nowadays, the client/server model is a representative distributed computing model, its most typical use being for web systems. The services provided by web applications are continually being developed to provide higher-level functions, which is creating the danger of Denial-of-Service attacks. We therefore propose a DoS attack-resistant framework using the client–server model. The focal point of this research is load reduction of the servers through hint information sent from clients to servers. We made a script generator that generates server-side and client-side scripts from one common script code. We implemented two client–server systems using the proposed script generator and evaluated the efficiency of the systems developed by the proposed framework.

Keywords Distributed computing · Client/server model · Denial-of-service attack · Software verification · Term rewriting

S. Tabata (✉) · S. Nishizaki
Department of Computer Science, Tokyo Institute of Technology, 2-12-1-W8-69, Ookayama, Meguro-ku, Tokyo, 152-8552, Japan
e-mail: shintaro.tabata@lambda.cs.titech.ac.jp

S. Nishizaki
e-mail: nisizaki@cs.titech.ac.jp

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_6,

6.1 Introduction

6.1.1 Denial-of-Service Attack

A denial-of-service (DoS) attack is an attempt to make a computer service unavailable to its users. The first study of the formalization of DoS attacks on communications protocols, and how to resist such attacks was performed by Meadows [1]. She extended the Alice-and-Bob notation by annotating the computational costs in processing data packets. Although the property was deeply related to operational behavior, cost annotation was assigned to each communication operation independently of the operational behavior. We therefore proposed another formal framework called spice calculus; this is based on process calculi where the cost estimation mechanism is linked to operational behavior [2, 3]. We can use this calculus successfully to formalize DoS attack resistance; however, it can only handle point-to-point communication, not broadcast communication.

As noted in [1, 2] and [4], an imbalance between a server and its clients creates vulnerability to DoS attacks.

6.1.2 Archive Server and Estimate-Attaching Method

In the paper [6], Nishizaki and Tamano studied the design of an equation archive, which is a storage system for equational systems. The main features of an equation archive are integrity (every equational system accepted into the archive is guaranteed to be complete) and 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, the 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. To ensure accessibility, potential vulnerability to denial-of-service attacks must be carefully examined.

As mentioned above, it is important to reduce vulnerability to denial-of-service at-tacks. They proposed a method in which a client system attaches the estimated computational cost of a request to the equation-archive server. Instead of imposing an upper limit of computational cost, the client system declares the computational cost of the request to the server. This method also controls the total load on the server and enables more appropriate resource allocation. The method is called the *estimate-attaching method*.

6.1.3 Research Purpose

In this paper, we propose a distributed computation model which is DoS attack-resistant, by extending and generalizing the estimate-attaching method proposed in the previous work [6].

6.2 DoS Attack-Resistant Design for Server-Side Computation

6.2.1 Overview

As mentioned in [1, 2, 6], the vulnerability to denial-of-service attacks is caused by the server-side having a heavier load than the client-side. We therefore consider how to reduce the server-side computational cost. The outline of our method is as follows:

1. The client generates information that guarantees safety with respect to DoS attacks.
2. The client sends the information to the server;
3. The server receives the information and carries out the computation, referring to the received information.

The information generated by the client in Step 1 gives the upper limit of the computational borne in the server and helps the computation in the server as a *hint.*

In the previous work [6], Nishizaki designed and implemented a theorem archive server. A client-side prover submits a theorem with its proof to the theorem archive server and then the server checks the proof sent from the client, instead of proving the theorem again. Since the proving procedure is essentially heavier than the checking procedure, avoiding this proving is important to improve the DoS attack-resistance of the server.

The DoS attack-resistant design proposed in this paper is an extension of the design of the theorem archive server as a general framework for DoS attack-resistant client-server systems whose servers are open to anonymous clients and which provide resource-consuming services.

We assume that the target client-server systems are in “proving-and-checking” style: in the client-side part, resource-consuming computation (“proving” in the previous work) is executed and the results are sent to the server-side part. In the server-side part, the results are checked for correctness. The checking could be naively done as re-proving.

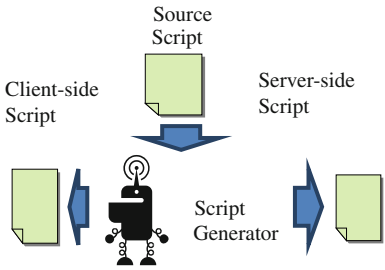


Fig. 6.1 Script Generator

6.2.2 Script Generator, Measuring Function and Proving Function

The client-side and server-side scripts in the system are generated from one source script by Script Generator. In the source script, two kinds of function are inserted as appropriate: measuring functions and proving functions.

Measuring Function. Measuring function calls are associated with calls of the function determining the computational cost. For example, in term-rewriting processing, a measuring function is associated with the substitution function's call.

Proving Function. Proving functions in the client-side scripts provide resource-consuming computation, typically, automated theorem proving. On the other hand, we should reduce the computational cost for checking the result, since the computational cost in the server causes DoS attack vulnerability. Therefore, the proving functions in the server-side scripts are implemented checking the results received from the client, instead of computing them again in the server ("re-proving"). Proving functions in the client-side scripts generate information for server-side checking, and those in the server-side scripts check the results of the client by using the information as a kind of "hint".

Script Generator. Since the corresponding server-side script and client-side script share the same algorithmic structure, these two kinds of script are generated from one source script by the Script Generator. In the source script, calls of measuring functions and proving functions are explicitly designated. The Script Generator rewrites each function call appropriately. For example, the measuring functions in the client-side script count how many times they are called in the client and send the number of times to the server; those in the server-side check that the received number is the same as the number of times that the functions are called in the server (Figs. 6.1 and 6.2)

The generated part should compute the essential processing in the system; the interface part and the communication part should be given individually for both the client and the server.

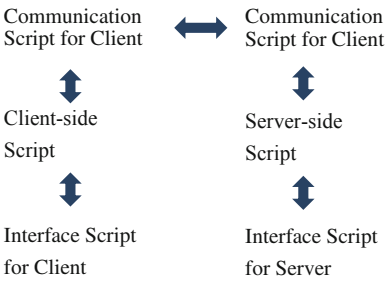


Fig. 6.2 Client-Server System

6.3 Implementation and Evaluation of Prototype System

In order to evaluate our proposed framework, we implemented the following two systems based on the framework.

- Theorem Archive for Equational Logic and Term Rewriting System: This system checks the equivalence between a given equation system and a given term rewriting system, and verifies the completeness of the term rewriting system. This is implemented in Nishizaki et al.'s previous work [6] and we use a similar one in the new framework. The proving function is assumed to be the Knuth-Bendix completion procedure [7]. The measuring functions are assigned to a substitution function of terms and a testing predicate for termination, which checks that a critical pair is found during the completion procedure.
- Interactive theorem prover for first-order predicate logic: This system is an interactive proof checker. A user makes a proof script interactively via a web-based interface, which is assumed to be the proving function of this system. The measuring function is not provided in this system, since the termination of the server-side processing is guaranteed by the finiteness of the sizes of the proofs sent from the client.

We evaluated our proposed framework by comparison with the previous style of implementation [6], that is, the first implementation in the above items. We considered an equation system that formalizes the Ackermann function. We give the following equation system to the system:

a(0,X) = s(X), a(s(X), 0) = a(X, s(0)), a(s(X), s(Y)) = a(X, a(s(X), Y)),
f(X) = a(s(s(0)), **n**)

The function symbols a(−, −) and s(−) represent the Ackermann function and the successor function. The natural numbers are represented based on Peano's encoding. For example, 2 is encoded as s(s(0)). The first three equations represent the Ackermann function. During the Knuth-Bendix completion procedure, the term a(s(s(0)), n) is computed and requires a great amount of computational cost. We gave several variations of the equation system, instantiating the variable n, such as 0, 3, 6, 9, 12, 15,… Then we obtained the result in Table 6.1, in which we list the number of times that the substitution and the unification functions are called in the server-side script of the paper's and the previous implementations.

Table 6.1 Measuring function of substitution and unification calls

	This paper's implementation		Previous implementation	
n	Substitution calls	Unification calls	Substitution calls	Unification calls
0	61	62	33	36
3	100	72	111	58
6	175	72	261	58
9	286	72	261	58
12	433	72	483	58
15	616	72	777	58
18	835	72	1,143	58
21	1,090	72	1,581	58
24	1,381	72	2,673	58
27	1,708	72	3,327	58
30	2,071	72	4,053	58

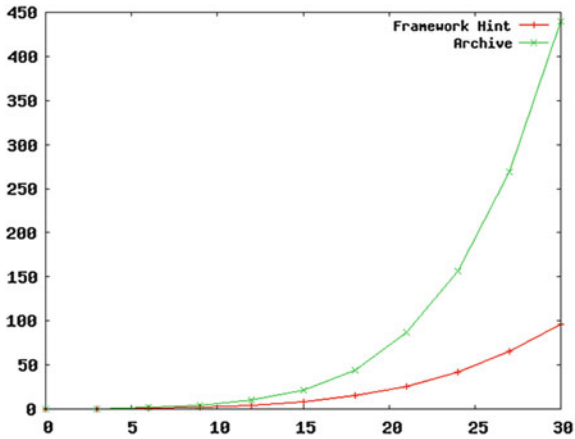


Fig. 6.3 Execution times between the two implementations

It is known that although the values of the measuring function of substitution increase if the values of **n** are increased, the number of unification calls in our implementation does not change as it does in the previous implementation, which means that the information sent to the server is effective as a hint, and consequently the number of times that the unification function is called is controlled.

In Fig. 6.3, we make a comparison between the execution time of the previous implementation and that of this paper's implementation. The X-axis represents the value of **n** and the Y-axis represents execution time in seconds. The performance of this paper's implementation is improved in comparison with the previous implementation. The hint information sent from the client to the server contributes to the improvement in execution time.

6.4 Concluding Remarks

We proposed a DoS attack-resistant framework for a client-server model, which is an improved and generalized extension of the previous study [6] on the open equation archive server. We plan several future studies as part of this research. For example, at the moment we have to write a frontend interface section and backend a reasoning section separately when we develop a server-client system using the proposed framework. However, this reduces the efficiency of program development. We should therefore construct a development environment that enables us to develop the client–server systems more synthetically.

Acknowledgments This work was supported by Grants-in-Aid for Scientific Research (C) (24500009).

References

1. Meadows C (1999) A formal framework and evaluation method for network denial of service. In: Proceedings of the 12th IEEE computer security foundations workshop, pp 4–13
2. Tomioka D, Nishizaki S, Ikeda R (2004) A cost estimation calculus for analyzing the resistance to denial-of-service attack. Software security—theories and systems. Lecture notes in computer science, vol 3233. Springer, New York, pp 25–44
3. Ikeda R, Narita K, Nishizaki S (2011) Cooperative model checking and network simulation for cost analysis of distributed systems. Int J Comput Appl 33(4):323–329
4. Aura T, Nikander P, Leiwo L (2001) DoS-resistant authentication with client puzzles, security protocols the 8th international workshop on security protocols. Lecture notes in computer sciences vol 2133. pp 170–177
5. Denial of service attacks (1997) http://www.cert.org/tech_tips/denial_of_service.html, CERT
6. Nishizaki S, Tamano H (2012) Design of open equation archive server resistant against denial-of-service attacks. "AIM 2012", in printing
7. Baader F, Nipkow T (1999) Term rewriting and all that. Cambridge University Press, New York

Chapter 7
Effect of Forwarding Strategy on the Life Time of Multi-Hop Multi-Sink Sensor Networks

Kaushik Ghosh and Pradip K. Das

Abstract Lifetime of a sensor network can be extended by judicious energy expenditure. Energy consumed is primarily a function of inter-nodal distance and thus effect of forwarding technique can no longer be overlooked while trying to enhance the lifetime of sensor network. Lifetime of a network has been defined differently in different papers depending upon the nature and application of the sensor network under consideration. In this paper we have proposed a forwarding scheme and have compared the same with greedy forwarding and residual energy based forwarding while finding the lifetime for the sensor network.

Keywords Sensor network lifetime · Multi sink · Fermat point · Residual energy · Internodal distance

7.1 Introduction

Battery replenishment in sensor nodes of a sensor network has been a matter of great concern due to the nature and deployment scenarios of these networks. Thus, measures are required to be taken to enhance the lifetime of sensor networks instead. Sensor network lifetime is dependent directly upon energy consumption of the network more than anything else. In a sensor network the total energy consumed is a

K. Ghosh (✉)
Department of Computer Science and Engineering,
Mody Institute of Technology and Science, Lakshmangarh,
Rajasthan, India

P. K. Das
Faculty of Engineering and Technology, Mody Institute of Technology and Science,
Lakshmangarh, Lakshmangarh, Rajasthan, India

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_7,

summation of the energy required for: (i) sensing, (ii) processing, (iii) transmitting, (iv) receiving and (v) listening.

Of all the components, transmitting and receiving data packets consume the lion's share of energy as compared to the remaining three taken together. Again, between transmission and reception, the former consumes considerably more energy than the latter. So, in a word we can say that energy consumed during transmission/reception is the primary determining factor of the life time for any given sensor network. Thus, selecting an energy efficient data forwarding scheme can be one of the ways of enhancing network lifetime in a WASN. Transmitting/receiving energy is again dependent upon inter nodal distance. Energy expenditure of a network increases exponentially with increase in the total distance traveled by data packets.

Of the different forwarding schemes, greedy forwarding is widely used in sensor and ad hoc networks. Greedy forwarding performs reasonably well for delay sensitive networks. But in sensor networks topology change is less frequent and even nil at most of the time. So using greedy forwarding as the forwarding technique can lead to energy drainage of certain nodes faster than many of their neighbors which in turn would demand new route discovery. That in itself is an overhead and is sure to consume some amount of energy without doing something meaningful. Moreover, this non-uniform depletion of energy results in *holes* [1] within the network for which measures like perimeter routing [2] has to be adopted. Another approach thus adopted in sensor networks is to find out the node with maximum residual energy among a group of other nodes and select it as the forwarding node. Periodic energy beacons can let a node know about the residual energy of its neighbors and thereby select a suitable node for packet forwarding. In [3] the authors have followed a similar kind of an approach using the *energy map* technique. But this obviously may not lead one selecting a forwarding node located at an optimal position such that the transmission distance and thereby the transmission energy is minimized. Since reducing the transmitting distance ensures enhancement in network lifetime for WASNs, Fermat point based packet forwarding schemes are also common in the said type of networks when it comes to energy conservation in a multi-hop multi-sink scenario [4–8].

In this paper we propose a Fermat point based packet forwarding scheme which gives importance both to (i) the distance of a node from the destination and (ii) its residual energy while selecting it as the next forwarding node. A weight age factor of a node is thus calculated based on both the factors and a node selects among its neighbors the one with highest value for this factor as the forwarding node.

The next section contains some of the related works. Section 7.3 explains our forwarding scheme and Sect. 7.4 presents the results. The final section comprises conclusion and scope for future work.

7.2 Related Works

The lifetime of a sensor network is the total time spent by the network functioning properly since it was installed. This time is directly dependent upon the battery life of the nodes. Different citations in the literature reveal that energy consumption is dependent upon the transmitting distance and the volume of data to be transmitted. Since data volume is something we cannot reduce considerably while maintaining a benchmark for quality of service, reduction in the transmission distance has been seen by the researchers as a viable solution for enhancing the network lifetime. It is known that a WASN once deployed, hardly ever changes its topology during its lifetime. So, Fermat point based forwarding schemes [4–8] have gained popularity for the stated purpose. That is because, in this type of forwarding scheme minimum distance traversal is guaranteed for data packets from source to destination.

Authors of [4, 5] were to our knowledge the first to introduce the concept of Fermat point based forwarding to gain energy efficiency in geocast routing protocols. Ghosh et al. [6, 7] have followed them to introduce an alternate novel method of finding the Fermat point when the numbers of geocast regions are more than two. While defining sensor lifetime, authors in [9] have done extensive work in subdividing sensor lifetime into different categories viz. lifetime based on number of alive nodes [10–12], lifetime based on coverage [13–16], lifetime based on coverage and connectivity [17, 18] and lifetime based on QoS [19–23]. Talking about lifetime based on number of alive nodes, some are of the opinion that lifetime is the time from deployment till the time the first node was drained out of its energy (n out of n) [10]. The definition, though simple, is clearly an impractical one. People thus altered the definition to m out of n form i.e. while m nodes in the network are alive, one is free to consider the network as functional. A variant of this definition was followed in [11]. In their view, a network is alive till the first cluster head is drained out of energy. [12] proposes a definition which is ready to consider a network alive till the last node is functional.

Coming to the coverage part, [13] came up with the idea of k-coverage i.e., consider a network as alive till an area of interest is covered by k nodes. Authors in [14, 15] however were of the opinion that a network is considered to be alive till the whole area is covered by at least one node. Some again considered both connectivity and coverage while defining lifetime of a network [16]. Giridhar and Kumar [17] considered the number of successful data gathering trips as the determining parameter for lifetime but authors in [18] decided upon the total number of transmitted messages for the same.

Lifetime of a sensor network has also been decided taking QoS into account [19]. According to this line of thought network can be treated as alive till an acceptable event detection ratio was maintained. [20–23] have marked a network alive till it satisfies the application requirement.

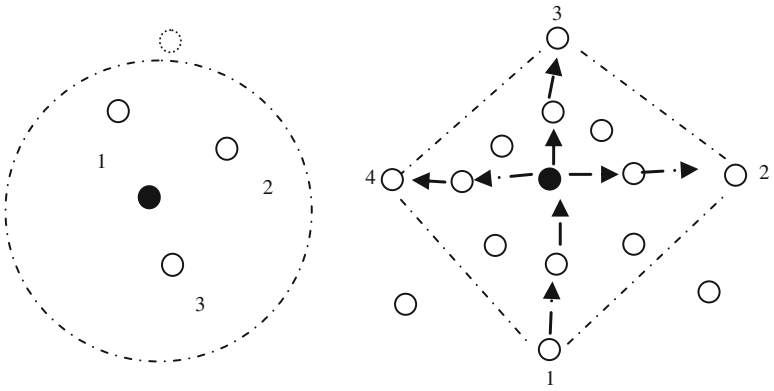


Fig. 7.1 **a** Different possible source-sink multi-hop transmissions. **b** Transmission to multiple sinks via Fermat point

7.3 Proposed Scheme

We propose a Fermat point based forwarding scheme where both residual energy and distance from the sink are taken into consideration while selecting the next forwarding node. Before explaining the scheme, let us mention the assumptions we have made throughout the text.

Assumptions

- Every sensor node may either act as a sender or a relay node.
- Mobility of the nodes has been considered to be *practically* nil.
- All the nodes are homogenous as far their transmission and computation capabilities are concerned.
- The nodes are deployed in a pseudo-random manner and at the time of deployment all of them are assumed to possess same energy.

All nodes are GPS enabled and are capable of knowing the location and residual energy of their neighbors through periodic beacon messages.

According to the scheme, a node while selecting a forwarding node from its neighbors needs to keep a note of the residual energy of its neighbors along with their respective distances from the sink. Let us take the scenario depicted in Fig 7.1a. The BLACK node is the source with its neighbors inside the dotted circle. The circle represents the transmission range of the said node.

But since we are taking residual energy of a node into consideration along with its distance from the sink, the selection criterion wouldn't be that simple. Every node is weighed on two different parameters. One, its geographical location and the other its present residual energy. This would result in the possibility of selection of a different forwarding node after every transmission. That is because, after every transmission/reception residual energy of a node changes and that in turn changes the possibility for it to continue as the forwarding node for any

source-sink pair. A sensor node would first try to transmit a data packet to the Fermat node [6, 7] and from there it is the responsibility of the Fermat node to relay the packet to different sinks (Fig. 7.1b). The circle marked 1 represents a sender and the circles marked 2, 3 and 4 are the different sinks. The black circle is the Fermat node—the node closest to the theoretical Fermat point for the enclosed region formed by the sender node and different sinks. All other dots are acting as relay nodes for the time being. At some other time, any one of these relay nodes may act as sender resulting in calculation for a new Fermat point and thereby a new Fermat node. That is because the enclosed region now formed has one vertex different from the existing one. In the present forwarding scheme every node in addition to keeping a track on the number of neighbors also maintains the **forwarding potential (κ)** of the respective neighbors. κ for any node is calculated as

$$\kappa = res\_energ / dist.$$

Where,

res_energy	Residual power of a node in mWatts
dist	Distance of the node from the sink in meters
κ	Forwarding potential of a neighboring node in **mWatts/meter**

Here, same weightage is given to the nodes for their geographical vicinity to the destination and the amount of residual energy present in them at any particular instance. However, a variant of this forwarding scheme can be obtained by having different weightages for ***res_energy*** and ***dist***.

Same will be the result when we consider the node with maximum residual energy as the best possible option for forwarding a message in case we need to go in for multi-hop routing. Say in our example node 2 is selected by the source as forwarding node after calculating κ for all its neighbors. So, for the time being, κ for node 2 has the highest value among all the other nodes. Considering topology change to be most unlikely of an event, we find **res_energy** of a node to be the parameter that would change after every transmission/reception. The value of κ thus also changes for the nodes resulting in possible selection of a new neighbor as the forwarding node. The immediate effect of this approach is elimination of the possibility of energy drain out of any particular node due to excessive usage. If after **t** transmissions some new node is selected for forwarding, then the previous forwarding node has a chance to rejuvenate its battery either by energy harvesting (**GEBRES**) [24] or by the natural bounce back capability of the battery.

The radio model decided for the scheme is in fact a result of constant evolution upon the equation proposed in [25]. It was modified by Hwang and Pang [26] and was again used in [7] as well. The radio model used in [26] has a serious drawback of considering the transmission medium to be free space i.e. they have made energy consumed for transmission proportional to **d** with its power being raised to 2. Where, d is inter nodal distance. But in practice that is hardly going to be the case. In fact the power of d can assume any value between 2 and 6 [27]. So it is better that we keep that flexibility in our radio model. To generalize the equation one can take

a middle path to assign 4 as the value of n. Again, from Frii's free space equation we know that the power consumed by a receiving node is not independent of the power consumed for transmission and is a factor of the distance between the sender and receiver [6]. Findings in [28] show that power consumed by a receiving node due to the presence of transmitting node at a distance of 100 meters from itself is 10^{-8} W. Since in our case we have considered the transmission range of a node to be 100 m, the radio model proposed in this paper is as follows

$$\boldsymbol{P_{TX}(m', d)} = \boldsymbol{m' * E} + \boldsymbol{m * \underline{\varepsilon} * d^n}$$

$$P_{RX}(m') = m' * E + 10^{-8}$$

where,

$\boldsymbol{P_{TX}}$	Power required for transmission
$\boldsymbol{P_{RX}}$	Power required for reception
$\boldsymbol{m}$	Data rate
$\boldsymbol{\dot{\varepsilon}}$	Permittivity of the transmitting medium
$\boldsymbol{E}$	Minimum energy consumed for transmission/reception irrespective of the total number of bits or inter nodal distance
$\boldsymbol{d}$	Distance between nodes
$\boldsymbol{n}$	path loss exponent whose value lies between 2 and 6

The above equation can be used when **n** has a value of 2. For higher values of **n** the equation needs to be converted into dBm form. The generalized radio model for any value of **n** is therefore of the following form

$$\boldsymbol{P_{TX}(m', d)dBm} = \boldsymbol{10log_{10}(m' * E} + \boldsymbol{m * \varepsilon * d^n})$$

$$\mathrm{P_{RX}}(m')d\mathrm{B}m = 10log_{10}\left(m' * E + 10^{-8}\right)$$

We have considered the operation scenario of the WSN to be like [5, 6] and [7]—where there are multiple geocast regions/sinks and data forwarding takes place via Fermat point for the triangular/polygonal region with sinks and source as the vertices. Before ending this section we would like to represent the forwarding scheme in algorithmic form. Following are the steps for our proposed forwarding scheme.

Input: Node_id of the destination
Output: Boolean.
D_ID: Node_id of the destination.
N_ID: Node_id of the node presently holding the packet.
NN: Number of neighbors of the forwarding node.
neigh[]: Node_id of all the neighbors for a node.
res-energy[]: Residual energy of all the neighbors for a node.
Neigh_ID = Node_id of a neighbor.
flag = 0

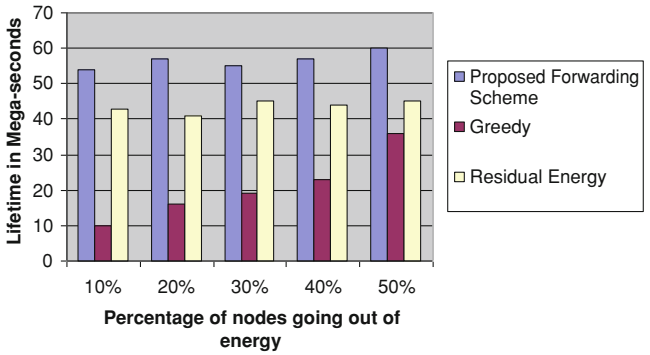


Fig. 7.2 Lifetime comparison for different forwarding schemes with similar deployment pattern

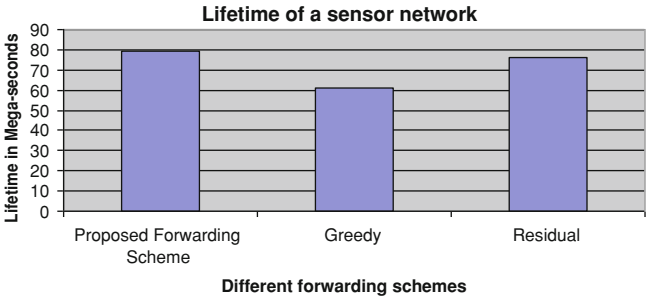


Fig. 7.3 Average lifetime of a network for different forwarding schemes with 25 % fade out

```
for(i = 0; i < NN; i++)
{d = dist (N_ID, D_ID);
if (d <= 100)
flag = 1;
if (flag == 0)
next_hop = KAPPA(D_ID, NN, neigh[], res_energy[]);
else
next_hop = D_ID;}
```

Function	Definition
dist (N_ID, D_ID)	**Input:** Node id of the destination, Node id of the node presently holding the packet **Output:** Distance between the nodes in meters
KAPPA (D_ID,NN,neigh[],res-energy[])	**Input:** Node id of the destination, Number of neighbors, Node id of all the neighbors, Residual energy of all the neighbors **Output:** Node id of the neighbor with highest value for κ within the neighbor list

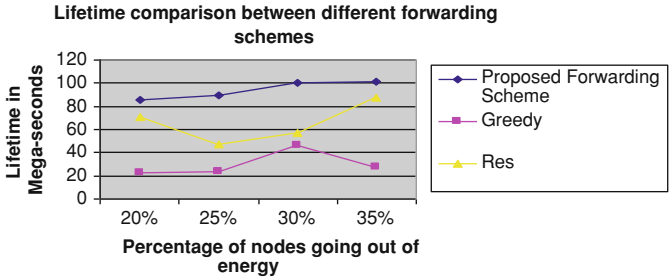


Fig. 7.4 Average lifetime of a network for different forwarding schemes with readings within 80 % window

7.4 Results

The experimental scenario comprises a rectangular 350 × 350 m^2 area with sinks at its corners. Within the said area there are 100 homogeneous nodes deployed in pseudo-random manner with transmission radius of each node equal to 100 m. All these nodes, as stated earlier, sense and transmit data other than acting as relay nodes. The initial energy assigned to the nodes is 2,400 J and the data rate is 256 kbps. The following table contains the different experimental parameters.

Of all the definitions discussed in Sect. 7.2, we pick up the one that seemed to us to be the most relevant in the context of our application—network lifetime based on the number of live nodes i.e. m out of n. When some distinct numbers of nodes are dead in a network it gives rise to routing holes. When these holes are too big, it may cause network partitioning and the network may be considered dead even if a considerable number of nodes are still alive. Now obviously it is a matter of discussion that what would be the exact number for m. It would surely depend upon the node density in the network, the transmission range of the nodes and would vary from application to application. We have thus taken different percentage values for m i.e. the percentage of dead nodes and have found out the performance of our proposed scheme with that of the greedy forwarding and residual energy based protocols for the same number of nodes and same deployment pattern. Network lifetime has been calculated in mega-seconds ($\times 10^6$). Figure 7.2 is the corresponding graph for above findings.

Since the experimental data are the outcome of a simulation program, in order to get better results we again have taken a series of results for these three forwarding schemes and have found out the average lifetime of a network in mega-seconds taking 40 readings for each scheme and have taken m as 25 %.

This time for each run we have taken different deployment patterns though (Fig. 7.3). Yet another way of finding the lifetime in simulation environment with some higher degree of precision is to clip off a certain percentage of readings from the top and bottom of the list. The results for the same are plotted in Fig. 7.4.

7.5 Conclusion

The proposed scheme performs better as compared to greedy forwarding and energy based forwarding when it comes to lifetime of a sensor network. Moreover, it takes the best of both greedy and residual energy based forwarding schemes. As future work we are planning to assign different weight age values for geographical location and residual energy of a node while calculating κ. This may throw further light on the energy consumption pattern and thus the lifetime of WASNs.

Acknowledgments We are grateful to the management of Mody Institute of Technology and Science for the facilities provided to the authors for carrying out this work.

References

1. Ahmed N, Kanhere SS, Jha S (2005) The holes problem in WSNs : a survey. ACM SIGMOBILE Mobile Comput Commun Rev 9(2):4–18
2. Karp B, Kung H (2000) GPSR: Greedy perimeter stateless routing for wireless networks. In: Procedings of the ACM/IEEE Mobicom, Aug 2000
3. Mini R, Machado M, Loureiro AF, Nath B (2005) Prediction-based energy map for WSNs. Ad Hoc Netw 3:235–253
4. Song Y-M, Lee S-H, Ko Y-B (2005) FERMA: an efficient geocasting protocol for wireless sensor networks with multiple target regions In: Enokido T et al (ed) EUC workshops 2005, LNCS vol 3823, pp. 1138–1147, © IFIP International Federation for Information Processing 2005
5. Lee SH, Ko YB (2006) Geometry-driven Scheme for geocast routing in mobile adhoc networks. IEEE conference on Vehicular technology-Spring 2006
6. Ghosh K, Roy S, Das PK (2009) An alternative approach to find the fermat point of a polygonal geographic region for energy efficient geocast routing protocols: global minima scheme. AIRCC/IEEE NetCoM 2009
7. K.Ghosh, S Roy and P. K. Das, "I-Min: An intelligent Fermat point based energy efficient geographic packet forwarding technique for wireless sensor and ad hoc networks", International journal on applications of graph theory in wireless ad hoc networks and sensor networks (GRAPH-HOC) vol2 no2 June 2010, pp 34–44
8. Ssu KF, Yang C-H, Chou C-H, Yang A-K (2009) Improving routing distance for geographic multicast with Fermat points in mobile ad hoc networks. Int J Comput Telecommun Netw 53(15):2663–2673
9. Dietrich I, Dressler F (2009) On the lifetime of wireless sensor networks. ACM Trans Sensor Netw 5(1)5.1–5.39
10. Chang JH, Tassiulas L (2000) Energy conserving routing in wireless ad-hoc networks. In: Proceedings of the 19th IEEE Conference on Computer Communications (INFOCOM)
11. Chiasserini C-F, Chlamtac I, Monti P, Nucci A (2002) Energy efficient design of wireless ad hoc networks. In: Proceedings of the 2nd IFIP networking, vol. LNCS 2345, 2002
12. Tian D, Georganas ND (2002) A coverage-preserving node scheduling scheme for large wireless sensor networks. In: Proceedings of the 1st ACM international workshop on wireless sensor networks and applications (WSNA), 2002
13. Mo W, Qiao D, Wang Z (2005) Mostly-sleeping wireless sensor networks: connectivity, k-coverage, and alpha-lifetime. In: Proceedings of the the 43rd annual allerton conference on communication, control, and computing 2005

14. Bhardwaj M, Garnett T, Chandrakasan AP (2001) Upper bounds on the lifetime of sensor networks. In: Proceedings of the IEEE international conference on communications (ICC), vol. 3, pp 785–790
15. Bhardwaj M, Chandrakasan AP (2002) Bounding the lifetime of sensor networks via optimal role assignments. In: Proceedings of the 21st IEEE Conference on Computer Communications (INFOCOM), vol. 3, pp 1587–1596
16. Cardei M, Wu J (2004) Coverage in wireless sensor networks. In: Ilyas M (ed) Handbook of sensor networks. CRC Press, West Palm Beach
17. A. Giridhar, and P. Kumar, "Maximizing the functional lifetime of sensor networks", Proceedings of the 4th International Symposium on Information Processing in Sensor Networks (IPSN), 2005
18. Yu Y, Govindan R, Estrin D (2001) Geographical and energy aware routing: a recursive data dissemination protocol for wireless sensor networks. Tech. Rep. UCLA/CSD-TR-01-0023, UCLA Computer Science Department
19. Tian D, Georganas ND (2002) A coverage-preserving node scheduling scheme for large wireless sensor networks. In: Proeedings of the 1st ACM international workshop on wireless sensor networks and applications (WSNA), pp 32–41
20. Blough MD, Santi P (2002) Investigating upper bounds on network lifetime extension for cell-based energy conservation techniques in stationary ad hoc networks. In: Proceedings of the 8th ACM international conference on mobile computing and networking (MobiCom)
21. Kumar S, Arora A, Lai TH (2005) On the lifetime analysis of always-on wireless sensor network applications. In: Proceedings of the IEEE international conference on mobile ad-hoc and sensor systems (MASS)
22. Tilak S, Abu-Ghazaleh NB, Heinzelman W (2002) A taxonomy of wireless micro-sensor network models. ACM SIGMOBILE Mobile Comput Commun Rev 6:2
23. Wieselthier JE, Nguyen GD, Ephremides A (2002) Resource management in energy limited, bandwidth-limited, transceiver-limited wireless networks for session-based multicasting. Comput Netw Int J Comput Telecommun Netw 39(5):113–131
24. Zeng K, Lou W, Ren K, Moran PJ (2006) Energy efficient geographic routing in environmentally powered WSNs. In: Procedings of military communications conference, MILCOM 2006
25. Heinzelman WB, Chandrakasan AP, Balakrishnan H (2002) An application-specific protocol architecture for wireless micro sensor networks. IEEE Trans Wirel Commun 1(4):2002
26. Hwang I-S, Pang W-H (2007) Energy efficient clustering technique for multicast routing protocol in wireless adhoc networks. IJCSNS 7:8
27. http://www.utdallas.edu/∼torlak/courses/ee4367/lectures/lectureradio.pdf
28. Bhattacharya PP, Banerjee PK (2006) User velocity dependent call handover management. Int J HIT Trans ECCN 1(3):150–155

Chapter 8
Concept Similarity and Cosine Similarity Result Merging Approaches in Metasearch Engine

K. Srinivas, A. Govardhan, V. Valli Kumari and P. V. S. Srinivas

Abstract Metasearch engines provide a uniform query interface for Internet users to search for information. Depending on users need, they select relevant sources and map user queries into the target search engines, subsequently merging the results. In this paper, we have proposed a metasearch engine, which have two unique steps (1) searching through surface and deep web, and (2) Ranking the results through the designed ranking algorithm. Initially, the query given by the user is given to the surface and deep search engines. Here, the surface search engines like Google, Bing and Yahoo are considered. At the same time, the deep search engine such as, Infomine, Incywincy and CompletePlanet are considered. The proposed method will use two distinct algorithms for ranking the search results, which are concept similarity and cosine similarity.

K. Srinivas (✉)
Department of IT, Geethanjali College of Engineering and Technology, Cheeryal(V), Keesara(M), Ranga Reddy, Andhra Pradesh 501301, India
e-mail: katkamsrinu@gmail.com

A. Govardhan
Jawaharlal Nehru Technological University, Kukatpally, Hyderabad, Andhra Pradesh 500 085, India
e-mail: govardhan_cse@yahoo.co.in

V. Valli Kumari
Department of CS and SE, Andhra University College of Engineering, Andhra University, Visakhapatnam, Andhra Pradesh 530003, India
e-mail: valli_kumari@rediffmail.com

P. V. S. Srinivas
Department of CSE, Geethanjali College of Engineering and Technology, Cheeryal(V), Keesara(M), Ranga Reddy, Andhra Pradesh 501301, India
e-mail: pvssrinivas_70@yahoo.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_8,

Keywords Metasearch engine · Concept · Cosine similarity · Deep web · Surface web

8.1 Introduction

In our work, we tried to develop an advanced metasearch engine. The proposed visible and invisible web-based Metasearch engine is divided into two major steps, (1) searching through surface and deep web, and (2) Ranking the results through the designed ranking algorithm. Initially, the query given by the user is given to the surface and deep search engines. Here, the surface search engines like Google, Bing and Yahoo can be considered. At the same time, the deep search engines such as, Infomine, Incywincy and CompletePlanet can be considered [1]. Once more number of pages are obtained from the visible and invisible web, the ranking of those pages is carried out to provide the most relevant pages. The ranking of those pages is carried out using the proposed algorithm that considers the similarity of input query to those web pages as well as the inter-similarity among the web pages retrieved. Finally, the experimentation is done to prove the efficiency of the proposed visible and invisible web-based Metasearch engine in merging the relevant pages [2].

The main contributions of our proposed approach are the two distinct algorithms that we have adapted for the search in the web. The algorithms are based on similarity measures, one algorithm is based on concept similarity and other is based on cosine similarity. We use the search results from both surface web search and deep web search as the input to the two algorithms. The proposed approach has given significant result in the experimentation phase [3, 4].

8.2 Proposed Metasearch Engine Algorithms

The proposed method is concentrated mainly on two algorithms. The basic architecture is build up from the search results obtained from the deep web search and the surface web search. The user has full control on giving the query to the proposed metasearch engine. We have proposed two algorithms for the processing in the metasearch engine [5].

1. Algorithm-1: Concept similarity based metasearch engine
2. Algorithm-2: Cosine similarity based metasearch engine

8.2.1 Algorithm-1: Concept Similarity Based Metasearch Engine

A concept is a keyword which has some relation with the documents, and has some particular characteristics. This concept is related to documents and is also related with other concepts in the domain, which it belongs. The set of keywords are the input in this first stage of concept map extraction. Consider that we have a domain which consists of a set of concepts,

$$D = k_1, k_2, \ldots\ldots\ldots, k_n$$

Where, D is the domain and $k_i$is concept that belongs to the domain. The aim of this step is to find the relation between the keywords and hence finding concepts to the domain. We adopt a sentence level windowing process, in which the window moves in a sliding manner. The text window formed is four term window which enclosed in a sentence. Initially, we find the highest frequent word and then, the approach finds the dependency of this word to other and other words to this [6].

$$freq(x) = \frac{X_n}{N_k}, X \in k.$$

Here, we find the most frequent element using the above mentioned equation. X_n represents the number of x present in the domain, where x is the element which is subjected for the frequency finding. N_k is the total number of elements in the domain. After finding the most frequent keyword, we have to find whether it belongs to a concept in the concept map [7, 8]. The selection that keyword to a concept is done by finding the inter relation between that keyword and other keywords. The bond between two keywords are obtained through finding the probability of occurrence of the keywords, we adopts a conditional probability for finding the relation between the keywords. The value of the dependency is used to extract the concept. If the keyword shows higher dependency between others, then it is considered as concept. Analysis of the method shows that the more the dependency the more the concept gets extracted from the text corpora. The dependency of the terms can be calculated through the following way,

$$dep(x : y) = \frac{P(x|y)}{P(y)}, x, y \in D.$$

$$P(x|y) = \frac{P(y \cap x)}{P(x)}$$

The function $dep(x : y)$ is used for finding the dependency between the terms and thus extract the concept which is required for the concept map extraction. The terms x and y represents the terms from the domain D. The function $P(.)$ is the probability of each word present in the domain. Both, the conditional probability and the probability are used in the proposed approach. Thus, the concept belong to each document is found from the further processes. The concept is then considered

as a term and the term belongs to the set of documents is obtained after the search [9, 10]. The concepts is prominent character for the documents which posses it. The main advantage of the concept is, it is composed of one or more top N keywords extracted from the documents. The concepts possess prominent part in the proposed metasearch engine. The next phase of the proposed approach is building an N × M matrix. In which the term frequency of the concept is identified in the documents in accordance with the search query.

Example 1 For a query like *"data mining"*.

Step 1 find the frequency of *"data"*, i.e. P(data)
Step 2 find the frequency of *"mining"*, i.e. P(mining)
Step 3 find the frequency of *"mining"* and *"data"*, i.e. P(min $ing \cap data$)
Step 4 find $\frac{P(\min ing \cap data)}{P(\min ing)}$, i.e. $P(\min ing|data)$
Step 5 find $\frac{P(\min ing|data)}{P(data)}$, i.e. dep(mining : data)
Step 6 dep values are passed to N × M formation.

In the similar way, concept in the every document is extracted and that terms are subjected for the N x M matrix calculation.

8.2.1.1 N × M Matrix Formation

The *dep* values of the document are arranged in an N x M matrix for the final calculation. In the N × M matrix the N is the number of concepts and M is the number of documents, which obtained from the search engines. All the dependency values of the terms in the documents are calculated using the $dep(x : y) = \frac{P(x|y)}{P(y)}, x, y \in D.$ formulae. Then a row wise sum operation is initiated on each document to find its relevance to the search based on the terms it posses.

$$\begin{array}{cccccc} & d_1 & d_2 & & d_n & \sum values \\ c_1 & dep(c_1,d_1) & dep(c_1,d_2) & & dep(c_1,d_n) & \sum_{n=1}^{N} dep(c_1,d_n) \\ & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_n & \cdots & \cdots & \cdots & \cdots & \sum_{n=1}^{N} dep(c_n,d_n) \end{array}$$

N × M matrix

Where d_1, d_2 …., $d_n \in D$, and $c_1, c_2, \ldots, c_n \in C$, C is the set concepts. The $\sum values$ are calculated and then it is sorted in descending order and a threshold is set for the $\sum values$. The $\sum values$ those are higher than the thresholds are selected as the search results and those documents are retrieved to the user as the final search result.

8.2.2 Algorithm-2 Cosine Similarity Based Metasearch Engine

The next phase of the proposed method deals with the term frequency (TF) and the inverse document frequency (IDF). The TF-IDF function is the centre of the proposed method, i.e. their values controls the flow of the proposed method. The term frequency–inverse document frequency is a numerical statistic which reflects how important a word is to a document in a collection or corpus. It is often used as a weighting factor in information retrieval and text mining. The TF-IDF value increases proportionally to the number of times a word appears in the document, but is offset by the frequency of the word in the corpus, which helps to control for the fact that some words are generally more common than others. Variations of the TF-IDF weighting scheme are often used by search engines as a central tool in scoring and ranking a document's relevance to given user query. In the proposed method the use of the TF-IDF through a specific formula is done [11, 12].

The term count in the given document is simply the number of times a given term appears in that document. This count is usually normalized to prevent a bias towards longer documents to give a measure of the importance of the term t within the particular document D. Thus we have the term frequency,

$$TF(t, D) = No\ of term\ t\ in\ document\ D$$

The inverse document frequency is a measure of whether the term is common or rare across all documents. It is obtained by dividing the total number of documents by the number of documents containing the term, and then taking the logarithm of that quotient.

$$IDF(t, D) = \log \frac{|D|}{|t \in d : d \in D|}$$

Where, $|D|$ is the cardinality of D, or the total number of documents in the corpus and the expression is the number of documents where the term t appears [7].

The proposed method formulates the TF-IDF weightage with specific formulae, which can be given by,

$$TF - IDF(t, d, D) = \sum\nolimits_{t \in d} \frac{(TF \times IDF)_1 \times (TF \times ID)_2}{\sqrt{\sum\limits_{t \in d_1} (TF \times IDF)_1^2 \times \sum\limits_{t \in d_2} \sqrt{(TF \times ID)_2^2}}}$$

This value of TF-IDF is calculated for all the terms in the document and the resultant values are passed to processes the N x N matrix.

Table 8.1 Analysis factors

Search engines	Keywords
Google	Data mining
Bing	Network security
Infomine	Data replication
Incywincy	Image processing

8.2.2.1 N × N Matrix Formation

The TF-IDF values of the document are arranged in an N x N matrix for the final calculation. In the N × N matrix the N is the number of documents, which obtained from the search engines. All the TF-IDF values of the terms in the documents are calculated using the *TF-IDF(t,d,D)* formulae. Then a row wise sum operation is initiated on each document to find its relevance to the search based on the terms it posses [8].

$$\begin{array}{c|cccc} & d_1 & d_2 & & d_n \\ d_1 & [TF-IDF(t,d_1)+ TF-IDF(t,d_1)] & [TF-IDF(t,d_1)+ TF-IDF(t,d_2)] & \cdots & [TF-IDF(t,d_1)+ TF-IDF(t,d_n)] \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ d_n & \cdots & \cdots & \cdots & \cdots \end{array}$$

$$\begin{array}{c} \sum values \\ \sum_{n=1}^{N} [TF-IDF(t,d_1)+ TF-IDF(t,d_n)] \\ \cdots \\ \sum_{n=1}^{N} [TF-IDF(t,d_n)+ TF-IDF(t,d_n)] \end{array}$$

N × N matrix

Where $d_1, d_2 \ldots, d_n \in D$. The $\sum values$ are calculated and then it is sorted in descending order and a threshold is set for the $\sum values$. The $\sum values$ those are higher than the thresholds are selected as the search results and those documents are retrieved to the user as the final search result.

8.3 Results and Performance Analysis

In this section the performance is evaluated of different search engines and with the proposed metasearch engine. The evaluation is done for different search queries and their responses to the evaluation function. The performance of the proposed system is different for different keywords given to it [13]. The behavior of the search engines and the proposed method is evaluated according to the given keywords. In this process we consider the following search engines, Google and Bing as surface search engines and Infomine and Incywincy as deep web search engines [14]. The performance of the above mentioned search engine are compared with the proposed metasearch engine based on the two algorithms, i.e. Concept based metasearch engine and TF-IDF weight age based metasearch engine (Table 8.1).

Table 8.2 TSAP values for "data mining"

Search engine	N = 10	N = 20
Google	0.40	0.55
Bing	0.30	0.35
Infomine	0.60	0.60
Incywincy	0.53	0.60
Proposed algorithm 1	0.75	0.80
Proposed algorithm 2	0.76	0.82

Consider the analysis based on keyword data mining. In the proposed method we have two algorithms to process with. So the data mining is given as input query to the search engines, according to the algorithm one it will generate some documents related to data and mining. Top n keywords from the documents are selected and then the concept is generated as "Data Mining" with the help of dependency value. Then from the N × M matrix, the relevant web sites or documents are selected. The responses of the keyword "Data Mining" is given below. The TREC-style average precision (TSAP) values of the concept data mining are plotted in the below table [15].

The analysis from the Table 8.2 showed that the most ranked results are generated for the two proposed algorithms. The value obtained are 0.76 and 0.75 @N = 10 and 0.82 and 0.80 @N = 20 respectively for algorithm 2 and algorithm 1, which is higher value than the other search engines considered in the evaluation process. It can be stated that, the results are more feasible with the proposed methods. Similarly all other keywords are processed with the above stated search engines.

The analysis from the Table 8.3 showed that the most ranked results are generated for the two proposed algorithms. The value obtained are 0.76 and 0.78 @N = 10 and 0.83 and 0.81 @N = 20 respectively for algorithm 2 and algorithm 1, which is higher value than the other search engines considered in the evaluation process. The response to the second keyword is little bit higher than the first keyword.

The analysis from the Table 8.4 showed that the most ranked results are generated for the two proposed algorithms. The value obtained are 0.75 and 0.77 @N = 10 and 0.84 and 0.84 @N = 20 respectively for algorithm 2 and algorithm 1, which is higher value than the other search engines considered in the evaluation process. In this case, the Algorithm one performs little more sensitive to the give keyword than the other search measures.

The analysis from the Table 8.5 showed that the most ranked results are generated for the two proposed algorithms. The value obtained are 0.69 and 0.74 @N = 10 and 0.80 and 0.82 @N = 20 respectively for algorithm 2 and algorithm 1, which is higher value than the other search engines considered in the evaluation process.

Table 8.3 TSAP values for "network security", analysis based on keyword "network security"

Search engine	N = 10	N = 20
Google	0.35	0.45
Bing	0.50	0.45
Infomine	0.55	0.50
Incywincy	0.63	0.65
Proposed algorithm 1	0.78	0.81
Proposed algorithm 2	0.76	0.83

Table 8.4 TSAP values for "data replication", analysis based on keyword "data replication"

Search engine	N = 10	N = 20
Google	0.40	0.45
Bing	0.38	0.43
Infomine	0.60	0.65
Incywincy	0.68	0.75
Proposed algorithm 1	0.77	0.84
Proposed algorithm 2	0.75	0.84

Table 8.5 TSAP values for "image processing", analysis based on keyword "image processing"

Search engine	N = 10	N = 20
Google	0.57	0.59
Bing	0.48	0.53
Infomine	0.70	0.75
Incywincy	0.78	0.85
Proposed algorithm 1	0.74	0.82
Proposed algorithm 2	0.69	0.80

The evaluation of the four key words states that our metasearch engine is sensitive to the user input and it has upper hand over the other methods in different search criteria [16].

In Fig. 8.1, the comparison of the TSAP values of different keywords are plotted with respect to the concept similarity based algorithm and the cosine similarity based algorithm. The Fig. 8.1 shows that the cosine similarity algorithm performs little less as compared to the concept similarity based algorithm. Even though, by neglecting their individual performance the proposed algorithm performs a way higher than the traditional search engines [17, 18].

The plotting in Fig. 8.2 shows the performance of the proposed approach with all other search engines considered in the experiment.

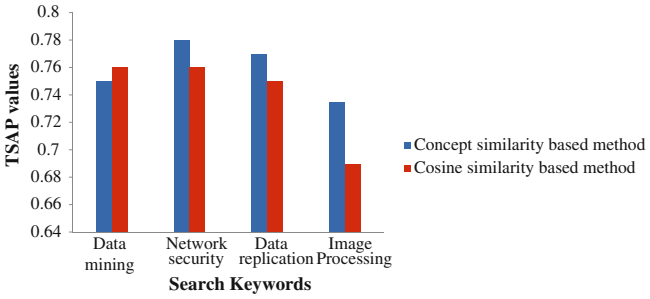


Fig. 8.1 TSAP value comparison 1

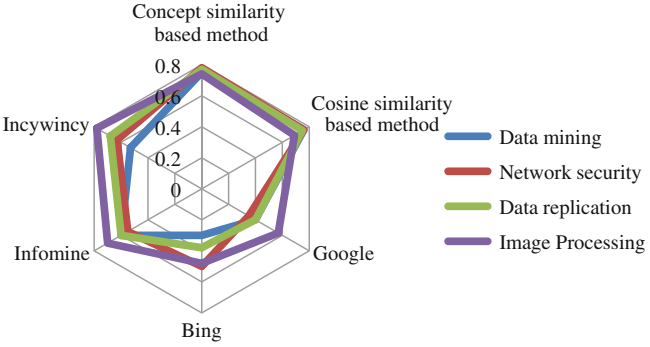


Fig. 8.2 TSAP value comparison 2

8.4 Conclusion

Now-a-days the search engines have been replaced with metasearch engines for getting more accurate and precise outputs. The metasearch engines are used because they are capable of overcome the limitations faced by the normal search engines. The proposed method introduces two algorithms, which improves the metasearch engine results. The proposed method defines two algorithms, they are concept similarity based method and cosine similarity based method. The first one considers the keyword as a concept and finds its relevance to the search criteria, on the other hand, cosine similarity makes use of the term frequency and inverse document frequency. The experimental results have shown that the proposed algorithm out performs the other search engines. The evaluation criteria we used in the proposed algorithms is TSAP. The futuristic advancements can be done by incorporating different evaluation parameters to the proposed methods.

References

1. Jianting L (2011) Processing and fusion of meta-search engine retrieval results. In: International conference on electric information and control engineering, pp 3442–3445
2. Ghaderi MA, Yazdani N, Moshiri B (2010) A social network-based metasearch engine. International symposium on telecommunications, pp 744–749
3. Krüpl B, Baumgartner R (2009) Flight meta-search engine with metamorph. Trans ACM J 1069–1070
4. Bravo-Marquez F, L'Huillier G, R′ios SA, Vel′asquez JD, Guerrero LA (2010) DOCODE-Lite: a meta-search engine for document similarity retrieval. In: Proceedings of the 14th international conference on knowledge-based and intelligent information and engineering systems, pp 93–102
5. Rasolofo Y, AbbaFci F, Savoy J (2001) Approaches to collection selection and results merging for distributed information retrieval". In: Proceedings of the tenth international conference on information and knowledge management, pp 191–198
6. Aslam JA, Montague M, (2001) Models for metasearch. In: Proceedings of the 24th annual international ACM SIGIR conference on research and development in information retrieval, pp 276–284
7. Zaka B (2009) Empowering plagiarism detection with a web services enabled collaborative network. J Inform Sci Eng 25:1391–1403
8. Selberg E, Etzioni O (1997) The metacrawler architecture for resource aggregation on the web. IEEE Expert J 11–14
9. Keyhanipour AH, Moshiri B, Kazemian M, Piroozmand M, Lucas C (2007) Aggregation of web search engines based on users'preferences in WebFusion. Knowl-Based Sys 20:321–328
10. Keyhanipour AH, Moshiri B, Piroozmand M, Lucas C (2006) WebFusion: fundamentals and principals of a novel meta search engine. IJCNN'2006, pp 4126–4131
11. Smyth B, Freyne J, Coyle M, Briggs P, Balfe E, Building T (2003) I-spy: anonymous, community-based personalization by collaborative web search. In: Proceedings of the 23rd SGAI international conference on innovative techniques and applications of artificial intelligence, pp 367–380
12. Chignell MH, Gwizdka J, Bodner RC (1999) Discriminating meta-search: a framework for evaluation. Inform Proces Manage 35:337–36
13. Dreilinger D, Howe AE (1997) Experiences with selecting search engines using metasearch. ACM Trans Inform Sys 15:195–222
14. Jansen BJ, Spink A, Koshman S (2007) Web searcher interaction with the dogpile.com metasearch engine. J Am Soc for Inform Sci Technol 58:744–755
15. Howe AE, Dreilinger D (1997) SavvySearch: a meta-search engine that learns which search engines to query. AI Mag 18
16. Gauch S, Wang G, Gomez M (1996) ProFusion: intelligent fusion from multiple, distributed search engines. J Univ Comput Sci 2
17. Selberg E, Etzioni O (1997) The MetaCrawler architecture for resource aggregation on the Web. IEEE Expert 12:11–1
18. Buzikashvili N (2002) Metasearch: properties of common document distributions. Lect Notes Comput Sci 2569:226–231

Chapter 9
Aspect Dependency Analyzer Framework for Aspect Oriented Requirements

K. Santhi, G. Zayaraz and V. Vijayalakshmi

Abstract Aspect Oriented Software Development (AOSD) is an emerging software development technology that seeks new modularizations of software systems in order to isolate broadly based functions from the main program's business logic. AOSD permits multiple concerns to be expressed separately and automatically unified into working systems. However, the complexity of interactions between aspects and base modules and among different aspects may reduce the value of aspect-oriented separation of cross-cutting concerns. This framework exploits the dependencies generated by the operators such as before, after, around and replace. It uses the specification of composition of aspects and if a conflicting situation emerges in a match point, it uses dominant candidate aspects to produce rules for composition which may be used to guide the process of composition. The proposed work generates a composition rule for each match point.

Key Words Aspect · AOSD · Cross-cutting concern · Framework · Match point

K. Santhi (✉) · G. Zayaraz · V. Vijayalakshmi
Department of CSE, Department of ECE, Pondicherry Engineering College,
Puducherry 605 014, India
e-mail: santhikrishnan@gmail.com

G. Zayaraz
e-mail: gzayaraz@pec.edu

V. Vijayalakshmi
e-mail: vvijizai@pec.edu

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_9,

9.1 Introduction

Traditional software development focuses on decomposing systems into units of primary functionality, while recognizing that there are other issues of concern like distribution, fault tolerance and synchronization that do not fit well into the primary decomposition. AOSD focuses on the identification, specification and representation of cross-cutting concerns and their modularization into separate functional units as well as their automated composition into a working system [1]. From the modularity, adaptability point of view, the separation of aspects in the base modules reduces the dependency between modules and improves modularity [2]. Some interactions may lead to the expected behavior while others are spring of unanticipated inconsistencies. Thus, it is desirable to detect interactions and potential inconsistencies, as early as possible in the life cycle, preferably at the requirement phase [2, 3]. Resolving conflicts at requirements phase is cheaper, faster, and desirable than carrying out necessary code modifications later on-the-fly [4, 5].

The rest of the paper is organized as follows: Sect. 9.2 introduces some background information related to this research. Section 9.3 briefly summarizes the framework and the main idea of the research is applied on a case study and Sect. 9.4 draws some conclusions and suggests directions for future work.

9.2 Background

The motivation for aspect-oriented programming approaches is to modularize crosscutting concerns. The implementation of a concern is scattered if its code is spread out over multiple modules. The concern affects the implementation of multiple modules. Its implementation is not modular [4, 6]. The implementation of a concern is tangled if its code is intermixed with code that implements other concerns. The module in which tangling occurs is not cohesive. Scattering and tangling often go together, even though they are different concepts.

9.3 Aspect Dependency Analyzer Framework

In this paper we propose a framework that allows the user to analyze interaction between aspects, identify aspects interactions, detect and resolve the conflicts between them based on the search of Hamiltonian path in which enhanced edge removal [6] and enhanced edge addition algorithm [6] are used to find the optimal path. The composition specification of aspect specifies its composition, i.e. where and how it will be attached at join points. The D-ACT approach for analyzing the interaction in one join point [2] is shown in Fig. 9.1. The analysis activity includes

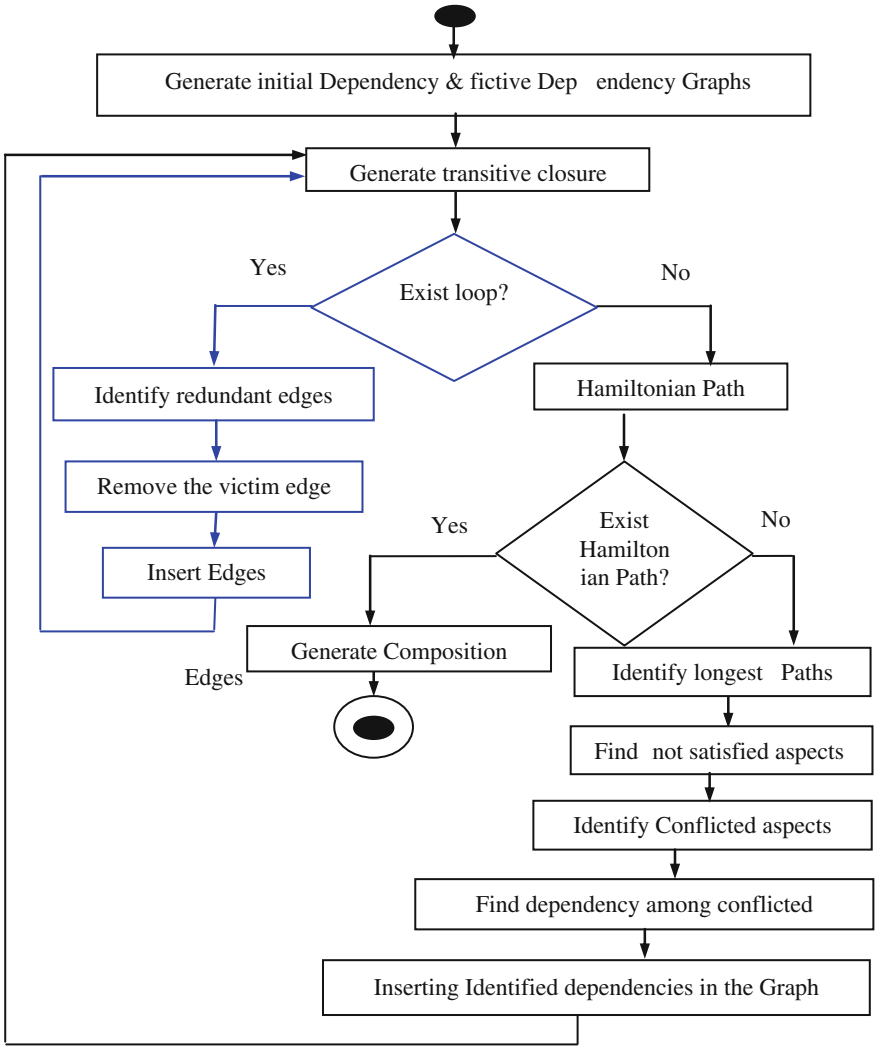


Fig. 9.1 D-ACT (diminishing aspects of cyclic conflicts using transitive closure)

the following tasks—Detecting interactions between aspects; detecting dependencies; Removing the conflicting dependencies and adding required dependencies; Reasoning and resolving conflicts; Generating composition rule.

Initially the Dependency graph is generated and the fictive dependencies (artificial dependency) if any are identified. Then transitive closure of the graph must be identified. If there is no loop but Hamiltonian path (A Hamiltonian path in an undirected graph is a path that visits each vertex exactly once) exist, the composition rule can be generated. If the Hamiltonian path does not exist, then the algorithm finds the longest path by analyzing the conflicts in the graph.

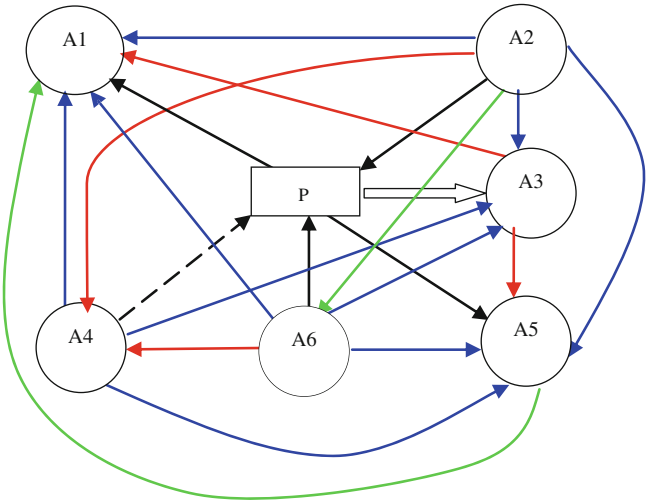


Fig. 9.2 Dependency graph with resolution dependencies (←)

Consider the example [2] where the candidate's aspects A1, A2, A3, A4 and A5 affected the match point (join point) P.A match point is where the crosscutting concern should join the behavior of the functional concern [7]. Suppose that: Aspect A1 overlaps before the match point, aspect A2 overlaps after the match point, aspect A3 wraps the match point, aspect A4 substitutes the match point, aspect A5 overlaps before the match point, aspect A6 overlaps after the match point.

For the Before operator the match point is never satisfied before the satisfaction of the aspects and the satisfaction of P depends on the satisfaction of aspects A1 and A5, P → A1 and P → A5.For the After operator the match point must be satisfied before satisfying the aspects and the behavior of aspect A2, A6 must be attached after P, A2 → P and A6 → P.For the Around operator the behavior of aspect A3 must be satisfied in parallel with the behavior of the join point P, P => A3.For the Replace operator the operator substitutes the behavior of P by the behavior of A4, A4 —> P. To locate the fictive dependencies around and replace operators are used. Operator Around: the behavior of aspect A3 must be satisfied in parallel with the behavior of the join point P, A3 → A1, A3 → A5 are identified. Operator Replace: Aspect A4 modifies the behavior of the join point P. Therefore, it is concluded that, there exists a concrete probability that all aspects depending on the join point P become dependent on the aspect A4. The fictive dependencies A6 → A4, A2 → A4 are identified.

The dependency is a transitive relationship for aspects Ai, Aj, Ak: Ai depends on Aj and Aj depend on Ak implies that Ai depends on Ak (Ai → Ak). They allow to generate the possible solutions on a certain degree of probability and to focus the analysis on a reduced set of dependencies (Fig. 9.2).

Table 9.1 Longest path

Longest paths	Analyze the longest paths
CH1-A2A4PA3A5	A6, A1 are not satisfied
CH2-A2A4PA3A1	A6, A5 are not satisfied
CH3-A6A4PA3A5	A2, A1 are not satisfied
CH4-A6A4PA3A1	A2, A5 are not satisfied
Synthesis of conflicts analysis	Conflicts between(A1, A5) Conflicts between(A6, A2)

If there is no loop as well as no Hamiltonian path in the transitive closure, then conflict exists. In this case, the longest paths are shown in Table 9.1.

A1 has a higher priority than A5 and A6 has a higher priority than A2, A5 → A1 and A2 → A6. The new dependency graph is generated which includes resolution dependencies. The transitive closure of dependency graph is also generated. At last, the Hamiltonian path: Ch = A2A6A4PA3A5A1 ← is found again. The composition rule is (according to the direction of the small arrow above the path) : A1 before P A5 before P A3 around P A4 replace P A6 after P A2 after P.

In the case of existence of any loop, the redundant edges are identified and the victim edge is removed using the enhanced edge removal algorithm. This algorithm removes the path between two nodes if they are bi-directional and also call the enhanced edge addition algorithm for all the parent nodes. The enhanced edge addition algorithm adds a path between two nodes if they are different and there is no path between those nodes. After performing these operations, again there is a check for loop and the above procedure continues to find the composition rule.

Algorithm 1: Enhanced Edge Removal

```
1. //a → b: = Edge to be removed
2. if ∃ path from a → b && b → a
3. if a > b then remove a → b
4. else b → a end if
5. end if
6. for all parents p of a do
7. add p → b with algorithm 2
8. end for
9. for all children c1 of a do
10. for all children c of b do
11. add c1 → c with algorithm 2
12. end for
13. end for
14. if ∃ path from a → b && b → a then
15. if a → b is a fictive dependency then remove b → a
16. else if b → a is a fictive dependency then remove a → b
17. else if a > b then remove a → b else remove b → a end if
```

18. **end if**

Algorithm 2 Enhanced Edge addition

1. //a → b: = Edge to be added
2. if (a = = b) then do nothing
3. **if** ¬∃ path from a → b **then**
4. add a → b
5. **end if**

9.3.1 Case Study

The problem: a Movie Theaters Chain asks for a system providing a massive use of functionality for public access [7]. It should offer mainly integrated on-line tickets sales facility. Tickets can be bought, reserved or payed on-line. Main facilities considered by the system are: on-line ticket bookings and ticket payment. Consider the aspects, Response Time (RT) around Ticket Payment (TP), Security. Availability (S.AV) before Ticket Payment (TP), Precision after Ticket Payment (TP), Security. Accuracy (S.AC) around Ticket Payment (TP), Ticket Booking (TB) before Ticket Payment. On processing these aspects as per the proposed approach, fictive dependencies and transitive closure were identified and after applying the algorithm the optimal dependency graph was generated and the composition rule [8] for the Movie Theatre: S.AV ≫ Ticket Booking ≫ (Ticket Payment ‖ S.AC ‖ RT) ≫ Precision was gererated.

This composition rule expresses the sequential order into which each candidate aspect must be composed. Availability, a sub concern of security must be satisfied first before the ticket booking aspect. Then ticket payment must be satisfied in parallel with accuracy, a sub concern of security, and response time. Finally precision must be satisfied (Fig. 9.3).

9.4 Conclusion and Future Works

This framework describes to support crosscutting concerns during requirements engineering. It extends the work developed in [2], providing new ideas for the identification, specification and composition of crosscutting concerns using Hamiltonian technique in which enhanced edge removal and enhanced edge addition algorithms were used, it minimized the time complexity needed to find the composition rules. The paper proposes the following tasks to be done in the future: define composition rules at a finer level of granularity, i.e. compose crosscutting actions; study the level of granularity at which a conflicting situation can be handled; how to use the Petri net [9] to determine the join points with impressive aspects (after, before, around and replace) to logical entities.

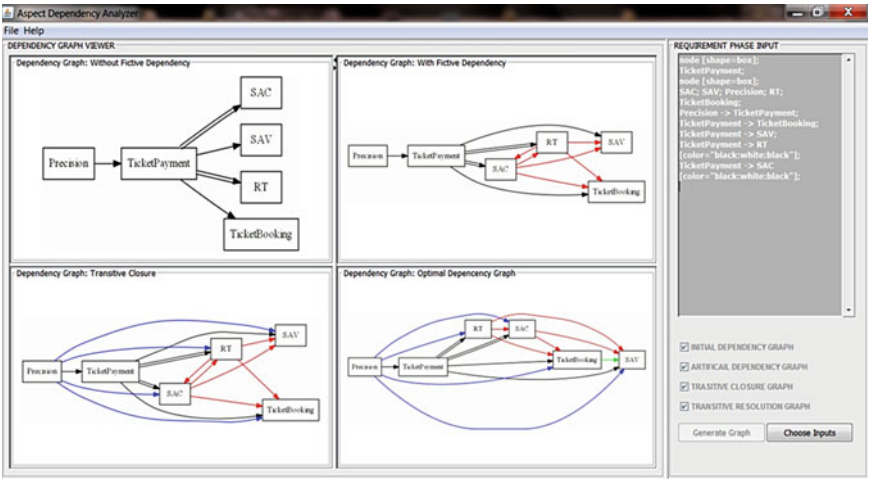


Fig. 9.3 Aspect dependency analyzer framework snapshot

Acknowledgments We thank Ms A. Shanthakumari for her help and support.

References

1. http://en.wikipedia.org/wiki/Aspect- oriented_ software _development.
2. Boubendir A, Chaoui A (2010) Towards a generic technique for analyzing interactions between aspects at requirement phase. Digital information management (ICDIM). Lake head University, Thunder Bay
3. Mehner K, Monga M, Taentzer G (2006) Interaction analysis in aspect-oriented models. In: 14th IEEE international requirements engineering conference (RE'06), pp 69–78
4. Fuentes L, Sanchez P (2007) Towards executable aspect-oriented UML models. In: Proceedings of the 10th international workshop on aspect-oriented modeling. ACM Press, New York, pp 28–34
5. Beniassad E, Clements PC, Araujo J, Moriera A, Rachid A, Tekmerdogan B (2006) Discovring early aspects. IEEE Softw e3(1):61–70
6. Case ML (2006) Online algorithms to maintain a transitive reduction. CS 294-8 class project, fall 2006. www.case-home.com/publications/caseCS2948.pdf
7. Cech S (2009) UML extensions for aspect oriented software development. J Object Technol 8(5):85–104
8. Brito I, Moreira A (2003) Advanced separation of concerns for requirements engineering. VIII jornadas de ingenieria de Software y bases de datos (JISBD), Alicande, Spain, 12–14 Nov 2003
9. Abdelzad V, Aliee FS (2010) A method based on petri nets for identification of aspects. Inform Sci Technol Bull ACM Slovakia 2(1)43–49
10. Santhi K, Zayaraz G, Vijayalakshmi V (2012) Diminution of interactions among aspects at requirement phase using hamiltonian path technique. In: National conference on information technology-2012, Feb 15th Pondicherry Engineering College, Puducherry

Chapter 10
Hybrid Extremal Optimization and Glowworm Swarm Optimization

Niusha Ghandehari, Elham Miranian and Mojtaba Maddahi

Abstract Glowworm Swarm Optimization algorithm is applied for the simultaneous capture of multiple optima of multimodal functions. In this paper, we have attempted to create a Hybrid Extremal Glowworm Swarm Optimization (HEGSO) algorithm. Aiming at the glowworm swarm optimization algorithm is easy to fall into local optima, having low accuracy, and to be unable to find the best local optima. However for solving these problems, the present algorithm has been increased the probability of choosing the best local optima. Moreover we want to use this method to have a best movement for agents in Glow worm optimization algorithm. Simulation and comparison based on several well-studied benchmarks demonstrate the effectiveness, efficiency and robustness of the proposed algorithms.

Keywords Extremal optimization · Glowworm swarm optimization · Hybrid algorithm

10.1 Introduction

Glowworm Swarm Optimization Algorithm (GSO) is a novel algorithm, which belongs to the group of algorithms based on swarm intelligence. The development of the optimization algorithm called Glowworm Optimization Algorithm has been

N. Ghandehari (✉) · E. Miranian · M. Maddahi
Department of Information Technology, Sufi Razi Institute of Higher Education, Zanjan, Iran
e-mail: Niusha682012@gmail.com

E. Miranian
e-mail: Elham.Miranian@gmail.com

M. Maddahi
e-mail: Mojtaba.maddahi@yahoo.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_10,

based on the behavior of glowworms (insects, which are able to modify their light emission and use the bioluminescence glow for different purposes).

GSO algorithm is especially useful for a simultaneous search of multiple optima, usually having different objective function values. Otherwise, only one (local or global) optimum will be found. In GSO agents exchange information locally. Moreover, their movements are non deterministic.

It should be pointed out, that GSO algorithm computes multiple optima in parallel during one program run, what provides a set of alternative solutions to the user. This is especially beneficial when search space represents parameters of e.g. a real production process, where some conditions are easier to set up or simply financially cheaper [1, 2].

Extremal optimization is an exploratory algorithm inspired from Bak-Sneppen's model. As pts name shows, this algorithm applies for resolving the optimization problems. In this algorithm, the cells of Bak-Sneppen's model are aligned with the solution of problem. The solution in each step is enhanced by replacing the worst cell amount with the other one. In this algorithm in contrast with Bak-Sneppen's model the adjoin cells is not changed by changes in one cell [3, 4]. However the extremal optimization algorithm provides some proper solutions for the given problems, then there is a high probability of remaining in the local optimum points and locus other than the solution. For solving this problem, Boettcher and Percus added the innovative parameter named τ to the extremal optimization algorithm and the new algorithm named extremal optimization with τ has been introduced. In this article, we use the parameter τ instead of the selection of agent. So we choose the best agent and increase the performance of algorithm [5] [6] [7].

The paper is organized as follows. Original GSO algorithm is presented in Sect. 10.2. Extremal optimization algorithm with parameter τ is present in Sect. 10.3. Modified EGSO algorithm is presented in Sect. 10.4. Test functions, performance measure and model constants and result of simulations are given in Sect. 10.5. Conclusions are enclosed in the Sect. 10.6.

10.2 Basic Glowworm Swarm Optimization Algorithm

In Basic Glowworm Swarm Optimization Algorithm (GSO), each glowworm distributes in the objective function definition space. These glowworms carry own luciferin respectively, and has the respective field of vision scope called local-decision range. Their brightness concerns with in the position of objective function value. The brighter the glow, the better is the position, namely has the good target value. The glow seeks for the neighbor set in the local-decision range, in the set, a brighter glow has a higher attraction to attract this glow toward this traverse, and the flight direction each time different will change along with the choice neighbor.

Each glowworm i encodes the object function value J (x_i (t)) at its current location x_i (t) into a luciferin value l_i and broadcasts the same within its

neighborhood. N_i (t) is the set of neighbors around glowworm i. Local-decision range update:

$$r_d^i(t+1) = \min\{r_s, \max\{o, r_d^i(t) + \beta(n_t - |N_i(t)|)\} \tag{10.1}$$

And r_d is the glowworm's local-decision range at the $t + 1$ iteration, r_s is the sensor range, n_t is the neighborhood threshold, the parameter β affects the rate of change of the neighborhood range. The number of glow in local-decision range:

$$N_i(t) = \{j : \|x_j(t) - x_i(t)\| < r_d^i; l_i(t) < l_j(t)\} \tag{10.2}$$

X_j (t) is the glowworm $_i$'s position at the t iteration l_j (t) is the glowworm $_i$'s luciferin at the t iteration, the set of neighbors of glowworm i consists of those glowworms that have a relatively higher luciferin value and that are located within a dynamic decision domain whose range r_d^i is bounded above by a circular sensor range r_s ($0 < r_d^i < r_s$). Each glowworm i selects a neighbor j with a probability p_{ij} (t) and moves toward it. Probability distribution used to select a neighbor:

$$p_{ij}(t) = \frac{l_j(t) - l_i(t)}{\sum\limits_{k \in N_i(t)} l_k(t) - l_i(t)} \tag{10.3}$$

Movement update:

$$x_i(t+1) = x_i(t) + s\left(\frac{x_j(t) - x_i(t)}{\|x_j(t) - x_i(t)\|}\right) \tag{10.4}$$

Luciferin-update

$$l_i(t) = (1-\rho)l_i(t-1) + \gamma J(x_i(t)) \tag{10.5}$$

$L_i(t)$ is a luciferin value of glowworm i at the t iteration, $\rho^1(0,1)$ leads to the reflection of the cumulative goodness of the path followed by the glowworms in their current luciferin values, the parameter only scales the function fitness values, $J(x_i\ (t))$ is the value of test function.

10.3 Extremal Optimization with Parameter τ

Although the extremal optimization algorithm provides some proper solutions for the given problems, then there is a high probability of remaining in the local optimum point and locus other than the solution. For solving this problem, proposed algorithm added the innovative parameter named τ to the extremal optimization algorithm and the new algorithm named extremal optimization with parameter τ has been introduced. The variation in the new algorithm and the

extremal optimization algorithm is related to the selection of replaceable cells. In this algorithm for finding the replaceable cells we first order, progressively the cells in terms of merits. Then, it is assigned to each cell the value p_k which is shown is Eq. (10.6). In this equation, k is ordered in terms of cell sequence:

$$P_k \propto k^{-\tau} \tag{10.6}$$

After assigning the p_k to each cell, the replaceable cell is chosen as follows:

1. In a loop, one of the cells is selected stochastically
2. A random number between 0 and 1 is selected
3. If the selected number is smaller than p_k, the cell will be selected as replaceable cell and the loop is closed, otherwise the loop is continued .

The general steps in extremal optimization algorithm with parameter τ as follows:

1. The original solution is regarded as the current and also the best solution
2. The cells is ordered progressively in terms of merits
3. Given the Eq. (10.1), one cell is selected as the replaceable cell and re-valued
4. The generated solution is regarded as the current solution
5. If the current solution is better than the best solution, it will be selected as the best one
6. The step 2 is continued until the required solution is obtained
7. The best solution is returned

It is important to note that the given algorithm and the way in which the replaceable cell is selected is the simple case of the limit optimization with parameter τ.

10.4 Hybrid Extremal Glowworm Swarm Optimization

As we mentioned, in the glowworm algorithm, the local factors is placed on the Luciferin and selection of the best factor is very critical for moving toward it. Then we use the extremal optimization with parameter τ (τ-EO) for choosing the best factor which leads to increasing the high performance. The basis of the algorithm is on the parameter τ. We use the element in the formula $1/k^{\tau}$ such that we place the factors will be selected in the row matrix and order them progressively in term of Luciferin level.

The random k ordered based on the indices of the factor is computed and placed in the formula $1/k^{\tau}$. We need the other number ρ which is a random number in [0, 1] If $1/k^{\tau} < \rho$, then the factor is selected, otherwise we continue until we find the factor. Note that if τ becomes large, then $1/k^{\tau}$ will be a small number and it very likely will be selected. For this reason, if k becomes large, then the selection possibility will be go up and it means that the possibility of selecting the factor

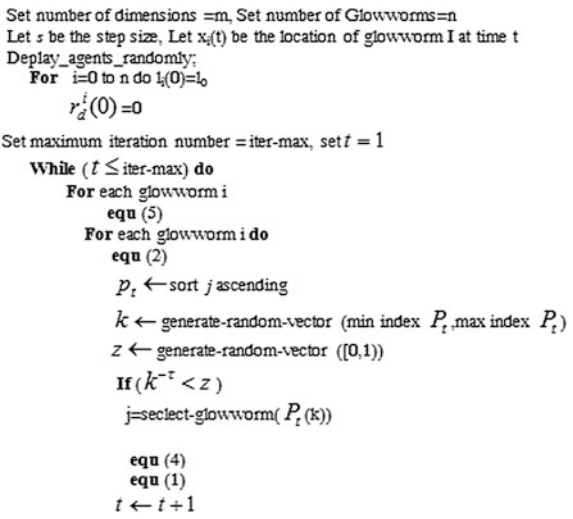


Fig. 10.1 Pseudo code of the proposed algorithm

with high Luciferin is increased. Note that the selection of τ has significant effects on the performance of algorithm (Fig. 10.1).

After an initial random distribution of possible solutions in the search space every solution tries to find a better state, i.e. a better solution, uses information carried by other solutions. The generic glowworm i is characterized by the position in the search space **x**, the light intensity (luciferine) l-$_i$ and by a neighborhood range r_i [1].

10.5 Experimental Results Comparison Between GSO and HEGSO

In this section, the efficiency of proposed algorithm is presented. For this purpose, Standard functions borrowed from Ref. [4] are used to show proficiency of suggested algorithm. In the below table, function with their characteristics are described (All experiments have been done on 10 dimension functions).

The set of GSO and HEGSO parameters are as below: $n = 100$, max of iteration max $t = 1000$, $\rho = 0.4$, $\gamma = 0.6$, $\beta = 0.08$, moving step $s = 0.03$, $n_t = 5$ and initialization of luciferin $l_0 = 5$.

Compare HEGSO and GSO with iteration t = 1000 that shows in Fig 5.1 (Fig. 10.2).

We want to calculate difference functions that mention in Table 10.1 in 1,000 iteration, to show effects of parameter τ to decrease the Lucifren level (Figs. 10.3, 10.4, 10.5, 10.6).

Fig. 10.2 Compare GSO with HEGSO on different function in $\tau = 2$

Table 10.1 Benchmark Functions and Their Characteristics

Function	Formulation	Range	Optimum
Sphere	$f_1(x) = \sum_{i=1}^{n} x_i^2$	$[-2, 2], i(1, \ldots, n, n = 10)$	0
Rosenberg	$f_2(x) = \sum_{i=1}^{n} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$	$[-2, 2], i(1, \ldots, n, n = 30)$	0
Schwefel	$f_3(x) = \sum_{i=1}^{i=n} (\sum_{j=1}^{i} x_j)^2$	$[-2, 2], i(1, \ldots, n, n = 40)$	0
Rastrigin	$f_4(x, y) = 20 + (x^2 - 10\cos(2\pi x) + y^2 - 10\cos(2\pi y))$	$[-2, 2], i(1, \ldots, n, n = 60)$	0

Fig. 10.3 The performance of the proposed algorithm

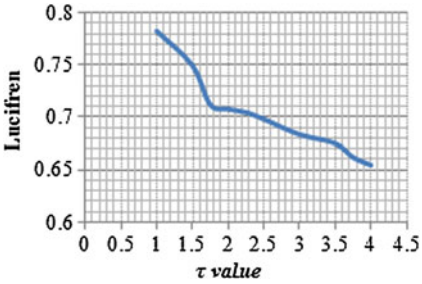
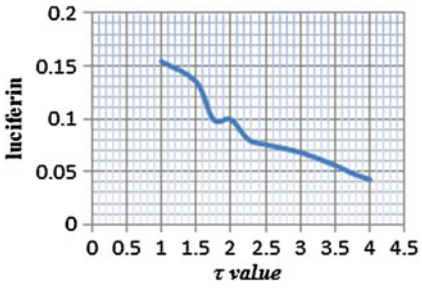


Fig. 10.4 The performance of the proposed on F_1 function algorithm F_2 function



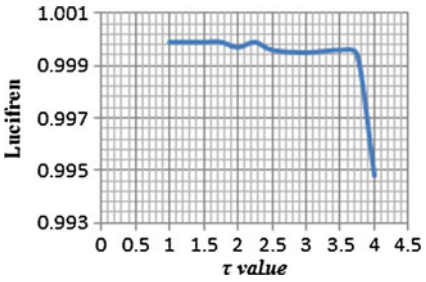


Fig. 10.5 The performance of the proposed algorithm

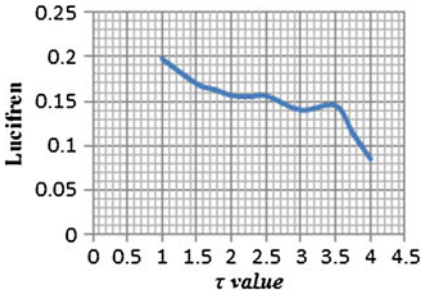


Fig. 10.6 The performance of the proposed on F_3 function algorithm on F_4 function

10.6 Conclusions

For Glowworm swarm optimization (GSO) the Probability to choose neighbors is not good any time and this is a disadvantage of GSO. In this paper, we present hybrid extremal glowworm swarm optimization algorithm (HEGSO). This algorithm chose better neighbors to move toward them and it increased the performance of algorithm The algorithm influenced by parameter τ. As it is shown in experiments, the presented algorithm has got excellence compared to previous ones.

References

1. Krishnanand KND, Ghose D (2009) Glowworm swarm optimization: a new method for optimizing multi-modal functions. Comput Intell Stud 1(1):93–119
2. Liu J, Zhou Y, Huang K, Ouyang Z, Wang Y (2011) A glowworm swarm optimization algorithm based on definite update search domains. J Comput Inform Sys 7:10, 3698–3705
3. Boettcher S, Percus AG Extremal optimization: an evolutionary local-search algorithm. http://arxiv.org/abs/cs.NE/0209030.
4. de Sousa FL, Vlassov V, Manuel Ramos F (2007) Centralized extremal optimization for solving complex optimal design problems. Lect Notes Comput Sci 2723:375–276
5. Liu H, Zhou Y, Yang Y, Gong Q, Huang Z (2010) A novel hybrid optimization algorithm based on glowworm swarm and fish school. J Comput Inform Sys 6(13):4533–4541
6. Yang Y, Zhou Y, Gong Q (2010) Hybrid artificial glowworm swarm optimization algorithm for solving system of nonlinear equations. J Comput Inform Sys 6(10):3431–3438
7. Krishnanand KND, Ghose D (2008) Theoretical foundations for rendezvous of glowworm-inspired agent swarms at multiple locations. Robot Auton Sys 56(7):549–569

Chapter 11
Implementation of New Technique for the Capacity Improvement in Wavelength Division Multiplexing Networks

N. Kaliammal and G. Gurusamy

Abstract In WDM network, the multicast routing and wavelength assignment (MC-RWA) problem is for maximizing the number of multicast groups admitted or for minimizing the call blocking probability. In this technique, the incoming traffic is sent from the multicast source to a set of intermediate junction nodes and then, from the junction nodes to the final destinations. Then, paths from source node to each of the destination nodes and the potential paths are divided into fragments by the junction nodes and these junction nodes have the wavelength conversion capability. By simulation results, it is proved that the proposed technique achieves higher throughput and bandwidth utilization with reduced delay.

Keywords Multicast routing and wavelength assignment (MC-RWA) problem · Total network capacity (TNC) · Least influence group (LIG) algorithm

N. Kaliammal (✉)
Department of ECE, N.P.R college of Engineering and Technology,
Dindigul, Tamil nadu, India
e-mail: kala_gowri@yahoo.co.in

G. Gurusamy
Prof/Dean/EEE, FIE Bannariamman Institute of Technology,
Sathyamangalam, Tamilnadu, India
e-mail: hodeee@bitsathy.ac.in

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_11,

11.1 Introduction

11.1.1 Wavelength-Division-Multiplexing (WDM) Networks

On-demand provisioning of wavelength routed channels within the transport layer has become more essential due to the recent emergence of high bit rate IP network applications. This is determined by wavelength division multiplexing (WDM) technologies. Due to the availability of ultra long-reach transport and all-optical switching, the deployment of all-optical networks has been made possible [1].

The WDM network provides the capability of transferring huge amount of data at high speeds by the users over large distance [2].

The light path establishment requires same wavelength and it should be used along the entire route of the light path without wavelength conversion. This is commonly considered to the wavelength continuity constraint [3].

11.1.2 Multicasting in WDM Networks

A network technology which is used for the delivery of information to a group of destinations is called as multicast. This simultaneously uses the most efficient strategy to deliver the message over each link of the network only once. Moreover, it creates the copies only when the links to the multiple destinations split [4].

In conventional data networks, in order to allow a multicast session, a multicast tree which is rooted at the source is constructed with branches spanning all the destinations [5].

Several network applications which require the support of QoS multicast such as multimedia conferencing systems, video on demand systems, real-time control systems, etc. [6].

11.1.3 Routing and Wavelength in WDM

In WDM network, for maximizing the number of multicast groups admitted or for minimizing the call blocking probability with certain number of wavelengths, the multicast routing and wavelength assignment (MC-RWA) problem is studied [7, 8]. The problem of the node architecture is that they were designed without having into account power efficiency, neither complexity of fabrication [9].

Bandwidth required for traffic paths depends on the ingress–egress capacities, and the traffic split ratios. The traffic split ratio is determined by the arrival rate of ingress traffic and the capacity of intermediate junction nodes [10].

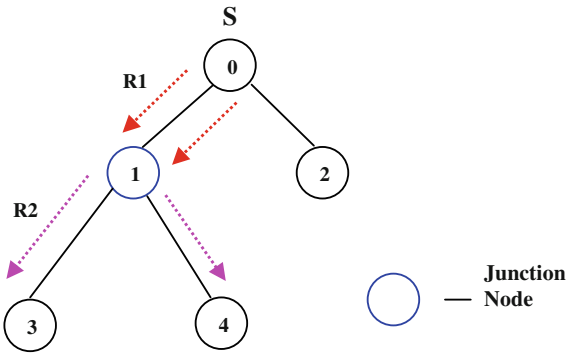


Fig. 11.1 Multicast routing process

11.2 Multicast Routing

The above diagram (Fig. 11.1) shows the routing process. The bandwidth requirement for the routing paths R1 and R2 is derived. Consider a node i with maximum incoming traffic I_i. Node i sends $\delta_j I_i$ amount of this traffic to node j during R1 routing for each jεN and thus the traffic demand is $\delta_j I_i$. Now, node i have received $\delta_i I_k$ traffic from any other node k. Out of this, the traffic destined for node j is $\delta_i r_{kj}$ since all traffic is initially split without regard to the final destination. The traffic that needs to be routed from node i to node j at R2 routing is given below:

$$\sum_{k \in N} \delta_i r_{kj} \leq \delta_i E_j.$$

Thus, the traffic demands from node i to node j at the end of R2 routing is $\lambda_i E_j$.

Hence, the maximum demand from node i to node j as a result of R1 and R2 routing is $\delta_j I_i + \delta_i E_j$.

Let $M = [m_{ij}] = [\delta_j I_i + \delta_i E_j]$ be the fixed matrix which can handle the traffic variation. It depends only on aggregate ingress–egress capacities and the traffic split ratios $\delta_1, \delta_2 \ldots. \delta_n$, Thus the routing scheme is unaware to the changes in traffic distribution.

11.3 Multicast Wavelength Assignment

11.3.1 Grouping the Paths

Assume the set $R_i = \{R_i^1, R_i^2 \ldots R_i^j \ldots\}$ to represent all fragments of the path from source to the ith destination in the multicast tree. R_i^j is the jth fragment of the path i. If AWR_i^j is the set of available wavelengths of the jth fragment of path i, then the number of wavelengths in AWR_i^j is regarded as the capacity of this fragment.

For the potential request paths, the set $Pi = \{p_i^1, p_i^2 \ldots\}$ is defined to indicate all fragments of the ith potential request.

The actual capacity of a path is basically determined by its fragment(s) with the least capacity. The fragment(s) with the least capacity of a path is named the critical fragment of that path. Let CP_i and CPP_i be the path capacity of the path i of the multicast tree, and the potential path i, respectively, then

$$CP_i = \min_{1 \le j \le r_i+1} SCPR_i^j \tag{11.1}$$

and

$$CP_{pi} = \min_{1 \le j \le r_i+1} SCPp_i^j \tag{11.2}$$

A path may have more than one critical fragment. Let $F_i = \{f_i^1, f_i^2 \ldots\}$ be the set of the critical fragments in the potential path i. Then F_i can be used to indicate whether the potential path is affected or not during the wavelength assignment of the multicast tree. Fragments which come from multicast tree with common links are coupled using grouping. A group is composed of fragments whose links are overlapped.

$$G = \{G1, G2 \ldots Gm \ldots GY\} \tag{11.3}$$

where G is the set of all groups in a multicast tree, G_m is the set of all fragments in the mth group.

Let AWG_m be available wavelength of all in the mth group. The group capacity, CG_m, is defined as the number of wavelengths in AWG_m.

11.3.2 Total Network Capacity Estimation

The influence of network capacity is examined by checking whether the links of potential paths overlap with those of the multicast groups. If the overlap occurs at the critical fragments of the potential path and the assigned wavelength is the one of the available wavelengths in that critical fragment, the path capacity of the potential path will be affected. Let $C_m(pi, \lambda)$ be the capacity of pi being influenced when the wavelength λ is assigned in the mth group.

$$C_m(p_i, \lambda) = \begin{cases} 1 & if(\lambda \in x_w) \wedge x \in LS_{m,Fi} \\ 0 & Otherwise \end{cases} \tag{11.4}$$

The network capacity affected when λ is assigned for the mth group, $TNC_{m,\lambda}$, can be obtained by the summation of the influence of all potential paths as

$$TNC_{m,\lambda} = \sum_{p_i \in P} C_m(p_i, \lambda) \tag{11.5}$$

The Least Influence Group (LIG) algorithm selects the wavelengths for groups to maximize the network capacity. The idea behind LIG algorithm is that the

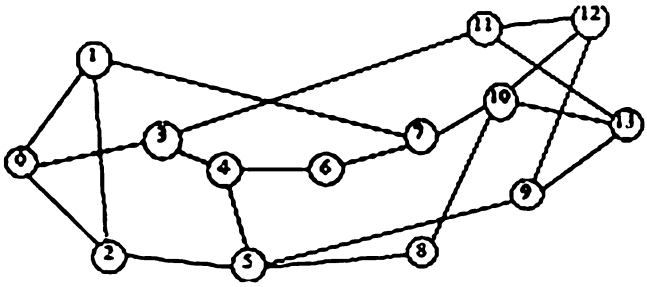


Fig. 11.2 NSF network of 14 nodes

wavelength having the least effect on the potential paths is chosen for that group. The LIG algorithm is illustrated below:

$AW_{Gm} = \{\lambda_1, \lambda_2, \lambda_3 \ldots.\}$

1. Find all p_b whose links overlap in the links of group m
2. For each $\lambda \in AW_{Gm}$

$$TNC_{m,\lambda} = \sum_{p_b \in P'} C_m(p_b, \lambda)$$

3. Assign λ in which $TNC_{m,\lambda}$ is minimum in group m

11.4 Simulation Model and Parameters

In this section, the performance of multicast routing and wavelength assignment technique is simulated using ns-2 network simulator [11]. The optical WDM network simulator (OWNs) patch in ns2 is used to simulate a NSF network (Fig. 11.2) of 14 nodes.

In this simulation, the results of MRWA are compared with "Resource Efficient Multicast Routing (REMR) protocol [10]".

11.5 Performance Metrics

In this simulation the blocking probability, end-to-end delay and throughput is measured.

- **Bandwidth utilization**: It is the ratio of bandwidth received into total available bandwidth for a traffic flow.

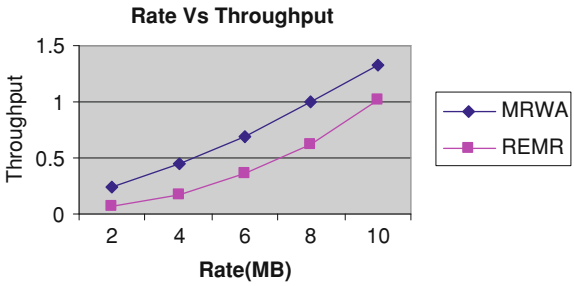


Fig. 11.3 Rate versus throughput

- **Average end-to-end delay**: The end-to-end-delay is averaged over all surviving data packets from the sources to the destinations.
- **Throughput**: It is the number of packets received successfully.

11.5.1 Results

11.5.1.1 Effects of Varying Traffic Rate

In the initial simulation, the traffic rate is varied as 2, 4, 6, 8 and 10 Mb and measures the throughput, end-to-end delay and bandwidth utilization.

Figures (11.3, 11.4, 11.5) shows that the throughput is more in the case of MRWA when compared to REMR, the delay of MRWA is significantly less than the REMR and MRWA shows better utilization than the REMR scheme.

11.5.1.2 Effect of Varying Traffic

In this simulation, the number of traffic sources is varied as 1, 2, 3, 4 and 5 and measure the throughput, end-to-end delay and bandwidth utilization.

Figures (11.6, 11.7, 11.8) shows that the throughput occurred when varying the number of the throughput is more in the case of MRWA when compared to REMR, the delay of MRWA is significantly less than the REMR and MRWA shows better utilization than the REMR scheme.

11.6 Conclusion

By simulation results, it is proved that the proposed technique achieves higher throughput (31 % increase) and bandwidth utilization (20 % increase) with

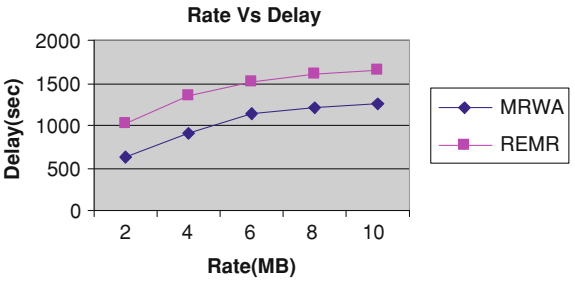


Fig. 11.4 Rate versus delay

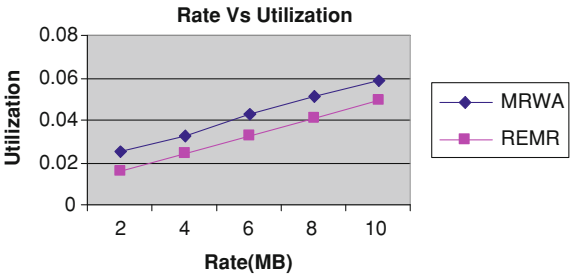


Fig. 11.5 Rate versus utilization

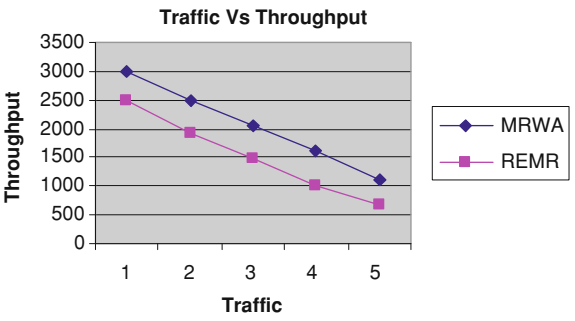


Fig. 11.6 Traffic versus through

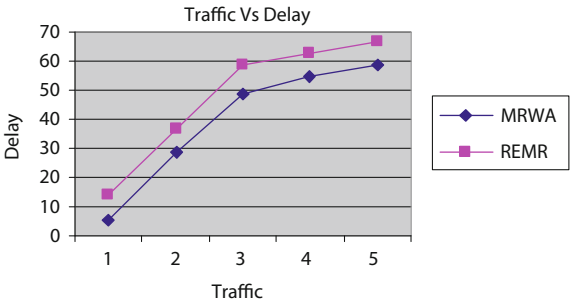


Fig. 11.7 Traffic versus delay

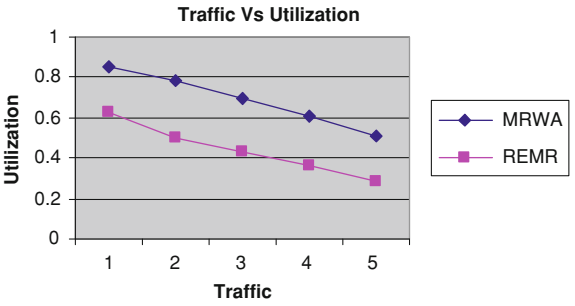


Fig. 11.8 Traffic versus utilization

reduced delay (29 % decrease) for varying rate & 24 % decrease in delay, 35 % increase in utilization and 24 % increase in throughput for varying traffic sources.

References

1. Rajkumar A, Murthy Sharma NS (2008) A distributed priority based routing algorithm for dynamic traffic in survivable WDM networks. IJCSNS Int J Comput Sci Netw Secur 8(11)
2. Canhui O, Zang H, Singhal NK, Zhu K, Sahasrabuddhe LH, Macdonald RA, Mukherjee B (2004) Sub path protection for scalability and fast recovery in optical WDM mesh networks. IEEE J Sel Areas Commun 22(9)
3. Le VT, Ngo SH, Jiang XH, Horiguchi S, Inoguchi (2005) A hybrid algorithm for dynamic lightpath protection in survivable WDM optical networks. IEEE
4. Multicasting: http://en.wikipedia.org/wiki/Multicasting
5. Fen Z, Miklos M, Bernard C (2008) Distance priority based multicast routing in WDM networks considering sparse light splitting. IEEE 11th Singapore International Conference on Communication Systems
6. Jia X-H, Du D-Z, Hu X-D, Lee M-K, Gu J (2001) Optimization of wavelength assignment for QoS multicast in WDM networks. IEEE Trans Commun 49(2)
7. He J, Gary Chan SH, Tsang DHK (2001) Routing and wavelength assignment for WDM multicast networks. In the proceedings of the IEEE GLOBECOM 2001
8. Wang A, Wu Q, Zhou X, Wang J (2009) A new multicast wavelength assignment algorithm in wavelength-converted optical networks. Int J Commun, Netw Syst Sci
9. Fernandez GM, Larrabeiti D (2009) Contributions for all-optical multicast routing in WDM networks. 16th International Congress of Electrical Engineering, Electronics and Systems, IEEE INTERCON
10. Kaliammal N, Gurusamy G (2010) Resource efficient multicast routing protocol for dynamic traffic in optical WDM networks. Eur J Sci Res
11. Network Simulator: www.isi.edu/nsnam/ns

Chapter 12
A Comparative Review of Contention-Aware Scheduling Algorithms to Avoid Contention in Multicore Systems

Genti Daci and Megi Tartari

Abstract Contention for shared resources on multicore processors is an emerging issue of great concern, as it affects directly performance of multicore CPU systems. In this regard, Contention-Aware scheduling algorithms provide a convenient and promising solution, aiming to reduce contention. By providing a collection of the scheduling methods proposed by latest research, this paper focuses on reviewing the challenges on solving the contention problem for UMA(Uniform Memory Access latency, single memory controller) and NUMA(Non Uniform Memory Access latencies, multiple memory controllers) types of system architectures. In this paper, we also provide a comparative evaluation of the solutions applicable to UMA systems which are the most extensively studied today, discussing their features, strengths and weaknesses. This paper aims to propose further improvements to these algorithms aiming to solve more efficiently the contention problem, considering that performance-asymmetric architectures may provide a cost-effective solution.

Keywords Uniform memory access (UMA) · Multicore CPU systems · Contention-aware scheduling · Non uniform memory access (NUMA) · Vector balancing scheduling · OBS-X scheduler · DIO scheduler · DINO scheduler · AMPS scheduler

G. Daci (✉) · M. Tartari
Faculty of Information Technology, Polytechnic University of Tirana,
Tirana, Albania
e-mail: gdaci@ieee.org

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_12,

12.1 Introduction

Contention for shared resources in multicore processors is a well-known problem. The importance of handling this problem is related with the fact that multicore processors are becoming so prevalent in desktops and also servers, that may be considered a standard for modern computer systems and also with the fact that this problem causes performance degradation. Let's consider a typical multicore system described schematically in Fig. 12.1, where cores share parts of memory hierarchy, that we call "memory domains", and compete for resources like last level cache (LLC), memory controllers, memory bus and pre-fetching hardware.

Preliminary studies considered cache contention [8] as the most crucial factor responsible for performance degradation. Driven by this assumption, they focused on finding mechanisms to reduce cache contention like Utility Page Partitioning [10] and Page Coloring [17].

Contention-Aware scheduling [2, 6, 9, 19] is proposed as a promising solution to this problem, because it reduces contention, by applying different thread migration policies. The major part of these studies, found solutions that could be applied only in UMA (Uniform Memory Access) systems, that are not suitable for NUMA (Non Uniform Memory Access).

So for UMA systems we will discuss DIO Scheduler, that uses thread classification schemes like SDC [13], LLC miss rate [19], Pain metric [19], Animal Classes [15] to take the best scheduling decision; For NUMA architecture, that still requires further research, is proposed DINO Scheduler. We will also discuss AMPS scheduler design for asymmetric-architecture multicore systems, that supports NUMA.

12.2 Contention-Aware Scheduling a Promising Solution1.1, Checking the PDF File

There are previous studies on improving thread performance in multicore systems were mainly focused on the problem of contention for the shared cache. Cache partitioning has a significant influence on performance closely relating with execution time. J. Lin, Q. Lu, X. Ding, Z. Zhang, X. Zhang, and P. Sadayappan [8], implemented an efficient layer for cache partitioning and sharing in the operating system through virtual-physical address mapping. Their experiments showed a considerable increase of performance up to 47 %, in the major part of selected workloads. A number of cache partitioning methods have been proposed with performance objectives [17] [10]. A. Fedorova, M. I. Seltzer, M. D. Smith [12] designed a cache-aware scheduler that compensates threads that were hurt by cache contention by giving them extra CPU time.

The difficulty faced from S. Zhuravlev, S. Blagodurov and A. Fedorova [19] in evaluating the contribution that each factor has on performance was that all the

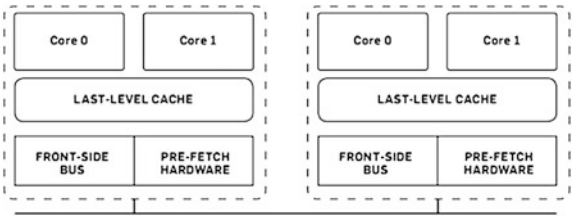


Fig. 12.1 A schematic view of a multicore system with 2 memory domains

degradation factors work in conjunction with each other in complicated and practically inseparable ways.

So the classification scheme serves to identify applications that must be co-scheduled or not. S. Zhuravlev, S. Blagodurov and A. Fedorova [19] help us with their contribution in analyzing the effectiveness of different classification schemes like:

- SDC (Stack Distance Competition), a well known method [13] for predicting the effects of cache contention among threads, based on the data provided from stack distance profiles, that inform us on the rate of memory reuse of the applications.
- Animal Classes is based on the animalistic classification of application introduced by Y. Xie and G. Loh [15]. It allows classifying applications in terms of their influence on each other when co-scheduled in the same shared cache.
- Miss Rate is considered as the heuristic for contention, because it gives information for all the shared resources contention.
- Pain Metric is based on *cache sensitivity* and *cache intensity*, where sensitivity is a measure of how much an application will suffer when cache space is taken away from it due to contention; intensity is a measure of how much an application will hurt others by taking away their space in a shared cache.

As a perfect scheduling policy, it is used an algorithm proposed by Y. Jiang, X. Shen, J. Chen, R. Tripathi [4].This algorithm is guaranteed to find an optimal scheduling assignment, i.e., the mapping of threads to cores, on a machine with several clusters of cores sharing a cache as long as the co-run degradations for applications are known. The optimal scheduling assignment can be found by solving a min-weight perfect matching-problem.

12.3 Proposed Schedulers for UMA Systems

Several studies investigated ways of reducing resource contention and as mentioned above in Sect. 12.2, one of the promising approaches that emerged recently is contention-aware scheduling [6, 9, 19].

We mention here the co-scheduling tasks proposed for SMP [19, 22] and for SMT [16]. These studies of contention-aware algorithms focused primarily on

UMA (Uniform Access Memory). In this section we will review and evaluate OBS-X, Vector Balancing scheduling policy, and DIO scheduler by discussing their features, merits, but also their gaps.

12.3.1 OBS-X Scheduling Policy Based on OS Dynamic Observations

According to R. Knauerhase, P. Brett, B. Hohlt, and S. Hahn [6], in a multicore environment the Operating System (OS) can and should make observations of the behavior of threads running in the system. Good policies can improve the overall performance of the system or performance of the application.

They ran two sets of experiments across four cores in two LLC groups. The first set of experiments consisted of four instances of cache-buster, an application that consumes as much cache as possible and four instances of spin-loop, that consumes CPU with a minimum of memory access. With the addition of OBS-X, cache-buster performance increased between 12 and 62 %, comparing with the default Linux default load balancing. The reason for the increase is that OBS-X distributed the cache-heavy tasks across LLC groups, thus minimizing the scheduling of heavy tasks together. To approximate real-world workloads, they ran OBS-X with a set of applications from the SPEC CPU 2000 suite run. The overall speedup increases to 4.6 %.

12.3.2 Vector Balancing Scheduling Policy

This policy reduces contention by migrating tasks, led by the information of *task activity vector* [9], that represents the utilization of chip resources caused by tasks. Based on the information provided form these vectors, it has been proposed from A. Merkel, J. Stoess and F. Bellosa [9] the scheduling policy that avoids contention for resources by co-scheduling tasks with different characteristics. They compute the cost for reading and it was proved that it required 54 CPU cycles.

Since these vectors update for every timer interrupt and task switch, this increase in the execution time is considered negligible. This policy can be easily integrated in the OS balancing policy, so we can exploit the existing strategies. This assumption is a weakness because it limits the space where the Vector Balancing can be applied successfully.

12.3.3 DIO (Distributed Intensity Online) Scheduler

S. Zhuravlev, S. Blagodurov and A. Fedorova [19] proposed DIO contention-aware scheduler. DIO scheduler continuously monitors the miss rates of applications, as we argued in Sector 2 that it was the best contention predictor, then finds the best performance case and separates threads. This makes DIO more attractive since the stack distance profiles, which require extra work to obtain online, are not required. DIO was experimented in AMD Opteron with 8 cores, 4 cores for each domain. DIO improved performance by up to 13 % relative to default. Another use of DIO is to ensure QoS (Quality of Service) for critical applications, since it ensures to never select the worst performance case of the scheduler.

12.4 Adaptation of Contention-Aware Schedulers for NUMA Systems

Research studies on contention-aware algorithms, were primarily focused on UMA (Uniform Memory Access) systems, where there are multiple shared last level caches (LLC). Modern multicore systems are using massively the NUMA (Non Uniform Memory Access) architecture, because of its decentralized and scalable nature. In these systems there is one memory node for each memory domain. Local nodes can be accessed for a shorter time than the distant ones, and each node has its own controller. According to S. Blagodurov, S. Zhuravlev, M. Dashti and A. Fedorova [2], when existing contention-aware schedulers designed for UMA architectures, were applied on a NUMA system (illustrated on Fig. 3 [2]), they did not effectively manage contention, but they also degraded performance compared with the default contention-unaware scheduler (30 % performance degradation).

12.4.1 DINO Contention-Management Algorithm for NUMA Systems

As argued above, previous contention-aware algorithms were valid only on UMA architectures, but when applied to NUMA architectures, used in today's modern multicore processors hurt their performance. To address this problem, a contention-aware algorithm on a NUMA system must migrate the memory of the thread to the same domain where it migrates the thread itself. In the study of S. Blagodurov, S. Zhuravlev, M. Dashti and A. Fedorova [2], they have designed and implemented Distributed Intensity NUMA Online (DINO).

DINO scheduler uses the same heuristic model for contention as the DIO (Distributed Intensity Online) scheduler discussed in Sect. 12.3.3, that uses the *LLC miss rate* criteria for predicting contention.

DINO can be extended to predict when co-scheduling threads on the same domain is more beneficial than separating them, using techniques described in [14] or [5]. DINO migrates the thread together with its memory. B. Goglin, N. Furmento [3] developed an effective implementation of the *move_pages* system call in Linux, which allows the dynamic migration of the large memory areas to be significantly faster than in previous versions of the OS. DINO organizes threads in broad classes according to their miss rates as shown in the research study of Y. Xie and G. Loh [15]. The classes in which threads get divided are:

- Turtles: less than 2 LLC miss rates for 1000 instructions
- Devils: 2-100 LLC misses for 1000 instructions
- Super_Devils: more than 100 LLC misses for 1000 instructions.

So the migrations will be performed only when threads change their classes, while they preserve their thread-core affinity relation as much as possible.

DINO in this situation should at least avoid memory migration back and forth, preventing so performance degradation. DINO achieves this by separating threads in classes as explained above. Results of DINO implementation showed that DINO achieved up to 30 % performance improvements for jobs in the MPI workload.

12.4.2 AMPS the Scheduling Algorithm for Performance-Asymmetric Multicore, System NUMA and SMP

Since industry is going towards multicore technology, and traditional operating systems are based on homogenous hardware, and performance-asymmetric architectures (or heterogeneous) [11]. As a first step towards this, T. Li, D. Baumberger, D. A. Koufaty and S. Hahn [7] designed the operating system scheduler AMPS, that manages efficiently both SMP and NUMA-style performance-asymmetric-architectures. AMPS contains three components:

- Asymmetry-aware-load-balancing, that balances threads to cores in proportion with their computing power
- Faster-core-first scheduling, that controls thread migrations based on predictions of their overhead.

Our evaluation demonstrated that AMPS improved stock Linux for asymmetric systems in the aspect of performance and fairness. AMPS uses thread-independent policies, which schedule threads independently regardless of application types and dependencies. This is considered a weakness that should be eliminated in the future. Thread-dependent policies mostly exists in research. H. Zheng, J. Nieh [20] dynamically detect process dependencies to guide scheduling.

12.5 Conclusions and Discussions

Based on the wide dissemination of multicore processors, we chose to handle the topic of contention for shared resources in such systems, as it affects directly their performance. Through this paper we reviewed and discussed the best solutions for this problem.

One of the major difficulty encountered during the design of such schedulers, was selecting the most effective thread classification scheme. The best predictor is *Miss Rate,* as it supplies us with information regarding contention for all resources.

To mitigate contention for shared resources, we discussed and reviewed the best scheduling algorithms and policies, that do not perform equally when applied to different multicore architectures. So for UMA systems, we reviewed OBS-X scheduling policy, that uses the operating system dynamic observations on tasks behavior to migrate threads.

For NUMA systems we have discussed the available solutions: DINO and AMPS scheduler. The most appropriate solution for NUMA systems is the DINO contention-aware scheduler, as it solves the performance degradation problem associated with the previous contention-aware solutions by migrating the thread along with its memory and also eliminates superfluous migrations. AMPS is the first scheduler proposed for the performance-asymmetric architectures, that supports both NUMA and SMP-style performance-asymmetric architectures, but it does not completely address contention, requiring further research in the future.

References

1. Blagodurov S, Zhuravlev S, Fedorova A (2010) Contention-aware scheduling on multicore systems. ACM Trans Comput Syst 28
2. BlagodurovS, Zhuravlev S, Dashti M, Fedorova A (2011) A case for NUMA-aware contention management on multicore systems. In: The 2011 USENIX annual technical conference, pp 1–9
3. Goglin B, Furmento N (2009) Enabling high-performance memory migration for multithreaded applications on Linux. In: Proceedings of IPDPS
4. Jiang Y, Shen X, Chen J, Tripathi R (2008) Analysis and approximation of optimal co-scheduling on chip multiprocessors. In: Proceedings of the 17th international conference on parallel architectures and compilation techniques (PACT '08), pp 220–229
5. Kamali A (2010) Sharing aware scheduling on multicore systems. Master's thesis, Simon Fraser University
6. Knauerhase R, Brett P, Hohlt B, Hahn S (2008) Using OS observations to improve performance in multicore systems. IEEE Micro 28(3):54–58
7. Li T, Baumberger D, Koufaty DA, Hahn S (2007) Efficient operating system scheduling for performance-asymmetric multi-core architectures. In: Proceedings of supercomputing, pp 1–4, 8–10
8. Lin J, Lu Q, Ding X, Zhang Z, Zhang X, Sadayappan P (2008) Gaining insights into multicore cache partitioning: bridging the gap between simulation and real systems. In: Proceedings of international symposium on high performance computer architecture, pp 1–5

9. Merkel A, Stoess J, Bellosa F (2010) Resource-conscious scheduling for energy efficiency on multicore processors. In: Proceedings of EuroSys, pp6–8, 11–13
10. Qureshi MK, Patt YN (2006) Utility-based cache partitioning: a low overhead, high-performance, runtime mechanism to partition shared caches. In: Proceedings of the 39th annual IEEE/ACM international symposium on microarchitecture, MICRO 39, pp 1–3
11. Shelepov D, Saez Alcaide JC, Jefferym S, Fedorova A, Perez N, Huang ZF, Blagodurov S, Kumar V (2009) A scheduler for heterogeneous multicore systems. SIGOPS Oper Rev 43(2)
12. Fedorova A, Seltzer MI, Smith, MD (2007) Improving performance isolation on chip multiprocessors via an operating system scheduler. In: Proceedings of the sixteenth international conference on parallel architectures and compilation techniques (PACT'07), pp 25–38
13. Chandra D, Guo F, Kim S, Solihin, Y (2005) Predicting inter-thread cache contention on a chip multi-processor architecture. In Proceedings of the 11th international symposium on high performance computer architecture, HPCA'05
14. Tam D, Azimi R, Stumm M (2007) Thread clustering: sharing-aware scheduling on SMP–CMP–SMT multiprocessors. In: Proceedings of EuroSys 2007
15. Xie Y, Loh G (2008) Dynamic classification of program memory behaviors in CMPs. In: Proceeding of CMP-MSI, pp 2–4
16. McGregor RL, Antonopoulos CD, Nikolopoulos DS Scheduling algorithms for effective thread pairing on irbid mutiprocessors. In: Proceedings of the 19th ieee international parallel and distributed processing symposium (IPDPS'05)
17. Zhang X, Dwarkadas S, Shen K (2009) Towards practical page coloring-based multicore cache management. In: Proceedings of the 4th ACM European conference on computer systems 2009
18. Zhang EZ, Jiang Y, Shen X (2010) Does cache sharing on modern CMP matter to the performance of contemporary multithreaded programs? In: Proceedings of PPOPP
19. Zhuravlev S, Blagodurov S, Fedorova A (2010) Addressing contention on multicore processors via scheduling. In: Proceedings of ASPLOS, pp 1–6

Chapter 13
Developing Embedded Systems from Formal Specifications Written in Temporal Logic

Shigeki Hagihara, Takahiro Arai, Masaya Shimakawa and Naoki Yonezaki

Abstract We propose a semi-automatic method for developing embedded systems using program code extraction from formal specifications written in temporal logic. This method consists of the following four steps. (1) Write a formal specification for a system. (2) Refine the specification to adapt to the structure and function of the hardware. (3) Obtain a transition system representing a program from the refined specification. (4) Assign program codes to atomic propositions used in the specification, and convert the transition system to the program. As a case study to demonstrate that the proposed method is practical, we generate a program which controls a robot as a line tracer.

Keywords Embedded system synthesis · Formal specification · Temporal logic

S. Hagihara (✉) · T. Arai · M. Shimakawa · N. Yonezaki
Department of Computer Science, Graduate School of Information Science and Engineering, Tokyo Institute of Technology, 2-12-1-W8-67 Ookayama, Meguro-ku, Tokyo 152-8552, Japan
e-mail: hagihara@fmx.cs.titech.ac.jp

T. Arai
e-mail: t_arai@psg.cs.titech.ac.jp

M. Shimakawa
e-mail: masaya@fmx.cs.titech.ac.jp

N. Yonezaki
e-mail: yonezaki@cs.titech.ac.jp

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_13,

13.1 Introduction

Recently, highly complex systems have been used to control hardware in safety critical systems such as nuclear plants and aviation systems etc. These kinds of systems must respond with appropriate timing to any requests of any timing. If the system fails to respond in unexpected situations, fatal accidents may occur. In order to develop safety-critical systems, a model checking technique [1, 2] is widely used. Although this technique can find errors in systems effectively, modeling systems requires a great deal of time. On the other hand, there is another approach to developing safety critical systems, which is reported in this paper. In this approach, we describe formal specifications in the temporal logic, and convert the specifications into transition systems representing programs [3, 4]. In this approach, we can obtain the transition systems automatically, and these systems are guaranteed to satisfy the specifications. If we can convert the transition system into programs correctly, the program is guaranteed not to contain errors. This approach has a disadvantage that conversion to a transition system is a highly complex process. However, recently there have been several studies on efficient conversion, we can obtain transition systems from simple specifications.

Based on this approach, we propose a semi-automatic method for developing embedded systems using program code extraction from formal specifications written in temporal logic. This method consists of the following four steps. (1) Write a formal specification for a system. (2) Refine the specification to adapt to the structure and function of the hardware. (3) Obtain a transition system representing a program from the refined specification. (4) Assign program codes to atomic propositions used in the specification, and convert the transition system to the program. As a case study to demonstrate that the proposed method is practical, we generate a program which controls a robot as a line tracer.

This paper is organized as follows. First, in Sect. 13.2, we present a method for developing embedded systems by program code generation from formal specifications, and apply our method to development of a robot program. Next, in Sect. 13.3, we discuss future works. Finally, we conclude our results in Sect. 13.4.

13.2 Developing a Program for a Line Tracer

We show our method of developing embedded systems by showing program generation from formal specifications written in linear temporal logic (LTL). Controlled by the generated program, the 3pi robot [5] moves and traces a line.

Fig. 13.1 The 3pi robot

13.2.1 The 3pi Robot

The 3pi robot (shown in Fig. 13.1) is controlled through a programmable chip microcontroller called Atmel AVR. It is easy to describe the control program by using libraries prepared by Pololu Robotics & Electronics. The 3pi robot has two wheels, one each on the left and right sides of the body. To control the speed of the left and right wheels, we can use the following functions respectively.

```
void set_m1_speed (int lspeed)
void set_m2_speed (int rspeed)
```

The robot is able to turn by setting different values for the speed of the left and right wheels. The robot has five sensors in front of the body. To recognize its position against a line, we can use the function `unsigned int read_line()`. When the 3pi robot is located on the right side of the line, `read_line()` returns a value between 0 and 1000. When the robot is located on the line, it returns a value between 1000 and 3000. When the robot is located on the left side of the line, it returns a value between 3000 and 4000.

13.2.2 Developing a Program for a Line Tracer

In this section, we show our method (Step 1–Step 4) for developing an embedded system and constructing a program which automatically controls the 3pi robot.

Step 1: Describe a Formal Specification We describe a formal specification for the system. We show the specification for line tracer in Fig. 13.2. We will rewrite this specification formally. We must consider what are input events and what are output events. In this case, environment is hardware. We cannot control what information is obtained from sensors, but can observe this information. Hence, the events by which the robot obtains information about its location are input events. We prepare the input event propositions *middle*, *right*, and *left* corresponding to

1. If the robot receives the information that it is located on the line, it will go straight.
2. If the robot receives the information that it is located on the right side of the line, it will turn left.
3. If the robot receives the information that it is located on the left side of the line, it will turn right.
4. Otherwise, it will stop.

Fig. 13.2 Specification of the line tracer

Fig. 13.3 The specification written in LTL

1. $G(middle \rightarrow X\, go\_straight)$
2. $G(right \rightarrow X\, go\_left)$
3. $G(left \rightarrow X\, go\_right)$
4. $G(\neg middle \wedge \neg left \wedge \neg right \rightarrow X\, stop)$

these input events. On the other hand, since the robot can control which direction it will move in, the events that cause it to go straight, those that cause it to turn left (right), and the ones that cause it to stop are output events. We prepare the output event propositions *go_straight*, *go_left*, *go_right*, and *stop* corresponding to these output events respectively. By using these event propositions, we can describe the specification in LTL as shown in Fig. 13.3, where G is a temporal operator representing 'always', and X is also a temporal operator representing 'at next'.

Step 2: Refine the Specification In Step 2, we refine the specification obtained in Step 1, to adapt it to the structure and function of the robot. As mentioned in Sect. 13.2.1, the robot has two wheels which are controllable individually, and we can control the direction of the robot by setting the speed of the wheels. From this fact, we introduce the new output event propositions l and r, which represent the rotation of the left and right wheels respectively. The output events propositions *go_straight*, *go_left*, *go_right,* and *stop* can be paraphrased into $r \wedge l$, $r \wedge \neg l$, $\neg r \wedge l$, and $\neg r \wedge \neg l$, respectively. Furthermore, the robot cannot obtain two different information inputs about its location simultaneously. Hence, we add an assumption about it. As a result of these refinements, the refined specification is shown in Fig. 13.4. The procedure for refining the specification written in LTL was summarized in our previous work [6].

Step 3: Convert the Specification into a Transition System In Step 3, we convert the refined specification into a transition system. The specification is required to satisfy realizability [3, 4, 7]. The realizability requires that there exists a transition system such that for any input events of any timing, the system produces output events such that the specification holds. A realizable specification can be converted into a transition system which represents a program. If a specification is not realizable, the specification must be modified. The procedure for modifying the specification was summarized in our previous work [8]. Here, we use a realizability checker Lily [9, 10], which checks whether a specification is

Assumption G(¬(*middle*∧*left*) ∧¬(*middle*∧*right*) ∧¬(*left*∧*right*))
1. $G(middle \rightarrow X(r \wedge l))$
2. $G(right \rightarrow X(r \wedge \neg l))$
3. $G(left \rightarrow X(\neg r \wedge l))$
4. $G(\neg middle \wedge \neg left \wedge \neg right \rightarrow X(\neg r \wedge \neg l))$

Fig. 13.4 The refined specification

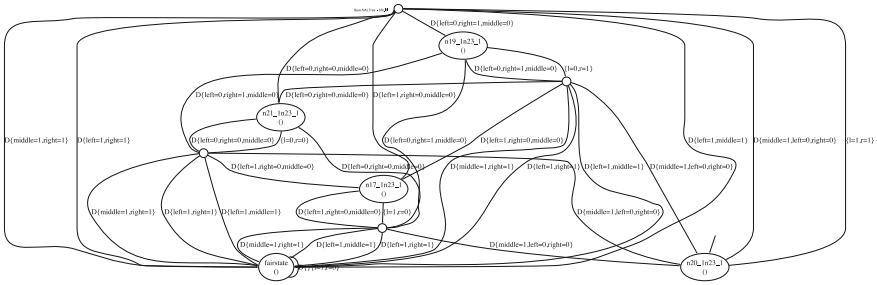


Fig. 13.5 The transition system representing program behavior

realizable or not, and generates a transition system from a realizable specification. The specification in Fig. 13.4 is realizable, and is converted into the transition system in Fig. 13.5 by Lily. The initial state of this transition system is shown at the lower right. This transition system is bipartite. Input events and output events are assigned to transitions. Small circles without characters are states of waiting for input events to occur. Circles with characters are states where output events occur. From a state of waiting for input events to occur, there are transitions corresponding to all the patterns of occurrence of input events, and the next states are determined by the occurrence of input events. On the other hand, from a state where output events occur, there is only one transition. Outputs events corresponding to the transition occur.

Step 4: Converting the Transition System into a Program In Step 4, we convert the transition system obtained in Step 3 into a program code. This program code follows transitions of the transition system and observes the input events and then executes the output events. This program has a variable (called the state variable) indicating the current state, and proceeds according to this variable. Output events correspond to program operations. Since the output event propositions *l* and *r* represent the rotation of left and right wheels respectively, we assign *l,* ¬*l, r,* and ¬*r* to the program operations `set_m1_speed(50)`, `set_m1_speed(0)`, `set_m2_speed(50)`, and `set_m2_speed(0)`, respectively. If a value of the state variable corresponds to a state where output events occur, the corresponding program operations are executed, and the state variable is updated to the next state.

On the other hand, input events correspond to conditional formulae. After the function `read_line` returns a value, the value is stored in the variable

`position`. Since the input event propositions *middle*, *right*, and *left* indicate that the robot gets information about its location from its sensor, we assign *middle*, *right*, and *left* to the following conditional formulae using position.

right `0 < position && position < 1000`
middle `1000 <= position && position < 3000`
left `3000 <= position && position < 4000`

If the value of the state variable corresponds to a state of waiting for the occurrence of input events, the corresponding conditional formulae are evaluated, and the state variable is updated to the next state according to the evaluation.

Through Step 1 to Step 4, we succeed in automatically generating C program code which controls the 3pi robot as a line tracer. The code is 122 lines in length. We confirmed that the program code moved the robot as a line tracer. As a result, we confirmed that our method for developing an embedded system is practical for an embedded system of this scale.

13.3 Future Work to Develop Practical Systems

Many practical embedded systems are required to satisfy real-time constraints, e.g. "The machine must respond in 0.5 s." Real-time constraints can be described in real-time logic such as MTL [11], MITL [12] and ECL [13]. Hence, by checking the realizability of specifications with real-time constraints, we can adapt our method to development of a real-time system. Although for many famous real-time logics such as MITL and ECL, the realizability problem is undecidable [14], the one for several subsets of the logics such as LTL $\triangleleft$ and MTL-B are decidable [14, 15]. We are required to implement a realizability checker and investigate whether our method is effective for developing real-time systems.

13.4 Conclusion

Our results clarified how to write formal specifications in LTL, how to refine the specifications, and how to generate program codes from the specifications. This is an important step towards developing safety critical embedded systems without faults using a formal approach. Future work to develop a practical system are also clarified. After these tasks are completed, it will be possible to develop large-scale practical embedded systems.

Acknowledgments This work was supported by a Grant-in-Aid for Scientific Research(C) (24500032).

References

1. McMillan KL (1993) Symbolic model checking. Kluwer Academic Publishers, Norwell
2. Tomita T, Hagihara S, Yonezaki N (2011) A probabilistic temporal logic with frequency operators and its model checking. In: Proceedings of the 13th international workshop on verification of infinite-state systems. EPTCS, vol 73. pp 79–93
3. Abadi M, Lamport L, Wolper P (1989) Realizable and unrealizable specifications of reactive systems. In: Proceedings of 16th international colloquium on automata, languages, and programming. LNCS, vol 372. Springer. pp 1–17
4. Pnueli A, Rosner R (1989) On the synthesis of a reactive module. In: Proceedings of the 16th ACM SIGPLAN-SIGACT symposium on principles of programming languages. pp. 179–190
5. Pololu 3pi robot user's guide, http://www.pololu.com/docs/pdf/0J21/3pi.pdf
6. Vanitha V, Yamashita K, Fukuzawa K, Yonezaki N (2000) A method for structuralisation of evolutional specifications of reactive systems. In: ICSE 2000. The third international workshop on intelligent software engineering. pp 30–38
7. Hagihara S, Yonezaki N (2006) Completeness of verification methods for approaching to realizable reactive specifications. In: Proceedings of 1st Asian working conference on verified software AWCVS'06. UNU-IIST, vol 348. pp 242—257
8. Hagihara S, Kitamura Y, Shimakawa M, Yonezaki N (2009) Extracting environmental constraints to make reactive system specifications realizable. In: Proceedings of the 2009 16th Asia-pacific software engineering conference. APSEC '09, IEEE Computer Society. pp 61—68
9. Lily: a LInear Logic sYnthesizer, http://www.iaik.tugraz.at/content/research/design_verification/lily/
10. Jobstmann B, Bloem R (2006) Optimizations for LTL synthesis. In: Formal methods in computer aided design, 2006 (FMCAD '06). pp 117–124
11. Koymans R (1990) Specifying real-time properties with metric temporal logic. Real-Time Syst 2(4):255–299
12. Alur R, Feder T, Henzinger TA (1996) The benefits of relaxing punctuality. J ACM 43(1):116–146
13. Raskin JF, Schobbens PY (1998) The logic of event clocks: decidability, complexity and expressiveness. Automatica 34(3):247–282
14. Doyen L, Geeraerts G, Raskin JF, Reichert J (2009) Realizability of real-time logics. In: Proceedings of the 7th international conference on formal modeling and analysis of timed systems. FORMATS '09, Springer-Verlag. pp 133–148
15. Maler O, Nickovic D, Pnueli A (2007) On synthesizing controllers from bounded-response properties. In: Proceedings of the 19th international conference on Computer aided verification (CAV'07). Springer-Verlag, pp 95–107

Chapter 14
Network Log Clustering Using K-Means Algorithm

Preeti Sharma and Thaksen J. Parvat

Abstract Network attacks are a serious issue in today's network environment. The different network security alert system analyse network log files to detect these attacks. Clustering is useful for wide variety of real time applications dealing with large amount of data. Clustering divides the raw data into clusters. These clusters contain data points which have similarity between themselves and dissimilarity with other cluster data points. If these clusters are given to these security alert systems, they will take less time in analysis as the data will be grouped according to the criteria the security system needs. This can be done by using k means clustering algorithm. In this first number of clusters are selected and then centroids are initialized. Then data points are assigned to the cluster with nearest centroid and mean of the centroid is calculated. This step is repeated till no data points are left. The objective is to cluster the network data log so as to make it easier for different security alert systems to analyse the data and detect network attacks.

Keywords Data clustering · K means algorithm · Centroid · Data points · Network attacks · Network traffic log

P. Sharma (✉) · T. J. Parvat
Department of Computer Engineering, Sinhgad Institute of Technology,
Lonavala 410401, India
e-mail: preeti5588@gmail.com

T. J. Parvat
e-mail: pthaksen.sit@sinhgad.edu

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_14,

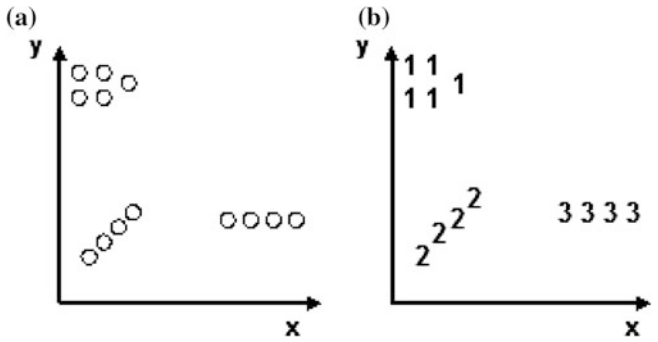


Fig. 14.1 Data clustering **a** raw data, **b** clustered data

14.1 Introduction

Clustering techniques are useful for a variety of real time applications dealing with large amount of data. Nowadays there are vast number of intrusions and other network attacks happening. For preventing and detecting these attacks many network security alert systems are used like firewall, intrusion detection system etc. These systems use raw data, that is, network log files for analysis. If these systems are provided with clusters they will take less time to detect attacks. After clusters are formed, they can be compared with existing labelled clusters to detect deviation from normal behaviour.

14.1.1 Data Clustering

Data clustering is grouping of data into its subsets by using unsupervised learning. These clusters then can be used for analysis, that is, for hypothesis formation or decision making. The data points within a cluster are more similar to each other than to data points belonging to other clusters. An example of data clustering is shown in Fig 14.1. The input data is shown in Fig. 14.1a, showing the data points are labeled similarly, and there are no clusters. In Fig. 14.1b, clusters are formed and the data points belonging to same cluster are similarly labeled.

Clustering has manly three stages: feature selection/extraction, data representation and finding similarity between the data points and then grouping them. This is depicted in Fig. 14.2.

Feature selection is selecting the most effective subset of the original features to be used in clustering. Feature extraction is to use one or more transformation of the input features to produce new salient features. Either or both of these techniques can be used to obtain an appropriate set of features to use in clustering.

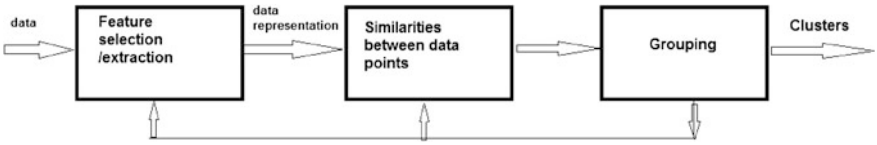


Fig. 14.2 Stages in data clustering

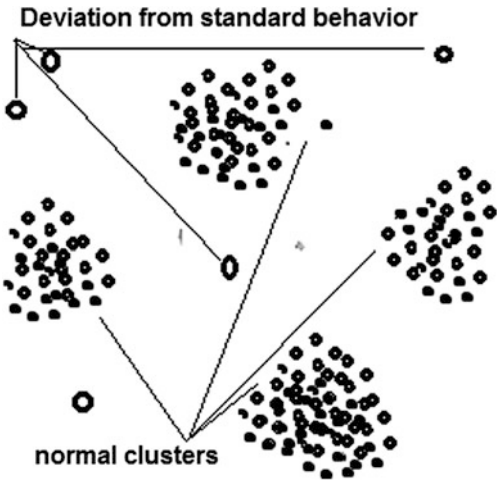


Fig. 14.3 Clusters formed with normal behavior and abnormal behavior

Data representation refers to the number of classes, the number of available patterns, and the number, type and scale of the features available to the clustering algorithm.

14.1.2 Network Traffic Log and Network Security

Nowadays almost all the systems are connected to networks. They access network for different purposes, for sending and receiving data from the network. All the activities and communication that the user does on network are recorded in logs. These logs are further used by different security systems like intrusion detection system, security alert systems, firewalls etc. to detect different network attacks. Some of the common network attacks are flooding, smurfing, denial of service, spoofing, Remote to local, unauthorized access to root etc.

Once data is clustered, as shown in Fig. 14.3 all of the instances that appear in small clusters are labeled as abnormal behavior because, the normal instances should form large clusters compared to intrusions and Malicious intrusions and normal instances are qualitatively different, so they do not fall into same cluster.

14.2 Related Work

In [1] clustering, unsupervised anomaly detection has been used for Intrusion detection system (IDS) aims to identify attacks with a high detection rate and a low false alarm rate. And to do so unsupervised k-means algorithm with labeling techniques was used.

In [2] unsupervised robust clustering technique is used to detect anomalous traffic flows, sequentially captured in a temporal sliding window basis. [3] Proposed anomaly detection approach which classifies data clusters from baseline and real traffic using the K-means combined with particle swarm optimization. In [4], paper focuses on detecting intrusions or anomalous behaviors in WLANs with data clustering techniques. In this, the wireless traces were converted into data records that can be used for clustering and then, an efficient online K-means clustering algorithm was used to group the data into clusters. Intrusive clusters were then determined by distance-based heuristics.

In [5] Using the reduction methods, the simplified network security assessment data set is established and the extraction by the decision-making rules is proposed and verified. Through computing dependence, non-discernibility of network security risk and management system parameters and relation of recognizing security system, the method based on *RS* is used to judge the network security, which is evaluated by using eigen value and more related system information are mined. In [6] the efficiency and effectiveness of current important data clustering techniques, partitioning and hierarchical, are evaluated for service analytics. In [7] this paper, an efficient iterative optimization algorithm is proposed for clustering distributed databases. In [8] a hybrid clustering method based on artificial fishes swarm algorithm and K-means so called KAFSA is proposed. [9] Defines a threshold distance for each cluster's centroid to compare the distance between data point and cluster's centroid with this threshold distance through which the computational effort can be minimized during calculation of distance between data point and cluster's centroid. It is shown that how the modified k-mean algorithm will decrease the complexity & the effort of numerical calculation, maintaining the easiness of implementing the k-mean algorithm.

Most of the experiments were done on standard datasets; in this paper we are trying to implement it for real time dataset and overcome the key challenges faced during implementing the same.

K-means clustering algorithm based on evidence accumulation has been used in intrusion detection [10], which has increased the detection rate and could reduce the false positive rate of attack detection. [11] Gives the following points which tell us why we should go for K-means algorithm:

- A partitioning method classifies the data into **k** groups, which together satisfy the requirements of a partition:
  - each group must contain at least one object
  - each object must belong to exactly one group.

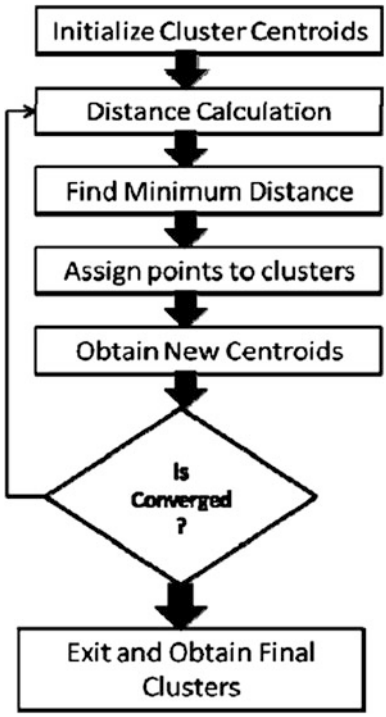


Fig. 14.4 Computational steps of k-means algorithm

- In large databases, assigning class labels to a large number of objects can be very costly.
- Solution is independent of order in which centroids are chosen, order in which points are processed.

14.3 K-Means Algorithm

K-means clustering is one of the unsupervised computational methods used to group similar objects into smaller partitions called clusters so that similar objects are grouped together. The algorithm aims to minimize the within cluster variance and maximize the intra cluster's variance which is shown in Fig. 14.4. The technique involves determining the number of clusters at first and randomly assigning cluster centroid to each cluster from the whole datasets; this step is called initialization of cluster centroid.

The distance between each point in the whole datasets and every cluster centroid is then calculated using a distance metric (e.g. Euclidean distance). Then, for every data point, the minimum distance is determined and that point is assigned to the closest cluster. This step is called cluster assignment, and is repeated until all

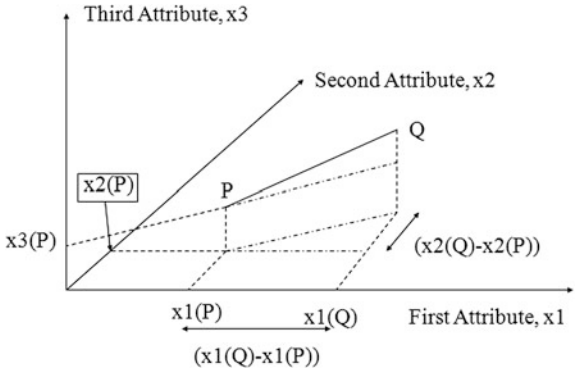


Fig. 14.5 Graphical data (3D)

of the data points have been assigned to one of the clusters. Finally, the mean for each cluster is calculated based on the accumulated values of points in each cluster and the number of points in that cluster. Those means are then assigned as new cluster centroid, and the process of finding distances between each point and the new centroid is repeated, where points are re-assigned to the new closest clusters. The process iterates for a fixed number of times, or until points in each cluster stop moving across to different clusters. This is called Convergence [11, 12].

Data matrix where x two p-dimensional data points and distance matrix which shows the distance between the data points can be represented respectively in matrix form as:

$$\begin{bmatrix} x_{11} & \dots & x_{1f} & \dots & x_{1p} \\ \dots & \dots & \dots & \dots & \dots \\ x_{i1} & \dots & x_{if} & \dots & x_{ip} \\ \dots & \dots & \dots & \dots & \dots \\ x_{n1} & \dots & x_{nf} & \dots & x_{np} \end{bmatrix} \qquad \begin{bmatrix} 0 & & & & \\ d(2,1) & 0 & & & \\ d(3,1) & d(3,2) & 0 & & \\ \vdots & \vdots & \vdots & & \\ d(n,1) & d(n,2) & \dots & \dots & 0 \end{bmatrix}$$

Figure 14.5 shows the data points depicted in the matrix with the distance between them [11]. Generally, the distance between two points is taken as a common metric to assess the similarity among the instances of a population. The commonly used distance measure is the **Euclidean metric** as shown in (14.1), which defines the distance between two points P = (x1(P), x2(P)…) and Q = (x1(Q), x2(Q)…) is given by [11] :

$$\begin{aligned} d(P,Q) &= \sqrt{\left((x1(P) - x1(Q))^2 + (x2(P) - x2(Q))^2 + \cdots \right)} \\ &= \sqrt{\left(\sum_{j=1}^{p} (xj(P) - xj(Q))^2 \right)} \end{aligned} \tag{14.1}$$

However, the Euclidean distance consumes a lot of computational resources when implemented in hardware due to the multiplication operation used for

obtaining the square operation; this reduces the amount of parallelism that can be exploited due to the need for calculating distances over and over. This is especially true when using a large number of clusters. Thus, an alternative distance metric called the Manhattan distance as shown below in (14.2) is used [11]:

$$d(i,j) = |x_{i1} - x_{j1}| + |x_{i2} - x_{j2}| + \cdots + |x_{ip} - x_{jp}| \tag{14.2}$$

where $i = (x_{i1}, x_{i2}\ldots x_{ip})$ and $j = (x_{j1}, x_{j2}\ldots x_{jp})$ are two p-dimensional data objects.

It performs faster than the Euclidean metric [11], because it does not require calculating the square offering better exploitation of parallelism and speed twice than that obtained by Euclidean distance. However, the accuracy of this distance measure was found to be slightly inferior to the Euclidean metric, but results were still within an acceptable error. The time needed to complete the clustering method for a whole datasets depends on the size of the set and the selected number of clusters: the larger they are, the longer it will take to compute the distances. Distance computation is the most computationally demanding part, and where most of the K-means processing time occurs [12]. K means algorithm is relatively efficient: O (tk (n)), where n is # instances, c is # clusters, and t is # iterations. Normally, k, t $\ll$ n. K means algorithm often terminates at a local optimum [11].

14.4 Experiments and Results

For experiments IDE Netbeans 6.9.1 is used with Jxl jar file, which is for reading and writing Microsoft office excel sheets. The data set used here is network traffic log report generated by Fortigate firewall of STES Lonavala campus network 45Mbps 1:1 lease line. It consists of 48 fields, out of which only 40 are selected for clustering like date, time, destination and, source port, destination and, source IP, received and, sent packet size, protocols, service type etc. every 24 h around four million tuples are generated.

As shown in Fig. 14.6, here in our proposed System we are using the Data captured by the network devices at our campus, STES, Lonavala, of existing System of Computer Network. This data is stored in Ms Excel file, i.e., the network log file, which we are using as input to our system. In this Technique we used following steps to achieve Clustering of network data

1. Single Parameter: In the beginning we read the Excel data and retrieve it in a java Data Structure object. Then this data is sent for securitization of single parameters like Source IP, Source Port, Destination IP, and Destination Port etc. Once this parameter matches with the raw data then we fetch complete row of the value and then adding into another Data object which yields us a Clustered value for single given parameter using K-means algorithm. The Fig. 14.7 shows a cluster formed when given single parameter source port as

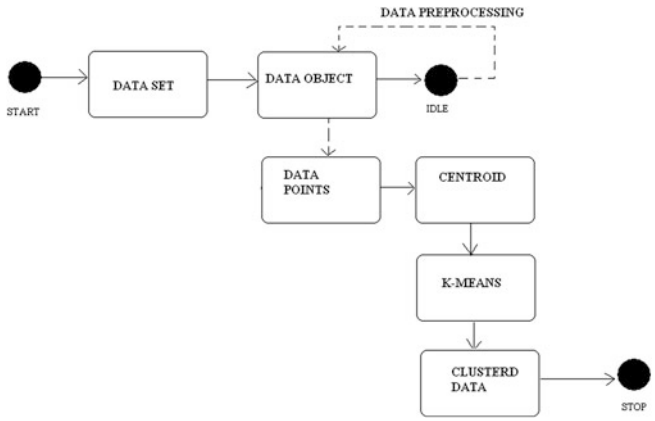


Fig. 14.6 State transition diagram

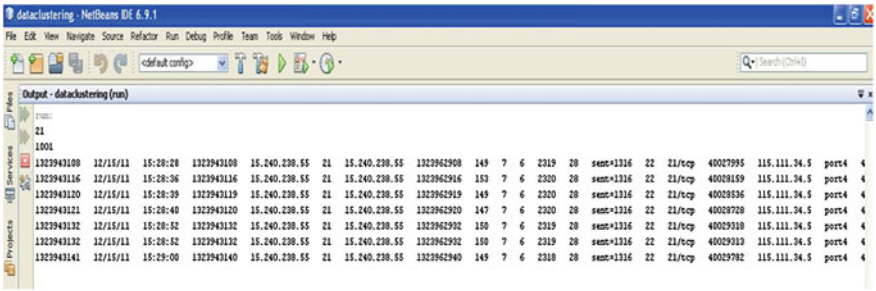


Fig. 14.7 Single parameter clustering

21. Like single source port, any single parameter can be given to the system to obtain clusters satisfying the condition of existence of that parameter.

2. Multi Parameter: In this Step we are using very strong and well know K- Means Algorithm for clustering when multi parameter are giving as input to the system. In the Initial K means, it detects the centroid of given parameter and then decides distance value for the given parameter. Then by checking for the converged values it is iterated till the formation of the desired cluster and end of data. This is depicted in Fig. 14.8. Here in multi value parameter we are clustering the data based on two values of the captured network data. For example: if we want to cluster the data based on Destination IP with Destination Port, so that we can get the cluster of which IP is hitting on which port, based on this we can come to know what kind of attack is happening. Like if the port is 21, we know that it is for FTP and if the port is 80 then, we know it is for HTTP.

So these multi-value parameters are helping us to track the exactness of the attacking clusters. Here in our Project K-Means provide two leases possible

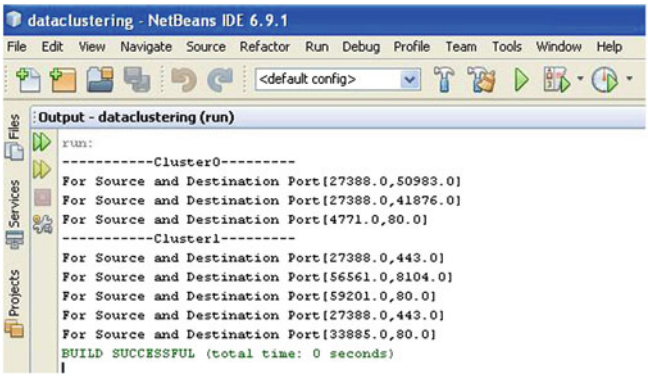


Fig. 14.8 Multiple parameter clustering

Fig. 14.9 Performance evaluation

clusters based on these value for more exactness. In Fig. 14.8, source port (27388.0) and destination port (50983.0) are given as parameters for forming clusters. We obtain two clusters, first is perfect cluster with these value and second is buffer cluster containing the nearest values to these given parameter values.

As depicted in Fig. 14.9, as the number of data records increases 10000 times still the time taken to form cluster changes very slitely. On the x-axis number of data records is taken and on y-axis execution time in miliseconds. For 10000 data records it took 511031 ms, likewise for 20000, 30000, 40000 it took 544000, 576281, 607765 ms respectively and so on. This shows that this experiment is time efficient.

14.5 Conclusion

Network security is needed everywhere to protect personal information and data. Traditional network security devices such as Intrusion Detection Systems (IDS), firewalls, and security scanners operate independently of one another, with

virtually no knowledge of the network assets they are defending. This lack of information results in numerous ambiguities when interpreting alerts and making decisions on adequate responses. So to enrich the Network security device intelligence we always need a strong Data base that should contain the threat producing data clusters. This is just like the data base of Antivirus which always contains virus names. Here in our proposed system we cluster the information of the data packets based on many parameters like source IP, Destination IP, Port, time etc.

The focus of this study was to cluster network traffic log successfully using k means algorithm. These clusters are formed depending on single parameter and multiple parameters taken from the data set. These clusters then can be used by different network security alert systems for analysis and detecting network attacks. So in short our model can act as a plug-in that can enrich the intelligence of the network devices. However work is going on to enhance the method and make it more efficient.

References

1. Das D, Sharma U, Bhattachacharyya DK (2008) An intrusion detection mechanism based on feature based data clustering. In: ICET 2008 4th international conference on emerging technology
2. Casas P, Mazel J, Owezarski P (2011) Steps towards autonomous network security: unsupervised detection of network attacks. In: 4th IFIP international conference on new technologies, mobility and security (NTMS)
3. Lima MF, Zarpelao BB, Sampaio LDH, Rodrigues JJPC, Abrao T, Proenca ML (2010) Anomaly detection using baseline and K-means clustering. In: International conference on software telecommunications and computer networks (SoftCOM)
4. Zhong S, Khoshgoftaar TM, Nath SV (2005) A clustering approach to wireless network intrusion detection. In: ICTAI 05, 17th IEEE international conference on tools with artificial intelligence
5. Qu Z, Wang X (2009) Study of rough set and clustering algorithm in network security management. In: NSWCTC '09, international conference on networks security, wireless communications and trusted computing
6. Zhao YW, Chi C-H, Ding C (2011) Analysis of data clustering support for service. In: IEEE 2nd international conference on software engineering and service science (ICSESS)
7. Elgohary A, Ismail MA (2011) Efficient data clustering over peer to peer network. In: 11th international conference on intelligent systems design and application (ISDA)
8. Yazdani D, Golyari S, Meybodi MR (2010) A new hybrid approach for data clustering. In: 5th international symposium on telecommunications (IST)
9. Singh RV, Bhatia MPS (2011) Data clustering with modified K-means algorithm. In: International conference on recent trends in information technology (ICRTIT)
10. Weng F, Jiang Q, Liang S, Wu N (2007) An intrusion detection system based on clustering ensemble. In: IEEE international workshop on 16–18 April 2007
11. Xu R, Wunsch D II (2005) Survey of clustering algorithms. IEEE Trans Neural Netw 16(3):648–666
12. Han J, Kamber M (2006) Data mining concepts and techniques, 2nd edn. Morgan Kaufmann Publishers, San Francisco
13. Jain AK, Murty MN, Flynn PJ (1999) Data clustering: a review. ACM Comput Surv 31(3):278–282

Chapter 15
Improvement Public Key Kerberos Using Identity-Based Signcryption

Hussein Khalid Abd-Alrazzaq

Abstract Several proposals have been developed that add public key cryptography to various stages of Kerberos to make the protocol work with large user communities and Public Key Infrastructures (PKI). But a man-in-the-middle attack on PKINIT allows an attacker to impersonate Kerberos administrative principals and end-servers to a client, hence breaching the authentication guarantees of Kerberos. It also gives the attacker the keys which an Authentication Server (AS) normally generates to encrypt the service requests of this client, hence defeating confidentiality as well. In this paper we provide alternative approach as Public crypto system instead of traditional public key infrastructure. This paper proposed used identity-based signcryption in Kerberos, that is eliminate need to public key certification that used in PKI by used identity of user as public key, and prevent the men-in-the-middle attacker from obtain the authentication key or impersonate Kerberos administrative principals. The identity-based signcryption used to sign and encrypt the message in a same algorithm in order to achieve authentication and confidentiality, also to avoid modified it during transmission.

Keywords Kerberos · Public key cryptography · Identity-based signcryption · Man-in-middle-attack

H. K. Abd-Alrazzaq (✉)
College of Administration and Economic-Ramadi, Anbar University, Anbar, Iraq
e-mail: hu_albasri@yahoo.co.uk

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_15,

15.1 Introduction

Kerberos is a network authentication protocol, it is an authentication service designed to allow clients to access servers in a secure manner over insecure networks. It is designed to provide strong authentication for client/server applications. One of the most important achievements of Kerberos is secure authentication.

The Kerberos's fundamental can be illustrated as: a KDC (key distribution center) consists of two logically separate parts: Authentication Server (AS) and a Ticket-Granting Server (TGS). The AS and TGS are responsible for creating and issuing tickets to the clients upon request. The AS and TGS usually run on the same computer, and are collectively known as the Key Distribution Center (KDC) [1]. When the client authenticates itself to the AS which forwards the username to a AS using a long-term shared secret. After receiving the user's request after check the user's ID the AS issues a Ticket Granting Ticket (TGT), which is time stamped and Session Key (used between Client and TGS), encrypts it using the user's password and returns the encrypted result to the user's workstation. When the client needs to communicate with TGS sends the TGT to it. After verifying the TGT is valid and the user is permitted to access the requested service, the TGS generates a session keys (for the communication between C and S) then issues a Ticket, which is returned to the client. The client then sends using a session key the request, which include a Ticket, to the Server (S). Later, when the client wants to contact some S, it can reuse TGT to get additional tickets from TGS which it uses to prove authentication to the S, without resorting to using the shared secret [2].

15.2 Public Key Cryptography for Initial Authentication in Kerberos

The main limitation of Kerberos is the scalability of network security infrastructures is becoming a growing. As the number of new applications and Internet is growing explosively, authentication schemes are needed which can be scaled to easily handle millions of principals within a single realm of trust.

In Ref. [3] defined extensions (PKINIT) to the Kerberos protocol in order to provide a method for using public key cryptography during initial authentication, this concern through the integration of public key cryptography with traditional Kerberos authentication. The advantages provided by public key cryptography include simplified key management; therefore adding Public Key cryptography will add a completely new dimension to Kerberos scalability as it eliminates the need to establish a large number of shared secrets.

In addition, if there is an unauthorized entry into the KDC's database, even if it is a read-only access, the privacy of the user's secret and the security of the KDC will be compromised. But with public key cryptography to violate the security one has to obtain the write access to the database [4].

There are existing three proposals in Internet drafts to add Public Key Cryptography to Kerberos:

1. Public key Cryptography for Initial Authentication in Kerberos (PKINIT) [5].
2. Public Key Cryptography for Cross Realm Authentication in Kerberos.
3. (PKCROSS) [6].
4. Public Key Utilizing Tickets for Application Servers (PKTAPP) [7].

15.2.1 PKINIT

PKINIT is an extension to Kerberos that uses public key cryptography to avoid shared secrets between a client and KAS; it modifies the AS exchange but no other parts of the basic Kerberos 5 protocol. The long-term shared key (k_C) in the traditional AS exchange is typically derived from a password, which limits the strength of the authentication to the user's ability to choose and remember good passwords.

In PKINIT, the client C and the AS possess independent public/secret key pairs are used for both signature and encryption, (pk_C, sk_C) and (pk_{AS}, sk_{AS}), respectively. Certificate sets $Cert_C$ and $Cert_{AS}$ issued by a PKI independent from Kerberos. In PKINIT the client must send a client's certificates $Cert_C$ and her signature (with her secret key) over a timestamp and another nonce to establish trust between the user and the. This information is included in the Kerberos pre-authentication fields-defined in the specification to support extensions to the protocol. The AS verifies the client's identity by verifying the digital signature. The AS replies to the client with a chain of certificates for the AS's public key, the AS's digital signature using his secret key, and the session key, with nonce; all of this is encrypted under Client's public key [8].

15.2.2 PKCROSS

PKCROSS is a logical extension of PKINIT. If a client wants certain services form a server, which is located remotely, then there has to be an authentication procedure, which relates the KDC of the client (Local KDC) with the KDC of the server (remote KDC). The primary reason why PKCROSS is used is to manage the Cross Realm Authentication. It simplifies the multiple realm authentications. The user needs to request a cross realm TGT request from its Local KDC so that it can access the remote server. The messages exchanged between local and remote KDC is similar to PKINIT, where the local KDC acts as a client. The local KDC sends a request comprising of the PKCROSS flag set to the remote KDC. The

remote KDC replies with a PKCROSS ticket and trusts the local KDC to issue the remote realm TGT to its client on behalf of the remote KDC.

15.2.3 PKTAPP

In traditional Kerberos system, the KDC issues all TGS, remote KDC, and server tickets in its realm. Thus, most authentication transactions should transit the KDC. Therefore, it can become a performance bottleneck. In PKTAPP is a more efficient protocol than traditional Kerberos from a message exchange perspective where the client may deal directly with the application server. The aim of PKTAPP specification is to eliminate this potential bottleneck and reduce communications traffic by implementing the authentication exchange directly between the client and the application server. This variation was originally introduced as Public key based Kerberos for Distributed Authentication [9], (PKDA) proposed by Sirbu & Chuang, describes PK based authentication that eliminates the use of a centralized key distribution center while retaining the advantages of Kerberos tickets.

15.3 Man-in-Middle-Attack on the Public-Key Kerberos

The client C sends a request to the KDC (K) which is intercepted by the attacker I, who constructs his own request message using the parameters from C's message. All data signed by C are sent unencrypted, therefore the client believed to be talking to KDC, is talking to I instead and this causes a failure of authentication problem, Also a failure of confidentiality. I forward the fabricated request to the KDC, who views it as a valid request for credentials if I is himself a legitimate client otherwise KDC would not talk to him; there is nothing to indicate that some of the data originated with C. KDC responds with a reply containing credentials for I, The TGT has the form $E_{k_{TGS}}\{AK, ID_I, t_{AS}\}$, because it is encrypted with the key k_{TGS} shared between KDC and the, TGS it is opaque to C (and I). The attacker knows the key AK (as well as k, which is not used other than to encrypt AK), and then C believes that KDC produced AK and k just for her, therefore the attacker can decrypt any message that C would protect with it, all the later request that client send into TGS or the server the attacker can read and replace it with his own message, then it can observe all communications between C and server or TGS and I can also pretend to C. The client is authenticated to server and TGS as I (not as C). The attacker I does not trick server or TGS to believe he is C [8], Fig. 15.1 illustrates the Man-In-The-Middle Attack on PKINIT exchange.

Where $GT = E_{K_T}\{AK, I, t_K\}$ Since the attacker learns AK in the AS exchange, he may either mediate C's interactions with the various servers while observing this traffic or simply impersonate the servers in the later exchanges. The C has AK

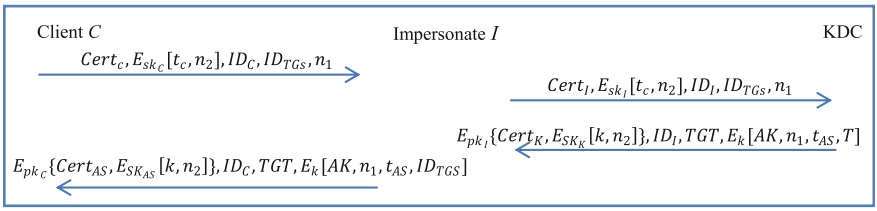


Fig. 15.1 Message flow in the man-in-the-middle attack on PKINIT

and a TGT; she would normally contact the TGS to get an *ST* for some application server S. This request contains an authenticator of the form $E_{AK}\{ID_C, t_C, T\}$. Because *I* knows *AK*, he may intercept the request and replace the authenticator with one that refers to himself $E_{AK}\{ID_I, t_I, T\}$. The reply from the TGS contains a freshly generated key *SK*; this is encrypted under *AK*, for *C* to read and thus accessible to *I*. this scenario repeated when client contact with server [8].

In another case, the attacker can intercept C's requests in the TG and CS exchanges and impersonate the involved servers instead of forwarding altered messages to them. In the TG exchange, *I* will ignore the TGT and only decrypt the portion of the request encrypted under AK. The attacker will then generate a bogus service ticket XST, which the client expects to be opaque, and a fresh key *SK* encrypted under *AK*, and sends these to *C* in what appears to be a properly formatted reply from the TGS. Also this scenario repeated when client contact with server.

15.4 Identity-Based Signcryption

The notion of identity-based (IB) cryptography was proposed by Shamir [10] as a specialization of public key (PK) cryptography which dispensed with the need for cumbersome directories, certificates, and revocation lists. The distinguishing characteristic of IB cryptography lies in its ability to use any string as a public key, such as the real name of a person. Because of this, IB systems implement an automatic directory with implicit binding, without the need for costly certification and public key publication steps.

An identity-based signcryption scheme, or IBSC, comprises four algorithms: Setup, Extract, Signcrypt, and Unsigncrypt. In a IBSC, the signcryption/unsigncryption algorithms are the composition of explicit subroutines: Signcrypt = Encrypt · Sign and Unsigncrypt = Verify · Decrypt. The Setup generates random instances of the common public parameters and master secret; Extract computes the private key corresponding to a given public identity string; Signcrypt produces a signature for a given message and private key, and then encrypts the signed plaintext for a given identity (note that the encryption routine may specifically require the signature as input); Decrypt decrypts a ciphertext using a given private key; Verify

checks the validity of a given signature for a given message and identity. Messages are arbitrary strings in $\{0,1\}^*$. the functions that compose a generic IBSC scheme are as follows [11]:

- **Setup**: produces a pair (msk, mpk), where msk is a randomly generated master secret and mpk the corresponding common public parameters.
- **Extract** (mpk, msk, ID): On input ID, computes a private key sk which corresponding to the identity ID under (msk, mpk).
- **Signcrypt** (mpk, IDS, IDR, skS, m): The sequential application of
  - **Sign** (mpk, ID_S, sk_S, m): On input (ID_S, sk_S, m), outputs a signature s, for sk_S, under mpk, and some ephemeral state data r .
  - **Encrypt** $(mpk, ID_R, sk_S, m, s, r)$: On input (ID_R, sk_S, m, s, r), outputs an anonymous ciphertext C, containing the signed message (m, s) encrypted for the identity ID_Runder mpk.
- **Unsigncrypt** (mpk, sk_R, C) : The sequential application of
  - **Decrypt** (mpk, sk_R, C): On input (sk_R, C), outputs a triple (ID_S, m, s) (containing the purported sender identity and signed message obtained by decrypting C by the private key sk_Runder mpk).
  - **Verify** (mpk, ID_S, m, s): On input (ID_S, m, s), outputs "true" or "false" (indicating whether s is a valid signature for the message m by the identity ID_S, under mpk).

15.5 The Proposed Improvement on Public Key Kerberos

15.5.1 Motivation

This paper presents new mode to solve two problems, first to eliminate need to the digital certification which is necessary in traditional public key infrastructure. Using Digital certification require providing many parts such as certification authority which issue the certificates to users, certification directory which store the issued certificates, and certification revocation list which contain on the expired certificates. Also there is need to maintain the certification list and certification revocation list for every public key continually. Second, to avoid applied the man-in-the-middle attack on the public key Kerberos by using signcryption scheme. Signcryption scheme aims at the combination of public key encryption and digital signatures in same algorithm. When the client sends request to AS, this request will sign (by his/her private key) to verify the identity of sender and encrypt (by receiver's identity) to sure the confidentiality.

The attack on public key Kerberos was possible because the attack can show the content of the request. Although a client can link a received response to a previous

request (nonce n_1 and n_2, and to the timestamp t_C), but the client cannot be sure that the AS generated the key AK and the ticket granting ticket TGT appearing in this response for her. Indeed, the only evidence of the principal for whom the AS generated these credentials appears inside the TGT, which is opaque to her.

In traditional public key that it uses in Kerberos, user before contact with the KDC he/she must get on the public key of it, also to ensure that public key is the right one he/she must get on certification of public key form Certification Authority. To use a public key that is contained in a certification, a user queries the public repository where the certificate can be found and retrieves the certificate. Because a public key may be valid for quite a while, it is often necessary to check such a public key for validity before using it. This may be by checking a list of invalid certificates or by querying an online service that returns the validity status of a certificate. After any necessary validity checking is done, the user then uses the public key to encrypt information to the owner of the public key.

This implies that the certificate approach can only work when all users are initially given at least one public key in a secure manner. This is a relatively minor assumption, since such a key can be pre-installed by a manufacturer and by a system administrator.

Therefore, in this paper propose using Identity-based signcryption scheme instead dependence on the traditional public key cryptography, that its goal to eliminate need to digital certifications in public key Kerberos, which is necessary to public key register. Also in order to simplify the key management, where the public key is any public information (such as email address or phone number) but it must be unique, also to revoke a user, the PKG will simply stop issuing her new keys. The client C and the AS possess independent private key which is derivation from the identity of each one, therefore, not need to send certification of public key to another side, this achieve better scalability to key management. This proposed is approach to making the public key immune to this attack, namely to require the AS to include the identity of this principal in a component of the response that the client can verify. The proposed mode performed on the PKINIT, PKCROSS, and PKTAPP.

15.5.2 Public Key Kerberos Using Identity-Based Signcryption

In this paper will use the identity-based signcryption as analogous in three proposal of internet draft:

1. Identity-Based SignCryption for Initial Authentication in Kerberos (IBSC-INIT).
2. Identity-Based SignCryption for Cross Realm Authentication in Kerberos (IBSC-CROSS).
3. Identity-Based SignCryption Utilizing Tickets for Application Servers (IBSC-TAPP).

In an identity-based signcryption scheme there are four algorithms that are used to create and use a public–private key pair. These are traditionally called setup, extraction, signcryption, and unsigncryption. Two algorithms of them will execute by the trust third party (TTP) to the interest of users and Kerberos. Setup is the algorithm with which the parameters needed for IBE calculations are initialized, including public parameters and master secret key. The trust third party distributes the public parameters while the master secret key keeps secretly. In addition, the TTP execute the second algorithm, is called extraction, using the system parameters, a user identity *id* and the master secret *msk* to generate the private key *usk* of the corresponding user that requests his own private.

15.5.2.1 Identity-Based SignCryption for Initial Authentication in Kerberos (IBSC-INIT)

In the initial authentication using the identity-based signcryption in the Kerberos to enable securely transmit credentials, which including TGT and session key, to the client. The client sent his request to KAS encrypt with identity (which is a public key) of KAS to ensure for user no one can read his request or modify it. Also sign user sign his request with his private key (which derivation form his identity), to authenticate user in KAS. The signature and encryption are achieve in same algorithm, where the use inputs two key his private key and KAS's public key (identity). KAS will respond with a session key and nonce_2 encrypted by user's public key and signed with KAS's private key. The user and KAS don't use the certification of public key because the identity-based eliminate need to it; this exchange is illustrated in Fig. 15.2.

If the attacker attempt using the man-in-the-middle attack to intercept the transmission in order to constructs his own request message, he must achieve two things: first, he must recover the (t_c, n_2) to enable re-encryption it with his key. This process require use two keys in unsigncryption algorithm, once is the public key of user which is known publicly, the other the secret key of KAS which is no one can know it only KAS, for that, the attacker cannot recover the encrypted content. Second the attacker must replace the identity C with his identity I before send message, but if he does that, the KSA use the identity of I as public key in unsigncryption algorithm which will return an error message because the key is incorrect. If the attacker generates his own message with replace the identity C to I without recover the original message and send to KAS (using his secret key) $SC_{SK_I, PU_{KAS}}[t_I, n_I], I, T, n_1$, after the attacker receives the response of KAS; as $SC_{SK_{KAS}, PK_I}[k, n_I], I, TGT, E_k[AK, n_1, t_k, T]$, then he must the resend message to original user. The attacker has two possible; first, he resend message to user without change, and then the user uses his private key to retrieve the message but will get the error message because it encrypts with attacker's identity. Second, he uses his private key and user's public key to generate the message to user; as $SC_{SK_I, PK_C}[k, n_I]$, but when user attempt retrieve them message he will get to error message because he use the KAS' public key to verifying form the signature in unsigncryption algorithm,

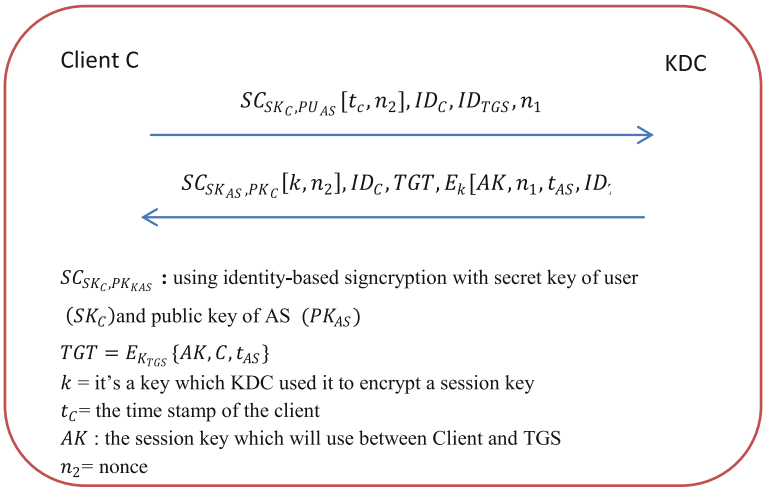


Fig. 15.2 IBSC-INIT message exchange

in addition the nonce n_I is not same which user sent. Therefore the attacker cannot modify the sent message. The Initial Authentication in Kerberos using identity-based signcryption performs without need to certification.

15.5.2.2 Identity-Based SignCryption for Cross Realm Authentication in Kerberos

IBSC-CROSS is extension of IBSC-INIT (as in PKCROSS). IBSC-CROSS is take place only between pairs of Key Distributed Center when a client wants certain services form a remote server, then the authentication procedure achieves between the KDC of the client (Local) with the remote KDC. This communication between local KDC and remote KDC will be transparent to end-users which requesting cross-realm tickets. The messages exchanged between local and remote KDC is similar to IBSC-INIT. Where the local KDC acts as a client and the local KDC sends a request comprising of the IBSC-CROSS flag set to the remote KDC. The remote KDC replies with an IBSC-CROSS ticket and trusts the local KDC to issue the remote realm TGT to its client on behalf of the remote KDC, this exchange is illustrated in Fig. 15.3.

Where AS_r is Authentication Server in remote Kerberos, TGS_l is Ticket Granting Server in local Kerberos and TGS_r is Ticket Granting Server in remote Kerberos.

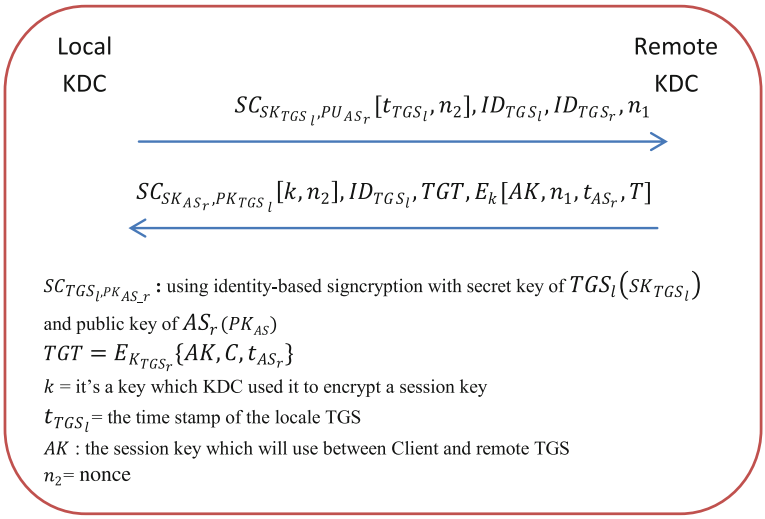


Fig. 15.3 IBSC-CROSS message exchange

15.5.2.3 Identity-Based SignCryption Utilizing Tickets for Application Servers

In traditional PKTAPP provides direct client to server authentication, thus eliminating the need for an online key distribution center. PKTAPP consist three exchange steps, the first step, the client requests a certificate from the server then the server returns response contain its certificate to the client. Therefore this step dedicated to exchange certificates between Client and server. Since, in this paper used identity-based signcryption where the identity used as public key, for that, this step removed because it has become useless. The client in IBSC-TAPP directly sends a request for a service ticket to the server; IBSC-TAPP is illustrated in Fig. 15.4.

15.6 Conclusions

Public-key cryptography enhancements to the traditional Kerberos standard incorporate a public-key infrastructure into the scope of underlying systems trusted by Kerberos. The extension employs public key cryptography to facilitate initial authentication directly between Kerberos servers and clients, this extension has two drawbacks: the first, it must use the digital certification to register the public key. The second the man-in-the-middle attack can intercept the exchange message to obtain the session key, as illustrate in [8]. This paper solves those problems by using new type of public key which depend on the identity of user to be public key

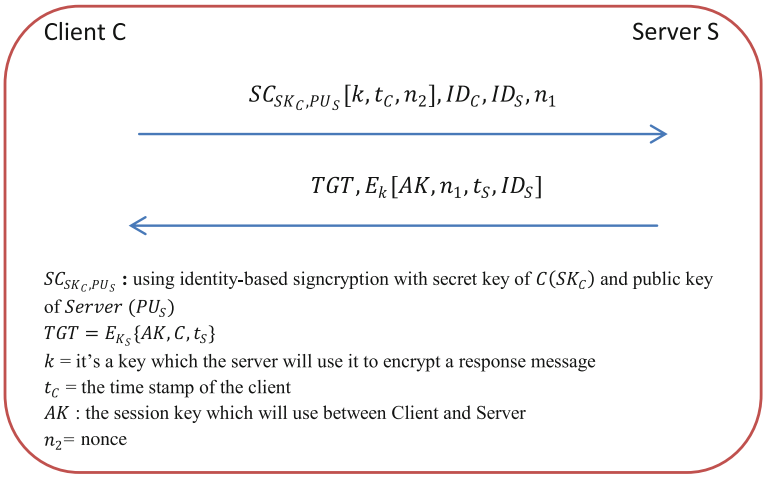


Fig. 15.4 IBSC-TAPP message exchange

and derivation the private key from this identity. This paper employs the identity-based signcryption to eliminate need to certification; also to prevent the attacker discovered the content of message.

This paper proposed three identity-based signcryption enabled Kerberos protocols IBSC-INIT, IBSC-CROSS, and IBSC-TAPP. They all improve Kerberos security and scalability by utilizing non-need to certification as in public key infrastructures and simplifying key management system. IBSC-INIT specification defines how identity-based signcryption can be used to secure the initial authentication procedure with removal danger man-in-the-middle attack. IBSC-CROSS improves the scalability of Kerberos in large multi-realm networks where many application servers may be participating in the authentication process. IBSC-TAPP enhances Kerberos scalability by distributing the authentication workload from the centralized KDC to the individual principals on the network with removal the first step which was dedicated to obtain the certification of server.

References

1. Pathan SK, Deshmukh SN, Deshmukh RR (2009) Kerberos authentication system–a public key extension. Int J Recent Trend Eng 1(2):15–18
2. Wen L, Hai C, Xingjian L, Hong Z (2010) An improved kerberos scheme based on dynamic password. Int J Inf Technol Comput Sci, MECS 2(2):33–39
3. Tung B, Neuman C, Hur M, Medvinsky A, Medvinsky S, Wray J,Trostle J (1997) Public key cryptography for initial authentication in kerberos, RFC 1510.
4. Farhana S Munnee, Jonnavitula A (2007), Kerberos using public key cryptography, GMU-ECE 646

5. Tung B et al (2001) Public key Cryptography for initial authentication in kerberos, draft-ietf-cat-kerberos-pk-init-12.txt, RFC 1510
6. Tung B et al (1998) Public key cryptography for cross-realm authentication in kerberos, draft-ietf-cat-kerberos-pk-cross-04.txt, RFC 1510
7. Medvinsky A, et al (2001) Public key utilizing tickets for application servers (PKTAPP)", draft-ietf-cat-kerberos-pk-tapp-03.txt
8. Cervesato I, Jaggard AD, Scedrov A, Tsay J-K, Walstad C (2007) Breaking and fixing public-key kerberos, pp 311–358
9. Sirbu MA, Chung-I Chuang J (1997) Distributed authentication in kerberos using public key cryptography. In:Symposium on Network and distributed system security, San Diego, CA, 10–11 Feb 1997
10. Shamir A (1984) Identity-based cryptosystems and signature schemes. In: Blakley GR, Chaum D (eds) Advances in cryptology–crypto'84, volume 196 of Lecture Notes in Computer Science. Springer, Berlin, pp 47–53
11. Dent AW, Zheng Y (2010) Practical Signcryption. Springer, Berlin

Chapter 16
Intrusion Detection Using Keystroke Dynamics

Mahalaxmi Sridhar, Teby Abraham, Johnelle Rebello, Winchell D'souza and Alishia D'Souza

Abstract In an effort to confront the challenges brought forward by the increased need for access control, we present an improved technique for authorized access to computer system resources and data via keystroke dynamics. A database of keystrokes of login ids and passwords collected from 38 users is constructed. From the samples collected, signatures were formed using three membership functions of Fuzzy Logic. Users were authenticated by comparing the typing pattern to their respective signatures. We have included the usage of the SHIFT and the CAPS LOCK keys as part of the feature sets. We analyzed the performance of the three membership functions of Fuzzy Logic based on features like FAR and FRR to evaluate the efficiency of the detection algorithms. The paper presents the results of the analysis thereby providing an inexpensive method of intrusion detection as compared to other behavioral biometric methods.

Keywords Keystroke dynamics · Intrusion detection · Fuzzy logic · Computer security

16.1 Introduction

One of the primary means of authenticating users and providing security to computers are textual passwords. Passwords are convenient and require no specialized hardware. However, users frequently share password with others,

M. Sridhar (✉) · T. Abraham · J. Rebello · W. D'souza · A. D'Souza
Don Bosco Institute of Technology, Kurla, Mumbai, Maharashtra, India
e-mail: mahalaxmi90sridhar@gmail.com

T. Abraham
e-mail: projectkd2011@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_16,

forget passwords, and select poor passwords that may be easily defeated. Compromised passwords and shared accounts are frequently exploited by both external attackers and insiders.

One idea to overcome this is to use keystroke dynamics. It is a novel approach in which a legitimate user's typing patterns such as durations of keystrokes, latencies between keystrokes etc. are combined with the user's password to generate a hardened password that is convincingly more secure than conventional passwords.

16.2 Literature Survey

Keystroke Dynamics is the manner and rhythm in which an individual types characters on a keyboard or keypad. It gives the detailed timing information that describes exactly when each key was pressed and released as a person is typing. Ever since Forsen et al. [1] investigated for the first time whether users could be distinguished by the way they type many different techniques for keystroke dynamics have been proposed.

In almost every technique the common feature sets used to form the signatures are:

- Enter: the Enter key is considered to be part of the password.
- KeyUp–KeyUp: Time between the key releases of consecutive keys is used as a feature.
- KeyUp–KeyDown: Time between the release of one key and the press of the next is used.
- KeyDown–KeyDown: Time between the key presses of consecutive keys is used as a feature.

16.2.1 Anomaly Detectors for Password Timing

Our main focus is on developing an intrusion detection system using the static login method. Various studies have been done on the use of anomaly detectors to analyze password-timing data.

Table 16.1 summarizes some of the anomaly detectors along with their results relevant to our study. False accept rate (FAR) denotes the rate that an imposter is allowed access. Similarly False reject rate (FRR) denotes the rate that the legitimate user is denied access. After thoroughly studying various anomaly detectors summarized in the Table 16.1 we concluded that fuzzy logic has a reasonable balance between FRR and FAR errors. Hence we planned to implement it using various membership functions.

Table 16.1 Comparison of various anomaly detectors and their error rates [4]

Algorithms	Feature sets			Results	
	Enter key	Keydown–keydown	Keyup–keydown	FRR	FAR
Euclidean		Y		2.8	8.1
Manhattans	Y	Y		0.25	16.4
Mahalanobis		Y		2.8	8.1
Neural-network		Y		0.2	0.22
Fuzzy-logic		Y		0.11	0.19
z-score		Y		0.02	0.13
K-means			Y	3.8	3.8

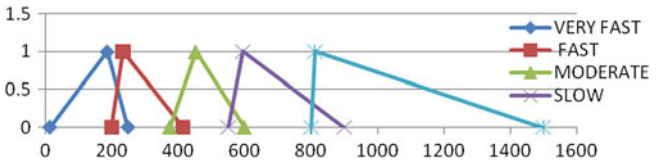


Fig. 16.1 Fuzzy sets for triangle membership function

16.3 Design

Fuzzy logic is a form of many-valued logic or probabilistic logic; it deals with reasoning that is approximate rather than fixed and exact. In contrast with traditional logic theory, where binary sets have two-valued logic: true or false, fuzzy logic variables may have a truth value that ranges in degree between 0 and 1. Every element requires a degree of membership to determine how strongly it belongs to a certain class. Degree of membership is calculated by a Membership Function. Of the many membership functions for fuzzy logic we have selected three membership functions for our study:

- Triangle
- Trapezoidal
- Gaussian

16.3.1 Formation of Intervals

A dedicated software module was designed to collect the features of 35 volunteers. Data collected from these volunteers were stored in a database and used to form the intervals of various classes; where each class represents different typing speeds. Different classes of typing speed that we decided for our project are: Very Fast, Fast, Moderate, Slow and Very Fast.

Based on the sample collected, the intervals for the three membership function mentioned before were designed as follows. (Figs. 16.1, 16.2, 16.3, 16.4).

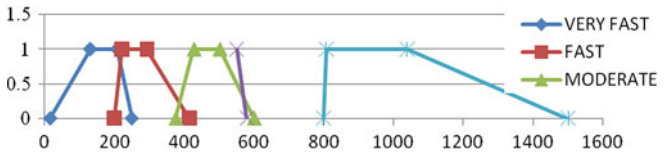


Fig. 16.2 Fuzzy sets for trapezoidal membership function

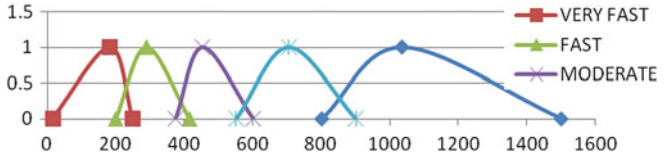
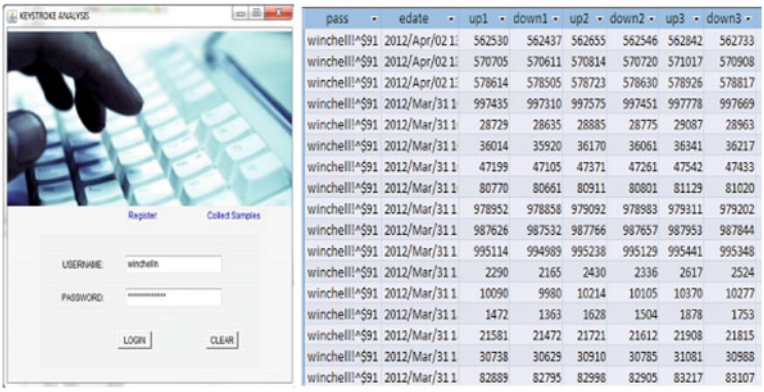


Fig. 16.3 Fuzzy sets for Gaussian membership function



pass	edate	up1	down1	up2	down2	up3	down3
winchelll^$91	2012/Apr/02 1:	562530	562437	562655	562546	562842	562733
winchelll^$91	2012/Apr/02 1:	570705	570611	570814	570720	571017	570908
winchelll^$91	2012/Apr/02 1:	578614	578505	578723	578630	578926	578817
winchelll^$91	2012/Mar/31 1	997435	997310	997575	997451	997778	997669
winchelll^$91	2012/Mar/31 1	28729	28635	28885	28775	29087	28963
winchelll^$91	2012/Mar/31 1	36014	35920	36170	36061	36341	36217
winchelll^$91	2012/Mar/31 1	47199	47105	47371	47261	47542	47433
winchelll^$91	2012/Mar/31 1	80770	80661	80911	80801	81129	81020
winchelll^$91	2012/Mar/31 1	978952	978858	979092	978983	979311	979202
winchelll^$91	2012/Mar/31 1	987626	987532	987766	987657	987953	987844
winchelll^$91	2012/Mar/31 1	995114	994989	995238	995129	995441	995348
winchelll^$91	2012/Mar/31 1	2290	2165	2430	2336	2617	2524
winchelll^$91	2012/Mar/31 1	10090	9980	10214	10105	10370	10277
winchelll^$91	2012/Mar/31 1	1472	1363	1628	1504	1878	1753
winchelll^$91	2012/Mar/31 1	21581	21472	21721	21612	21908	21815
winchelll^$91	2012/Mar/31 1	30738	30629	30910	30785	31081	30988
winchelll^$91	2012/Mar/31 1	82889	82795	82998	82905	83217	83107

Fig. 16.4 Snapshots of software used to collect keystroke samples and the associated database

16.3.1.1 Triangle Membership Function

$$
\begin{aligned}
\Lambda(u{:}\alpha,\beta,\gamma) \quad &= 0 && u<\alpha \qquad \text{Where} \quad \alpha - \text{Lower Bound Value}\\
&= (u-\alpha)/(\beta-\alpha) && \alpha<=u<=\beta \qquad \beta - \text{Modal Value}\\
&= (\alpha-u)/(\beta-\alpha) && \beta<=u<=\gamma \qquad \Gamma - \text{Higher Bound Value}\\
&= 0 && u>\gamma
\end{aligned}
\tag{16.1}
$$

16.3.1.2 Trapezoidal Membership Function

$$\begin{aligned} F(x,\ a,\ b,\ c,\ d) &= 0 && \text{when } x < a \text{ and } x > d \\ &= (x - a)/(b - a) && \text{when } a <= x <= b \\ &= 1 && \text{when } b <= x <= c \\ &= (d - x)/(d - c) && \text{when } c <= x <= d \end{aligned} \tag{16.2}$$

16.3.1.3 Gaussian Membership Function

$$G\ (u:\ m, \sigma) = \exp[-\{(u - m)/\sqrt{2\sigma}\}2] \quad \text{Where m---Mean Value} \tag{16.3}$$

16.4 Implementation

16.4.1 Sample Collection and Signature Formation

Inter-key delays were collected using dedicated software and the samples were stored in a simple database.

Once sufficient samples were collected, a minimum of 15 samples for each user were collected, intervals generated and algorithms for each of the three membership functions are generated. An example of a simple algorithm [2] implementing the triangle membership function as an anomaly detector is shown below. Similar algorithms were developed by us for the other two membership functions.

If (Input < LowerBound OR Input > UpperBound)

Then 0

Else If (Input < Midvalue)

Then (Input − LowerBound)/(Midvalue − LowerBound)

Else If (Input = Midvalue)

Then 1

Else (UpperBound − Input)/(UpperBound − Midvalue)

The Feature Sets used for our study are listed below in Table 16.2. We included the SHIFT and the CAPS LOCK key in the feature sets of our fuzzy logic. It is often observed that the tendency to use the RIGHT_SHIFT or the LEFT_SHIFT or CAPS LOCK to type special characters and upper-case letters differ from user to

Table 16.2 Feature sets

Algorithm	Feature sets			
	Keydown-keyup	Keyup–keyup	Shift key	Caps lock
Fuzzy logic	No	Yes	Yes	Yes

user [3]. This variation can thus be used as an additional parameter to validate legitimate users from imposters.

A membership function calculates the degree of membership to each class for each inter-key delay (KeyUp–KeyUp) given as input.

Based on the input a signature for a particular user is determined. One such signature formed is shown in Table 16.3.

16.4.2 Signature Comparison

In the working phase the real time signature of a user is compared with the stored signature. If both signatures match up to a certain limit (in this case it is up to 70 %) then the user is verified as the genuine user and granted access; else they are not granted access.

16.5 Testing

To increase the confidence in the correctness (accuracy) of specified membership function of Triangular, Trapezoidal and Gaussians, we conducted testing by supplying typical test inputs (request) and subsequently checking test output (responses) against expected ones to enhance the correctness of specified algorithm (Fig. 16.5)

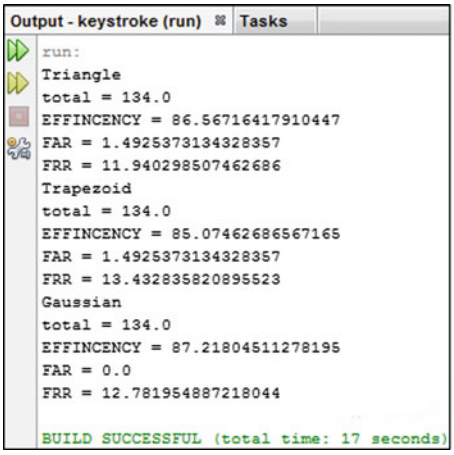
As we can see from Table 16.4, comparison of the FAR and FRR of all the three membership function shows that Gaussian function yields the best results as compared to the other two membership functions.

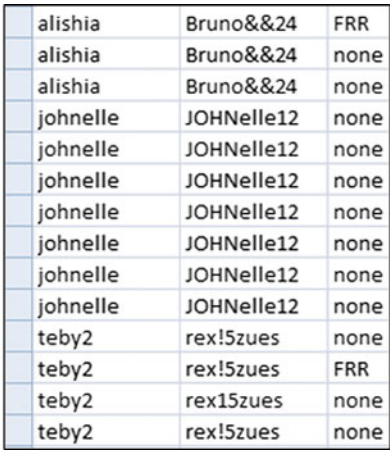
16.6 Conclusion and Future Scope

We believe keystroke dynamics can be used effectively to safeguard against unauthorized access of computer as well as mobile resources [2]. When implemented in conjunction with traditional schemes, it allows for the design of more robust authentication systems than traditional password based alternatives alone.

Table 16.3 Signature formation and comparison

Signature	1st	2nd	3rd	4th	5th	6th	7th	8th	Shift left	Shift right	Caps lock
Stored	Very fast	Fast	Moderate	Very fast	Fast	Fast	Very fast	Very fast	1	0	0
Detected	Fast	Fast	Moderate	Very fast	Slow	Fast	Very fast	Very fast	0	1	0





	alishia	Bruno&&24	FRR
	alishia	Bruno&&24	none
	alishia	Bruno&&24	none
	johnelle	JOHNelle12	none
	johnelle	JOHNelle12	none
	johnelle	JOHNelle12	none
	johnelle	JOHNelle12	none
	johnelle	JOHNelle12	none
	johnelle	JOHNelle12	none
	johnelle	JOHNelle12	none
	teby2	rex!5zues	none
	teby2	rex!5zues	FRR
	teby2	rex15zues	none
	teby2	rex!5zues	none

Fig. 16.5 Snapshots showing the experimental results generated, stored and evaluated

Table 16.4 Performance measure of membership functions

	FAR	FRR	Accuracy (%)
Triangle	1.4	11.19	87.41
Trapezoidal	1.4	12.59	86.01
Gaussian	0.38	11.97	88.03

In this project we compared triangular, trapezoidal and Gaussian membership functions of fuzzy logic to authenticate users based on their typing speed and proved that among the three, Gaussian membership function is the most effective means of implementing an intrusion detection system using Fuzzy logic. To implement such a system, the code developed by us in Java could be used as a plug-in for intrusion detection, once the database has been created for the authentic users.

The approach of using keystroke dynamics in our project was limited only to passwords. This can be extended to include all the text typed by a user during his

work. This way, not only will there be monitoring at the login stage but also during the entire active session for a particular user.

References

1. Maxion RA, Killourhy KS (2010) Keystroke biometrics with number-pad input, Computer Science Department, Carnegie Mellon University
2. Haider S, Abbas A, Zaidi AK (2000) A multi-technique approach for user identification through keystroke dynamics. In: IEEE international conference on systems, man and cybernetics
3. Killourhy KS, RA Maxion Comparing anomaly-detection algorithms for keystroke dynamics
4. Forsen G, Nelson M, Staron R Jr (1977) Personal attributes authentication techniques. Technical Report RADC-TR-77-333, Rome Air Development Center
5. Ahmed Awad EA, Traore I Detecting computer intrusions using behavioural biometrics
6. Monrose F, Rubin AD (1999) Authentication via keystroke dynamics
7. Killourhy KS (2012) A scientific understanding of keystroke dynamics
8. Joyce and G. Gupta (1990) Identity authentication based on keystroke latencies. Commun ACM
9. Ahmed AAE, Traore I Department of Electrical and Computer Engineering, University of Victoria. Detecting computer intrusions using behavioural biometrics
10. Lane Department of Computer Science and Electrical Engineering (2005) Morgantown, West Virginia, Username and password verification through keystroke dynamics
11. Duda RO, Hart PE, Stork DG (2001) Pattern classification, 2nd edn. Wiley
12. Mandal SN, Choudhury JP, De D, Chaudhuri SRB (2008) Roll of membership functions in fuzzy logic for prediction of shoot length of mustard plant based on residual analysis

Chapter 17
MRI–PET Medical Image Fusion Technique by Combining Contourlet and Wavelet Transform

Ch. Hima Bindu and K. Satya Prasad

Abstract This paper proposes the application of the hybrid Multiscale transform in medical image fusion. The multimodality medical image fusion plays an important role in clinical applications which can support more accurate information for physicians to diagnosis diseases. In this paper, a new fusion scheme for Magnetic Resonance Images (MRI) and Positron Emission Tomography (PET) images based on hybrid transforms is proposed. PET/MRI medical image fusion has important clinical significance. Medical image fusion is the important step after registration, which is an integrative display method of two images. The PET image indicates the brain function and a low spatial resolution; MRI image shows the brain tissue anatomy and contains no functional information. Hence, a perfect fused image should contains both more functional information and more spatial characteristics with no spatial and color distortion. Firstly, the image is decomposed into high and low frequency subband coefficients with discrete wavelet transform (DWT). On these coefficients apply contourlet transform individually before going for fusion process. Later the fusion process is performed on contourlet components for each subband, for fusion the spatial frequency method is used. Finally, the proposed algorithm results are compared with different Multiscale transform techniques. According to simulation results, the algorithm holds useful information from source images.

Keywords Image fusion · Discrete wavelet transform · Contourlet transform

Ch. Hima Bindu (✉)
Department of Electronics and Communications, QIS College of Engineering and Technology, Ongole, Andhra Pradesh, India
e-mail: hb.muvvala@gmail.com

K. Satya Prasad
Department of Electronics and Communications, JNTU College of Engineering, Kakinada, Andhra Pradesh, India
e-mail: prasad_kodati@yahoo.co.in

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_17,

17.1 Introduction

Image fusion is to integrate the information of two or more source images in order to obtain the more accurate, comprehensive and reliable description of the same body part. Its aim is to combine complementary information from multiple images, such that the obtained new image is more suitable for human visual and machine perception or further image processing and analysis tasks. There are many schemes for the multi-focus image fusion, which include the probabilistic method, PCA method [1], multi-scale method [2] and the multi-resolution methods [3–6]. The wavelet-based fusion scheme cannot preserve the salient features in source images and will probably introduce some artifacts and inconsistency in the fused images [7]. Contourlet transform has drawn much attention because it can efficiently handle several important problems. Contourlet is a flexible multi-resolution, local, and directional image expansion using contour segments. It can provide a Multiscale and directional decomposition for images, which is more suitable for catching the features in images that are abundant in complex contours, edges and textures. A most often referred implementation scheme of contourlet is Laplacian Pyramid (LP) followed by directional filter bank (DFB) [5]. The MRI image shows the brain tissue anatomy and contains no functional information. The PET image indicates the brain function and has a low spatial resolution. The MRI–PET image fusion by Sabalan Daneshvar et al. based on combining HIS and retina models to improve the functional and spatial information content [1].

The rest of the paper is organized as follows: Sect.17.2 reviews the multi scale transform techniques. Section 17.3 presents the generic fusion model. Section 17.4 explains the proposed medical image fusion algorithm. Section 17.5 gives the experimental results. In the laconic section the paper is concluded.

17.2 Multi Scale Transforms

17.2.1 Contourlet Transform

Wavelet bases present some limitations, because they are not well adapted to the detection of highly anisotropic elements such as alignments in an image. Recently Do and Vetterli [8] proposed an efficient directional multi resolution image representation called the contourlet transform. Contourlet transform has better performance in representing the image salient features such as edges, lines, curves and contours than wavelet transform because of its anisotropy and directionality. The contourlet transform consists of two steps which is the sub band decomposition and the directional transform. A Laplacian pyramid is first used to capture point discontinuities, then followed by directional filter banks to link point discontinuity into lineal structure. Contourlet are implemented by using a filter bank that decouples the multiscale and the directional decompositions. In Fig. 17.1, Do and Vetterli show a conceptual filter bank setup that shows this decoupling.

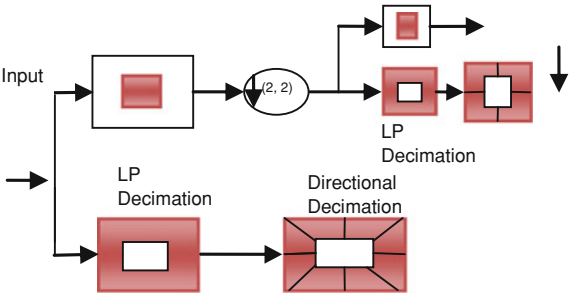


Fig. 17.1 Filter bank for contourlet transforms

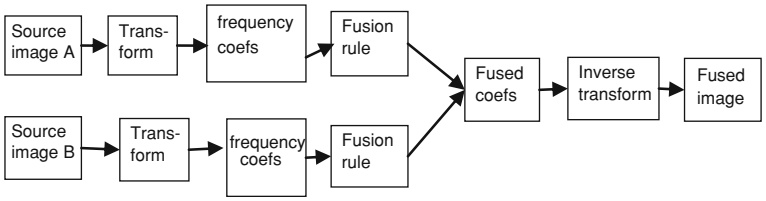


Fig. 17.2 Block diagram of generic model of multi scale image fusion

17.3 Generic Image Fusion Model

The importance of information offered by the medical images for diagnosis support can be increased by combining images from different compatible medical devices. The fusion process allows combination of salient feature of these images. In this subsection, to better understand the concept and procedure of the wavelet based fusion technique, a schematic diagram is first given bellow [9]. The main idea of our algorithm is as follows and shown in the Fig. 17.2.

- The two images are respectively decomposed into the sub images using forward transform, which have the same resolution at same levels and different resolution among different levels.
- Information fusion is performed based on the high frequency sub images of decomposed images and finally the result image is obtained inverse transform.

17.4 Proposed Scheme

The previous section explains generic fusion process using multi scale transformation techniques individually are encouraging. Each transform has its own advantages may be useful for fusion. So, here we proposed hybrid method considering both the advantages of the wavelet and contourlet transform. The fusion process, as shown in Fig 17.3, is accomplishes by the following steps:

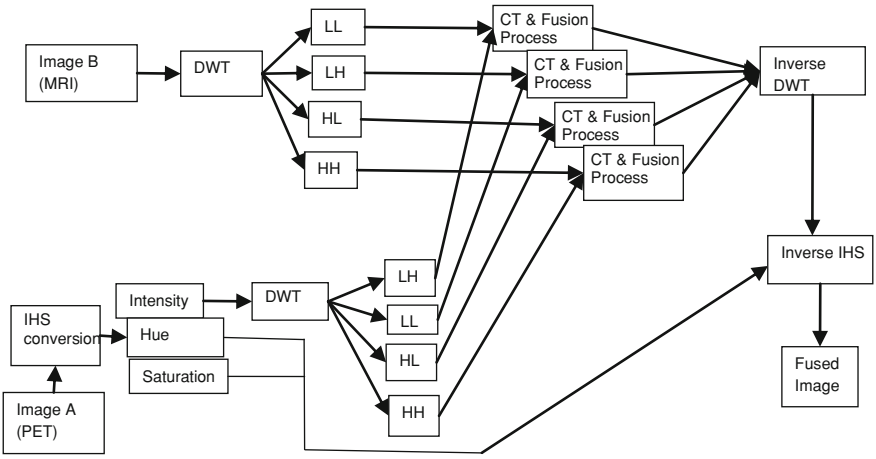


Fig. 17.3 Proposed fusion method

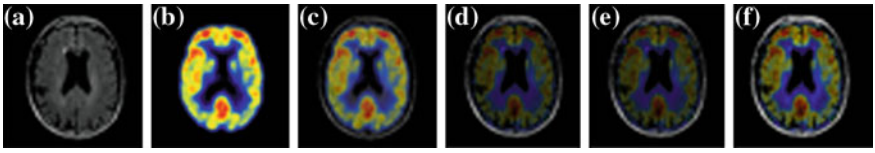


Fig. 17.4 New: mild Alzheimer's disease MRI and PET images (**a**, **b**). **c** Ground truth image, fused image using **d** DWT method, **e** contourlet transform method, **f** proposed method

Step-I. Read the MRI and PET scan images of same part of the body information are to be fused.

Step-II. Each of the source images are decomposed by DWT into one low-frequency approximate component and three high-frequency detail components, LL_1, LH_1, HL_1, HH_1 and LL_2, LH_2, HL_2, HH_2.

Step-III. Each pair of the DWT components are decomposed into contourlet coefficients (C_1, C_2, … C_n) and then apply Fig 17.2 fusion process.

Step-IV. Fuse the transform coefficients, C_1, C_2, … C_n using fusion rule. The fusion rule starts with calculation of Spatial Frequency (SF) value from the B × B block transform coefficients. Then select those block coefficient with the largest SF value as the coefficient at that location in fused image. The approach of fusion process as:

$$\text{Fused} = \begin{cases} C_1 & SF_1 > SF_2 + TH \\ C_2 & SF_1 < SF_2 + TH \\ \left(\frac{C_1 + C_2}{2}\right) & otherwise \end{cases} \tag{17.1}$$

Here the B × B block size is selected as 8 × 8 and threshold (TH) value is 3. The fused coefficients are obtained by using the inverse contourlet transform.

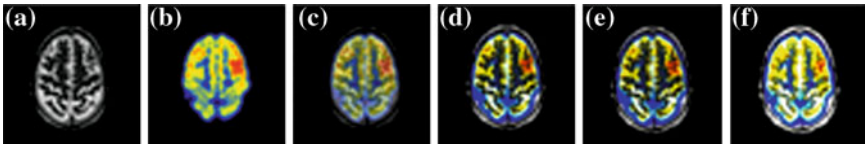


Fig. 17.5 New: Mild Alzheimer's disease MRI and PET images (**a**, **b**). **c** Ground truth image, fused image using **d** DWT method, **e** contourlet transform method, **f** proposed method

Table 17.1 Evaluation of different methods for two set of MRI–PET images

Algorithm	Image1		Image2	
	PSNR (db)	Entropy (bits/pix)	PSNR (db)	Entropy (bits/pix)
DWT	60.38	2.211	63.09	2.211
Contour let transform	60.39	2.965	63.14	2.613
Hybrid method	61.02	3.267	63.26	3.267

Step-V. Apply Inverse DWT Transform to resultant coefficients in Step IV to get the fused image F.

17.5 Results and Discussion

The proposed algorithm for the fusion of MRI and CT images is tested and compared with the traditional wavelet and contourlet fusion algorithm. For the evaluation purpose, the visual quality of the obtained fusion result as well as the quantitative analysis is used.

17.5.1 Performance Measures

In this paper, we use the Peak Signal to Noise Ratio (PSNR) and entropy (EN) are used to evaluate the performance of the proposed fusion method.

$$PSNR = 10\log\left|\frac{255^2}{\frac{1}{MxN}\sum\sum\left(G(m,n) - Z(m,n)\right)^2}\right| \tag{17.2}$$

It is expressed in dB. Its value will be high when the fused and reference images are similar. Higher value implies better fusion.

17.5.2 Fusion Results

The test data consist of color PET and high resolution MRI images. The spatial resolution of MRI and PET images are 256 × 256 and 128 × 128 pixels. The color PET images were registered to the corresponding MRI images (Figs. 17.4, 17.5, Table 17.1).

17.6 Conclusion

In this paper, we proposed a new approach for PET/MRI image fusion by using hybrid Multiscale transform techniques using the wavelet and the contourlet transform. For fusion process the spatial frequency values are calculated for block by block coefficient values. These values are compared to generate new fusion coefficients. This eliminated the influence of image imbalance and blurred phenomenon of fusion image, improved the clarity and provides more reference information for doctors. Then statistically analyze tools such as PSNR and entropy were concluded that the proposed algorithm did considerably increase spatial information content than basic transform methods of DWT and contourlet transform methods.

References

1. Daneshvar S, Ghassemian H (2010) MRI and PET image fusion by combining his and retina-inspired models, Inf Fusion 11:114–123
2. Jia YH (1998) Fusion of landsat TM and SAR images based on principal component analysis. Remote Sens Technol Appl 13(1):46–49
3. Mukhopadhyay S, Chanda B (2001) Fusion of 2D gray scale images using multi-scale morphology. Pattern Recognit 34(12):1939–1949
4. Petrovic VS, Xydeas CS (2004) Gradient-based multi-resolution image fusion. IEEE Trans Image Process 13(2): 228–237
5. Liu Y, Yang J, Sun J (2010) PET/CT medical image fusion algorithm based on multiwavelet transform. In: 2nd international conference on advanced computer and control, pp 264–268
6. Barron DR, Thomas ODJ (2001) Image fusion through consideration of texture components. IEEE Trans Electron Lett 37(12):746–748
7. Yang L, Guo BL, Ni W (2008) Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform. Neurocomputing 72:203–211
8. Yang L, Guo BL, Li W (2008) Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform. Neurocomputing 72(1–3):203–211
9. Chien JT, Wu CC (2002) Discriminant wavelet faces and nearest feature classifiers for face recognition. IEEE Trans PAMI 24:1644–1649
10. Bloch I (1996) Information combination operators for data fusion: a review with classification. IEEE Trans SMC (Part A) 26(1):52–67
11. Daneshvar S, Ghassemian H (2010) MRI and PET image fusion by combining his and retina-inspired models. Inf Fusion 11:114–123

12. Zhang J, Ma S, Han X (2006) Multiscale feature extraction of finger-vein patterns based on curvelets and local interconnection structure neural network. In: Proceedings of international conference on pattern recognition, vol 4, pp 145–148
13. Zhao M, Li P, Liu Z (2008). Face recognition based on wavelet transform weighted modular PCA. In Proceedings of the congress in image and signal processing
14. LI S, Wang JT (2001) Combination of images with diverse focuses using the spatial frequency. Inf Fusion 2:169–176

Author Biographies

Ch. Hima Bindu is currently working as an Associate Professor in ECE Department, QIS College of Engineering and Technology, Ongole, Andhra Pradesh, India. She is working towards her Ph.D. at JNTUK, Kakinada, India. She received her M.Tech. from the same institute. She has 9 years of experience in teaching the undergraduate and post graduate students. She has published eight research papers in International journal and more than three research papers in National and International Conferences. Her research interests are in the areas of image segmentation, image fusion, image Feature Extraction and Signal Processing.

Dr. K. Satya is currently Rector and Professor in ECE Department, JNTUK, Kakinada, India. He received his Ph.D. from IIT, Madras. He has more than 31 years of experience in teaching and 20 years of R&D. He is an expert in Digital Signal Processing. He produced four PhD's and guiding ten PhD scholars. He authored Electronic Devices and Circuits text book. He held different positions in his carrier like Head of the Department, Vice Principal, and Principal for JNTU Engg College. He published more than 40 technical papers in national and International journals and conferences.

Chapter 18
Comparison of Routing Protocols in Mobile Ad-hoc Network

Shubhangi M. Mahamuni, Vivekanand Mishra and Vijay M. Wadhai

Abstract In case of wireless Mobile Ad-hoc Networks, routing protocols plays very important role. Routing protocols like AODV and DSDV are the protocols which are used for the high mobility to improve packet delivery ratio. AODV and DSDV are compared in this paper in terms of throughput, end to end delay and packet delivery ratio varying number of nodes, speed and time. In case of DSDV, routing protocol routing table at node is required to maintain which will not be the need in case of AODV. DSDV Routing protocol consumes more bandwidth, because of High mobility results in frequent link failures and the overhead involved in updating all the nodes with the new routing information. AODV use on-demand route discovery, but with different routing mechanisms. Our simulation result in NS-2 shows the performance of AODV is better under high mobility than DSDV.

Keywords AODV · DSDV · MANET · CBR

S. M. Mahamuni (✉)
Department of Electronics and Telecommunications, MAE, Alandi(D), Pune, Maharastra, India
e-mail: shubhangim11@gmail.com

V. Mishra
Department of Electronics, SVNIT, Surat, Gujrat, India
e-mail: vive2009@gmail.com

V. M. Wadhai
MITCOE, Kothrud, Pune, Maharashtra, India
e-mail: Wadhai.vijay@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_18,

18.1 Introduction

Lower utilization and tremendous increase in the use of radio resources needs efficient utilization of radio resources. Now days the concept of cognitive radio is more challenging area of research as well as is the need of the new era. We proposed a spectrum sensing using routing protocol AODV. In this paper we were compared routing protocols like AODV and DSDV. As routing is closely attached with the spectrum assignment to determine frequency band. We have been proposed our work by comparing to other typical approaches, our protocol provides better adaptability to the dynamic spectrum and multi-flow environment, and incurs much lower cumulative delay [1]. Energy-Aware Performance Metric for AODV and DSDV Routing Protocols is explained in Mobile Ad-Hoc Networks [2–4]. Dynamic topologies, Bandwidth-constrained links, Energy constrained operation, and limited physical securities are studied from [5]. Performance Comparison of AODV, DSDV and I-DSDV Routing Protocols in Mobile Ad Hoc Networks is based on [6]. The concept of mobile Ad-hoc network at different routing methods can be considered from [7, 8]. To minimize routing overhead when nodes construct network topology instead of how to save routing table storage [9] used during the design of AODV protocol.

18.1.1 Classification of Ad-hoc Routing Protocol

Routing protocol in MANET can be classified into several ways depending upon their network structure, communication model, routing strategy, and state information and so on but most of these are done depending on routing strategy and network structure [1]. Based on the routing strategy the routing protocols can be classified into two parts: 1. Table driven and 2. Source initiated (on demand) while depending on the network structure these are classified as flat routing, hierarchical routing and geographic position assisted routing [2]. Flat routing covers both routing protocols based on routing strategy (Fig. 18.1).

18.1.2 On-Demand Distance Vector Routing Protocol

The Ad hoc On-Demand Distance Vector (AODV) [3] algorithm enables dynamic, self-starting, multihop routing between participating mobile nodes wishing to establish and maintain an ad hoc network. AODV allows mobile nodes to obtain routes quickly for new destinations, and does not require nodes to maintain routes to destinations that are not in active communication. AODV allows mobile nodes to respond to link breakages and changes in network topology in a timely manner. The operation of AODV is loop-free, and by avoiding the Bellman-Ford “counting

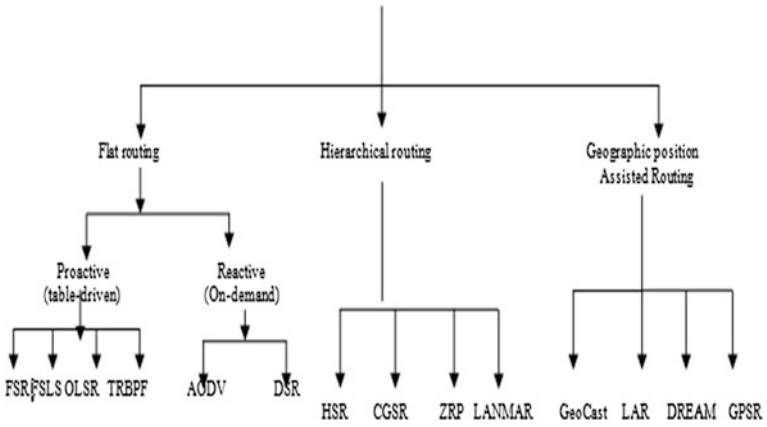


Fig. 18.1 Classification of routing protocols

to infinity" problem offers quick convergence when the adhoc network topology changes (typically, when a node moves in the network). When links break, AODV causes the affected set of nodes to be notified so that they are able to invalidate the routes using the lost link. Route Requests (RREQs), Route Replies (RREPs) and Route Errors (RERRs) are message types defined by AODV [4].

18.1.3 Destination-Sequenced Distance-Vector Routing

Destination-Sequenced Distance-Vector Routing (DSDV) is a table-driven routing scheme for adhoc mobile networks based on the Bellman-Ford algorithm. It was developed by C. Perkins and P. Bhagwat in 1994. It eliminates route looping, increases convergence speed, and reduces control message overhead. In DSDV, each node maintains a next-hop table, which it exchanges with its neighbors. There are two types of next-hop table exchanges: periodic full-table broadcast and event-driven incremental updating. The relative frequency of the full-table broadcast and the incremental updating is determined by the node mobility. In each data packet sent during a next-hop table broadcast or incremental updating, the source node appends a sequence number. This sequence number is propagated by all nodes receiving the corresponding distance-vector updates, and is stored in the next-hop table entry of these nodes. A node, after receiving a new next-hop table from its neighbor, updates its route to a destination only if the new sequence number is larger than the recorded one, or if the new sequence number is the same as the recorded one, but the new route is shorter. In order to further reduce the control message overhead, a settling time is estimated for each route. A node updates to its neighbors with a new route only if the settling time of the route has expired and the route remains optimal [5].

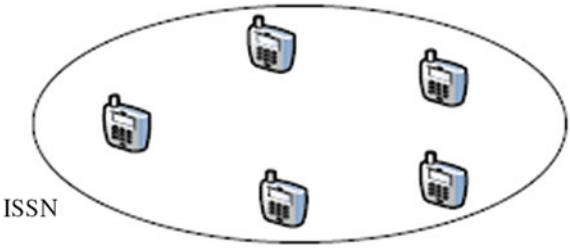


Fig. 18.2 Mobile Ad-hoc network

18.1.4 Mobile Ad-hoc Network

An ad-hoc network is a collection of wireless mobile hosts forming a temporary network without the aid of any stand-alone infrastructure or centralized administration. Mobile Ad-hoc networks (MANET) are self-organizing and self-configuring multihop wireless networks where, the structure of the network changes dynamically. This is mainly due to the mobility of the nodes. Nodes in these networks utilize the same random access wireless channel, cooperating in a friendly manner to engaging themselves in multihop forwarding. The node in the network not only acts as hosts but also as routers that route data to/from other nodes in network [6]. Each device in a MANET is free to move independently in any direction and therefore change its links to other devices frequently. Each must forward traffic unrelated to its own use, and therefore be a router. Routing in ad-networks has been a challenging task ever since the wire- less networks came into existence. The major reason for this is the constant change in network topology because of high degree of node mobility. A number of protocols have been developed for accomplish this task. Routing is the process of selecting paths in a network along which to send network traffic. In packet switching networks, routing directs packet forwarding, the transit of logically addressed packets from their source toward their ultimate destination through intermediate nodes. An ad hoc routing protocol is a convention, or standard, that controls how nodes decide which way to route packets between computing devices in a mobile ad-hoc network. In ad-hoc networks, nodes do not start out familiar with the topology of their networks. The basic idea is that a new node may announce its presence and should listen for announcements broadcast by its neighbors. Each node learns about nodes nearby and how to reach them, and may announce that it, too, can reach them. Wireless ad-hoc networks have gained a lot of importance in wireless communications. Wireless communication is established by nodes acting as routers and transferring packets from one to another in ad-hoc networks. Routing in these networks is highly complex due to moving nodes and hence many protocols have been developed. In this paper we have selected two main and highly proffered routing protocols for analysis of their performance. Figure 18.2 represents mobile Ad–hoc Network (MANET).

18.2 Performance Metrics

The following metrics are used in this paper for the analysis of AODV and DSDV routing protocols.

(i) Packet Delivery Ratio
(ii) Average End to End Delay
(iii) Throughput

18.2.1 Packet Delivery Ratio

The packet delivery ratio in this simulation is defined as the ratio between the number of packets sent by constant bit rate sources (CBR, "application layer") and the number of received packets by the CBR sink at destination.

18.2.2 Routing Overhead

It is the number of packet generated by routing protocol during the simulation and can be defined as: overhead = i.

Where overhead is the control packet number generated by node i. The generation of an important overhead will decrease the protocol performance.

18.2.3 Average end-to-end Delay of Data Packets

There are possible delays caused by buffering during route discovery latency, queuing at the interface queue, retransmission delays at the MAC, and propagation and transfer times. Once the time difference between every CBR packet sent and received was recorded, dividing the total time difference over the total number of CBR packets received gave the average end-to-end delay for the received packets. This metric describes the packet delivery time: the lower the end-to-end delay the better the application performance [7].

Avg E2E delay = $\sum$ CBR PKT Received by CBR sinks/$\sum$ CBR PKT Received by CBR sources.

Table 18.1 Parameters considered for the simulation

Parameter	Value
Simulator	Ns 2.34
Channel	Wireless channel
Propagation	Radio Propagation
Network interface	Phy/wireless phy
MAC type	Mac/802_11
Interface queue	Queue/drop tail/pri-queue
Link layer type	LL
Antenna model	Antenna/Omni antenna
Number of mobile nodes	23
Routing protocol	AODV/DSDV

18.2.4 Throughput of the Mobile Ad-hoc Networks

It is one of the dimensional parameters of the network which gives the fraction of the channel capacity used for useful transmission selects a destination at the beginning of the simulation i.e. information whether or not data packets correctly delivered to the destinations.

18.3 Simulation Model and Results

18.3.1 Simulation Model

Parameters required for the formation of simulation model in NS-2 are as shown in Table 18.1.

18.3.2 Algorithm

Comparison of AODV and DSDV routing Protocols

1. Setting the Distance Variables
2. Defining Node Configuration
3. Creating The Wireless Nodes
4. Setting The Initial Positions of Nodes Giving Mobility to Nodes
5. Setting The Node Size
6. Setting The Labels For Nodes
7. Setting Color For Server
8. Establishing Communication
9. Mobile Node Movements

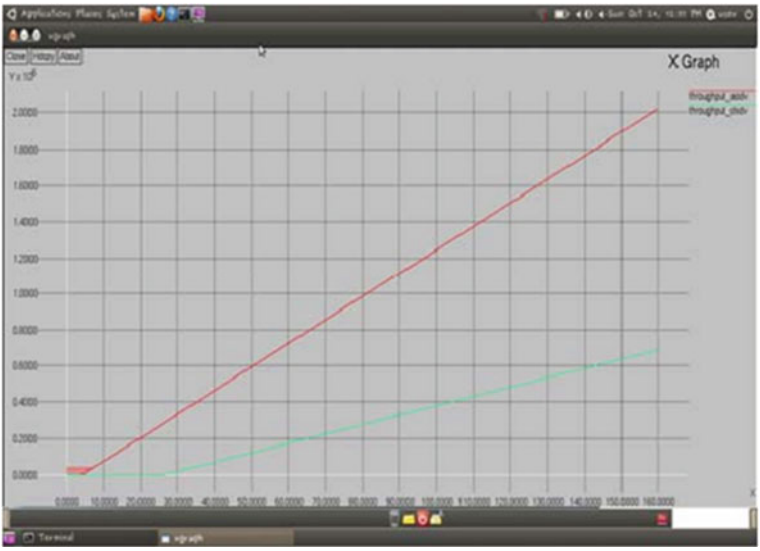


Fig. 18.3 Throughput Xgraph of AODV and DSDV with pause time set to 4 ms

10. Node27 Cache The Data From Server
11. Packet Loss At Node27
12. Node10 Cache The Data
13. Calculation of throughput, End to End delay, Packet Delivery Ratio for different packet size and time interval
14. Comparison of AODV and DSDV protocols.

In this paper we have been used ns-2 as a simulator tool for the determination of throughput, end to end delay and packet delivery ratio as shown in Figs. 18.3, 18.4, and 18.5 respectively using both routing protocols AODV and DSDV [8, 9]. At packet size 100, 500 and 1000 bytes, number of packet received is 74. Performance of DSDV protocol at different packet size is shown in Fig. 18.5. The number of packets received in AODV is decreasing with increase in packet size so the packet delivery ratio (PDR) is also decreasing. Average delay between packet sending is also decreasing with increasing packet size. Throughput is also decreasing with increasing packet size. The performance of AODV protocol is decreasing with increase in packet size as shown in performance matrices. In the analysis part packet delivery ratio (PDR) is decreasing with increase in packet size.

Throughput of DSDV protocol is decreasing as packet size is increasing. Routing overhead is also increasing with increase in packet size. Average end to end delay is decreasing with increase in packet size. It means the performance of DSDV protocol is high at less packet size except the routing load. As shown in Tables 18.2 and 18.3, it clearly shows that AODV routing protocol is better than DSDV in terms of throughput, packet delivery ratio. But if you consider the end to and delay, DSDV is better as compared to AODV routing protocol.

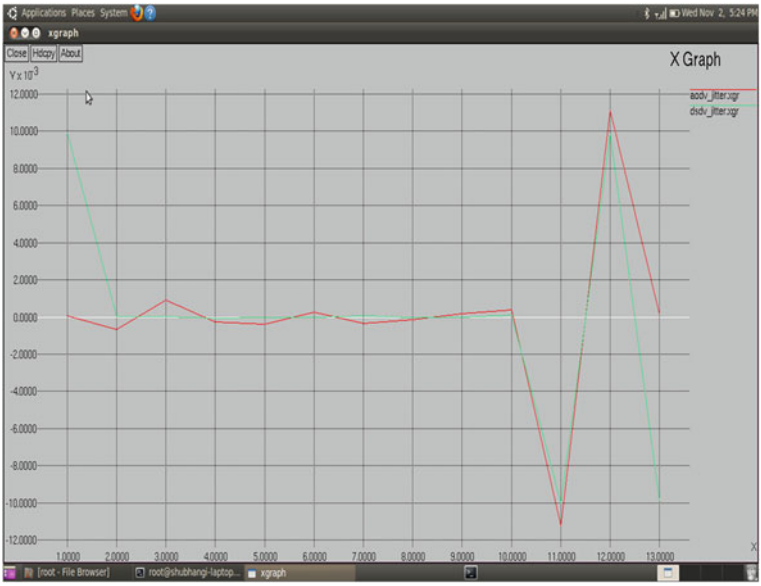


Fig. 18.4 Screen shot of end to end delay

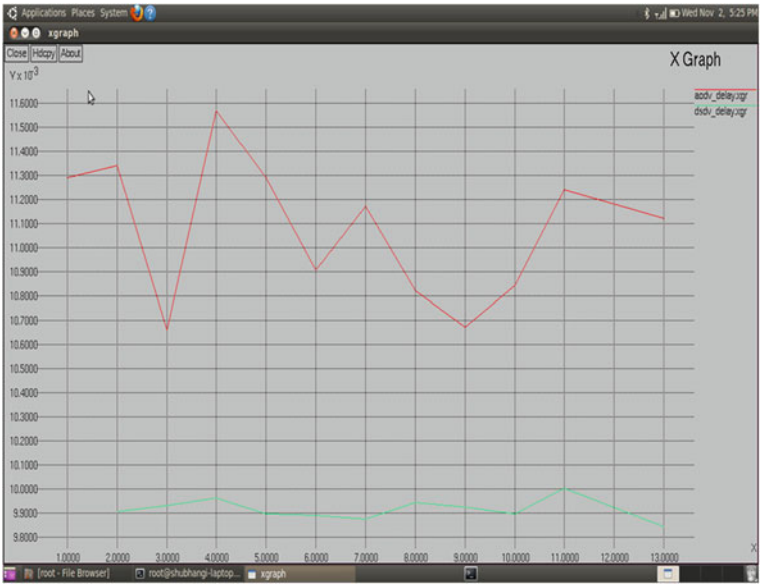


Fig. 18.5 Screen shot of packet delivery ratio fraction

From the Figs. 18.4 and 18.5, the on-demand protocols, DSDV and AODV performed well delivering the greater % of the originated data. When the no. of

Table 18.2 Performance matrix packet size = 500 bytes interval of 0.01 s

Protocol	Packets sent	Packets received	Average end-to-end delay	Throughput	Routing load
AODV	8	8	0	2.2	11.20
DSDV	8	6	−10	0.6	9.85

Table 18.3 Performance Matrix Packet Size of 100 bytes interval of 0.015 s

Protocol	Packets sent	Packets received	Average end-to-end delay	Throughput	Packet delivery ratio (%)
AODV	500	320	2.54	30.11	64
DSDV	500	214	2.26	12.22	42.8

nodes increases AODV performs better as nodes attain a stable path and become stationary. DSDV performance drops because more packets drop due to link breaks. Variation in speed of nodes has less impact on AODV protocol. DSDV produces more sent packet as it recovers from dropped packets.

Acknowledgments Towards this work, we are very thankful to research lab, MIT Kothrud, Pune and SVNIT, Surat. This work would be useful for the further work in handoff delay minimization in case of cognitive radio communication.

References

1. Cheng G, Liu W, Li Y, Cheng W (2007) Spectrum aware on-demand routing in cognitive radio networks. Proceedings of IEEE DySpan, 2007, vol 1, no 4, pp 364–371. 1-4244-0663-3/07/$20.00 ©2007 IEEE (Huazhong University of Science & Technology, China)
2. Vijayalakshmi P, Saravanan V, Ranjit Jeba Thangiah P, Abraham Dinakaran J (2011) Energy-aware performance metric for AODV and DSDV routing protocols in mobile Ad-Hoc networks. IJCSI Int J Comput Sci Issues, 8(4)(1). ISSN (Online):1694-0814 www.IJCSI.org (Karunya University)
3. Brouwer F, De Graaf M, Nikookar H, Hoeksema F (2004) Adaptive Ad-hoc free band wireless communications. Project plan for AAF, Twente, 19 May 2004
4. Kaushik SS, Deshmukh PR (2009) Comparison of effectiveness of AODV, DSDV and DSDV routing protocols in mobile Ad hoc networks. In J Inform Technol Knowl Manag 2(2):499–502
5. Rahman AHA, Zukarnain ZA (2009) Performance comparison of AODV, DSDV and I-DSDV routing protocols in mobile Ad hoc networks. Eur J Sci Res 31(4):566–576 (ISSN 1450-216X © Euro Journals Publishing, Inc. 2009) (Rahman AHA, Department Of Communication Technology and Network ,University Putra Malaysia, Zukarnain ZA, Department Of Communication Technology and Network Putra Malaysia)
6. Karavetsios RC, Economides AA (2004) Performance comparison of distributed routing algorithms in ad hoc mobile networks. 6th international conference Teleinfo-mobile computing and mobile Networks, 12–14 May 2004, WESEAS (Egnatia Street, Thessaloniki 54006 GREECE)
7. Cheng G, Liu W, Li Y, Cheng W (2007) Department Of Electronics and Information Engineering, 1-4244-0663-3/07/$20.00 ©2007 IEEE

8. Ju S, Evans JB (2010) Scalable cognitive routing protocol for mobile Ad-hoc networks. IEEE communications society subject matter experts for publication in the IEEE Globecom 2010 proceedings (University of Kansas)
9. Ju S, Evans JB (2009) Mobility-aware routing protocol for mobile cognitive networks. Accepted for publication at IEEE CogNet 2009 workshop, Dresden, Germany, June 2009

Chapter 19
Analysis of Image Segmentation Algorithms Using MATLAB

Sumita Verma, Deepika Khare, Ravindra Gupta and Gajendra Singh Chandel

Abstract Image segmentation has played an important role in computer vision especially for human tracking. The result of image segmentation is a set of segments that collectively cover the entire image or a set of contours extracted from the image. Its accuracy but very elusive is very crucial in areas as medical, remote sensing and image retrieval where it may contribute to save, sustain and protect human life. This paper presents the analysis and implementation using MATLAB features and one best result can be selected for any algorithm using the subjective evaluation. We considered the techniques under the following five groups: Edge-based, Clustering-based, Region-based, Threshold-based and Graph-based.

Keywords Image segmentation · N-cut · Mean-shift · Fuzzy-C mean · Image analysis

S. Verma (✉) · D. Khare · R. Gupta · G. S. Chandel
SSSIST, Sehore, Madhya Pradesh
e-mail: avantika08@gmail.com

D. Khare
e-mail: deepika.united@gmail.com

R. Gupta
e-mail: ravindra_p84@rediffmail.com

G. S. Chandel
e-mail: hod.cseit@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_19,

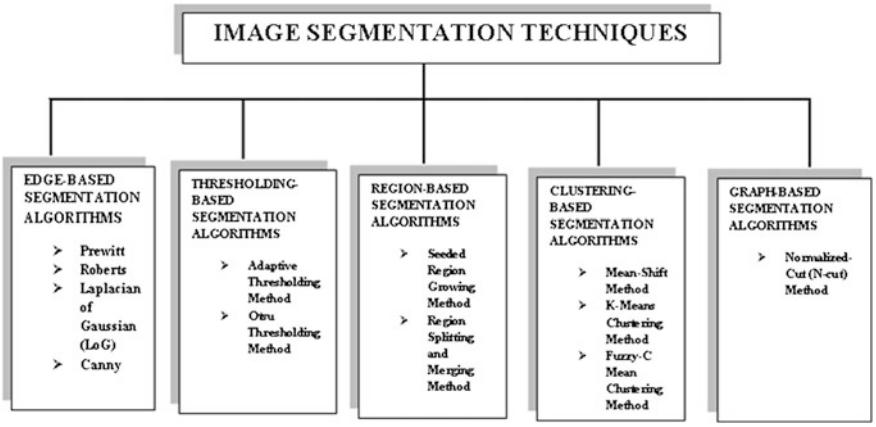


Fig. 19.1 Classification of image segmentation techniques

19.1 Introduction

The main goal of image segmentation is domain independent partitioning of an image into a set of disjoint regions that are visually different, homogeneous and meaningful with respect to some characteristics or computed property(ies), such as grey level, texture or color to enable easy image analysis (object identification, classification and processing). Discontinuity and similarity/homogeneity are two basic properties of the pixels in relation to their local neighborhood used in many segmentation methods. The segmentation methods that are based on discontinuity property of pixels are considered as boundary or edges based techniques and that are based on similarity or homogeneity are region based techniques. Unfortunately, both techniques often fail to produce accurate segmentation results [2].

This paper analyzes the results of various segmentation algorithms, using the subjective evaluation, on the different types of images and particularly on gray level images. This paper will be organized as follows:

- MATLAB
- Segmentation Algorithms and its Results
- Implementation of the proposed system
- Performance Evaluation
- Conclusion.

Figure 19.1 indicates the classification of image segmentation techniques we have considered in this paper. The methods explained and used to segment the image in Figs. 19.4, 19.5, 19.6, 19.7 and 19.8 were used only to clarify the segmentation methods.

Fig. 19.2 Prewitt mask

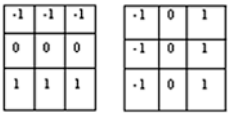
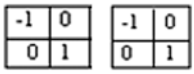


Fig. 19.3 Roberts mask



19.2 MATLAB

Matlab (MATrix LABoratory) is a tool to do numerical computations, display information graphically in 2D and 3D, and solve many other problems in engineering and science. Developed by MathWorks, MATLAB allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs written in other languages, including C, C++, Java, and Fortran [3]. Matlab is an interpreted language for numerical computation. Matlab allows its users to accurately solve problems, produce graphics easily and produce code efficiently [4].

19.3 Segmentation Algorithms and its Results

19.3.1 Edge-Detection Based Segmentation Algorithms

Edge-based segmentation generally indicates the segmentation method based on the edge in an image. The simple methods apply some edge detection methods before segmentatio.

Prewitt Detection: (Fig. 19.2).

Roberts Detection: (Fig. 19.3).

Laplacian of Gaussian (LoG) Detection.

The Laplacian of a Gaussian (LoG) function was defined as:

$$^{2}\mathrm{h(r)} = -\left[\frac{\mathrm{r}^2 - \sigma^2}{\sigma^4}\right]\mathrm{e}^{\mathrm{r}^2/2\sigma^2} \tag{19.1}$$

where, $\mathrm{h(r)} = -\mathrm{e}^{-\mathrm{r}2/2\sigma 2}$, $\nabla^2 = \mathrm{x}^2 + \mathrm{y}^2$, σ = standard deviation, x = direction in x-axis and y = direction in y-axis.

Canny Detection.

The algorithm of Canny detection is:

Step1: Smooth image with a Gaussian.
Step2: Optimizes the trade-off between noise filtering and edge localization.
Step3: Compute the Gradient magnitude using approximations of partial derivatives 2 × 2 filters.

Fig. 19.4 Edge-detection based algorithms and its segmentation results

Fig. 19.5 Thresholding-based algorithms and its segmentation results

Step4: Thin edges by applying non-maxima suppression to the gradient magnitude
Step5: Detect edges by double thresholding (Fig. 19.4).

19.3.2 Thresholding-Based Segmentation Algorithm

Thresholding becomes then a simple but effective tool to separate objects from the background. The output of the thresholding operation is a binary image whose gray level of 0 (black) will indicate a pixel belonging to a print, legend, drawing, or target and a gray level of 1 (white) will indicate the background. Two algorithms are used:

Adaptive Thresholding Method.

In adaptive thresholding, a criterion function is devised that yields some measure of separation between regions. A criterion function is calculated for each intensity and that which maximizes this function is chosen as the threshold.

Otsu Thresholding Method.

Otsu's thresholding method involves iterating through all the possible threshold values and calculating a measure of spread for the pixel levels each side of the

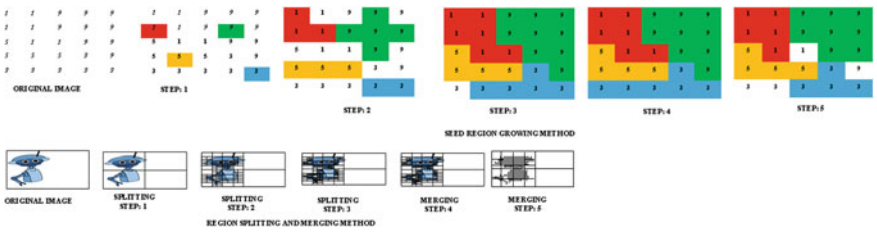


Fig. 19.6 Region-based algorithms and its segmentation results

threshold, i.e. the pixels that either falls in foreground or background. The aim is to find the threshold value where the sum of foreground and background spreads is at its minimum (Fig. 19.5).

19.3.3 Region-Based Segmentation Algorithms

Region-based methods mainly rely on the assumption that the neighboring pixels within one region have similar value. It compares one pixel with its neighbors. If a similarity criterion is satisfied, the pixel can be set belong to the cluster as one or more of its neighbors. Two algorithms are used:

Seeded Region Growing Method.

The algorithm of Seeded Region Growing method is:

Step1: Choose number of seed points which have been clustered into n clusters.
Step2: If the neighboring pixels of the initial seed points are satisfy the criteria such as threshold, they will be grown. The threshold can be intensity, gray level texture, and color…etc.
Step3: Repeat Step2 until all pixels in image have been allocated to a suitable cluster.

Region Splitting and Merging Method.
The algorithm of Region Splitting and Merging method is:

Step1: Splitting steps: For any region Ri, split it into four disjoint quadrants.
Step2: Merging steps: When no further splitting is possible, merge any adjacent regions Rj and Rk.
Step3: Stop only if no further merging is possible (Fig. 19.6).

19.3.4 Clustering-Based Segmentation Methods

Clustering is one of methods widely applied in image segmentation and statistic. The main concept of clustering is to use the centroid to represent each cluster and base on the similarity with the centroid of cluster to classify. Three algorithms are used:

K-Means Clustering Method.

The algorithm of K-Means Clustering method is:

Step1: Decide the numbers of the cluster N and choose randomly N data points (N pixels or image) in the whole image as the N centroids in N clusters.
Step2: Find out nearest centroid of every single data point (pixel or image) and classify the data point into that cluster the centroid located. After doing step 2, all data points are classified in some cluster.
Step3: Calculate the centroid of every cluster.
Step4: Repeat step 2 and step 3 until it is not changed.

Fuzzy C-Means (FCM) Clustering Method.
The algorithm of Fuzzy C-Means (FCM) method is:

Step1: Choose a number of clusters.
Step2: Assign randomly to each point coefficients for being in the clusters.
Step3: Repeat until the algorithm has converged (that is, the coefficients' change between two iterations is no more than the given sensitivity threshold).

Step3.1: Compute the centroids c_k for each cluster which is the mean of all points x (i.e., set of coefficients), weighted by their degree of belonging $w_k(x)$ to the cluster k such that

$$c_{k=} \sum_{x} w_k(x)\ x\ / \sum_{x} w_k(x) \tag{19.2}$$

Step3.2: For each point, compute its coefficients of being in the clusters, using the Eq. 19.2.

Mean Shift Method.
The algorithm of Mean Shift method is:

Step1: Determine the number of clusters we want in the final classified result and set the number as N.
Step2: Classify each pattern to the closest cluster centroid. The closest usually represent the pixel value is similarity, but it still can consider other features.
Step3: Recompute the cluster centroids and then there have N centroids of N clusters as we do after Step1.

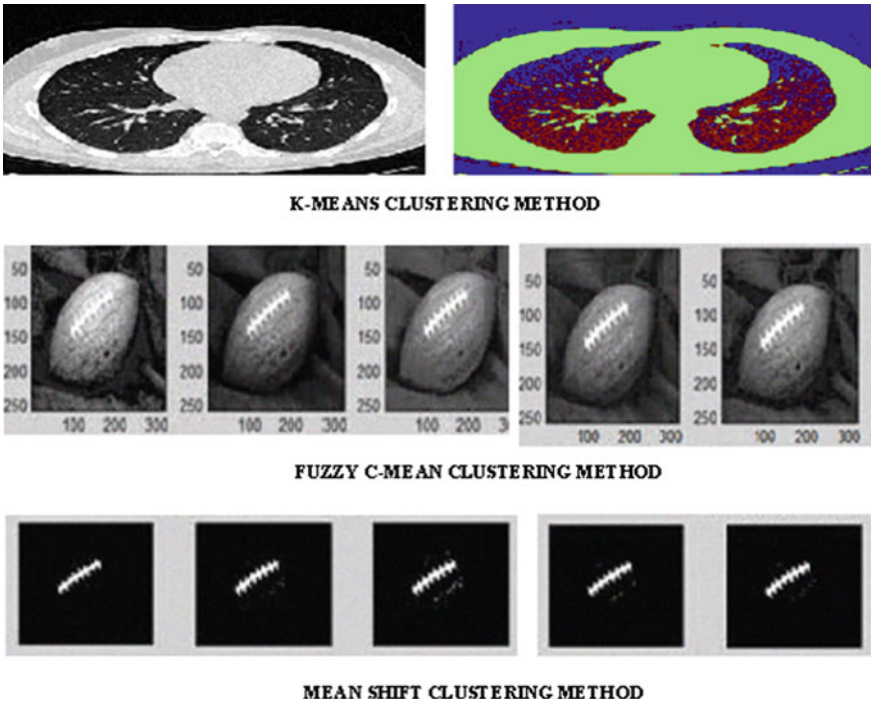


Fig. 19.7 Clustering-based algorithms and its segmentation results

Step4: Repeat the iteration of Step 2 to 3 until a convergence criterion(no reassignment of any pattern from one cluster to another, or the minimal decrease in squared error) is met (Fig. 19.7).

19.3.5 Graph Based Segmentation Algorithm

The graph based image segmentation is based on selecting edges from a graph, where each pixel corresponds to a node in the graph. Weights on each edge measure the dissimilarity between pixels.

Normalized-Cut (N-Cut) Method.
The algorithm of N-Cut method is:

Step1: The input is a graph

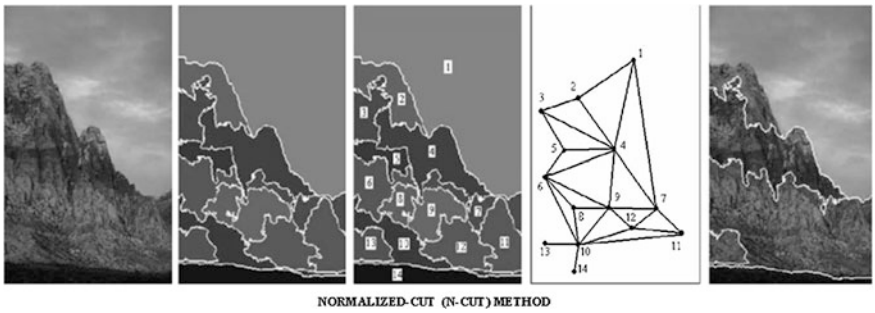


Fig. 19.8 Graph-based algorithms and its segmentation results

$$G = (V, E) \tag{19.3}$$

where V are the n vertices and E are m edges. Each edge has a corresponding weight.

Step2: Perform the segmentation.
Step3: If the weight of the edge connecting two vertices in adjacent components is small compared to the internal difference of both the components, then merge the two components, otherwise do nothing.
Step4: Repeat Step3 (Fig. 19.8).

19.4 Implementation of the Proposed System

Step1: Initially an image is selected and then converted into Binary and Gray level image of size 256 × 256.
Step2: The original and resized image is displayed.
Step3: The value of the concerned parameters, if any, is given.
Step4: The result of each algorithm is displayed in the same figure window
Step5: The segmented image is displayed at a particular position according to the range of the selected value of the parameter.

Above step is repeated with various values of parameters for the same algorithm and the results are obtained in the same figure window. With the help of those results in a figure window, a best segmented Image can be selected on the basis of visual inspection and the value of the parameters for that segmented image can be chosen as a result. Similarly, all the algorithms are applied on an image and for every segmentation algorithm; the result is displayed on same figure window.

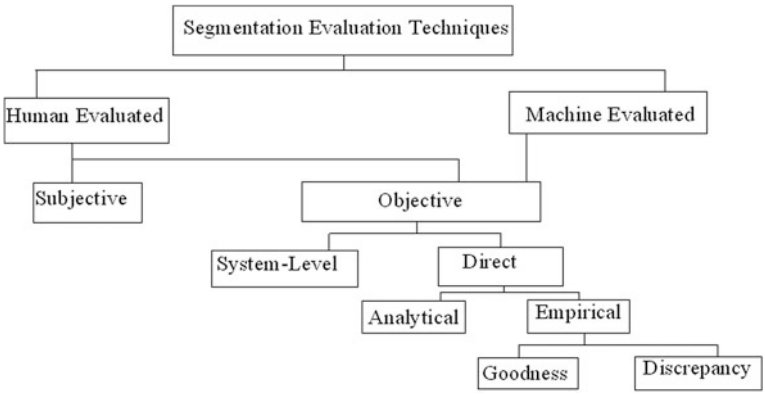


Fig. 19.9 An overview of evaluation techniques

19.5 Performance Evaluation

There have been many image segmentation methods created and being created using many distinct approaches and algorithms but still it is very difficult to assess and compare the performance of these segmentation techniques [5]. Researchers would evaluate their image segmentation techniques by using one or more of the following evaluation methods in Fig. 19.9.

The full description of the above evaluation methods can be found from [6]. Most of these methods ideally should be domain independent but in reality they are domain dependent. Both the subjective and objective evaluation have been used to evaluate segmentation techniques but within a domain dependent environment [6]. These methods have been used to adjust parameters of the segmentation techniques in order to solve the following problems in segmentation area:

- The segmented region might be smaller or larger than the actual
- The edges of the segmented region might not be connected
- Over or under-segmentation of the image (arising of pseudo edges or missing edges).

It is very sad that [7] concluded that there is no segmentation method that is better than the other in all domains.

19.6 Conclusion

After analysis of various image segmentation algorithms and the comparison of the results of each algorithm separately with different parameter's value using MATLAB, the conclusion is that: In Edge Based Segmentation Algorithms, the Canny Algorithm produced the best segmentation in comparison of Roberts,

Prewitt and LoG. In Thresholding Based Algorithms, the Adaptive Thresholding produced the good edges and Otsu Thresholding recognized the object very well. In Region Based Algorithms, the split and merge method produced the better result. In Clustering Based Segmentation, the mean shift method produced the good result. When the K-means and Fuzzy-C means methods are compared, the Fuzzy C-means is better than the K-means method. In Graph Based Algorithms, Normalized-cut is used to cut an image into specified number of cuts. Other methods are interactive methods. In N-cut, the foreground and background area is selected by a user.

To produce a good result with a single technique for the images of all the applications, the further research is required and from the proposed system it can be concluded that the further research should concentrate on such techniques in which the user's interaction is involved so that the segmented result can be improved after automatic segmentation.

References

1. Gonzalez RC, Woods RE (2002) Digital image processing. 2nd Prentice-Hall Inc, Upper Saddle River
2. Shapiro LG, Stockman GC (2001) Computer vision. Prentice-Hall Inc., Upper Saddle River, pp 279–325
3. http://en.wikipedia.org/wiki/MATLAB
4. http://www.math.utah.edu/ ~ eyre/computing/matlab-intro/
5. Zhang H, Fritts JE, Goldman SA (2008) Image segmentation evaluation: a survey of unsupervised methods. Comput Vis Image Underst 10(2):260–280
6. Polak M, Zhang H, Pi M (2009) An evaluation metric for image segmentation of multiple objects. Image Vis Comput 27(8):1223–1227
7. Hu S, Hoffman EA, Reinhardt JM (2001) Automatic lung segmentation for accurate quantization of volumetric X-ray CT images. IEEE 20(6):490–498
8. Zhang YJ (2001) A review of recent evaluation methods for image segmentation. Paper presented at the International Symposium on Signal Processing and its Applications (ISSPA), Kuala Lumpur
9. Udupa JK, Leblanc VR, Zhuge Y, Imielinska C, Schmidt H, Currie LM et al (2006) A framework for evaluating image segmentation algorithms Comput Med Imaging Graph 30:75–87
10. Varshney SS, Rajpal N, Purwar R (2009) Comparative study of image segmentation techniques and object matching using segmentation. Paper presented at the international conference on methods and models in computer science
11. Wang L, He L, Mishra A, Li C (2012) Active contours driven by local Gaussian distribution fitting energy. Signal Process 2(3):737–739
12. Wang Y, Guo Q, Zhu Y (2007) Medical image segmentation based on deformable models and its applications. Springer, p 2
13. Boucheron LE, Harvey NR, Manjunath BS (2007) A quantitative object-level metric for segmentation performance and its application to cell nuclei. Springer, pp 208–219
14. Padmavathi G, Subashini P, Sumi A (2010) Empirical evaluation of suitable segmentation algorithms for IR images. IJCSI Int J Comput Sci 7(4)(2):
15. Mobahi H, Rao SR, Yang AY, Sastry SS, Ma Y. Segmentation of natural images by texture and boundary compression

Chapter 20
Tumor Mass Identification Based on Surface Analysis and Fractal Dimensions

Medhavi Verma and Mahua Bhattacharya

Abstract In present paper we have utilized wavelet transform and fractal dimensions to analyze tumor mass for breast cancer screening using mammogram. Boundary based features from shape of the tumor have taken into consideration as these represent one of the very important property for tumor mass analysis. In present work surface analysis using imaging of tumor mass for analysis of the lesions has been accomplished.

Keywords Wavelet · Fractal dimenisions · Sufrace analysis · Mammogram · Boundary feature

20.1 Introduction

Breast cancer is major from of cancer for women; they have better chance to survival if it is diagnosed early stages with proper treatment, failing to do so can lead to disastrous consequences [1, 2]. With the advancement in computer assisted medical technology and medical imaging system can provide implicit knowledge and information systems for diagnosis [3]. For diagnosis and screening of tumor mass structural Irregularity of tumor boundary is a significant feature. Many experts, researchers and students have studied contour irregularity of tumors [4]. By analysis of boundary irregularity and surface characteristics we can analysis the characteristic of tumor mass. As benign and malignant tumor have different

M. Verma · M. Bhattacharya (✉)
Department of Information Technology, Indian Institute of Information Technology and Management, Gwalior, India
e-mail: bmahua@hotmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_20,

characteristic of appearance as the former is being mostly homogeneous/smooth and the later shows heterogeneous/rougher texture [5]. In this work we have performed tumor mass identification based on surface analysis using fractal dimensions and boundary feature analysis on mammograms. Fractals are irregular and fragmented patterns which could be reduced into size copy. Fractals objects are self-similar under some change in scale, either strictly or statistically [6]. Fractal analysis has many applications; one such application is estimation of surface roughness. Fractals analysis is very useful in medical image analysis since Medical images typically have a degree of randomness associated with the natural random nature of structure [7].

20.2 Literature Review

Various studies have been done on tumor mass analysis. In [8] Author has provided an automated diagnostic procedure for breast cancer using multiresolution and FCM based clustering algorithm with further region growing approach is applied for considerably large tumor calcifications. Authors have applied wavelet transform and classified tumor using SVM and RBF function as a kernel function to assist ultrasound diagnosis of solid breast tumors in which the contour feature quantified by boundary based corner counts [9]. Localized texture analysis is an important feature breast tissue mass. Fractal analysis of extracted textural features is done and use of advance classifier is done. Results are compared with one used by clinical radiologist [10]. Regression-line analysis and multiscale filtering of mammographic images is reported in [11] where image analysis is done at different scale level. Tumors detection is done on finest resolution on the abstracted plane. After detection regressions line analysis of part of mammary glands is done which estimates the degree of concentration by the analysis of average minimal distance to the concentration point. Study of variation of density with the help of edge characteristics of region of interest (ROI) by studying the sharpness of the boundaries of tumors to discriminate between benign and malignant tumors also they have studied various shape features such as compactness, Fourier descriptors, moments, and chord-length statistics to distinguish between the tumors [12]. Analysis of mammograph for the suspicious location of the breast masses is done with the help of shape based feature. Fourier descriptors are used for shape feature extraction and for segmentation of ROI fuzzy C-means clustering technique is used [1]. Fractal Analysis with direction analysis is used for feature extraction where fractal analysis is performed in frequency domain and Classification is done using ANN [13, 14].

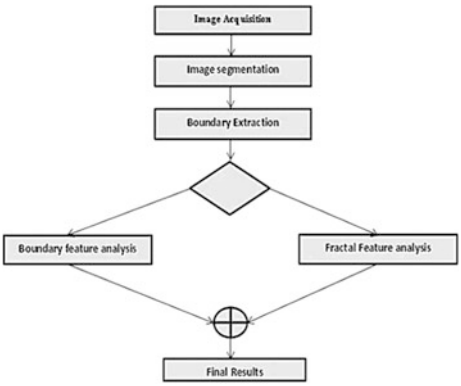


Fig. 20.1 The flowchart of methodology

20.3 Methodology

Here we have used Fractal dimensions (FD) for surface analysis with boundary feature analysis for better results (Fig. 20.1).

20.3.1 Boundary Extraction

For boundary feature extraction the image is converted into grayscale image and then into a binary image with the help of global threshold where pixel values are set 0 for intensity value less than threshold and 1 for more than threshold value. From the binary image the boundary is extracted using 8 pixel connected component analysis.

20.3.2 Wavelets

For surface analysis we have first applied wavelet analysis of image which we have derived from pervious steps. Wavelet transformation (WT) is mathematical function through which one can divide a continuous time signal into various different scale components [15]. Wavelet decomposition gives us more accurate information of physical properties such as roughness, discontinuities and can be analysed accurately [16]. We have used Daubechies Wavelet filter. Wavelet function is shown in Eq. 20.1.

$$\psi_{x,y}(t) = \frac{1}{\sqrt{x}}\psi\left(\frac{t-x}{y}\right), \quad x, y \in R, \quad x \neq 0, \tag{20.1}$$

where 'x' is scaling parameter 'y' a translation parameter which determines the location of the wavelet, and t is the space variable. ψ(t) is the parent wavelet. Discrete wavelet transform (DWT) is shown Eq. 20.2 which used form wavelet transforms here

$$C_f(x,y) = \int_{-\infty}^{+\infty} f(t)\overline{\psi_{x,y}(t)}\mathrm{d}t \tag{20.2}$$

Multi resolution signal decomposition (MSRD) algorithm is used to for DWT In MSRD the orthogonal wavelets filters are applied on the original signal, the signal is split into high frequency D_n^i components (details) and low frequency C_n^i components (details) as shown in Eqs. 20.3 and 20.4 [17].

$$D_n^i = \sum_{i=n-N}^{n} D_i^{k-1} h_{j-n} \tag{20.3}$$

$$C_n^i = \sum_{i=n-N}^{n} C_i^{k-1} g_{j-n}, \quad i = 1,2,3,\ldots N \tag{20.4}$$

N is the sampling number, n is the sampling position, k is the scale level. Low pass and high pass filter are h and g respectively. They are written in relation to the wavelet function as

$$A_j = \int_{-\infty}^{\infty} 2^{-1}\psi\left(2^{-1}x\right)\phi(x-j)dx \tag{20.5}$$

where ϕ is the scaling of the corresponding wavelet function ψ. High frequency components obtained for fractal analysis.

20.3.3 Fractal Analysis

Fractal analysis is a useful tool for surface roughness. Most of the analysis of fractals is done in terms of fractal dimension. Fractal dimensions are an important quantitative factor which helps us to characterize the degree of surface roughness [16, 18, 19]. The image after 2D DWT results in 2D sub–image after decompositions. By calculating the slope of plot of graph of log magnitude versus log of frequency we estimated FD.

$$S(x) \propto x^{-a} \tag{20.6}$$

$$a = 8 + 2D \tag{20.7}$$

Here D denotes FD. To calculate power spectral density (PSD) we convert 2D sub image into frequency domain using Fourier transform Power spectral density

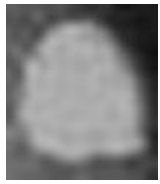


Fig. 20.2 Orginal mammogram

of 2D of an M_xXM_y image can be calculated with 2D FFT the Eq. 20.8 as PSD is calculated using 20.9.

$$f(u,v) = \sum_{x=0}^{M_{x-1}} \sum_{y=0}^{M_{y-1}} z(x,y) e^{-j2\pi\left(\frac{ux}{M_x} + vy/M_y\right)} \tag{20.8}$$

$$S(k) = |F(u.v)|^2 \tag{20.9}$$

Thus by calculating the slope we can compute the fractal dimension which gives an idea of surface roughness. RMS value of the slope is taken to give more ideal result.

20.3.4 Centroid Calculation of the Boundary and Distance Variation

We have calculated centroid of tumor mass form extracted boundary using connected component analysis. The distance of each pixel of the boundary to the centroid is calculated using the formula in Eq. 20.10

$$dis = \sqrt{(x_c - x_b)^2 + (y_c - y_b)^2} \tag{20.10}$$

where x_c, y_c represents location of the Centroid and x_b, y_b is location of the boundary form where the distance is calculated. Slope of the plot gives us variation in distance.

20.4 Results

Simulations were carried out on more than 30 images out of which most significant 16 are considered. Here we have results of both fractal and boundary feature analysis.

From Figs. 20.2, 20.3, 20.4 show the extraction of boundary form mammogram Fig. 20.5 is the centroid of the tumor Fig. 20.6 is the plot of Distance of centroid, slope gives us variation.

Fig. 20.3 Binary image

Fig. 20.4 Extracted boundary

Fig. 20.5 Calculated centorid

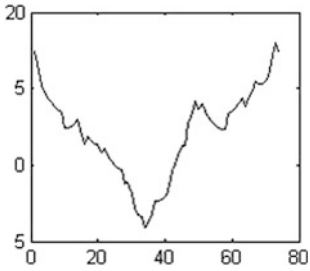


Fig. 20.6 Plot of distance vartion w.r.t centorid

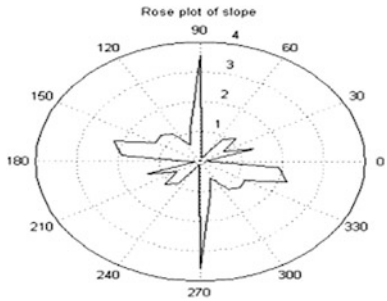


Fig. 20.7 Plot of slope

Figures 20.7 and 20.8 is the plot of slope and plot of intercept. By computing RMS value of slope we get surface roughness.

RMS surface roughness of the tumor masses is shown in Table 20.1. As the surface roughness increases as we move form benign stage to malignant stage.

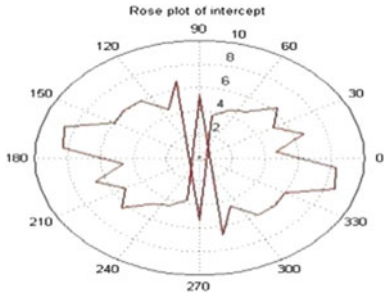


Fig. 20.8 Plot of intercept respectively

Table 20.1 RMS surface roughness

S.NO 1–8	2.0037	3.7298	4.6087	1.8712	1.9218	3.9906	2.0101	2.2183
9–16	3.8664	3.2357	2.9511	3.3356	3.9262	3.6446	3.078	3.8179

Table 20.2 Classification criteria

Surface roughness	Centroid distance variation	Stage of tumor
High	High	Malignant
High	Low	Tendency malignant
Low	Low	Benign

Table show the analysis outcome of rate of change of boundary distance with surface roughness (Table 20.2).

20.5 Conclusion

Results of the simulation we observe surface analysis using FD with boundary based feature shows how the variation surface roughness can be used to classify tumor.

References

1. Bhattacharya M, Das A (2009) Soft computing based decision making approach for tumor mass identification in mammogram. Int J Bioinformatics Res 1(2):37–46. ISSN: 0975–3087
2. Sankar D, Thomas T (2009) Analysis of mammograms using fractal features. India nature & biologically inspired computing, 2009. NaBIC 2009. World congress on issue date: 9–11 Dec 2009, pp 936–941
3. Saritha S, Santhosh Kumar G (2011) Interestingness analysis of semantic association mining in medical images. Commun Comput Inf Sci 157:1–10

4. Qin B, Ma L, Xu W (2010) Comparative study on boundary structural irregularity using local FD and curvature analysis. Bioinformatics and Biomedical Engineering (iCBBE), 2010 4th international conference on 18–20 June 2010, pp 1–4
5. Mu T, Nandi AK, Rangayyan RM (2008) Classification of breast masses using selected shape, edge-sharpness, and texture features with linear and kernel-based classifiers. J Digit Imaging 21(2):153–169
6. Jampala S (1992) Fractals: classification, generation and applications. Circuits and systems, 1992, Proceedings of the 35th Midwest symposium, pp 1024–1027
7. Iftekharuddin KM, Jia W, Marsh R (2003) Fractal analysis of tumor in brain MR images. Machine vision and applications. Springer, Heidelberg, pp 352–362
8. Bhattacharya M, Das A (2010) Identification of tiny and large calcification in breast: a study on mammographic image analysis. Int J Bioinformatics Res Appl 6(4):418–434
9. Zuo Y, Lin J, Chen K, Peng Y (2009) Boundary-based feature extraction and recognition of breast tumors using support vector machine. 2009 International forum on information technology and application, pp 89–92
10. Mavroforakis ME, Georgiou HV, Dimitropoulos N, Cavouras D, Theodoridis S (2006) Mammographic masses characterization based on localized texture and dataset fractal analysis using linear, neural and support vector machine classifiers. Artif Intell Med 37:145–162
11. Hirano S, Tsumoto S (2008) A method for detecting suspicious regions in mammograms based on multiscale image filtering and regression-line analysis. Automation congress, 2008. WAC 2008. World 09 Dec 2008, pp 1–6
12. Rangayyan RM, El-faramawy NM, Leo Desautels JE, Alim OA (1997) Measures of acutance and shape for classification of breast tumors. IEEE Trans Med Imaging 16(6):799–810
13. Tourassi GD, Eltonsy NH, Graham JH, Floyd CE, Elmaghraby AS (2005) Feature and knowledge based analysis for reduction of false positives in the computerized detection of masses in screening mammography. 2005 IEEE engineering in medicine and biology 27th annual conference Shanghai, China, 1–4 Sept 2005, pp 6524–6527
14. Chung D, Revathy K, Choi E, Min D (2009) A neural network approach to mammogram image classification using fractal features. Intelligent computing and intelligent systems, 2009. ICIS 2009, pp 444–447
15. http://en.wikipedia.org/wiki/Wavelet#Wavelet_transforms
16. Lin Y, Xiao XR, Li XP, Zhou XW (2005) Wavelet analysis of the surface morphologic of nanocrystalline TiO2 thin films. Surf Sci 579:37–46
17. Mallat SG (1989) A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans Pattern Anal Mach Intell 11(5):674–693
18. Borodich FM, Onishchenko DA (1999) Similarity and fractality in the modelling of roughness by a multilevel profile with hierarchical structure. Int J Solids Struct 36:2585–2612
19. Dubuc B, Quiniou J, Roques-Carmes C, Tricot C, Zucker SW (1989) Evaluating the fractal dimension of profiles. Phys Rev A 39:1500–1512

Chapter 21
Mean-Shift Algorithm: Verilog HDL Approach

Rahul V. Shah, Amit Jain, Rutul B. Bhatt, Pinal Engineer and Ekata Mehul

Abstract The abstract should summarize the contents of the paper and should Object tracking algorithms, when it comes to implementing it on hardware ASIC, it becomes difficult task, due to certain limitations in hardware. This paper shows how mean- shift algorithm is implemented in HDL along with the description of ports and interfaces.

Keywords Object tracking · Complexity in hardware ASIC · Mean Shift algorithm · Histogram · Bhattacharya coefficient

R. V. Shah (✉) · A. Jain · E. Mehul
ASIC Department, Einfochips, Ahmedabad, India
e-mail: rahul.shah@einfochips.com

A. Jain
e-mail: amit.jain@einfochips.com

E. Mehul
e-mail: ekata.mehul@einfochips.com

P. Engineer
Electronic & Communication Department, SVNIT, Surat, India
e-mail: pje@svnit.ac.in

R. B. Bhatt
Electronic & Communication Department, SVIT,
Gujarat Technology University, Vasad, India
e-mail: er.rbb10@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_21,

21.1 Introduction

Object tracking in its simplest form can be defined as the problem of estimating the trajectory of an object in an image plane as it moves around the scene. In other words the tracker assigns consistent labels to the tracked objects in different form of videos.

Object tracking is the important part in computer vision technology, but till time it is mainly dealt with the high grade softwares like MATLAB, ADOBE etc. But when it comes to implement it on ASIC chips it becomes difficult when many mathematical operations are mainly operated through adders and subtracters and not with multipliers and dividers. These object tracking algorithms when implemented on ASIC covers good amount of chip area as many such mathematical operations are involved in it, which are complex to implement with adders and subtracters. Mean shift algorithm when implemented on HDL it become really useful to check the functions of the object tracking. The Mean shift algorithm requires Histogram, Bhattacharya coefficient, centre calculation which are implemented with small mathematical operations like multipliers, Dividers and square root etc. That makes algorithm to implement on HDL easy.

21.2 Object Tracking Controller

The Object Tracking algorithm is designed for tracking of objects in video or in real-time. It is based on Mean shift algorithm. Interface with external video source i.e. camera for input video, HOST interface for providing inputs, interface with external memory for storing video frames and interface with monitor for video output are some of the interfaces required for object tracking. The input video interface is used to provide video input to the algorithm. In this algorithm, we are using 24-bit RGB format for input frames. After getting video from the source, it is passed on to the processing unit as well as external memory. Host Interface provides inputs to the processing units for performing the operations on the input video. Processing unit consists of different functional sub-blocks like histogram, Kernel, Mean shift, Bhattacharyya Coefficient etc. After completion of the operations, outputs are passed on to the monitor for display.

External Memory interface is used for storing video frames depending upon the clock frequency and video speed. Since video is continuous, so it is required to save frames in memory and simultaneously perform processing on continuous flowing frames. Output interface with monitor is required to show the output video frames after doing the operations on them and verify the result. Figure 21.1 shows the detailed Processing Unit diagram with different sub-blocks like histogram, Mean shift, rect_box, Bhattacharyya coefficient. According to functionality, video frames are stored in the memory. The number of frames to be stored depends on clock frequency and video speed. According to inputs **ot_centerx_i, ot_centery_i,**

ot_hx_i, ot_hy_i, and **ot_bins_i** reference histogram **(ot_qu)** is calculated for first frame. For finding reference histogram **(ot_qu)** for first frame, cropping is done on the image and kernel value is calculated for that. Then depending upon the R, G, B values hist_model is calculated.

Now to track object in next frames, reference histogram **(ot_qu)** and other inputs are given to Mean-Shift block. In Mean-Shift block, histogram **(ot_1pu)** is calculated for second frame using old center values **(ot_centerx_i, ot_centery_i)** and then Bhattacharyya coefficient **(ot_1rho)** is calculated using **ot_qu** and **ot_1pu**. Then weighted array **(ot_w)** and normalized row and column arrays **(ot_1pl and ot_2pl)** are calculated for second frame. Then by using weighted array and normalized row and column arrays, new center values **(ot_new_centerx and ot_new_centery)** for second frame are calculated. The reason for calculating new center values is that we assume that object has displaced from its original position as compared to frame 1. Then new histogram **(ot_2pu)** using new center values and new Bhattacharyya coefficient **(ot_2rho)** using **ot_qu and ot_2pu** are calculated. **Iteration loop** depends upon comparison results of **ot_1rho** and **ot_2rho.** According to iteration loop, final center value for 2nd frame is obtained. With final **center values**, **ot_hx_i** and **ot_hy_i,** rectangular box is created around the targeted object in second frame. The iteration loop is looped 20 times, which is maximum value for this algorithm. Iteration is done because at first instant it doesn't give final center for current frame and moreover, object is displaced in next consecutive frames but to know how much it moved and in which direction it moved, iteration value is taken. For all other frames, this process is repeated in Mean-shift block and object is tracked in rest of the frames.

21.3 Histogram

The histogram is an important term in this algorithm. It is a combination of R, G, B and Spatial information. The inputs of rect_crop are images, ot_centerx, ot_centery, ot_hx, ot_hy. Binwidth is calculated by dividing 256 by bins. The image_model is generated by dividing pixel values by binwidth to give output image_model in normalized format. Kernel value is calculated using these inputs. Histogram array of size defined by bin value is initialized with zeros.

R, G, B values of normalized image_model are calculated using I and j logic. Then these R, G, B values are used for finding locations in histogram array where squared kernel value is added. Then this squared Kernel value is added to constant value "c". The histogram is divided by the final value of constant "c" to get normalized histogram (Fig. 21.2).

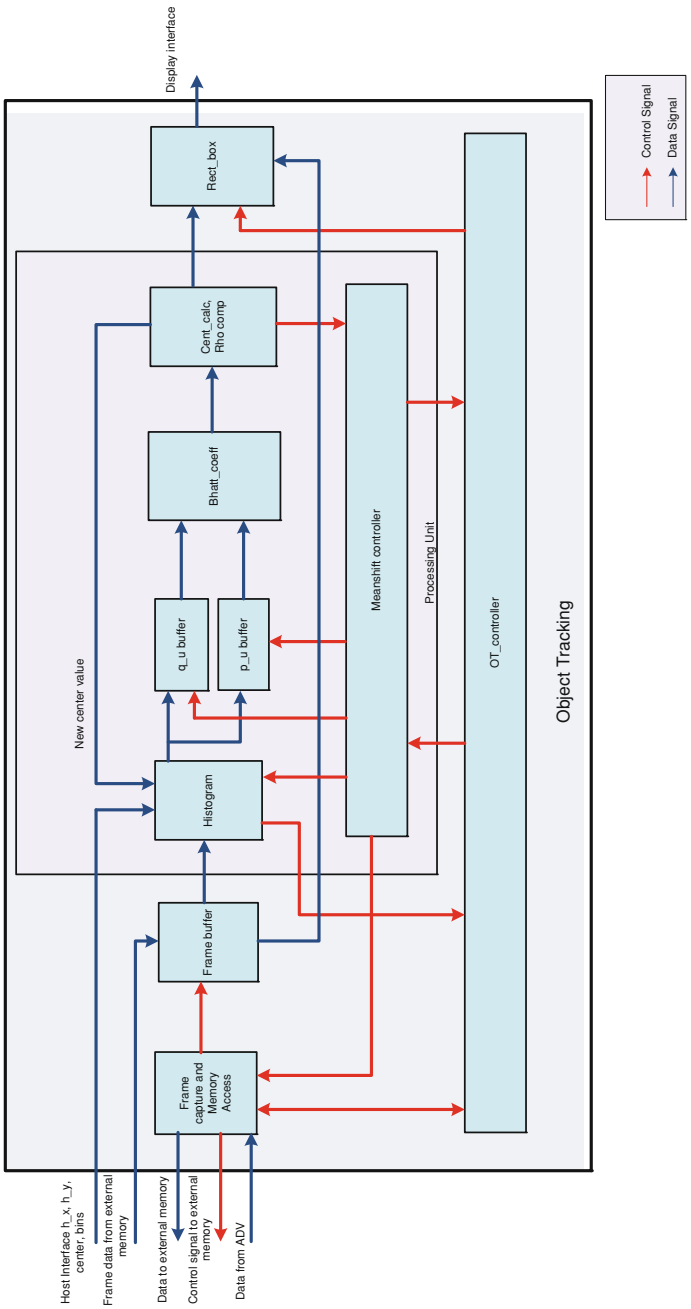


Fig. 21.1 Block diagram of algorithm

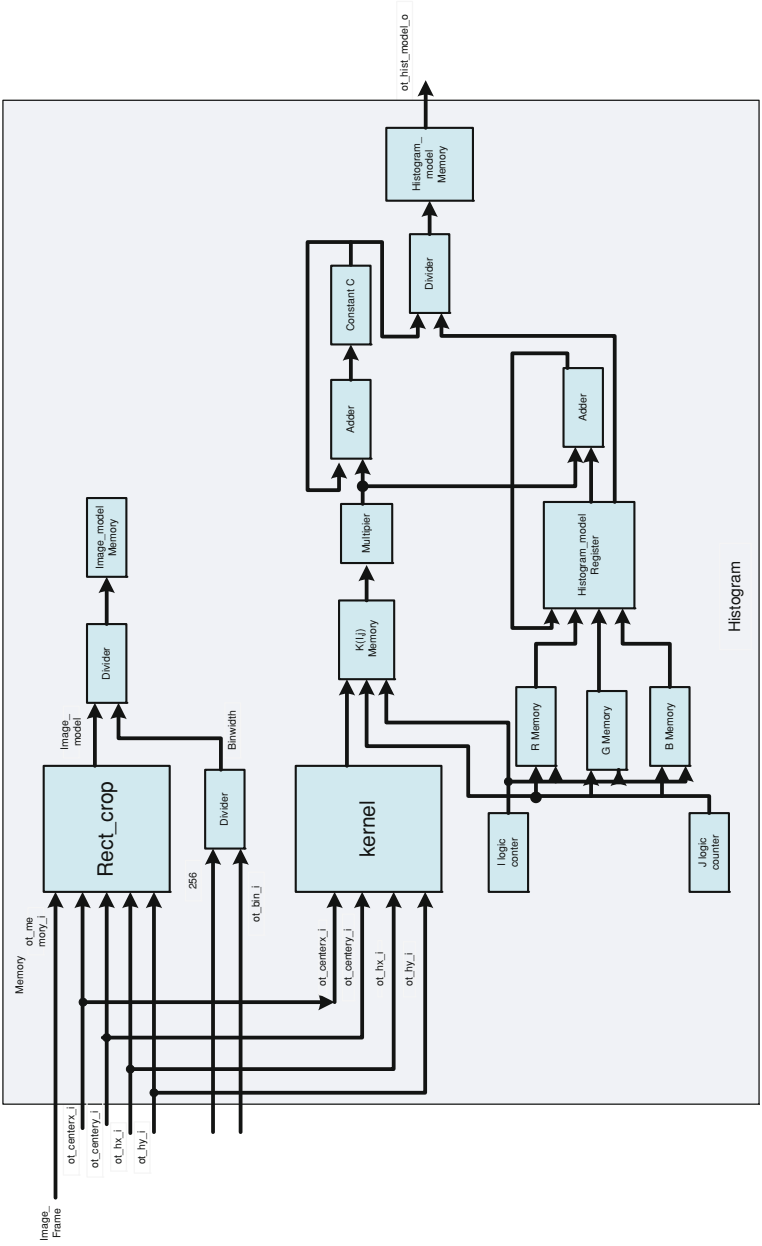


Fig. 21.2 Histogram block diagram

21.4 Bhattacharya Co-Efficient

According to i, j, k logic reference histogram ot_qu(i,j,k) and new histogram (ot_pu(i,j,k) are multiplied. Then the square root of result is added to the initialized rho value.

Bhattacharyya coefficient measures the similarity between two histograms. To calculate Bhattacharyya Coefficient, histograms of the reference image and histogram of current frame is required (Fig. 21.3).

21.5 Center Calculation and Rho Logic

21.5.1 Centre Calculation

The inputs of Center Calculation Block are ot_image_model_size1, ot_image_-model_size2, ot_old_centerx, ot_old_centery, ot_hx, ot_hy and ot_w. According to i and j values, which are dependent on ot_image_model_size1 & ot_image_-model_size2, all values of normalized row and column arrays are multiplied by weighted array values and tempx and tempy variables are calculated. Weighted array values are also added to temp_t variable.

At the end of the loop, temp_x and temp_y values are divided by temp_t variable and the final result is stored in temp_x and temp_y. For calculating new center values, temp_x and temp_y are multiplied by ot_hx and ot_hy and added to ot_old_centerx and ot_old_centery respectively. Then the result is stored in ot_center1 & ot_center2 (Fig. 21.4).

21.5.2 Rho Logic

The inputs of Rho Logic Block are ot_pu1, ot_pu2, ot_rho1, o_rho2, ot_centerx, ot_centery, ot_old_centerx, ot_old_centery. Bhattacharyya coefficient of 2nd frame for old center (Ot_rho1) and Bhattacharyya Coefficient of 2nd frame for new center (ot_rho2) are compared. If ot_rho1 is greater than ot_rho2, then ot_old_center is final center for this frame and iteration loop is broken. If ot_rho2 is greater then ot_rho1, then sum of squares of ot_center1—ot_old_center1 and ot_center2—ot_old_center2 is calculated. If result is less then 1, then final center is center value and iteration loop is broken. Both conditions are not satisfied then ot_pu2, ot_rho2 and center (ot_centerx & ot_centery) are feedback to iteration loop.

The iteration value is looped 20 times, which is maximum value for this algorithm. Iteration is done because at first instant it did not give final center for that frame and moreover, object is displaced in next consecutive frames but to know how much it moved and in which direction it moved, iteration value is taken (Fig. 21.5).

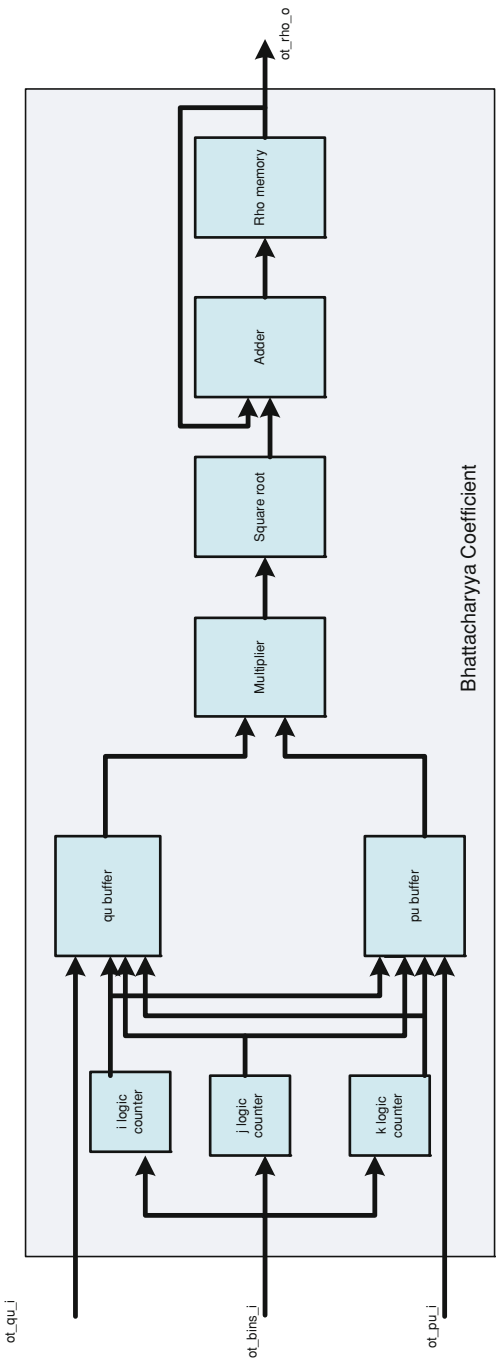


Fig. 21.3 Bhattacharya co-efficient

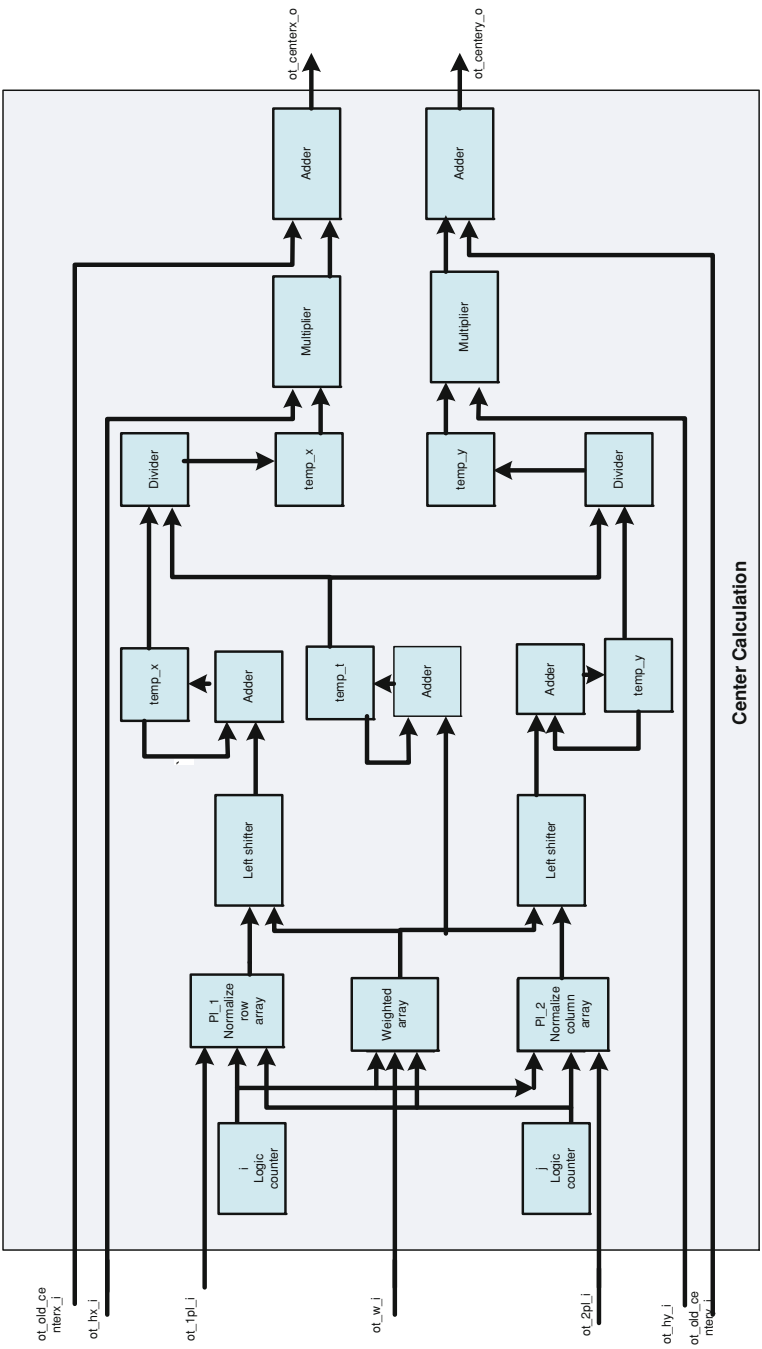


Fig. 21.4 Center calculation

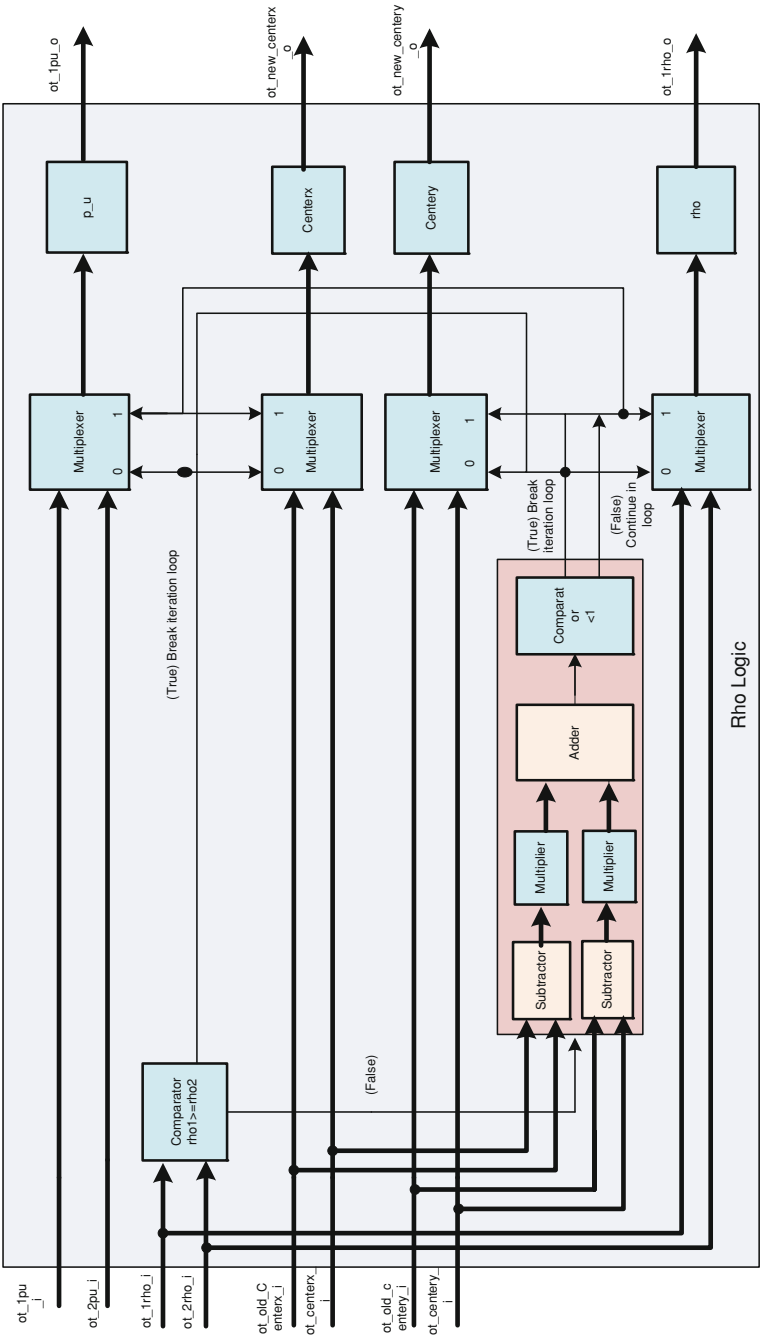


Fig. 21.5 Rho logic

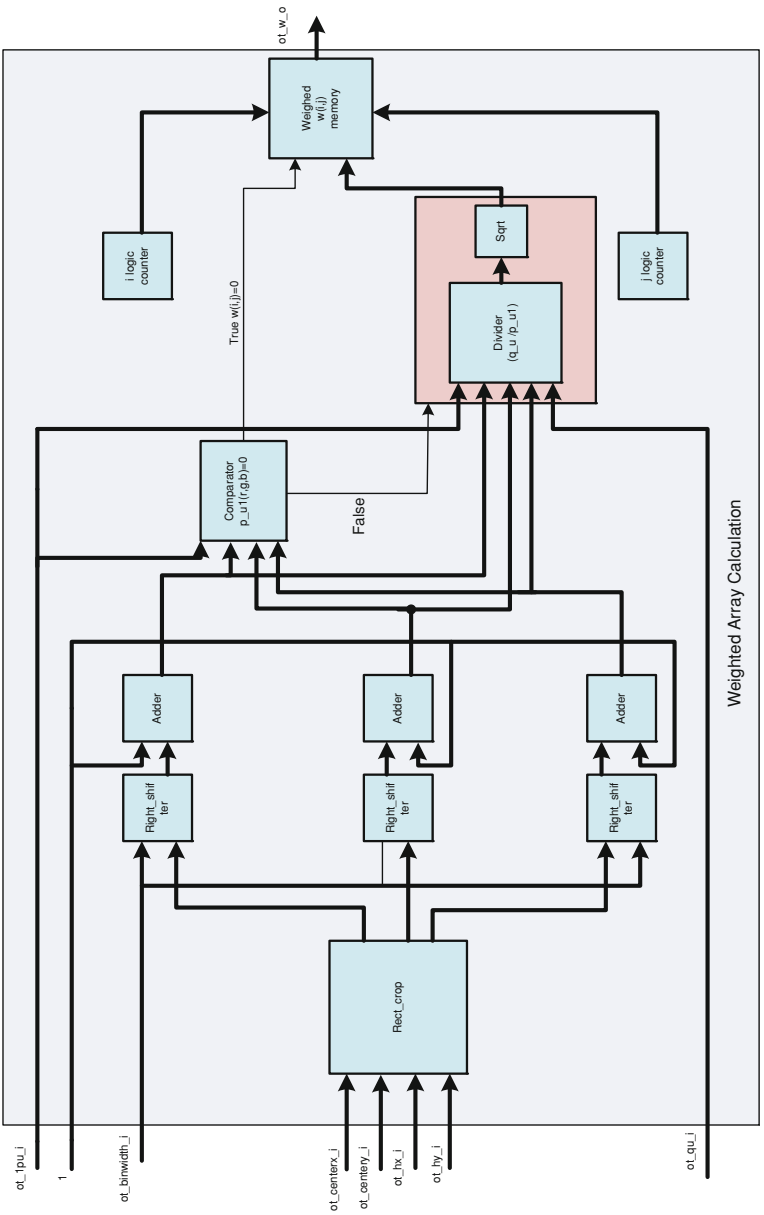


Fig. 21.6 Weighted array

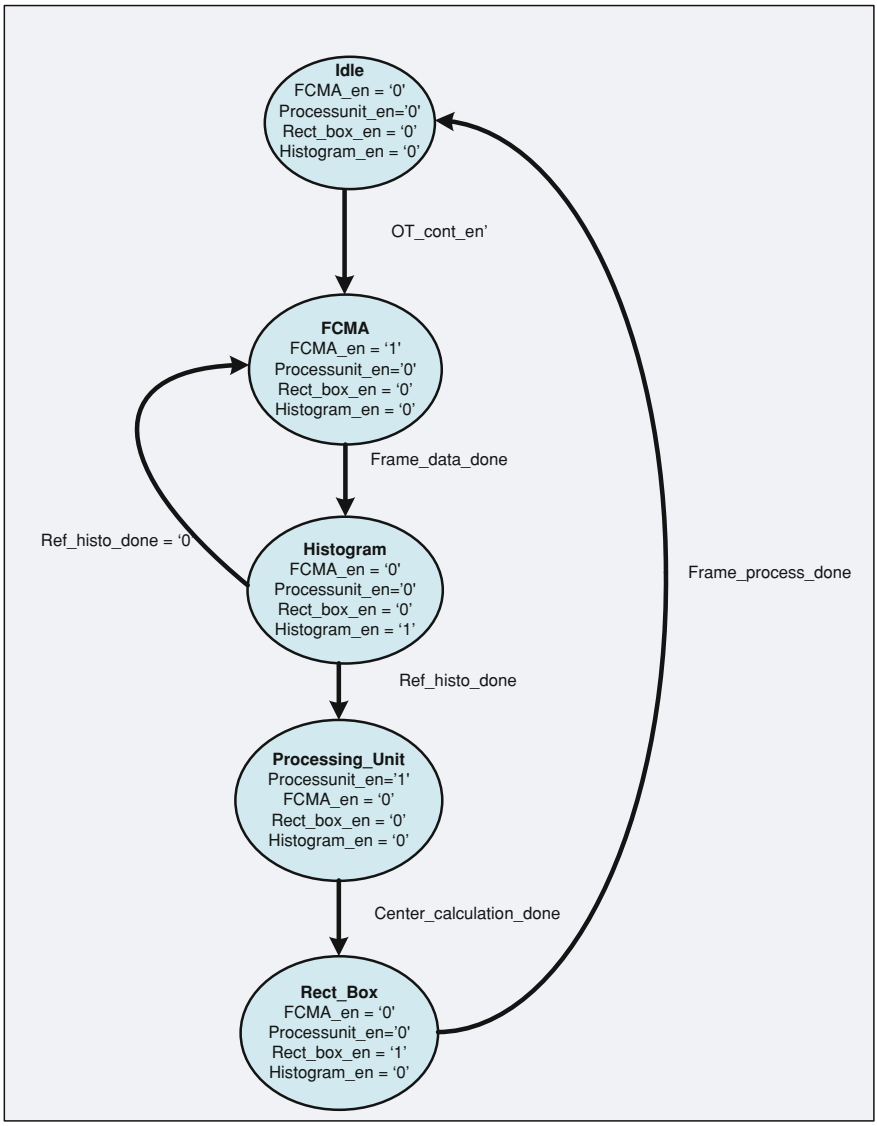


Fig. 21.7 State diagram of OT controller

21.5.3 Weighted Array

The inputs of weighted array calculation are ot_centerx, ot_centery, width (ot_hx), height (ot_hy), ot_bins, histogram of frames (ot_pu1) and reference Histogram (ot_qu). Image model from R, G, B values is obtained by using Rect_crop. Then, binwidth is calculated by dividing 256 by bins value, which is a input value. For

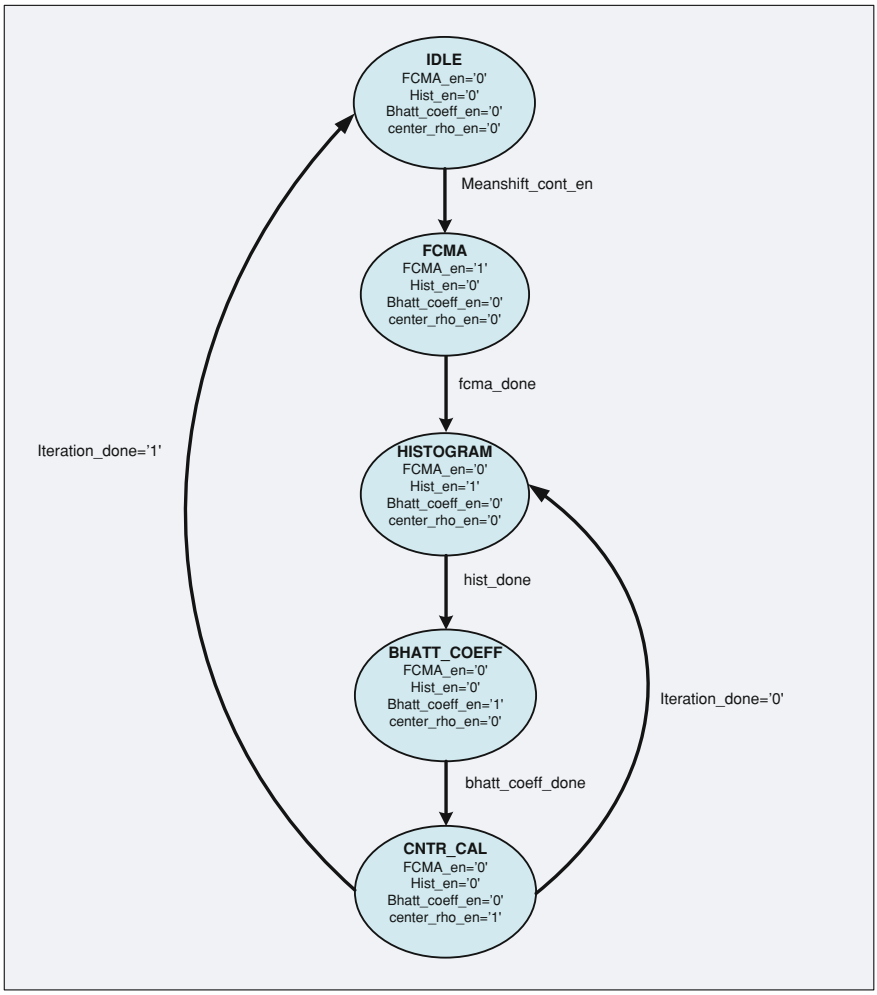


Fig. 21.8 State diagram of mean shift algorithm

normalization of R, G, B values, divide them by binwidth. By doing so, R, G, B values come into limit of bins value and decrease calculation cost.

Check histogram of frames (ot_pu1) for R, G, B values. If it is zero, then weighted array value is also zero. Otherwise, Weighted array value is calculated by taking square root of division result of reference histogram (ot_qu) by histogram of frame (ot_1_pu) (Figs. 21.6, 21.7).

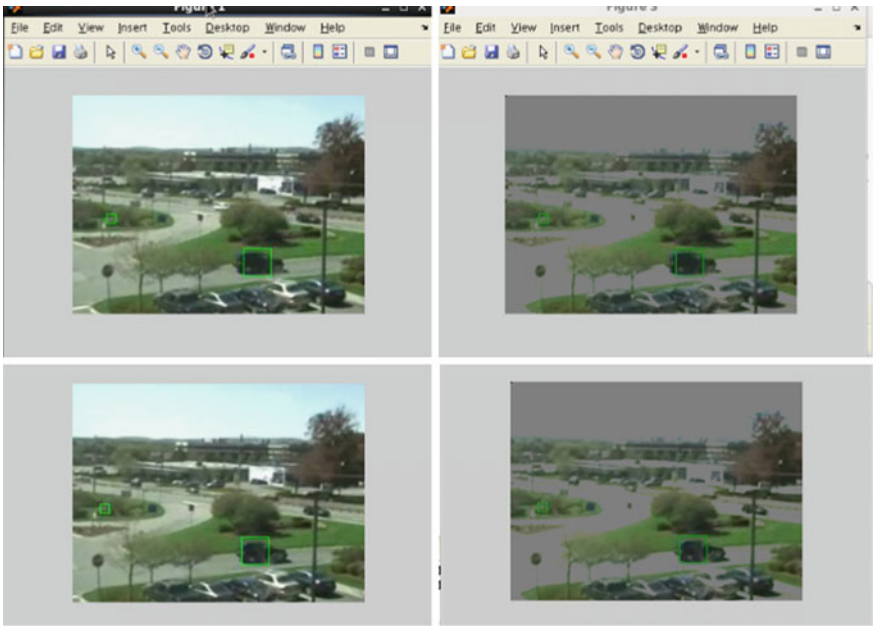


Fig. 21.9 Comparison of outputs

21.6 State Representation

The OT controller only deals with the FCMA, histogram, Processing unit(internal logic after histogram calculation) and Rectangular box. When OT controller is enabled the it sets the FCMA and the image from the series of images is loded into it which now FCMA checks the status of histogram if enabled the histogram of the image is calculated after the calculation of histogram the FCMA sends other image and the histogram of that image is also calculated and this is compared and through processing unit and center and Rho value is calculated after itrating it to to get good output the rectangular box is made based on the kernel needed and set during the histogram calculation (Fig. 21.8).

In mean shift algorithm state diagram the internal processing unit is explained in much detail after the calculation of histogram the the values are stored in buffer and then a type of comparison that is calculation of bhattacharya coefficient is done this bhattacharya coefficient act as the threshold for the image and based on this center is calculated. This process is iterated almost twenty times to get the perfect output of where the object is shifted and how it moves. When the comparison is done it again passes the control to the FCMA to load the next image and the comparison is done with the new image loaded.

21.7 Conclusion and Results

Considering all the above descriptions the Mean shift algorithm when implemented using verilog HDL the real problem was to build the HDL for multiplication, division and square-root through verilog the results which were compared through MATLAB was really proving the working of Mean-shift algorithm. The difference was only that the MATLAB output was with using floating point and verilog output was using fixed point (Fig. 21.9).

References

1. Yilmaz A, Javed O, Shah M (2006) Object tracking: a survey. ACM computing surveys, vol 38(4), Article 13, Publication date: Dec 2006
2. Ballard D, Brown C (1982) Computer vision. Prentice-Hall, Englewood Cliffs
3. Comanciu D, Ramesh V, Meer P (2003) Kernel-based object tracking. IEEE Trans Patt Anal Mach Intell 25:564–575
4. Wang J, Adelson E (1994) Representing moving images with layers. IEEE Image Process 3(5):625–638
5. Xu N, Ahuja N (2002) Object contour tracking using graph cuts based active contours. In: IEEE international conference on image processing (ICIP), pp 277–280
6. Palnitkar S (2003) Verilog HDL: a guide to digital design and synthesis. Sun Microsystems, Inc., Palo Alto
7. eInfochips. Design Specification document for the Object Tracking

Chapter 22
Optimal Node Selection Using Estimated Data Accuracy Model in Wireless Sensor Networks

Jyotirmoy Karjee and H. S. Jamadagni

Abstract One of the major tasks of wireless sensor network is to sense accurate data from the physical environment. Hence in this paper, we propose a new methodology called Estimated Data Accuracy Model (EDAM) for randomly deployed sensor nodes which can sense more accurate data from the physical environment. We compare our results with other information accuracy models which show that EDAM performs better than the other models. Moreover we simulate EDAM under such situation where some of the sensor nodes become malicious due to extreme physical environment. Finally using our propose model, we construct a probabilistic approach for selecting an optimal set of sensor nodes from the randomly deployed maximal set of sensor nodes in the network.

Keywords Data accuracy · Spatial correlation · Optimal sensor nodes · Wireless sensor networks

22.1 Introduction

Recent progress in wireless technology has made a drastic improvement over wireless sensor networks. In wireless sensor networks, nodes are deployed in the sensing region to sense the physical phenomenon of data for the event like seismic event, fire, temperature, humidity etc. from the environment [1] and transmit the

J. Karjee (✉) · H. S. Jamadagni
Department of Electronic Systems Engineering, Indian Institute of Science, Bangalore, India
e-mail: kjyotirmoy@cedt.iisc.ernet.in

H. S. Jamadagni
e-mail: hsjam@cedt.iisc.ernet.in

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_22,

data to the sink node. Data collected by the sensor nodes are generally spatially correlated [2] among them. These spatially correlated data sensed by the sensor nodes are directly transmitted to the sink node which estimates [3–6] the data (information) accuracy. In the literature [7–10], authors discuss the data accuracy under distributed conditions. In this paper, we take the same scenario discussed in [10] where we deploy sensor nodes randomly in the sensor region.

The main motivation of this paper is to develop a new methodology called Estimated Data Accuracy Model (EDAM) which can sense more accurate data from the physical environment and compare our results with the other information accuracy (distortion) models [4–6]. Finally a probabilistic model is proposed to select an optimal set of sensor nodes to sense the data from the randomly deployed maximal set of sensor nodes in the network using data accuracy function [7]. In literature [11], maximizing the network life time subjected to event constraint and in literature [12], total information gathered subjected to energy constraints are discussed without verifying the information accuracy. Hence gathering information without verifying the accuracy level cause problem if some of the sensor nodes get malicious [13]. If the sensor nodes get malicious, it can read inaccurate data. If the inaccurate data gets aggregated with the other correct data sensed by the sensor nodes, it causes incorrect data aggregation at the sink node. Hence sink node estimate the incorrect data reading for the network. Therefore it is essential to verify the data accuracy before data aggregation discussed in [7–10]. However to the best understanding of authors, this is the first time we perform results for data accuracy where some of the sensor nodes behaves live malicious nodes due to extreme physical environment such as heavy rain fall, heavy snow fall in the hilly region etc.

The rest of the paper is given as follows. Section 22.2 demonstrates the system model for EDAM. Section 22.3 performs the simulations for EDAM under various topological scenarios. Section 22.4 performs a probabilistic model to find the optimal set of sensor nodes in the network using data accuracy function and finally we conclude our work in the Sect. 22.5.

22.2 System Model

In this section, we discuss the system architecture for Estimated Data Accuracy Model (EDAM) in three phases. In the first phase, we clarify the nodes deployment strategy in the WSNs. In the second phase, we explore the foundation of EDAM in the network and finally in the third phase, we construct EDAM under spatially correlated data in the sensing region. The brief explanations are given bellow.

22.2.1 Deployment of Nodes in the Sensing Region

In the first phase, we assume U set of sensor nodes deployed randomly over a region Q such that $Q \subseteq R^2$ where $||U|| = u$ are the total numbers of sensor nodes. Suppose a source event S [10] has occurred in the sensing region Q. When S occurs in Q, a set of sensor nodes V wake up to sense the physical phenomenon of S. We define V as maximal set of sensor nodes which forms fully connected network [10] among them. We define $||V|| = v$ be the number of wake up sensor nodes where $v(v-1)/2$ is the direct one hop link to form a fully connected network with v sensor nodes.

22.2.2 Model for Estimated Data Accuracy in the Network

Since V set of sensor nodes wake up to sense the physical phenomenon of source event S, we construct a mathematical model to estimate the observed data at the sink node in the second phase of EDAM. Sink node is responsible for collecting the observation made by v sensor nodes to estimate $\hat{S}$ from S. Hence the error signals [14, 15] can be defined as

$$\tilde{S} \triangleq (S - \hat{S}) \tag{22.1}$$

We determine $\hat{S}$ by minimizing the mean square error from the expectation of $\tilde{S}^2$ as

$$\min_{\hat{S}} E(\tilde{S})^2 \tag{22.2}$$

Observation made by each sensor node i in the wireless sensor network is given as

$$X_i = S_i + N_i \qquad \text{where } i \in V \tag{22.3}$$

We assume uncoded transmission for the observed data sensed by sensor nodes in the network. Each sensor node i transmits a scaled version Y_i [4] of the observed data X_i to the sink node with power constraint P. Hence transmitted signal is given as

$$Y_i = \sqrt{\frac{P}{\sigma_{S_i}^2 + \sigma_{N_i}^2}} X_i = \alpha_i X_i, \quad \text{where } \alpha_i = \sqrt{\frac{P}{\sigma_{S_i}^2 + \sigma_{N_i}^2}} \text{ for } i \in V$$

The encoded signal Y_i transmitted by each sensor node i through additive white Gaussian noise (AWGN) channel [6, 16] is sent to the sink node. Sink node store the received signal in Y matrix for all sensor nodes as

$$Y = \alpha X \tag{22.4}$$

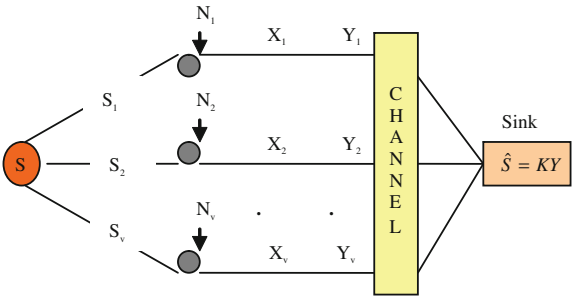


Fig. 22.1 Architecture for Estimated Data Accuracy Model (EDAM) in WSNs

$$\text{for } Y = \begin{pmatrix} Y_1 \\ Y_2 \\ . \\ Y_v \end{pmatrix}, \alpha = \begin{pmatrix} \alpha_1 & 0 & . & 0 \\ 0 & \alpha_2 & . & . \\ . & . & . & . \\ 0 & . & . & \alpha_v \end{pmatrix} \text{ and } X = \begin{pmatrix} X_1 \\ X_2 \\ . \\ X_v \end{pmatrix} \text{ where } i = 1, 2, \ldots \nu$$

Sink node decodes the signal to retrieve the estimate $\hat{S}$ of S. We define $\hat{S}$ as a random variable as a function of Y to recover the estimate $\hat{S}$ of all the observations done by sensor nodes at the sink node.

$$\hat{S} = h(Y) \tag{22.5}$$

Thus the mean square error becomes

$$\min_{h(Y)} E(\tilde{S})^2 \tag{22.6}$$

We choose $h(Y)$ for the subclass of affine functions [15] of Y as

$$\hat{S} = h(Y) = (KY + b) \tag{22.7}$$

where K is matrix and b is a scalar quantity. The affine estimator of S is unbiased, hence we get $E(\hat{S}) = 0$ and $E(\hat{S}) = KE(Y) + b = b$. For a linear estimator, we have $b = 0$ to get

$$\hat{S} = KY \tag{22.8}$$

Therefore, we find the optimal value of K at the sink as shown in Fig. 22.1 for $\hat{S}$ such that

$$\min_{K} E(S - KY)^2 \tag{22.9}$$

We calculate the optimal value of K for the estimate $\hat{S}$ using orthogonality principle. Y is orthogonal to the error signal $(\tilde{S})$ i.e. $Y \perp \tilde{S} = 0$. To get the optimal value of K for the estimator at the sink node, we define a linear model for (22.4) as

$$Y = \alpha(ZS + N) \tag{22.10}$$

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ . \\ Y_v \end{pmatrix} = \begin{pmatrix} \alpha_1 & 0 & . & 0 \\ 0 & \alpha_2 & . & . \\ . & . & . & . \\ 0 & . & . & \alpha_v \end{pmatrix} \left[\begin{pmatrix} Z_1 \\ Z_2 \\ . \\ Z_v \end{pmatrix} S + \begin{pmatrix} N_1 \\ N_2 \\ . \\ N_v \end{pmatrix} \right]$$

for zero mean random vector $\{S, Y\}$ for some matrix Z. N is a zero mean random noise vector with known covariance matrix $E(NN^T) \approx \sigma_N^2 I$. The covariance matrix of S is also known. $E(SS^T) \approx \sigma_S^2 I$ and $\{S, N\}$ are uncorrelated. The linear least mean square estimator [15] works according to orthogonality principal as $E\left(Y^T \tilde{S}\right) = E(Y^T(S - KY)) = 0$ to get

$$K = \frac{E(Y^T S)}{E(Y^T Y)} \tag{22.11}$$

Substituting (22.10) in (22.11) we get the expression as

$$K = \frac{\sigma_S^2 Z^T \alpha^{-1}}{\left(\sigma_S^2 Z^T Z + \sigma_N^2\right)} \tag{22.12}$$

Using (22.12) in (22.8), we get the linear least mean square estimator of S given Y is

$$\hat{S} = \frac{\sigma_S^2 Z^T \alpha^{-1}}{\left(\sigma_S^2 Z^T Z + \sigma_N^2\right)} Y \tag{22.13}$$

Put the value of (22.4) in (22.13), we get

$$\hat{S} = \frac{Z^T}{\left(Z^T Z + \sigma_N^2 / \sigma_S^2\right)} X$$

Therefore linear least mean square estimator of S given X for V sensor nodes in the network is illustrate as

$$\hat{S}(V) = \frac{1}{\left(V + \sigma_N^2 / \sigma_S^2\right)} \sum_{i=1}^{V} X_i = \frac{1}{\beta} \sum_{i=1}^{V} X_i \text{ where } \beta = \left(V + \frac{\sigma_N^2}{\sigma_S^2}\right) \tag{22.14}$$

We define mean square error between S and $\hat{S}(V)$ to find the data accuracy [9] for V sensor nodes in the network as

$$D(V) = E[S - \hat{S}(V))]^2 = E[S^2] - 2E[S\hat{S}(V)] + E[\hat{S}(V)^2] \tag{22.15}$$

The normalized [6, 10] data accuracy $D_A(V)$ for V sensor nodes in the network is given as

$$D_A(V) = 1 - \frac{D(V)}{E[S^2]} = \frac{1}{E[S^2]}[2E[S\hat{S}(V)] - E[\hat{S}(V)^2]] \tag{22.16}$$

The normalized data accuracy $D_A(V)$ for V sensor nodes in the sensor region can be implemented in spatial correlation model explained bellow.

22.2.3 EDAM Under Spatially Correlated Data in the Network

Finally third phase of EDAM demonstrates a mathematical model for the normalized data accuracy for spatial correlated data among V sensor nodes in the sensing region. We model a spatially correlated physical phenomenon of sensed data for V sensor nodes as a joint Gaussian random variable (JGRV's) [4, 5] as follows:

Step1: $E[S] = 0, E[S_i] = 0, E[N_i] = 0$ and $Var[S] = \sigma_S^2$, $Var[S_i] = \sigma_{S_i}^2$, $Var[N_i] = \sigma_{N_i}^2$

Step2: $Cov[S, S_i] = \sigma_S^2 Corr[S, S_i], Cov[S_i, S_j] = \sigma_S^2 Corr[S_i, S_j]$

Step3: $E[S, S_i] = \sigma_S^2 Corr[S, S_i] = \sigma_S^2 \rho(S, S_i) = \sigma_S^2 K_v(d_{S,S_i})$, $E[S_i, S_j] = \sigma_S^2 Corr[S_i, S_j] = \sigma_S^2 \rho(S_i, S_j) = \sigma_S^2 K_v(d_{S_i,S_j})$

We illustrate the covariance model [17] for Steps 2, 3 for spatially correlated data among sensor nodes in the network. Using covariance model, we have $K_V(d_{i,j})$ where $d_{i,j} = ||S_i - S_j||$ represents the Euclidian distance between node i and j. The covariance function is non-negative and decrease monotonically with the Euclidian distance $d_{i,j} = ||S_i - S_j||$ with limiting values of 1 at $d = 0$ and of 0 at $d = \infty$. We take the power exponential model [18] i.e. $K_V^{P.E}(d_{i,j}) = e^{-(d_{i,j}/\theta)}$ for $\theta > 0$ where θ is called as 'Range parameter'. 'Range parameter' controls the relation between the distance among sensor nodes (i,j) and the correlation coefficient $\rho(i,j)$. Thus from the correlation model, we get $\rho_{S_i,S} = e^{-(d_{S,i}/\theta)}$ and $\rho_{S_i,S_j} = e^{-(d_{i,j}/\theta)}$. Using (22.3) and (22.14) in (22.16), we get the normalized estimated data accuracy model (EDAM) for V sensor nodes in the network as

$$D_A(V) = \frac{1}{\beta}\left(2\sum_{i=1}^{V} e^{-(d_{S,i}/\theta)}\right) - \frac{1}{\beta^2}\left(\sum_{i=1}^{V}\sum_{j=1}^{V} e^{-(d_{i,j}/\theta)}\right) - \left(\frac{\left(\sum_{i=1}^{V}\sigma_{N_i}^2\right)}{\beta^2\sigma_S^2}\right) \tag{22.17}$$

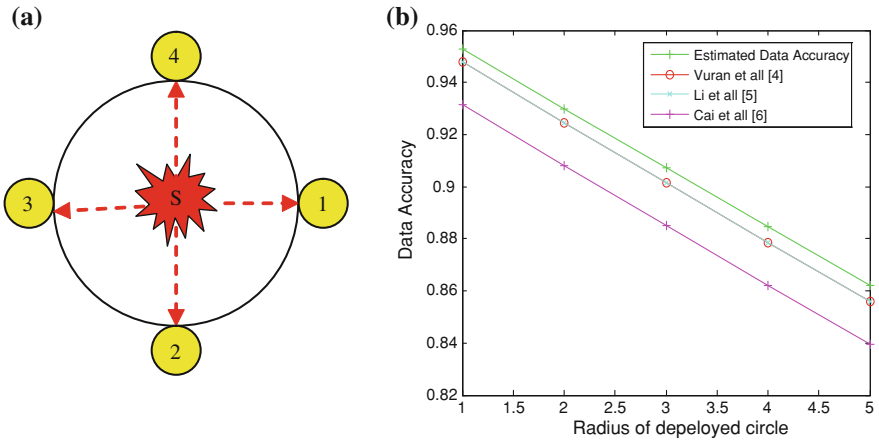


Fig. 22.2 **a** Circular topology for deployed sensor nodes. **b** Radius of circular topology versus data accuracy

22.3 Simulation Results

To perform simulations, we deploy U set of sensor nodes in the sensing region Q. When a source event S is detected, V set of sensor nodes wake up from U set of sensor nodes. In the first simulation, we vary the distance S from sensor nodes $(v = 4)$ i.e. $d_{s,i}$ where $i = 1, 2, 3, 4$ are equidistance in the sensing field as shown in Fig. 22.2a. Therefore we put $v = 4$ sensor nodes in a deployed circle with a S occurred at the centre of the deployed circle. As we keep on increasing the radius of the deployed circle for $d_{s,i}$ with same proportion, the data accuracy decreases as shown in Fig. 22.2b. We take $\theta = 70$ for our statistical data to calculate the normalized estimated data accuracy and compare the results from the literature [4, 5] and [6]. We conclude that as the radius of the deployed circle increases with same proportion our estimated data accuracy model (EDAM) always perform better compare to other models [4–6]. The accuracy models, [4] and [5] shows the same results for the normalized data accuracy with increasing deployed circle as shown in Fig. 22.2b.

In the second simulation, a sensing region of $2m \times 2m$ grid based sensor topology is taken with a sink node and a fixed source event S occurred in the center of the sensor network as shown in Fig. 22.3a. We deployed thirty-four nodes and a sink node in a grid based sensor topology as done in literature [5, 8]. When we simulate for this topology, we get the results for estimated data accuracy as shown in Fig. 22.3b which can sense more accurate data compare to other information (distortion) accuracy models [4–6]. Thus our propose EDAM can sense more accurate data than the other models [4–6], as we keep increasing the number of sensor nodes for $\theta = 70$. In Fig. 22.3b, it is clear that seven to ten sensor nodes are

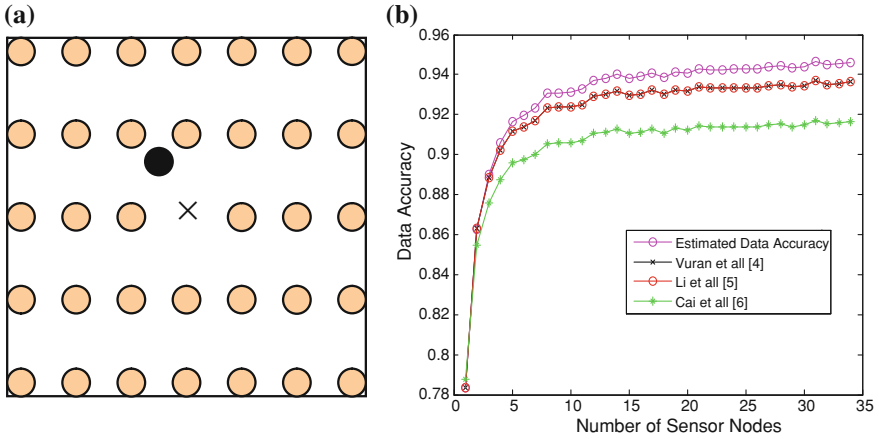


Fig. 22.3 **a** Wireless sensor network topology: ◯ means sensor node, ● means sink node, × means source event. **b** Number of sensor nodes versus data accuracy

sufficient for achieving the same estimated data accuracy level instead of deploying thirty-four sensor nodes in the field. Hence it is unnecessary to choose all the sensor nodes to perform data accuracy at the sink node. Therefore an optimal $||W|| = w$ number of sensor nodes can be selected for sensing source event and performs the data accuracy at the sink node.

In third simulation, we assume some of the sensor nodes get malicious due to extreme physical environment e.g. heavy rain fall. In such tropical situation, these sensor nodes read inaccurate data in the network. This means noise variances of malicious nodes are much higher compare to the noise variance of normal nodes. We define normal node as non malicious node or good node in the wireless network. We perform simulation for estimated data accuracy model (EDAM) and compare with the other information accuracy models [4–6] under malicious nodes condition. For the simplicity of our simulation, we initially deployed ten sensor nodes and assume out of ten sensor nodes, six sensor nodes are malicious as shown in Fig. 22.4a. We keep on adding the number of sensor nodes to thirty-four sensor nodes deployed randomly in the network. Finally we conclude that EDAM performs much better than other models still we introduce some malicious nodes in the network.

In Fig. 22.4b, we simulate for EDAM with respect to normal nodes and for introducing some malicious nodes in the sensing region. In this simulation setup, we compare two deployment scenarios. In first scenario, initially we deploy ten sensor nodes and keep adding nodes to thirty-four sensor nodes. These nodes are normal nodes. In another scenario, initially we deploy ten sensor nodes in similar way but out of ten nodes, six sensor nodes are malicious. We keep going on adding sensor nodes till we get thirty-four sensor nodes in the network. We compare these two deployment scenarios and conclude that sink node estimates more accurate data when there are normal nodes in the network. But if there are some malicious nodes in the network, the sink node estimates inaccurate data and performs poor

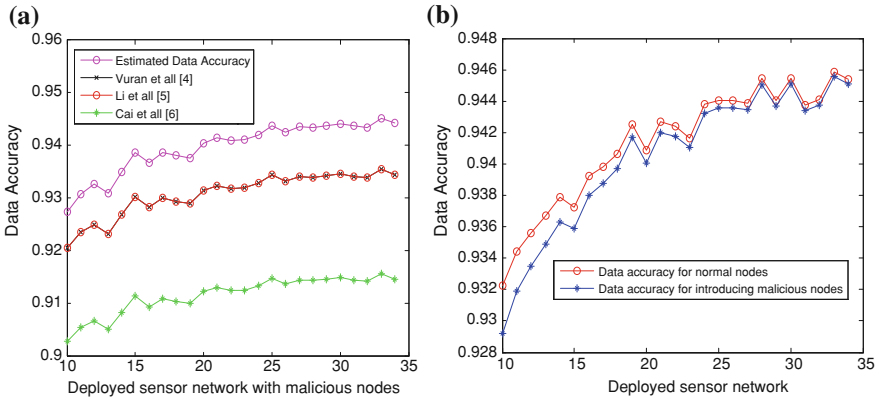


Fig. 22.4 **a** Data accuracy for malicious nodes in the network. **b** Comparison for data accuracy under normal nodes and malicious nodes condition in the network

data gathering for the deployed sensor nodes. Another conclusion we can draw from Fig. 22.4b is that the effect of noise variances of malicious nodes decreases as we keep increasing the number of sensor nodes in the network.

22.4 Probabilistic Models for Selecting Minimal Set of Sensor Nodes

In the previous section, we find EDAM for V set of sensor nodes to get an optimal W set of sensor nodes which are sufficient to achieve approximately the same data accuracy. Here we demonstrate a probabilistic approach for selecting W, an optimal set of sensor nodes from V set of sensor nodes, which are in active mode and keeping rest of the sensor nodes in sleep mode in the network. Therefore, we find the expectation of V set of sensor nodes in the network as follows

$$E[V] = \sum_{i=1}^{v} E[\chi_i] \tag{22.18}$$

where $\chi_i \in \{0, 1\}$ with $P(\chi_i = 1) = P$, for $\chi_i = 1$ denotes node i is selected. We perform all the combination of v sensor nodes taken w at a time given as ${}^{v}C_{w}$ to find data accuracy at the sink node. Hence probability by which selecting an optimal active sensor nodes subjected to estimated data accuracy function is given as

$$A(P) = \sum_{w=1}^{v} A(v = w)^{v} C_{w} P^{w} (1 - P)^{v-w} \tag{22.19}$$

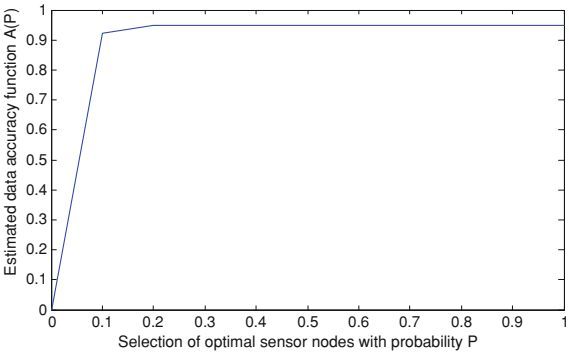


Fig. 22.5 Selecting optimal nodes with probability P versus estimated data accuracy function

Thus minimum probability $(P_{\min})$ [7] for achieving a required level of data accuracy with δ as a user dependent factor $(0<\delta<1)$ to find optimal set of sensor nodes in network is given as

$$P_{\min} = \arg\min_{P}\{A(P) \geq \delta A_{\max}\} \text{ where } A_{\max} = A(1) \tag{22.20}$$

In Fig. 22.5, we plot for the selection of optimal sensor nodes with probability P with respect to estimated data accuracy function $A(P)$. Here we take $P_{\min} = 0.2$ for achieving a required level of $A(P)$ for $\delta = 0.946$.We find $W(w = 7)$ optimal set of sensor nodes from $V(v = 34)$ set of wake up sensor nodes and rest $(v - w = 27)$ of the sensor nodes goes to sleep mode in the network. Thus we can increase the lifetime of the network by reducing the number of sensor nodes subjected to data accuracy.

22.5 Conclusions

In this paper, we propose estimated data accuracy model (EDAM) for the sensor nodes to sense more accurate data from the physical environment. Simulation results shows EDAM performs better and can sense more accurate data than other information accuracy models. Moreover we perform EDAM under malicious nodes condition and conclude that if some of the sensor nodes get malicious in the network, it read and transmits inaccurate data which results poor data gathering at the sink node. Finally a probabilistic model is developed using EDAM to find the minimal set of sensor nodes which is sufficient to perform the same data accuracy level achieve by the maximal set of sensor nodes.

References

1. Akyuildz IF, Su W, Sankarasubramanian Y, Cayirci E (2002) A survey on sensor networks. IEEE Commun Mag 40:102–104
2. Pradhan SS, Ramchandran K (2000) Distributed source coding: symmetric rates and applications to sensor networks. Proceedings of the data compressions conference, pp 363–372
3. Gastpar M, Vetterli M (2003) Source channel communication in sensor networks. Second international workshop on information processing in sensor networks
4. Vuran MC, Akan OB, Akyildiz IF (2004) Spatio-temporal correlation: theory and applications for WSNs. Comput Netw J 45(3):245–259
5. Li H, Jiang S, Wei G (2006) Information accuracy aware jointly sensing nodes selection in wireless sensor networks. LNCS. Springer, Berlin, pp 736–747
6. Cai K, Wei G, Li H (2008) Information accuracy versus jointly sensing nodes in WSNs. IEEE Asia Pacific conference on circuits and systems, pp 1050–1053
7. Karjee J, Jamadagni HS (2012) Energy aware node selection for cluster-based data accuracy estimation in wireless sensor networks. Int J Adv Netw Appl 3(5):1311–1322
8. Karjee J, Jamadagni HS (2011) Data accuracy model for distributed clustering algorithm based on spatial data correlation in wireless sensor networks. http://arxiv.org/abs/1108.2644 (under review)
9. Karjee J, Jamadagni HS (2011) Data accuracy estimation for spatially correlated data in wireless sensor networks under distributed clustering. J Netw 6(7):1072–1083
10. Karjee J, Jamadagni HS (2011) Data accuracy estimation for cluster with spatially correlated data in wireless sensor networks. IEEE international conference on information system and computational intelligence, China, vol 3, pp 284–291
11. Perillo MA, Heinzelman WB (2003) Optimal sensor management under energy and reliability constraints. Proceedings of IEEE wireless communications and networking, pp 1621–1626
12. Ordonez F, Krishnamachari B (2004) Optimal information extraction in energy limited wireless sensor network. IEEE J Sel Area Commun 22(6):1121–1129
13. Karjee J, Banerjee S (2008) Tracing the abnormal behavior of malicious nodes in MANET. Fourth international conference on wireless communications, networking and mobile computing, Dalian, China, pp 1–7
14. Poor V (1994) An Introduction to signal detection and estimation, 2nd edn. Springer, Berlin
15. Syed AH (2008) Adaptive filters. John, Hoboken
16. Goblick TJ (1965) Theoretical limitations on the transmission of data from analog sources. IEEE Trans Theory IT-11(4):558–567
17. Berger JO, de Oliviera V, Sanso B (2001) Objective Bayesian analysis of spatially correlated data. J Am Statist Assoc 96:1361–1374
18. De Olivera V, Kedan B, Short DA (1997) Bayesian prediction of transformed Gaussian Random Fields. J Am Stat Assoc 92:1422–1433

Chapter 23
Improvement of system stability margins using coordination control of Static Var Compensator (SVC) and Thyristor Controlled Series Capacitor (TCSC)

Venu Yarlagadda, K. R. M. Rao and B. V. Sankar Ram

Abstract The Thyristor Controlled Series Compensator (TCSC) and Static Var Compensator (SVC) are variable impedance Flexible AC Transmission Systems (FACTS) Controllers. A combination of the TCSC and the SVC installation is proposed to acquire superior performance for the power system. The coordination between the two pieces of equipment is designed with the SVC treated as the supplement of the TCSC. When operation of the TCSC is constrained by the inherent limitation of equipment, such as due to the firing-angle limitation of the thyristors, the adjustable SVC can supply the auxiliary support to improve the overall performance. The voltage and angle stability margins can be greatly improved with the compatible control schemes of the TCSC and the SVC.

Keywords TCSC · SVC · Co-ordination control of SVC and TCSC · Design of small scale TCSC model · Variable impedance FACTS controllers · Single machine two bus system · Voltage stability · P–V curves and P-δ curves

V. Yarlagadda (✉)
EEE Department, VNR VJIET, Hyderabad, India
e-mail: venuyar@gmail.com

K. R. M. Rao
EEE Department, MJCET, Hyderabad, India

B. V. Sankar Ram
EEE Department, JNTUH, Hyderabad, India

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_23,

23.1 Introduction

Voltage stability improvement demands different techniques, fixed compensation and the variable compensation. FACTS controllers which are the variable compensation devices are being used for more effective results. In this paper, the coordination control of Thyristor Controlled Series Capacitor (TCSC) and Static Var Compensator (SVC) is implemented practically in the laboratory. The TCSC is used as an auxiliary controller and SVC is used as the Master controller by which the Stability Margins have been enhanced tremendously which has been proved by the P–V and P-δ curves and bar charts.

23.2 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. The reliability of the power supply implies much more than merely being available.

Ideally, the loads must be fed at constant voltage and frequency at all times. In practical terms this means that both voltage and frequency must be held within close tolerances so that the consumers' equipment may operate satisfactorily.

23.3 Stability Indices

23.3.1 P-V Curve

As the power transfer increases, the voltage at the receiving end decreases. Finally, the critical or nose point is reached. It is the point at which the system reactive power is out of use. 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.

23.3.2 P-δ Curve

The relation between input power and the load angle is called power angle characteristics. The equation is given by, $P = EV\sin\delta/X$. The steady state stability limit is EV/X and it occurs at 90°.

23.4 Thyristor Controlled Series Capacitor

A capacitive reactance compensator which consists of series capacitor bank shunted by a thyristor controlled reactor in order to provide a smoothly variable series capacitive reactance.

A TCSC is a series-controlled capacitive reactance that can provide continuous control of power on the ac line over a wide range. From the system viewpoint, the principle of variable-series compensation is simply to increase the fundamental-frequency voltage across an fixed capacitor (FC) in a series compensated line through appropriate variation of the firing angle, α. A simple understanding of TCSC functioning can be obtained by analyzing the behavior of a variable inductor connected in parallel with an FC. The equivalent impedance, *Z*eq, of this *LC* combination is expressed as The impedance of the FC alone, however, is given by—j(l/ωC).

If $\omega C - (1/\omega L) > 0$ or, in other words, $\omega L > (1/\omega C)$, the reactance of the FC is less than that of the parallel-connected variable reactor and that this combination provides a variable-capacitive reactance are both implied.

23.5 Design of Thyristor Controlled Series Capacitor

Consider the Line reactance of the transmission line in per unit system. For 50 % compensation, the value of the capacitor in the TCSC will be 50 % of the line reactance.

Now for capacitive compensation, the value of inductive reactance must be greater than capacitive reactance, that is, $X_l > X_c$

$$X_{tcsc} = (X_l * X_c) / (X_l - X_c) \tag{23.1}$$

Total reactance of the line with TCSC is

$$X = X_l - X_{tcsc} \tag{23.2}$$

$$Q_{tcsc} = (I_c * I_c * X_c) - (I_{tcr} * I_{tcr} * X_l) \tag{23.3}$$

The variation of reactive power demand with load variations are obtained as

Now, if Q_{dmin} = minimum reactive power demand, Q_{dmax} = maximum reactive power demand

$$Q_{tcsc} = Q_{dmax} - Q_{dmin} \tag{23.4}$$

$Q_{ref} = Q_1 = 1$ p.u (corresponding to voltage value of 1 p.u)
$V = V_{ref} = 1.0$ p.u

$$Q_c = Q_{max} - Q_{ref} \tag{23.5}$$

$$Q_{tcr}(\alpha) = Q_{ref} - Q_{min} \tag{23.6}$$

$$Q_{tcsc} = Q_c - Q_{tcr}(\alpha) \tag{23.7}$$

Therefore

$$I_{tcr} * I_{tcr} * X_l = I_c * I_c * X_c - Q_{tcsc} \tag{23.8}$$

Hence

$$X_l = (I_c * I_c * X_c - Q_{tcsc}) / I_{tcr} * I_{tcr} \tag{23.9}$$

This is how the value of the inductive reactance in the TCSC circuit is calculated.

23.6 Design Static Var Compensator

The SVC 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. The design of SVC is based on our 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 $Q_1, Q_2,\ldots, Q_n$ 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.

$$Q_c = V^2 / X_C \tag{23.10}$$

Therefore

$$X_C = V^2 / Q_c \tag{23.11}$$

$$X_C = 1/2\pi fC \tag{23.12}$$

$$C = 1/2\pi X_C \tag{23.13}$$

In SVC circuit X_C is in parallel with X_l

$$Q_{tcr}(\alpha) = Q_{ref} - Q_{min} \tag{23.14}$$

$$X_l = = V^2 / Q_{tcr}(\alpha) \tag{23.15}$$

$$Q_{svc} = Q_c - Q_{tcr}(\alpha), \tag{23.16}$$

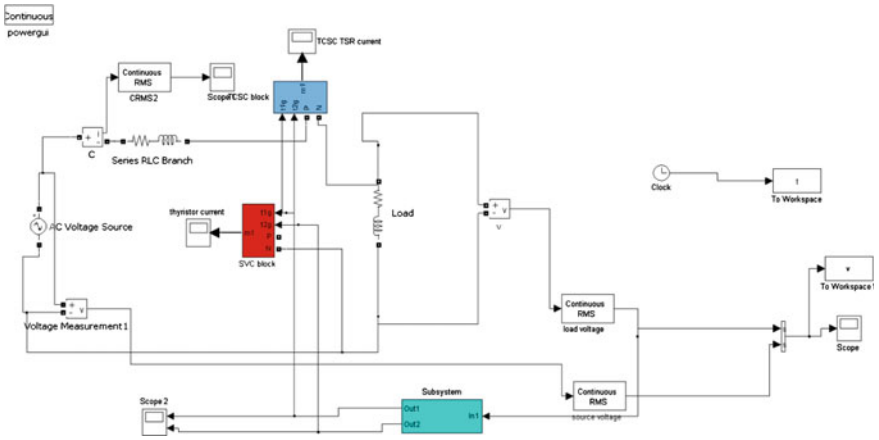


Fig. 23.1 SIMULINK model of test system with SVC & TCSC

23.7 MATLAB/SIMULINK Based Coordination Control of TCSC and SVC

The MATLAB/SIMULINK model is shown in Fig. 23.1 below, it consists of TCSC in series and SVC in shunt. The feed back system uses voltage feedback, PI controller, error detector to compensate error voltage, so the output voltage is fixed at reference voltage, the simulation diagram is shown below in Fig. 23.1 shows the simulation results without SVC and TCSC and Fig. 23.2 and 23.3 shows the simulation results without and with coordination control of SVC and TCSC respectively. Simulation results shows the improvement in the voltage profile.

23.8 Laboratory Based Co-Ordination Control of TCSC and SVC

The SMTB test system with a source feeding the RL load through a transmission line model and tested with and without Coordination control of TCSC and SVC. PV and P-δ curves have been drawn for both the cases. The system stability has been assessed with these curves. The test results show the improvement in the stability margins.The circuit arrangement for testing and coordination control of SVC and TCSC is shown Fig. 23.4 below.

This circuit consists of SVC and TCSC, TCSC is connected in series and TCR of TCSC is connected in parallel to series capacitor. SVC is connected in parallel to load. Shunt capacitor is connected through TPST switch. Initially firing angle is

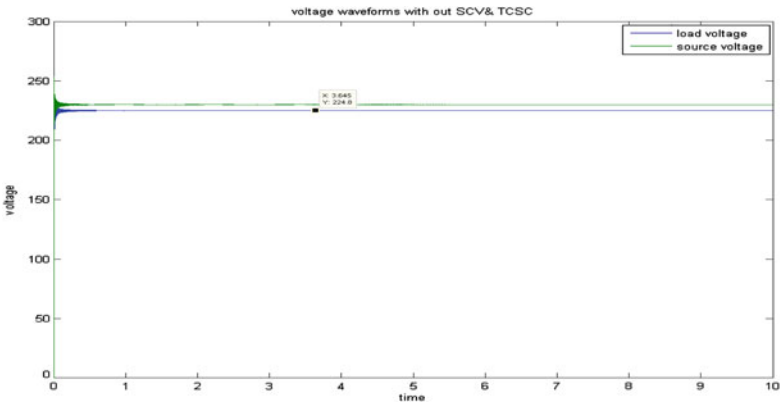


Fig. 23.2 Voltage waveforms without SVC & TCSC

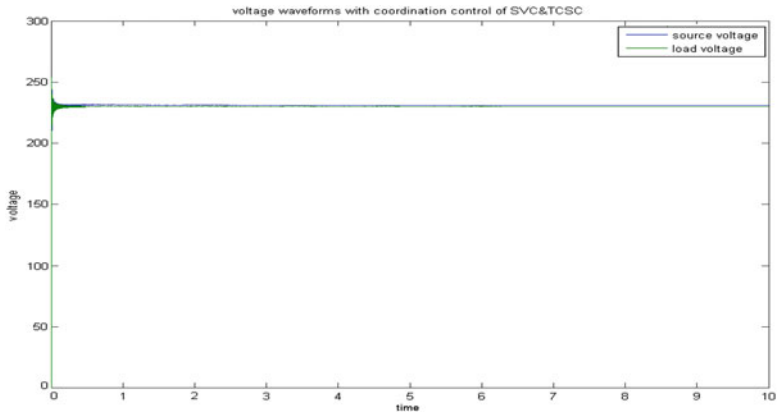


Fig. 23.3 Voltage waveforms with SVC

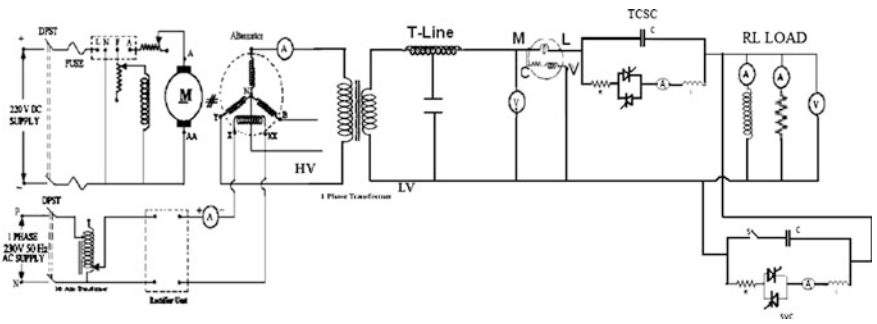


Fig. 23.4 Circuit diagram for coordination control of SM three bus test system (T-Line)

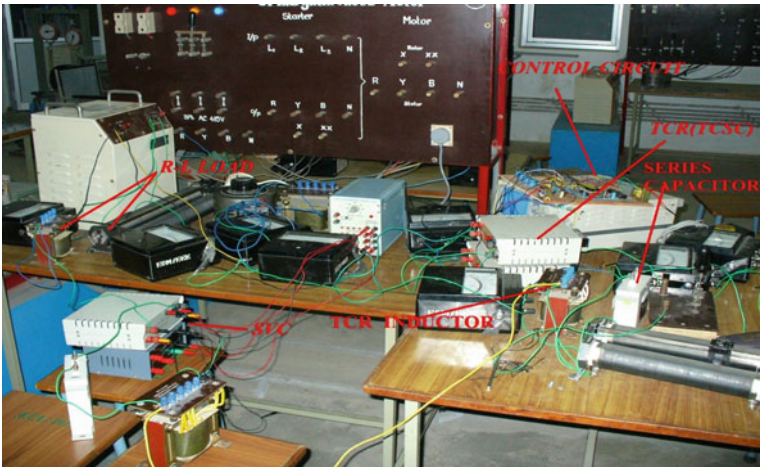


Fig. 23.5 Test system experimental setup in the laboratory

kept at 180°, if voltage is more than reference load voltage then firing angle to both TCR (SVC & TCSC) are decreased towards 90° and voltage is verified continuously. If voltage is find bellow reference voltage then firing angles of both TCR's increased towards 180°, after keeping firing angles at reference voltage, still voltage is less then shunt capacitor is turned on by closing TPST switch. The feedback is taken continuously and output voltage is maintained constant. Figure 23.5 shows the Test System Experimental setup in the Laboratory. Figure 23.6 shows the P-δ Bar Graphs of Coordination control of TCSC and SVC, Fig. 23.7 shows P-V Bar Graphs of Coordination control, Fig. 23.8 shows the P-V curves without and with Coordination control and Fig. 23.9 shows the P-δ curves without and with compensation. All of these graphs shows that the system stability margins have been enhanced tremendously.

23.9 Conclusions

A combination of the small scale models of TCSC and SVC has been developed in the laboratory to achieve superior performance of the system. A coordinated control scheme is implemented using TCSC and SVC to enhance the stability margin further.

The SVC serves as a master controller in supplementary to the TCSC, which acts as auxiliary controller. Whenever the TCSC is constrained, the SVC will automatically supply the auxiliary adjustment to the system under the proposed control scheme in order to improve the overall stability of the system. The test results show the effectiveness of the coordination control on improvement in system stability margins.

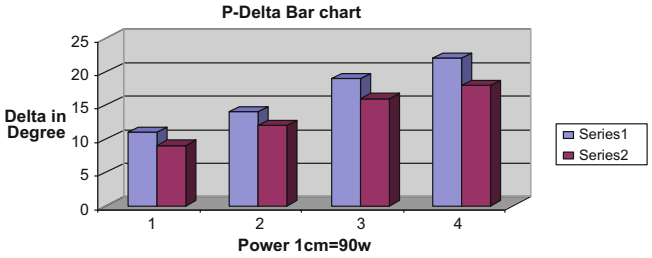


Fig. 23.6 P-δ bar graph of coordination control of TCSC and SVC

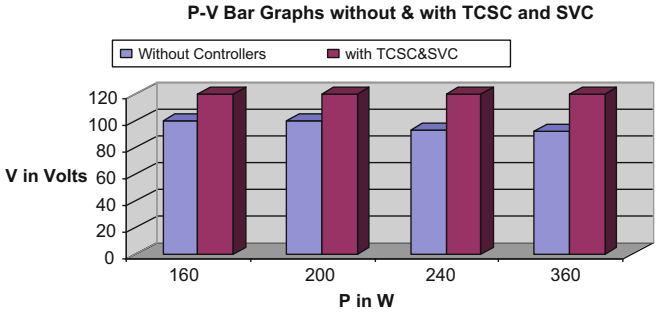


Fig. 23.7 P-V bar graph of coordination control of TCSC and SVC

Fig. 23.8 P-V curves without and with coordination control

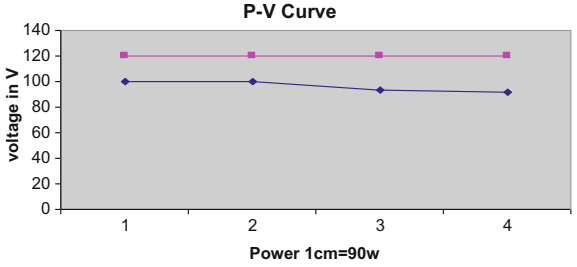
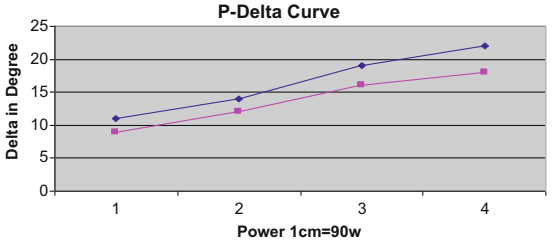


Fig. 23.9 P-δ curves without and with compensation



23.10 Future Scope

The coordination control of SVC & TCSC can be extended to the large rating machines and Large Interconnected Power Systems. The SVC & TCSC can also be fabricated by using IGBT's and testing can also be performed using DSP.

Acknowledgments We are immensely thankful to Management and Principal of VNR VJIET for providing R&D lab and other facilities to carryout this work.

References

1. Hingorani NG, Gyugyi L (1999) Understanding FACTS. IEEE Press, New York
2. Mathur RM, Varma RK (2002) Thyristor-based FACTS controllers for Electrical transmission systems. IEEE Press and Wiley Interscience, New York
3. IEEE Power Engg. (1995) Society/CIGRE, "FACTS overview", Publication 95 TP 108. IEEE Press, New York
4. Kundur P (1994) Power system stability and control. McGraw-Hill, Inc., New York
5. IEEE Power Engineering Society (1996) "FACTS applications", Publication 96 TP 116-0. IEEE Press, New York
6. Lyon WA (1954) Transient analysis of alternating current machinaery, chapter 2. Wiley, New York
7. Noroozian N, Ghandari M (1997) Improving power system dynamics by series connected FACTS devices. IEEE Trans Power Deliv 12(4):1635–1641
8. Yarlagadda V, Rao KRM, Sankar Ram BV (2011) Hardware circuit implementation of automatic control of static var compensator (SVC) using micro controller. IJICA 1(2)
9. Yarlagadda V, Rao KRM, Sankar Ram BV (2012) Power system generator and voltage stability enhancement by the hardware circuit implementation of 3-Ph static var compensator (SVC). In: CCPE, 27th and 28th April 2012
10. Venu Yarlagadda, Rao KRM, Sai Saroja P, Avinash P, Nikilesh P (2012) Dynamic stability improvement with combined fast acting automatic voltage regulator (AVR) and automatic load frequency control (ALFC) loops. In: ICMAET, 1st Jan 2012
11. Yarlagadda V, Rao KRM, Sankar Ram BV (2012) Voltage stability improvement using thyristor controlled series capacitor (TCSC) using Lmn and VCPI stability indices. IJSER 3(4)

Chapter 24
Effect of Parasitics of Feed-Forward Compensated OTA on Active-RC Integrators

S. Rekha and T. Laxminidhi

Abstract This paper analyzes the effect of parasitics of the Operational transconductance Amplifiers (OTAs) on Active-RC Integrators. The analysis is carried out for an Active-RC integrator built around a feed-forward compensated OTA designed in 180 nm CMOS technology to operate at a supply voltage of 0.5 V. A non-ideality factor (NIF) has been defined that accounts for the deviation of the response of the Active-RC integrator from the ideal. Simulations performed on the transistor level integrator justifies the mathematical analysis presented.

Keywords OTA · Feed-forward compensation · Non-ideality

24.1 Introduction

Continuous time filters are one of the important analog modules in a System on Chip. Basic building block of such analog filters is an integrator. Operational transconductance Amplifiers (OTAs) are typically used to realize the integrator. The conventional Miller compensation adopted in OTAs make them slow due to the limited Unity gain Bandwidth (UGB). This limitation can be overcome by using feed-forward frequency compensation [1] which compensates the loop by introducing a left half plane zero without significant change in OTA poles [2–4].

S. Rekha (✉) · T. Laxminidhi
National Institute of Technology Karnataka, Surathkal, India
e-mail: rsbhat_99@yahoo.com

T. Laxminidhi
e-mail: laxminidhi_t@yahoo.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_24,

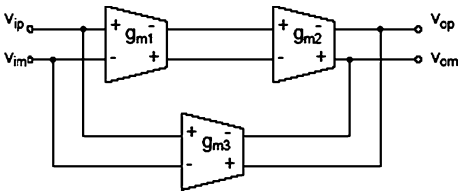


Fig. 24.1 Schematic of a feed-forward compensated OTA

In the regime of integrated circuits, parasitic capacitors and resistors are inherent to the circuit. The transistors are the major contributors to these parasitic capacitances and then follows the parasitic capacitances and resistors introduced by the layout. The circuit designers and layout designers take utmost care to minimize these parasitics. However, in sub-micron technology it becomes almost impossible to reduce these parasitics beyond certain point. In the case of continuous time filters, the filters are design centered to obtain a response that is close to ideal even in the presence of parasitics by using design centering techniques. One such technique is reported in [5], where the integrated capacitors are tuned. This may not be the case always. One example would be an active RC filter built using an OTA having feed-forward compensation [4]. A schematic of a fully differential feed-forward compensated OTA is shown in Fig. 24.1 where g_{m1}, g_{m2} and g_{m3} are the transconductors. There will be finite input/output and intermediate stage resistance and parasitic capacitance that will affect the performance of the integrator (filter) designed using such an OTA, specially the range of frequency over which the integrator can be operated satisfactorily. This paper analyzes how the range of frequency of operation is limited by the OTA non-idealities in the regime of low-voltage circuits. A bulk-driven OTA is designed to operate at 0.5 V for the analysis the details of which are given in Sect. 24.2. The effect of OTA non-idealities on an integrator designed using this OTA is given in Sect. 24.3. Results and conclusions are given in Sects. 24.4 and 24.5 respectively.

24.2 Pseudo Differential Bulk Driven OTA

The fully differential feed-forward compensated OTA shown in Fig. 24.1, is designed by using fully differential transconductors g_{m1}, g_{m2} and g_{m3}. g_{m1} and g_{m2} form the forward path and g_{m3} forms the feed-forward path. The g_m s are realized using PMOS input bulk-driven pseudo-differential transconductor, the schematic of which is shown in Fig. 24.2a. Pseudo-differential operation is preferred in order to have sufficient head-room when operating at 0.5 V supply. For a detailed information on the design, the reader is referred to [6]. M_1 and M_2 are the input transistors operating in saturation. M_3 and M_4 form the NMOS loads and operate in weak inversion. Gates of M_1 and M_2 are biased to carry a quiescent current of 10 μA. Gates of M_3 and M_4 are biased with a common mode feedback voltage which sets the output common mode voltage. M_{1d} and M_{2d} are the dummy

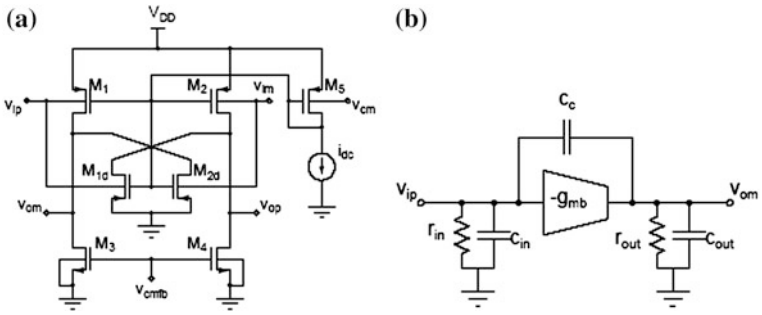


Fig. 24.2 Bulk driven pseudo differential Transconductor **a** Schematic **b** Single ended small signal model

transistors connected to reduce the effect of coupling capacitance between the input and output (Bulk and Drain) of the bulk driven transistors. The input and output common mode voltages of the transconductor are fixed at 0.25 V.

The transconductor is modeled to match the actual response and the single ended small signal equivalent circuit is shown in Fig. 24.2b. In the figure, g_{mb} is the bulk-transconducatnce of input transistor (M_1/M_2). c_{in} and c_{out} are the effective parasitic capacitances at input and output respectively. Similarly, r_{in} and r_{out} are the effective input and output resistances respectively. c_c is the effective coupling capacitance between the input and the output (bulk and drain) whose effect is minimized in the transconductor with the help of dummy transistors shown in Fig. 24.2a. Modeled values of the transconductor are as follows. $g_{mb} = 45.44\ \mu S$; $r_{in} = 14.36\ M\Omega$; $r_{out} = 385.80\ k\Omega$; $c_{in} = 443.57$ fF; $c_{out} = 189.63$ fF; $c_c = 0.682$ fF. The effect of c_c is neglected for the analysis as it is made small by design. The OTA, shown in Fig. 24.1, is then realized using the transconductor shown in Fig. 24.2a. The designed OTA has an open loop DC gain of 44 dB, 3-dB bandwidth of 601 kHz, UGB of 22 MHz and a phase margin of 68.5° and a gain margin of 54.89 dB under no load condition. Figure 24.3 shows the small signal model (single ended) of the OTA with input, output and intermediate stage parasitics. Note that g_ms shown in the figure actually represent the bulk transconductance (g_{mb}) shown in Fig. 24.2b. Transfer function of the OTA can be written as shown in (24.1).

$$\frac{Vo}{Vi} = \frac{r_{\text{int}} r_b g_m^2 + g_m r_b + g_m r_{\text{int}} r_b c_{\text{int}} s}{[1 + r_{\text{int}} c_{\text{int}} s][1 + r_b c_b s]} \tag{24.1}$$

The OTA has two poles and a zero in the left half s-plane. Because of these parasitic poles and the zero, the integrator designed using this OTA will be non-ideal. The effect of these parasitic poles and zero on the integrator is analyzed in the next section.

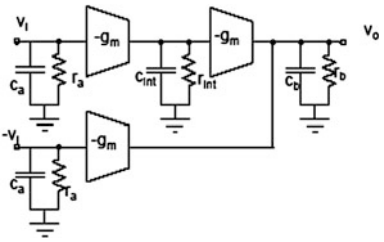


Fig. 24.3 Small signal model of the feed-forward OTA

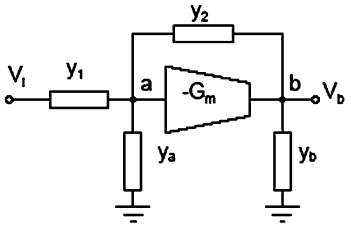


Fig. 24.4 Single ended Integrator topology

24.3 Integrator

Consider a general schematic of an active-RC integrator as shown in Fig. 24.4. Single ended circuit is shown for simplicity. y_1 represent the admittance of the integrating resistor. y_2 represent the feedback admittance (integrating capacitor). y_a and y_b are the input and output admittances of the OTA. G_m is the transconductance of the OTA. The transfer function of this circuit can be written as in (24.2).

$$\frac{V_b}{V_i} = -\frac{y_1}{y_2}\left[\frac{1}{1-F(s)}\right] \tag{24.2}$$

where

$$F(s) = \frac{y_a' y_b'}{y_2(y_2 - G_m)} \tag{24.3}$$

y_a' and y_b' are the net admittances at nodes 'a' and 'b' respectively. All the variables in (24.2) are functions of complex frequency 's'.

$$y_a' = y_1 + y_2 + y_a \tag{24.4}$$

$$y_b' = y_2 + y_b \tag{24.5}$$

In the case of feed-forward compensated OTA shown in Fig. 24.3, the transconductance G_m can be written as in (24.6).

$$G_m = \frac{g_{m1} r_1 g_{m2}}{1 + r_1 c_1 s} + g_{m3} \tag{24.6}$$

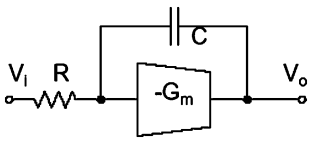


Fig. 24.5 Single ended schematic of the integrator

If F(jω) approaches zero, then (24.2) reduces to the form shown in (24.7), in the Fourier domain.

$$\frac{V_b(j\omega)}{V_i(j\omega)} = -\frac{y_1(j\omega)}{y_2(j\omega)} \tag{24.7}$$

which is what would have been obtained using an ideal OTA with $G_m(j\omega)$ tending to infinity. So, we can denote 1/(1-F(jω)) as a non-ideality factor (NIF) which must approach to one or in other words F(jω) must be negligible compared to unity. It is clear that F(jω) is frequency dependent which in-turn is decided by y_a, y_b and G_m. Therefore for a given OTA, there is a range of frequency over which the integrator can be expected to work satisfactorily. Conversely, if the integrator is to be designed for a given frequency range then the OTA needs to be designed such that its parasitics have minimum effect on the integrator response.

24.4 Results and Discussion

An integrator is designed to study the effect of non-idealities of the OTA. The single ended schematic of the integrator is shown in Fig. 24.5. OTA parasitics are not shown for simplicity. The integrator is designed for an Unity gain frequency of 380 kHz (assuming ideal integrator), with $R = 290\ k\Omega$ and $C = 1.44$ pF. The ideal transfer function of the integrator is as given in (24.8).

$$\frac{V_o(j\omega)}{V_i(j\omega)} = \frac{-1}{jwCR} \tag{24.8}$$

Considering the parasitics of OTA, the Non-ideality factor, NIF (1/(1-F(jω))) is computed using MATLAB. The magnitude and phase of the NIF are plotted in Fig. 24.6. Ideally one needs |NIF| = 1 and phase of (NIF) = 0. In Fig. 24.6 it can be seen that it is not so and is frequency dependent. From the magnitude plot of NIF, we can see that there is a zero at the origin of s-plane. This zero gets canceled with the pole (at origin) of ideal integrator producing a lossy integrator response. Magnitude plot of NIF also shows a pole at around 2.5 kHz which is the −3 dB frequency of the lossy integrator. For frequency beyond 1 MHz, |NIF| is found to increase and is highly frequency dependent for certain frequency range i.e., from 1 to 10 MHz. Similarly the phase plot of NIF shows an additional phase lag in this frequency range. This deviation is due to the non-ideality of OTA. The two poles and a zero of the OTA will be dispositioned and moved when the feedback is built

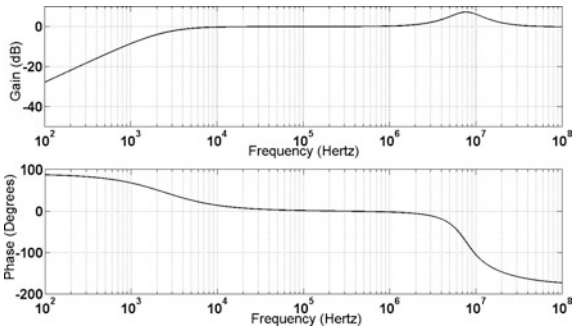


Fig. 24.6 Frequency response of NIF

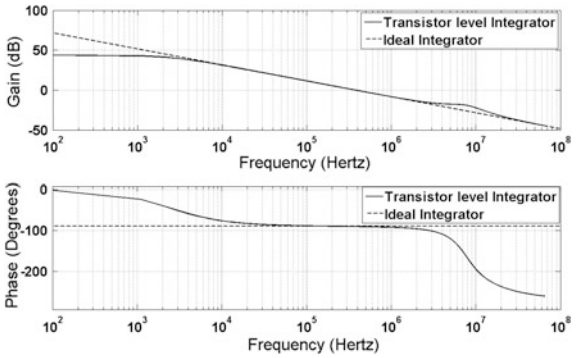


Fig. 24.7 Frequency response of integrator

around the OTA. In addition, a right half plane zero is introduced due to the integrating capacitor. To justify the analysis, the transistor level integrator is simulated and frequency response is shown in Fig. 24.7. It can be seen that beyond 1 MHz, the response is in error from ideal, justifying the mathematical analysis.

24.5 Conclusions

Effect of the non-idealities of feed-forward compensated OTA, particularly finite gain-bandwidth of the OTA on the integrator performance is studied. The non-ideality factor of the transfer function depends on the parasitic resistances and capacitances of the OTA. The results are proved with the help of a lossy integrator built around a feed-forward compensated OTA in 180 nm CMOS process operating at a supply voltage of 0.5 V. The simulation results show a good match with the analytical results.

References

1. Harrison JN (2002) Dynamic range and bandwidth of analog CMOS circuits. Ph D thesis, Macquarie University, Sydney, Australia
2. Harrison JN, Weste N (2002) 350 MHz Opamp-RC filter in 0.18 μm CMOS. In: IEE Electronics Letters, pp 259–260
3. Laxminidhi T, Prasadu V, Pavan S (2009) Widely programmable high frequency active RC filters in CMOS technology. In: IEEE transactions on circuits and systems, vol 56(2), pp 327–336
4. Thandri B, Silva Martinez J (2003) A robust feed-forward compensation scheme for multi-stage operational transconductance amplifiers with no miller capacitors. In: IEEE journal of solid state circuits, vol 38(2), pp 237–243
5. Laxminidhi T, Pavan S (2007) efficient design centering of high frequency Integrated continuous time filters. In: IEEE transactions on circuits and systems–I, Regular papers, vol 54(7), pp 1481–1488
6. Rekha S, Laxminidhi T (2011) A low power fully differential bulk driven OTA in 180 nm CMOS technology. In: Proceedings 3rd international conference on signal acquisition & processing, Singapore

Chapter 25
Modeling of Photovoltaic Charging System for the Battery Powered Wireless Sensor Networks

R. Hemalatha, R. Ramaprabha and S. Radha

Abstract Wireless Sensor Networks (WSN) requires energy harvesters to reduce the frequent replacement of the motes on field. This paper presents the modeling and design of a Solar Photovoltaic Charging (SPC) system with Incremental Conductance algorithm and Boost converter. Modeling of the chosen PV module (950 mW) is done and the characteristics are analyzed. The working of the Maximum Power Point Tracker (MPPT) is checked under arbitrarily varying irradiance and temperature conditions. The generated energy is stored in the 4.8 V, 150 mA NiMH battery. In this paper, mathematical modeling of WSN mote as a resistor based on the energy consumption of the mote in the active and sleep state is proposed. Series Charge regulation is used to improve the battery lifetime. The entire SPC system is developed using MATLAB/SIMULINK.

Keywords WSN · MPPT · PV module · Mote model · MATLAB · SIMULINK

Nomenclature

I, V	Solar current and solar voltage
I_{ph}	Light generated current
I_o	Reverse saturation current
q	Electron charge

R. Hemalatha (✉)
Department of ECE, SSN College of Engineering, Chennai, India
e-mail: hemas.lakshman@gmail.com

R. Ramaprabha
Department of EEE, SSN College of Engineering, Chennai, India
e-mail: ramaprabhar@ssn.edu.in

S. Radha
Department of ECE, SSN College of Engineering, Chennai, India
e-mail: radhas@ssn.edu.in

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_25,

R_{se}	Series resistance
n	Diode ideality factor
K	Boltzmann's constant
T	Temperature
T_{ref}	Reference temperature
$I_{sc\_}T_{ref}$	Short circuit current at T_{ref}
$V_{t\_}T_{ref}$	Thermal voltage at T_{ref}
$I_{o\_}T_{ref}$	Reverse saturation current at T_{ref}
ΔI, ΔV_c	Ripple current and voltage
k	Duty cycle

25.1 Introduction

A WSN consists of spatially distributed autonomous sensors to monitor physical or environmental conditions and to cooperatively pass their data through the network to a main location. The development of wireless sensor networks was motivated by military applications such as battlefield surveillance; today such networks are used in many industrial and consumer applications. The sensor nodes have limited energy supply (batteries). It is hard to replace or recharge nodes battery once deployed. To achieve longer operational lifetime of batteries energy harvesting can be used to increase the life time of the node.

Energy harvesting is the process by which energy readily available from the environment is captured and converted into usable electrical energy. There are several sources of energy harvesting among which solar energy is a mature technology [4] for large scale energy generation since it has maximum power density. The solar energy is a one form of an environmental energy and the pattern of energy will be different in space and frequently it varies with time.

Ambimax, an energy harvesting circuit and a supercapacitor based energy storage system for WSN is demonstrated in [11]. It performs MPPT autonomously, and uses supercapacitor as storage element. It also enables the composition of multiple energy harvesting sources including solar, wind, thermal, and vibration. Efficient multi-stage energy transfer system that reduces the common limitations of single energy storage systems is proposed in [10] to achieve near perpetual operation. The system uses super capacitors as primary buffer and a lithium rechargeable battery as secondary buffer. DuraCap [12] utilized supercapacitor as the energy storage element. A bound-control circuit for PFM regulator switching is used as the MPPT to achieve high conversion efficiency and minimal downtime.

In this paper the SPC system is designed with boost converter and Incremental Conductance algorithm to track the maximum power. Since the battery voltage is greater than the PV module voltage, to get better efficiency boost converter is used. NiMH Battery is used. Series Charge Regulation is done to increase the lifetime of

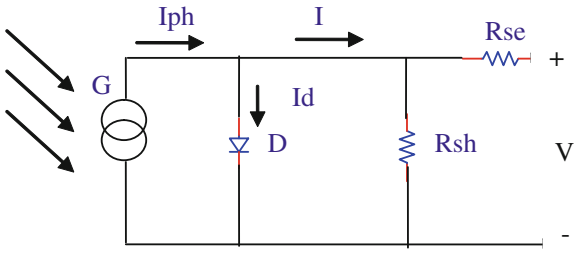


Fig. 25.1 Equivalent circuit of ideal PV module

the battery. To suit the WSN requirement solar panel in milli watt range is taken for consideration. The analysis is done with a resistive load representing the active and sleep state of the WSN mote.

25.2 Photovoltaic Modeling

25.2.1 PV Model

The photovoltaic effect is the creation of a voltage (or a corresponding electric current) in a material upon exposure to light. Before using the solar cells practically, it is essential to have a simulation model to analyze the behavior of solar cells under varying illumination and temperature conditions. From the several models available in literature [1–3]; the one diode model [2] is presented here. The ideal solar cell can be modeled as a current source in parallel with the diode as in Fig. 25.1. The output of the current source is directly proportional to the light falling on the cell.

The net current of the module is the difference of the light generated current, Iph and the normal diode current Id. The calculation is done with including the series resistance and neglecting the shunt resistance. The diode quality factor is set to a value between 1 and 2 to provide better curve match. In this paper it is chosen as 1.1. The following Eqs. (25.1–25.8) are used for modeling the solar module [2].

$$I = I_{ph} - I_o\left(e^{\frac{q(V+IRse)}{nkT}-1}\right). \tag{25.1}$$

$$I_{ph} = I_{sc\_Tref} \times \frac{G}{G_n} + k_1\left(T_{op} - T_{ref}\right). \tag{25.2}$$

$$k1 = \frac{I_{sc}\_T_2 - I_{sc}\_T_{ref}}{T_2 - T_{ref}}. \tag{25.3}$$

$$Vt\_Tref = \frac{AkTref}{q}. \tag{25.4}$$

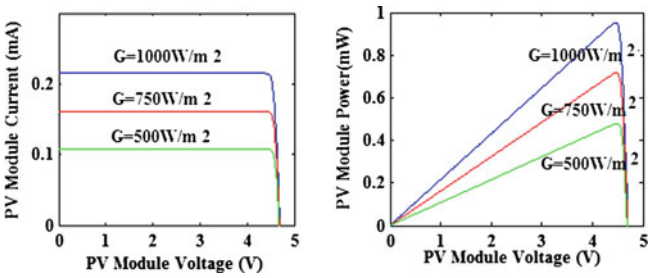


Fig. 25.2 Solar module V–I and power characteristics under varying irradiances at 25 °C

$$I_o\_T_{ref} = \frac{I_{sc}\_T_{ref}}{\left(\exp\left(\frac{V_{oc}\_T_{ref}}{V_t\_T_{ref}}\right) - 1\right)}. \tag{25.5}$$

$$I_o = I_o\_T_{ref} \times \frac{T_{ref^3}}{T_{ok^3}} e^{\frac{\frac{qVg}{Ak}}{\left(\frac{1}{Tref} - \frac{1}{Tok}\right)}}. \tag{25.6}$$

$$R_s = -\frac{dV}{dIvoc} - \frac{1}{X_V}. \tag{25.7}$$

$$X_V = \frac{Io\_Tref}{Vt\_Tref} \times e^{\frac{Voc\_Tref}{Vt\_Tref}}. \tag{25.8}$$

25.2.2 MATLAB Model of the PV Module

The Blue Solar SL8585 mm PV module has been chosen for modeling since it is well suited with the requirement range of WSN. It has 950 mW of nominal maximum power. It is of monocrystalline type. The electrical characteristic of the cell under Standard test conditions is given as: Maximum Power-950 mW (P_{max}), Voltage at max power-4.5 V (V_m), Current at maximum power-210 mA (I_m), Open circuit Voltage-4.7 V (V_{oc}) and Short circuit Current-215 mA (I_{sc}).

When the irradiance (G) changes at constant temperature, V_{oc} remains almost unchanged and there is variation in the module current. When G increases, the short circuit current also increases proportionally as shown in Fig. 25.2, this increases the power accordingly. When temperature changes at constant irradiance, I_{sc} remain almost unchanged and there is variation in the open circuit voltage V_{oc}. When temperature increases, the open circuit voltage decreases proportionally this reduces the power accordingly. The variation in temperature and irradiance affects the amount of power extracted from the module and the maximum power point. This necessitates the use of MPPT to track the changing peak [2, 5].

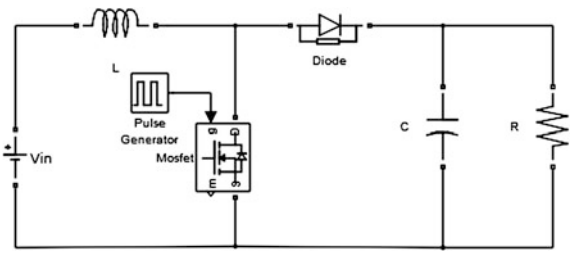


Fig. 25.3 Boost converter circuit

25.3 Modeling of DC–DC Boost Converter

DC converter can be considered as an ac transformer with a continuously variable turns ratio. It is used to control the power flow from the PV module to the battery and the load. The transistor switching loss increases with the switching frequency and as a result efficiency decreases. The core loss of the inductor limits the high frequency operation. The choice of the switching frequency is made by considering these losses. Boost converter as in Fig. 25.3, is used as MPPT for the SPC system proposed in this paper.

The basic assumptions made in this simulation are: All switches are ideal, Boost converter operates in continuous charging mode and the current is assumed to rise and fall linearly. The design equations of the boost converter are given through (25.9–25.11).

The relation of input and output voltage in the boost converter is given in (25.9)

$$\frac{V_{in}}{V_o} = (1 - k). \tag{25.9}$$

The inductor and the capacitor values are obtained using,

Considering the input voltage to the converter as the maximum power point voltage of the solar module, using (25.9–25.11) the parameters for the converter are designed and the values are given as: Input Voltage-4.5 V, Switching Frequency 10 kHz, Inductance 200 mH and Capacitance 200μF.

$$\Delta I = \frac{v_{in}k}{fL}. \tag{25.10}$$

$$\Delta V_c = \frac{I_o k}{fC}. \tag{25.11}$$

25.4 Maximum Power Point Tracking

The maximum power extracted from the PV module depends mainly on three factors: irradiance, load impedance and module temperature. When a PV module is directly connected to the load, the system will operate at the intersection of the

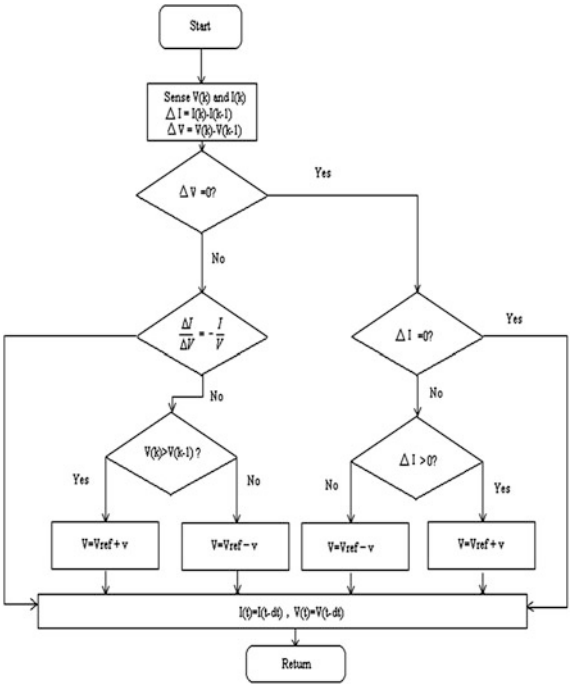


Fig. 25.4 Flowchart for incremental conductance algorithm

I–V curve and load line, which can be far from the Maximum Power Point (MPP). The MPP production is therefore based on the load-line adjustment under varying atmospheric conditions. The SPC system should be designed to operate at the maximum output power levels for any temperature and solar irradiation levels at all times. To adapt the load resistance to the PV modules and extract maximum power from them, the duty cycle is set to its optimal value which corresponds to its optimal operating point (V_m, I_m) using MPPT. The flowchart for the incremental conductance algorithm is shown in Fig. 25.4.

25.4.1 Incremental Conductance Algorithm

The output voltage and current from PV module are monitored to calculate the conductance and incremental conductance. Comparing the conductance values the MPP reference signal is generated by increasing or decreasing the photovoltaic voltage reference. Variation in the reference voltage is taken as 0.01 V to reduce the fluctuation voltage at maximum power point. When PV system is connected to load through boost converter with I&C algorithm, the irradiance change is detected correctly and the output power in the circuits is tracked effectively as shown in Fig. 25.5. The irradiance changes at 0.5 s in the simulation. The output power produced in this case is better than the converter circuit without MPPT.

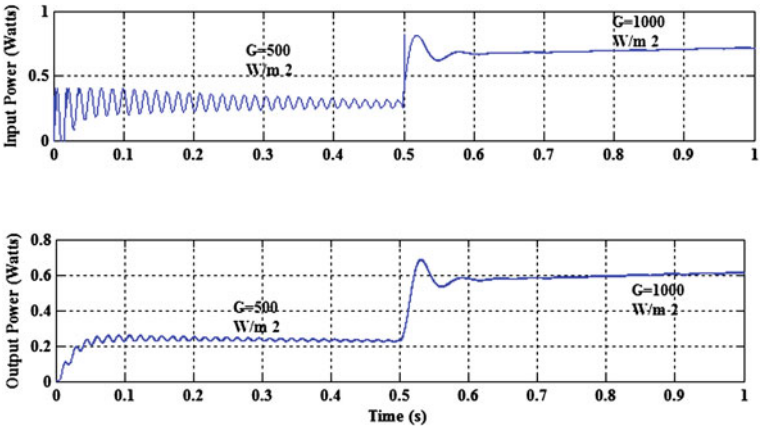


Fig. 25.5 PV system response to varying irradiance with incremental conductance algorithm

25.5 Battery Modeling

The 4.8 V 150mAh Nickel metal hydride rechargeable battery is used as the storage element. NiMH battery is chosen since it has the following advantages: better performance, Long battery life, extremely low battery cost, Stable performance due to flat discharge curve, No memory effect, easy charging and usage. The specifications used for the battery are: Type: V 110H, Nominal Voltage: 4.8 V, Nominal Capacity: 110 mAh, Typical Capacity: 120 mAh, Internal Resistance: 1Ω

Dynamic modeling of the battery is done with the Eqs. (25.12–25.13) considering both the charging and discharging states [6],

$$V_{batt} = E_o - R \cdot i - K\frac{Q}{Q - it} \cdot (it + i^*) + Exp(t). \tag{25.12}$$

$$V_{batt} = E_o - R \cdot i - K\frac{Q}{|it| - 0.1 \cdot Q} \cdot (i^*) - K\frac{Q}{Q - it} \cdot it + Exp(t). \tag{25.13}$$

25.6 Modeling of the Load

As the system is designed for the WSN, the load has been modeled based on the *active and sleep state energy consumption* of the motes as shown in Table 25.1. Resistive loads are used such that the current drawn by the load is equivalent to the sleep state and active state current of the mote. Power consumed by the mote in the active and sleep state are considered to calculate the resistance values in these two

Table 25.1 Mote—load modeling

Mote name	Mote state (at min volt 2.7 V)		Load resistance at 2.7 V		Load resistance at 3.3 V	
	Active (mW)	Sleep (μW)	Active (Ω)	Sleep (kΩ)	Active (Ω)	Sleep (kΩ)
Mica2 MPR400CB	95	44	76.73	165.68	114.63	247.5

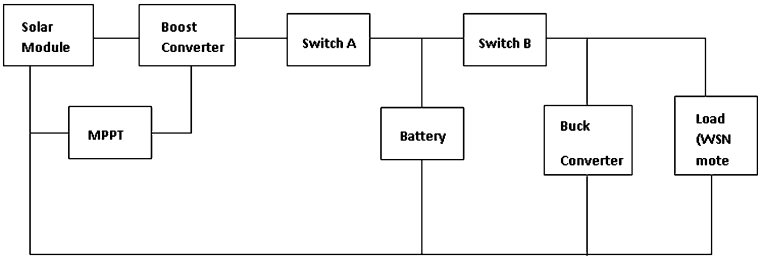


Fig. 25.6 Proposed solar photovoltaic charging system

states using (25.14). The values of the resistances calculated for the *Mica2 (MPR400CB)* is given in Table 25.1.

$$R = \frac{V^2}{P}. \tag{25.14}$$

25.7 Solar Photovoltaic Charging System

The solar photovoltaic charging system shown in Fig. 25.6 consists of the solar module, boost converter, battery, charge controller, buck converter to reduce the voltage down to the mote operating voltage and the load. The MPPT controller provides the peak operating point of the solar module whereas the boost converter provides the impedance matching with the battery and the solar module. The input and the output power of the SPC system are shown in Fig. 25.7 where the output power approximately equals the input power.

The switches provide the series charge regulation of the system. Whenever the battery terminal voltage is greater than the fully charged voltage (100 % State of Charge) or greater than the upper limit of the charging cycle (80 % State of Charge) the switch A is opened and prevents the overcharging of the battery and increases its lifetime. If the battery voltage is less than the lower critical limit (20 % State of Charge) and there is no power available from the solar module the switch B is opened to avoid the draining of the battery. Battery nominal voltage is 4.8 V and the fully charged voltage will be 5.62, to supply power to the mote the

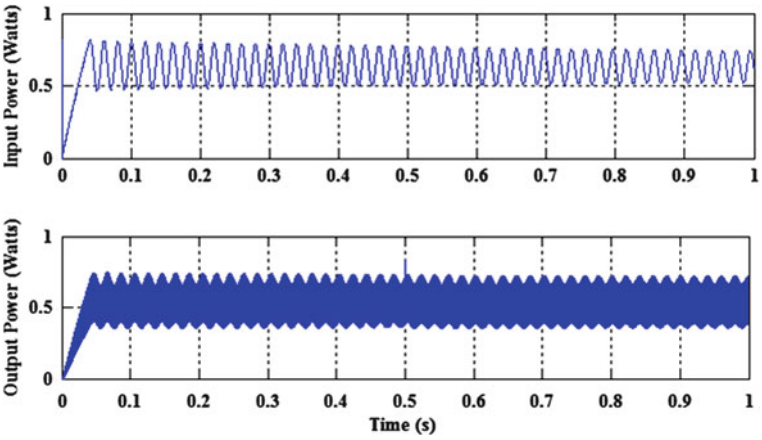


Fig. 25.7 Input and output power at the test condition

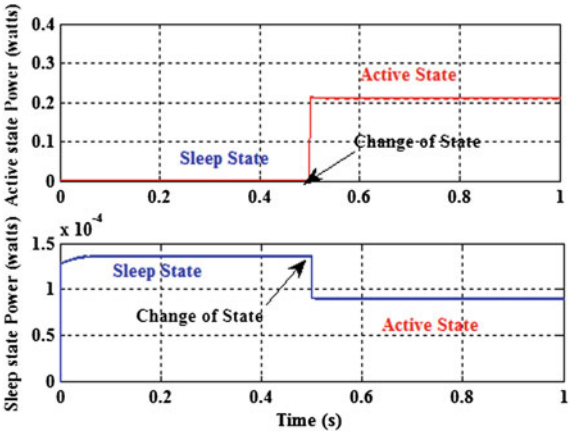


Fig. 25.8 Load power in active and sleep states

voltage is stepped down to 3.3 V using buck converter. The voltage is chosen from the actual working range of the Mica2 mote under consideration (2.7–3.3 V).

For validation the mote is assumed to be in sleep state for 0.5 s and in active state for 0.5 s. The initial state of charge (SOC) of the battery is considered to be 50 % in the simulation. The SOC is increased to 50.02 after simulation as shown in Fig. 25.9. The irradiance used is G = 1,000 W/m^2, so sufficient energy is available to supply to the load in both active and sleep states and the remaining is used to charge the battery. The power consumption of the mote at active (more) and sleep (less) state is shown in Fig. 25.8. The charging current of the battery is also shown in the Fig. 25.9. Less current is consumed by the mote in sleep state so the charging current to the battery up to 0.5 s is greater than the remaining duration

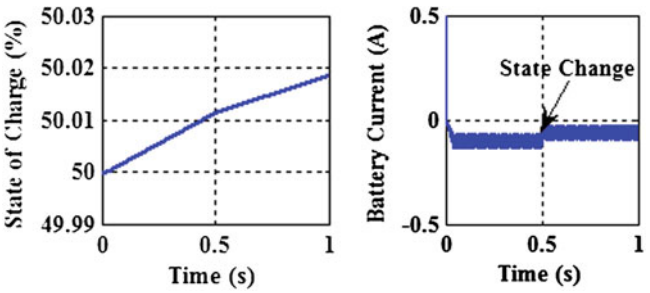


Fig. 25.9 SOC and current of the battery

where the current consumption by the load is high. The proposed PC system thus works properly at the varying states of the mote and varying irradiance condition.

25.8 Conclusion

In this paper SPC system for WSN mote with power flow control and battery charging control is presented. The system is simulated using MATLAB-SIMULINK. In this paper modeling of WSN mote as a resistor based on the energy consumption of the mote in the active and sleep state is done. From the simulated results the effectiveness of the proposed system is proved. In future work, closed loop charging system for the battery will be implemented and the inductor ripple current will be suppressed to get better efficiency.

References

1. Gow JA, Manning CD (1999) Development of a photovoltaic array model for use in power electronics simulation studies. In: IEE Proceedings on electric power applications, vol 146(2), pp 193–200
2. Walker G (2001) Evaluating MPPT converter topologies using a MATLAB PV model. J Electr Electron Eng, IEAust, 21(1):49–56 (Australia)
3. Villalva MG, Gazoli JR, Filho ER (2009) Comprehensive approach to modeling and simulation of photovoltaic arrays. In: IEEE transactions on power electronics, vol 24(5), pp 1198–1208
4. Hande A, Polk T, Walker W, Bhatia D (2007) Indoor solar energy harvesting for sensor network router nodes. J Microprocess Microsyst 31(6):420–432
5. Liu F, Duan S, Liu F, Liu B, Kang Y (2008) A variable step size INC MPPT method for PV systems. In: IEEE transaction on industrial electronics, vol 55(7), pp 2622–2628
6. Tremblay O, Louis Dessaint A (2009) Experimental validation of a battery dynamic model for EV applications. World Electr Veh J 3 ISSN 2032-6653-2009 AVERE
7. Salas V, Barrado A, Lazaro A (2006) Review of the maximum power point tracking algorithms for stand-alone photovoltaic systems. Sol Energy Mater Sol Cells 90:1555–1578

8. Oi A (2005) Design and simulation of photovoltaic water pumping system. Electrical Engineering, Master of Science in Electrical Engineering, California Polytechnic State University, San Luis Obispo
9. Jiang X, Polastre J, Culler D (2005) Perpetual environmentally powered sensor networks. In IPSN/SPOTS 2005
10. Park C, Chou PH (2006) Ambimax: autonomous energy harvesting platform for multi-supply wireless sensor nodes. In: 2006 3rd annual IEEE communications society on sensor and ad hoc communications and networks, pp 168–177
11. Chen C-Y, PH Chou (2010) DuraCap: a super capacitor-based, power-bootstrapping, maximum power point tracking energy-harvesting system. In: Proceedings of the 2010 international symposium on low power electronics and design, Austin, Texas, USA
12. NiMH battery datasheet. www.rapidonline.com
13. Blue Solar SL8585mm, Solar Panel. http://www.xscyz.com/
14. Mica2 Datasheet. www.xbow.com
15. www.mathworks.com

Chapter 26
Torque Computation of Induction Motor with VVVF Drive Subjected to Severe Torque Fluctuation

M. V. Palandurkar, J. P. Modak and S. G. Tarnekar

Abstract Load with severe torque fluctuations can be driven by an induction motor with control of input frequency using VVVF inverters. Situation of rise in load torque can be met by reduction in frequency at that instant. Situation of sudden reduction in load torque can be met by increase in frequency, at that instant. This paper deals with computations for such cases using the graphical relation between torque and speed, finally leading to a plot of instantaneous motor torque as a function of time.

Keywords Flywheel · Process machine · Demand torque · VVVF drive · T-s curve

26.1 Introduction

Process machines with tougher duty cycles are required to drive variable demand torque over one cycle of operation. These often need a bulky flywheel to smoothen out variations in the speed of a shaft caused by torque fluctuations. Many machines

M. V. Palandurkar (✉)
Department of Electrical Engineering, RCOEM, Nagpur, Maharashtra, India
e-mail: mpalandurkar@yahoo.com

J. P. Modak
Department of Mechanical Engineering, Nagpur, Maharashtra, India
e-mail: jpmodak@yahoo.com

S. G. Tarnekar
Department of Electrical Engineering, GHRCE, Nagpur, Maharashtra, India
e-mail: sgtarnekar@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_26,

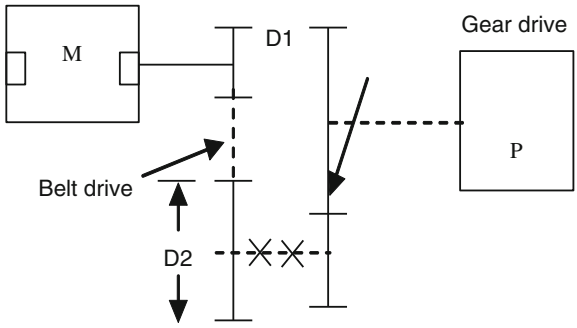


Fig. 26.1 Schematics of an arbitrary process unit, mechanical power transmission and 3 phase induction motor

have load patterns that cause the torque to vary cyclically. Adding bulky flywheel is a solution accepted traditionally. Internal combustion engines with one or two cylinders are a typical example. Piston compressors, punch presses, rock crushers etc. are the other systems that have bulky flywheels. Flywheel is an inertial energy-storage device which transacts mechanical energy and serves as a reservoir. It absorbs mechanical energy by increasing its angular velocity and delivers the stored energy by decreasing its velocity.

Figure 26.1 describes the schematics of an arbitrary process unit P along with usual mechanical power transmission system for torque amplification and speed reduction. In this figure, pulley D2 is a power transmission pulley also acts as a flywheel. Pulley D1 is driving pulley receive power from induction motor M. The arbitrary process machine P makes use of link mechanism or cam mechanism or combination of linkage, cam and gears. For such process unit, at every instant, demand torque changes with respect to time. The arbitrary demand torque characteristic of any process machine can be estimated based on cycle time of operation, process resistance and inertia resistance. These can be detailed based on intended operation and proposed details of partial mechanical design [1, 2].

Hence, this variation is cyclic and cycle time is commensurate with rpm of process unit. A typical torque time relation for arbitrary process machine (say mechanical punching press) is shown in the Fig. 26.2. Here, crank speed of input shaft of the process machine is chosen as 20 rpm. Therefore, time for complete cycle of operation should be 3000 m-Sec which gets completed in one rotation of the input link of the process machine. This figure shows that demand torque varies with time, which induction motor cannot generate. Hence, the flywheel is required to make up for the difference of the torque in all time intervals marked in Fig. 26.2. The portion of the system between D2 and process unit is subjected to severe torsional vibrations. Also presence of flywheel with high moment of inertia J in the process machine reduces acceleration, increases weight of engine. It is harder to start and causes fatigue to the components of power transmission thereby prolonging equipment functional failure [3]. Therefore, it is desirable to eliminate bulky flywheel from the design of any process machine in general.

With the advent of electric drives and power electronics circuitry using VVVF method, proper energy monitoring is possible to control the power supply to

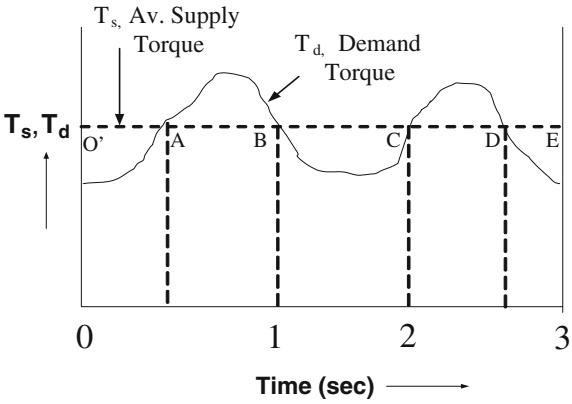


Fig. 26.2 Arbitrary demand torque characteristics

induction motor having low moment of inertia to generate supply torque closely matching with demand torque resulting in elimination of flywheel.

Among different control schemes, a constant volt per hertz principle is chosen to drive three phase induction motor as shown in Fig. 26.3. In this technique, a dynamic model of three phase induction machine is derived from two phase machine. [4–6]. The equivalence between three phase and two phase machine is based on the equality of the mmf produced in the two phase winding and three phase winding. The stator and rotor variables are transformed to a synchronously rotating reference frame that moves with the rotating magnetic fields. Finally, a dynamic machine model in synchronously rotating and stationary reference frame is developed in per unit by defining the base variables both in $a - b - c$ and the $d - q - o$ variables. Authors have already reported [9], that the above method can be analyzed. It uses *VVVf* based induction motor drive by controlling input side frequency for better performance, with much smaller system inertia. According to change in demand torque, varying cyclically with respect to time, the requirement of input frequencies to the main drive during different time intervals also changes in order to generate electromagnetic torque matching with demand torque. Hence, problems occurring due to the presence of large flywheel between induction motor and process machine are eliminated. It is observed that required effective energy transaction from rotational masses to shaft of the motor to match the change of load torque to peak value is also less when drive is controlled from input side by frequency control, using VVVF technique, with low moment of inertia [10].

In the present paper, based on above technique, energy is calculated graphically by plotting T-s characteristics at different required frequencies to meet demand load torque characteristics changes suddenly low to high value and vice versa.

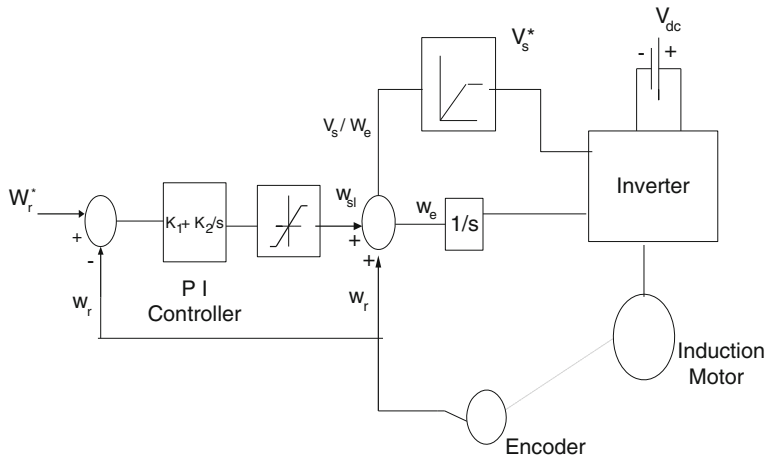


Fig. 26.3 Induction motor drive with closed loop volts/hertz control

26.2 Closed Loop Induction Motor Drive with Constant Volts per Hertz Control Strategy

In an attempt to simplify the analysis, and to test the proposed system, using standard logics in control system engineering, the demand torque variation of assumed process machine is say as shown in Fig. 26.4. In order to produce same demand torque, an implementation of the constant volts/hertz control strategy for the PWM inverter fed induction motor on per unit basis, is simulated in MATLAB simulink as shown in Fig. 26.5 with given mechanical load torque. In PWM inverter, the per unit voltage command through volts/hertz function generator is converted into three phase stationary reference frame variables $a - b - c$ which are further transformed into two phase stationary reference frame $d^s - q^s$ variables and then into synchronously rotating frame in $d^e - q^e$ variables. A PI controller is employed to regulate the slip speed of the motor to keep the motor speed at its set value with respect to frequency given to drive. The major blocks consist of PWM inverter, induction motor with mechanical load [7, 8]. In this scheme mechanical load is varying cyclically in an assumed pattern. As the load torque increases, the speed loop error generates the slip speed command w_{sl} through proportional-integral controller and limiter. The slip is added to the speed feedback signal w_r to generate the slip frequency command w_e. The slip frequency command generates the voltage command Vthrough a volts/hertz function generator. A step increase in slip frequency command $w_e$produces a positive speed error and the slip speed w_{sl} is set at the maximum value.

The drive accelerates due to changes in the frequency and current, producing the torque, matching with demand torque. The drives finally settle at a slip speed for which motor torque balances the load torque. Hence, for varying load torque

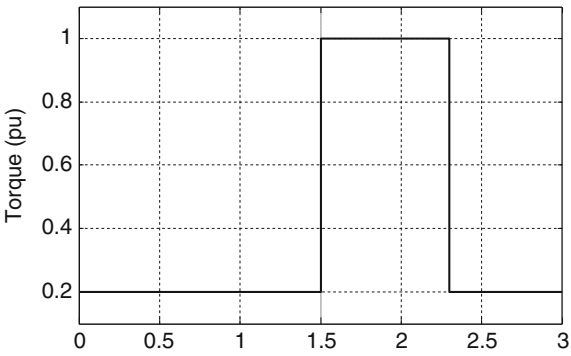


Fig. 26.4 Demand torque characteristic of a specific process machine

with respect to time, the drive generates electromagnetic torque which almost matches with demand torque of the process machine.

26.3 Locus of Operating Point

Here, two cases of varying load torque are considered as shown in Fig. 26.4. In Case 1, load torque suddenly changes from low to high value. In case 2, the load torque changes from high to low value.

26.3.1 Sudden Increase of Load Torque from Low to High Value

From t = 0–1.5 s as shown in Fig. 26.4, value of load torque is at low value, hence required frequency to generate electromagnetic torque matching with demand torque is 50 Hz, which is say at point A on torque slip characteristics plotted at 50 Hz shown in Fig. 26.6a, where induction motor is operating on motoring mode. At t = 1.5 s, the torque suddenly rises to a peak value. In order to meet sudden rise in load torque, required frequency for induction motor, using VVVF method is, say, f1 where f1 < 50 Hz. At this instant, load torque changes its position from A (on T-s curve at 50 Hz) to D point which is on T-s curve plotted at frequency f1. The path required to reach A to D point is from A to B, B to C and C to D. When load torque was at low value, induction motor was operating in motoring mode but as soon as load torque suddenly rises to peak value, induction motor shifts its motoring mode to generating or braking mode at frequency f1 which travels from A to B, B to C to meet the required load torque. C is a point where induction motor runs at synchronous speed corresponding to frequency f1 and changes its operation from generating mode to motoring mode to reach the point D to generate electromagnetic torque equal to load torque. Hence, by plotting T-s curve at 50 Hz and

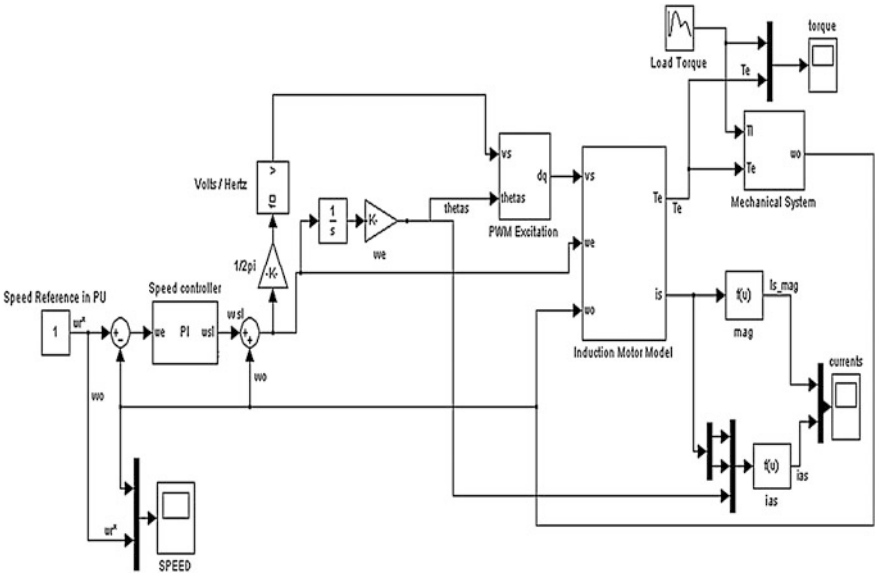


Fig. 26.5 Complete induction motor model with PWM excitation and mechanical system along with v/f control scheme in MATLAB simulink

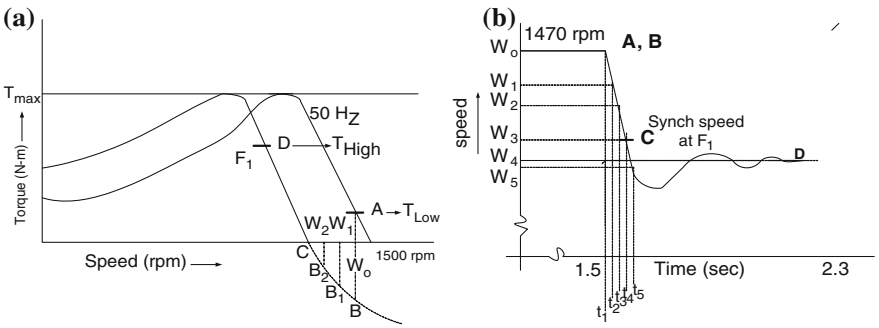


Fig. 26.6 **a** Flow of energy path of induction motor for sudden increase of load torque **b** Reduction of speed of induction motor for sudden increase of load torque

at frequency f1, generating and motoring electromagnetic torques are calculated with speeds varying from A or B to D.

Similarly from Fig. 26.6b, same variations of speeds are calculated at different instant of time when speed of induction motor suddenly reduces to low value because of rise in peak load. Finally, required energy is calculated graphically by plotting generating and motoring electromagnetic torque with respect to different instant of time based on two graphs, shown in Fig. 26.6a and b to meet demand torque characteristics replacing bulky flywheel.

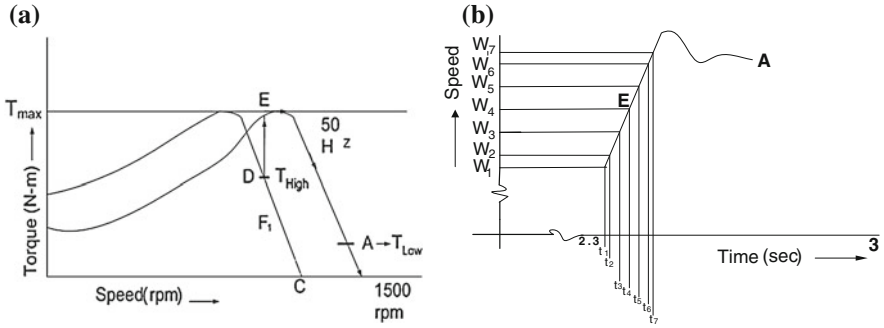


Fig. 26.7 **a** Flow of energy path of induction motor for sudden decrease of load torque **b** Increase of speed of induction motor for sudden decrease of load torque

26.3.2 Sudden Decrease of Load Torque from High to Low Value

At t = 1.5–2.3 s, shown in Fig. 26.4, value of load torque is at high value, hence required frequency to generate electromagnetic torque matching with demand torque is say at frequency, f1, which is a point D on torque slip characteristics plotted at that frequency shown in Fig. 26.7a. Here, induction motor is operating on motoring mode. At t = 2.3 s, the torque suddenly decreases to low value. In order to meet sudden decrease in load torque, required frequency given to induction motor, using VVVF method is say at 50 Hz where 50 Hz > f1. At this instant, load torque changes its position from D point (on T-s curve at F1) to A point (T-s curve plotted at 50 Hz). The path required to reach D to point A is from D to E and E to A. Here induction motor operates only in motoring. Hence by plotting T-s curves both at frequency f1 and at 50 Hz, electromagnetic torques are noted down at different speeds varying from D to A. Similarly, from Fig. 26.7b, same variations of speeds are calculated at different instants of time when speeds of induction motor suddenly increase to high value because of decrease in load torque. Finally, required energy is calculated graphically by plotting electromagnetic torque with respect to different instant of time based on two graphs, shown in Fig. 26.7a and b to meet demand torque characteristics.

26.4 Case Study

A process machine is so selected which comprises of some linkage mechanism as a main processor. The total cycle time of the process machine is 3000 m-Sec. The induction motor rating is three phase, 415 V, 1 hp with a synchronous speed of 1500 rpm. In this case, the average angular velocity of the input crank of the process unit is chosen to be 20 rpm. This gives torque amplification from motor shaft to the process unit shaft of the order of 1500/20 = 75 Induction motor

Table 26.1 1 HP induction motor data

HP	1 = 0.75 kW
Rated voltage	415 V, ±10 % tolerance
Winding connection	Star
Rated frequency	50 Hz
Pair of poles	2
Rated speed	1500 rpm
Stator resistance	12.5487 Ω
Rotor resistance	12 Ω
Stator leakage inductance	144.67 mH
Rotor leakage inductance	144.67 mH
Mutual inductance	545.78 mH
Moment of inertia	0.0018 kg m^2
Friction factor	0.01

generates average supply torque of 0.596 kgf-m (with given torque formula [4]). Thus, the supply torque at the process unit input shaft is 0.596 × 75 = 44.7 kgf-m. Hence, the hp demand of the process unit with a given formula is,

$$\begin{aligned} hp &= \frac{2 \cdot \pi \cdot N \cdot T}{4500} \\ &= \frac{2 \cdot \pi \cdot 20 \cdot (0.596 \times 75)}{4500} \\ &\approx 1.248 \end{aligned}$$

26.5 Simulation and Result

In order to get desired result, the demand torque characteristic as shown in Fig. 26.4, is imposed on VVVF based induction motor drive. The induction motor drive is simulated in the synchronously rotating reference frame per unit basis using MATLAB simulink [5, 9, 10]. The parameters of the sample induction motor are shown in Table 26.1. After simulation, it is observed that induction motor generates similar type of electromagnetic torque with respect to demand torque as shown in Fig. 26.8a. The required frequency to generate electromagnetic torque similar to load torque when it suddenly rises to peak value is 32.16 Hz. with drop in speed is 88.75 rad/sec as shown in Fig. 26.8b and c respectively. Similarly frequency of the supply to the induction motor changes from 32.61 to 50 Hz when load torque decreases to low value as shown in Fig. 26.8d, e shows rise in speed for fall in load torque.

Hence, for general case, knowing the variations in torque during different time intervals, input frequency for the induction motor should be changed suitably.

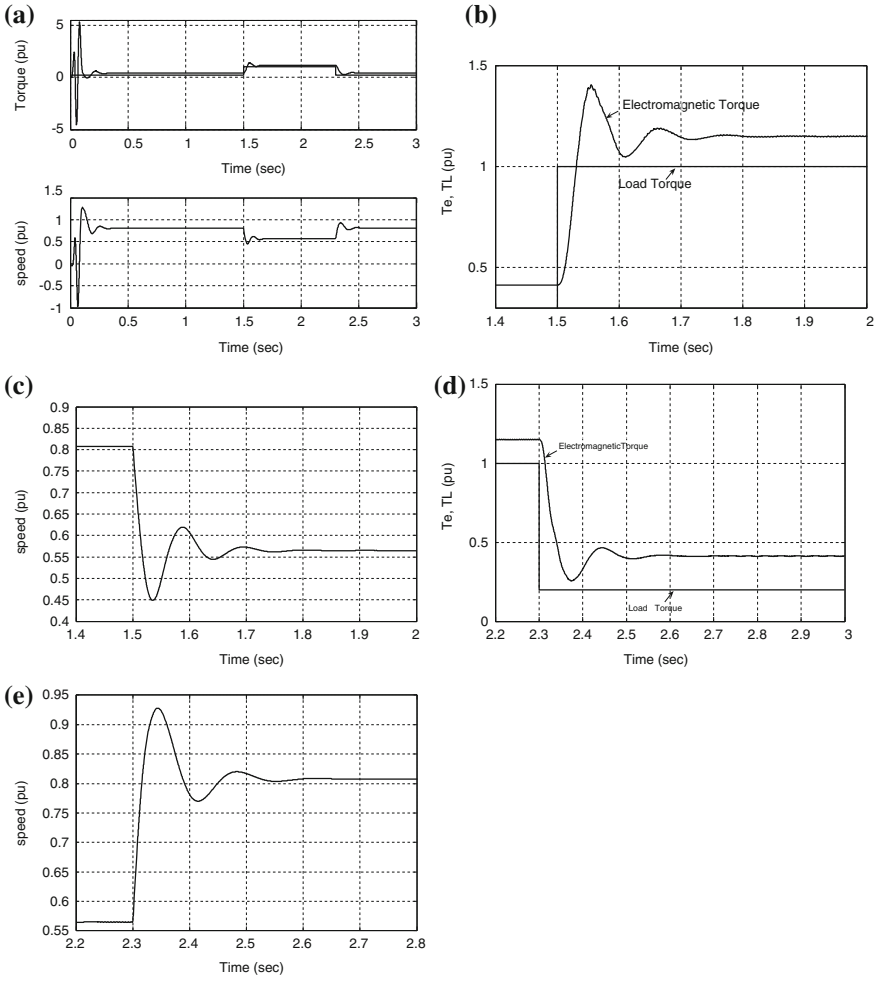


Fig. 26.8 **a** Electromagnetic torque and speed of induction motor for load torque **b** Electromagnetic torque with respect to rise in load torque **c** Drop in speed for increase value of load torque, 88.75 rad/sec **d** Electromagnetic torque with respect to fall in load torque **e** Speed of induction motor for fall in load torque

After knowing the required frequency given to induction motor for two cases, torque slip characteristics for 32.16 Hz (shown in Fig. 26.9) and 50 Hz are plotted by writing programme in M file in MATLAB software. It is noted that when load torque suddenly fluctuate to high value, frequency should be changed from 50 to 32.61 Hz. At that instant, induction motor which was operating in motoring mode (50 Hz) suddenly starts operating in generating mode (32.61 Hz) and follows the path: A to B, B to C. Then it operates in motoring mode from C to D to get desired electromagnetic torque with different values of speeds as shown in Fig. 26.6a

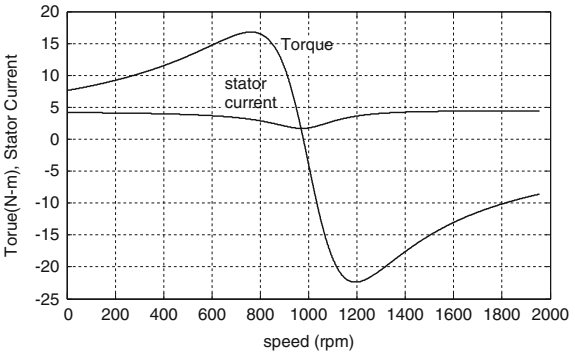


Fig. 26.9 Torque-slip characteristics at 32.61 Hz

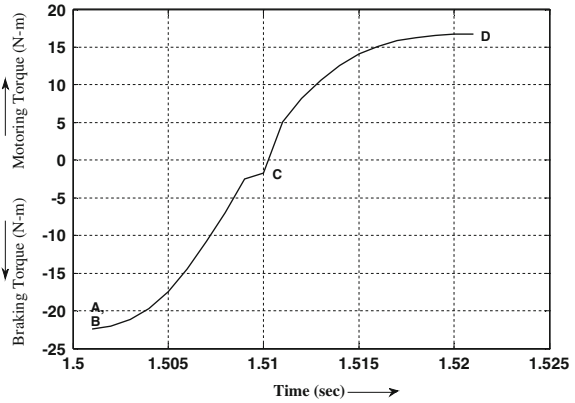


Fig. 26.10 Plot of energy graph, electromagnetic torque versus time for rise in torque

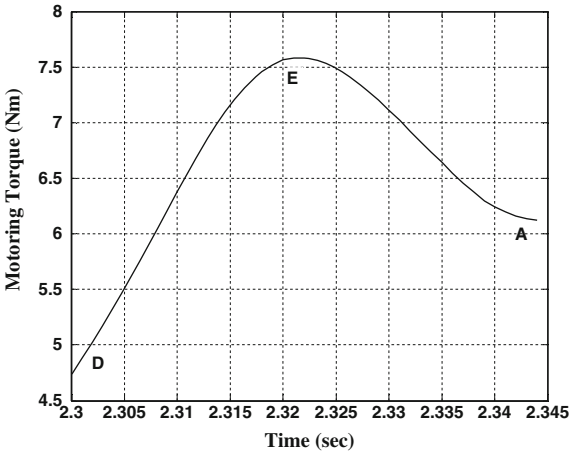


Fig. 26.11 Plot of energy graph, electromagnetic torque versus time for fall in torque

(simulation result is shown in Fig. 26.9). Further, these speeds are calculated at different time interval as shown in Fig. 26.6b (simulation result is shown in Fig. 26.8c). Finally based on two curves, energy graph is plotted graphically which is based on available electromagnetic torque at different time as shown in Fig. 26.10. Similar case happens when load torque changes from high to low value (based on torque slip characteristics at 50 Hz and Fig. 26.8e). Plot of energy graph at this case is shown in Fig. 26.11. Here, motor follows the operating point from D to E and E to A as shown in Fig. 26.7a. In this case, induction motor operates only in motoring mode while following this path. Hence, based on graphically plotted path, behavior of dynamic nature of load torque is studied.

26.6 Conclusion

In order to eliminate bulky flywheel from the process machine having wide fluctuation in load torque, now it is possible to control input side power and frequency of the main drive using VVVF technique, to generate electromagnetic torque characteristics almost matching with demand torque characteristics of the process machine. Based on above technique, the required energy of induction motor is calculated graphically by plotting two curves, one is T-s characteristics at different frequencies (based on load torque) and corresponding speed with respect to time at those frequencies. Induction motor changes its operating mode i.e. from motoring mode to generating mode and then generating to motoring mode to meet demand torque characteristics when load torque sudden changes from low to high value whereas, from high to low value of load torque, operation of induction motor remains in motoring mode. Hence, by graphically plotting instantaneous motor torque as a function of time, dynamic behavior can be analyzed.

References

1. Shigley JE, Mischke CR (1985) Mechanical engineering design, 2nd edn. Tata McGraw Hill, New Delhi
2. Shigley JE, Vicker JJ Jr (1995) Theory of machines & mechanisms, 2nd edn. Tata McGraw Hill International, New Delhi
3. James D, de Ven V Fluidic variable inertia flywheel. American Institute of Aeronautics Foster, and Astronautics, pp 1–6
4. Bose BK (2005) Modern power electronics and AC drives, 5th edn. Prentice Hall PTR, Upper Saddle River
5. Krishnan Bharat R (2008) Electric motor drives, modeling analysis and control. PHI Learning Pvt. Ltd., New Delhi
6. Ozpineci B, Tolbert LM (2003) Simulink implementation of induction model—a modular approach. In: IEEE transaction on industry applications, pp 728–734
7. Sarhann H, Issa R (2006) Improving mechanical characteristics of inverter induction motor drive system. Am J Appl Sci 3 8:1961–1966

8. Ogbuka CU, Nwosu CA (2009) Generalized vector method of induction motor Transient and steady state analysis. Pac J Sci Technol 10(1):52–58
9. Palandurkar MV, Modak JP, Tarnekar SG (2011) Elimination of a flywheel of a process machine by controlling power frequency of the main drive. Int J Eng Sci Technol (IJEST) 3:3580–3590
10. Palandurkar MV, Modak JP, Tarnekar SG (2012) Investigate to substitute large inertia by a combination of very small inertia driven by v/f controlled drive. Int J Sci Eng Res (IJSER—France) 3(4)
11. Shit KL, Chan TF, Wong YK, Ho SL (1999) Modeling and simulation of the three phase induction motor using simulink. Int J Electr Eng Educ 36:163–172, Manchester UP
12. Noaman MN (2008) Speed control for ifoc induction machine with robust sliding mode controller. Asian J Sci Res 4:324–333

Chapter 27
Performance Analysis of Different Current Controllers for Active Power Filter

Dipti A. Tamboli and D. R. Patil

Abstract Power Quality issues are becoming a major concern for today's power system engineers. Large scale incorporation of non-linear loads has the potential to raise harmonic voltages and currents in an electrical distribution system to unacceptable high levels that can adversely affect the system. Active power filter (APF) based on power electronic technology is currently considered as the most competitive equipment for mitigation of harmonics and reactive power simultaneously. Instantaneous power theory is used for generation of reference current. This paper presents a comparative study of the performance of three current control strategies namely ramp comparison method, hysteresis current controller (HCC) and Adaptive hysteresis current controller (AHCC) is carried out and superiority of AHCC is established. Simulation results for all the method are presented using MATLAB/SIMULINK power system toolbox demonstrating the effectiveness of using adaptive hysteresis band.

Keywords Active power filter · Harmonics · Instantaneous power theory · Hysteresis current control · AHCC

D. A. Tamboli (✉) · D. R. Patil
Department of Electrical Engineering,
Walchand College of Engineering, Sangli, India
e-mail: dipti_tamboli@rediffmail.com

D. R. Patil
e-mail: dadasorpatil@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_27,

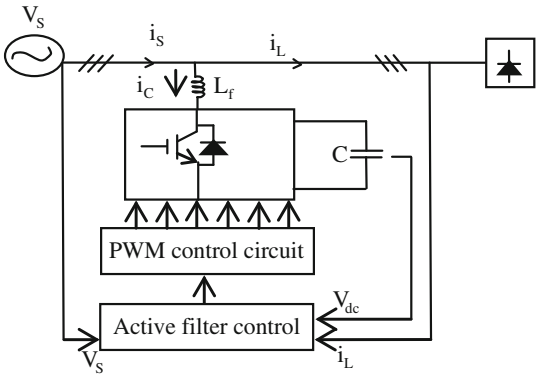


Fig. 27.1 APF control blocks

27.1 Introduction

Recent wide spread of power electronic equipment has caused an increase of the harmonic disturbances and excessive reactive power in the power systems. The harmonics causes problems in power systems and in consumer products such as equipment overheating, capacitor blowing, excessive neutral currents and low power factor. Extensive surveys [1, 2] have been carried out to quantify the problems associated with electric power networks having nonlinear loads. Without the drawbacks of passive harmonic filters, such as component aging, fixed compensation, large size and resonant problems, the active power filter appears to be a viable solution for reactive power compensation as well as for eliminating harmonic currents. The theories and applications of active power filters have become more popular and have attracted great attention since two decades ago [3]. APF consisting of voltage source inverters and a dc capacitor have been researched and developed for improving the power factor and stability of transmission systems. APF have the ability to adjust the amplitude of the synthesized ac voltage of the inverters by means of pulse width modulation or by control of the dc-link voltage, thus drawing either leading or lagging reactive power from the supply.

One of the peculiar features of shunt APFs is that it does not require energy storage units such as batteries or active sources in other forms for its compensation mechanism. To accomplish this function, it requires an effective reference compensation strategy for both reactive and harmonic power of the load. Generally, the performance of APF is based on three design criteria [4–9]: (i) design of power inverter; (ii) types of current controllers used; (iii) methods used to obtain the reference current. Many control techniques have been used to obtain the reference current [5–9]. Among these controllers the instantaneous real-power theory provides good compensation characteristics in steady state as well as transient states. Figure 27.1 shows a system control block for a three-phase shunt APF.

Similarly various current controller techniques proposed for APF configuration, such as triangular-current controller, periodical-sampling controller and hysteresis current controller. In ramp comparison method multiple crossings of the ramp by

the current error may become a problem when the time rate change of the current error becomes greater than that of the ramp. Therefore nowadays, hysteresis current controller method attracts researcher's attention due to unconditional stability, fast transient response, simple implementation and high accuracy. However, this control scheme exhibits several unsatisfactory features such as uneven switching frequency and switching frequency variation within a particular band. The adaptive-hysteresis current controller overcomes these demerits of HCC; adaptive- HCC changes the bandwidth according to instantaneous compensation current variation. This paper presents design and analysis of an active power filter that uses instantaneous power-theory with three types of current controller which generates switching pulses for APF. The shunt APF is investigated under non-linear load and found to be effective for harmonics and reactive power compensation according to IEEE standards.

The simulation is carried out using MATLAB/Simulink for non-linear loads at different firing angles. This entire simulation studies were carried out by choosing an 11 kV feeder providing supply to Walchand College of Engineering, Sangli, India.

27.2 Instantaneous Reactive Power Theory (p-q Theory)

In 3-phase circuits with balanced voltage, instantaneous currents and voltages are converted into instantaneous space vectors. In instantaneous reactive power theory, the instantaneous 3-phase currents and voltages are expressed as the following equations. These space vectors are easily converted into α-β coordinates [10, 11].

$$\begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} = C_{32} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} \tag{27.1}$$

$$\begin{bmatrix} e_\alpha \\ e_\beta \end{bmatrix} = C_{32} \begin{bmatrix} e_a \\ e_b \\ e_c \end{bmatrix} \; where \; C_{32} = \sqrt{2/3} \begin{bmatrix} 1 & -1/2 & -1/2 \\ 0 & \sqrt{3/2} & -\sqrt{3/2} \end{bmatrix} \tag{27.2}$$

and α, β are orthogonal coordinates. eα and iα are on α axis, eβ and iβ are on β axis. When the source supplies nonlinear loads, the instantaneous power delivered to the loads includes both active and reactive components. So, the current vector i was divided into active current component and reactive current component, which are ip and iq respectively, as shown in Fig. 27.2.

In the representation of electric quantities, instantaneous active and reactive powers are calculated as follows:

p = eip, q = eiq, Where, ip = icosφ, iq = isinφ make up Eq. (27.3):

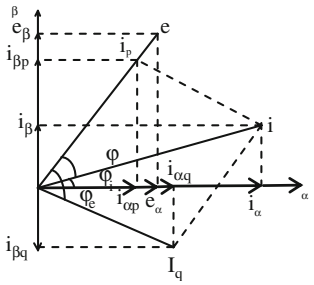


Fig. 27.2 Vector diagram of voltage and currents

$$\begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} \bar{p} + \tilde{p} \\ \bar{q} + \tilde{q} \end{bmatrix} = C_{pq} \begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} where\, C_{pq} = \begin{bmatrix} e_\alpha & e_\beta \\ -e_\beta & e_\alpha \end{bmatrix} \tag{27.3}$$

Here, "$-$" and "$\sim$" stand for dc and ac components, respectively. $\bar{p}$and $\bar{q}$ are the instantaneous active and reactive power (dc value) originating from the symmetrical fundamental (positive-sequence) component of the load current, $\tilde{p}$ and $\tilde{q}$ are the instantaneous active and reactive power (ac value) originating from harmonic and the asymmetrical fundamental (negative-sequence) component of the load current. These power quantities given above for an electrical system are represented in a-b-c coordinates and have the following physical meaning:

$\bar{p}$ = The mean value of the instantaneous active power—corresponds to the energy per time unit transferred from the power supply to the load, through a-b-c coordinates.

$\tilde{q}$ = Alternated value of the instantaneous active power—it is the energy per time unit that is exchanged between the power supply and the load through a-b-c coordinates.

$\bar{q}$ = Instantaneous reactive power—corresponds to the power that is exchanged between the phases of the load, but is responsible for the existence of undesirable currents, which circulate between the system phases.

$\tilde{q}$ = The mean value of the instantaneous (conventional) reactive power.

From Eq. (27.3), in order to measure the harmonic currents and reactive current component, fundamental active current corresponding to reactive power on α-β coordinates should be first calculated and then transformed into a-b-c reference frame shown in by Eqs. (27.4) and (27.5) respectively.

$$\begin{bmatrix} i_{\alpha f} \\ i_{\beta f} \end{bmatrix} = C_{pq}^{-1} \begin{bmatrix} p + p_{loss} \\ 0 \end{bmatrix} \tag{27.4}$$

$$\begin{bmatrix} i_{af} \\ i_{bf} \\ i_{cf} \end{bmatrix} = \sqrt{2/3} \begin{bmatrix} 1 & 0 \\ -1/2 & \sqrt{3/2} \\ -1/2 & -\sqrt{3/2} \end{bmatrix} \begin{bmatrix} i_{\alpha f} \\ i_{\beta f} \end{bmatrix} \tag{27.5}$$

Finally, the reference compensation currents are obtained by Eq. (27.6)

$$\begin{bmatrix} i_a^* \\ i_b^* \\ i_c^* \end{bmatrix} = \begin{bmatrix} i_{La} \\ i_{Lb} \\ i_{Lc} \end{bmatrix} - \begin{bmatrix} i_{af} \\ i_{bf} \\ i_{cf} \end{bmatrix} \tag{27.6}$$

The DC side voltage of APF should be controlled and kept at a constant value to maintain the normal operation as well as to recover the energy loss due to conduction and switching power losses associated with the diodes and IGBTs of the inverter in APF, which tend to reduce the value of Vdc across capacitor Cdc. A feedback voltage control circuit is used and an error function which is difference between the reference value, Vref and the feedback value (Vdc), passes through a PI controller and the output of the PI regulator is added in the alpha axis value of the fundamental current components.

27.3 Current Controller Techniques

27.3.1 Ramp Comparison Controller

The controller can be thought of as producing sine-triangle PWM with the current error considered to be the modulating function. The current error is compared to a triangle waveform and if the current error is greater(less) than the triangle waveform, and then the inverter leg is switched in the positive (negative) direction with the frequency of the triangular wave. Multiple crossings of the ramp by the current error may become a problem when the time rate change of the current error becomes greater than that of the ramp. However, such problems can be adjusted by changing the amplitude of the triangle wave suitably.

27.3.2 Hysteresis Controller

The hysteresis band current control technique has proven to be most suitable due to its unconditioned stability, very fast response, good accuracy, and inherent-peak current limiting capability [12].

The conventional hysteresis band current control scheme is shown in Fig. 27.3 decides the switching pattern of active power filter [15]. The reference line current of APF is referred to as I_{ca}^*, and actual line current I_{ca}. The switching logic is formulated as follows: If $i_{ca} < (I_{ca}^* - HB)$ upper switch is OFF and lower switch is ON for leg "a" :(SA = l). If $i_{ca} > (I_{ca}^* + HB)$ upper switch is ON and lower switch is OFF for leg "a" (SA = O). The switching functions SB and SC for phases B and C are determined similarly, using corresponding reference and measured currents and hysteresis bandwidth (HB). The switching frequency of the hysteresis band current control method described above depends on how fast the

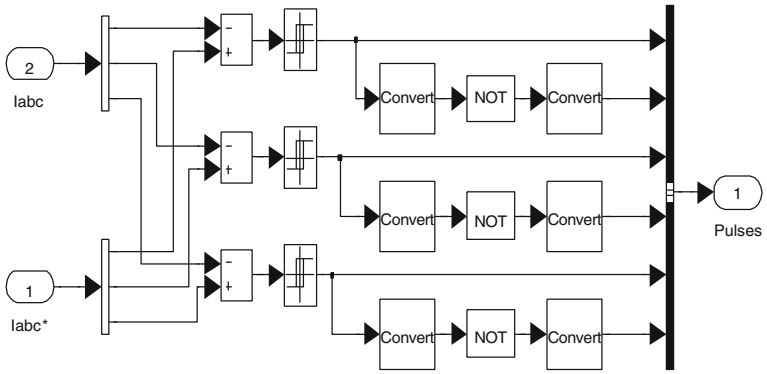


Fig. 27.3 Simulation diagram of HCC

current changes from the upper limit of the hysteresis band to the lower limit of the hysteresis band, or vice versa.

27.3.3 Adaptive Hysteresis Current Controller

Width of hysteresis band determines the allowable current shaping error to control the switching frequency of the inverter. The bandwidth should also be small enough to supply the reference current precisely keeping the view of switching losses and EMI related problems [12]. Therefore, the range of switching frequencies used is based on a compromise between these factors. Switching frequency of the hysteresis band current controller is depends on the rate of change of the actual APF current and therefore switching frequency varies along with the current waveform. Figure 27.4 shows the PWM current and voltage waves for phase a [13]. When the actual line current of the active power filter tries to leave the hysteresis band, the suitable IGBT is switched to ON or OFF to force the current to return to a value within the hysteresis band. Then the switching pattern will be trying to maintain the current inside the hysteresis band. Current i_{ca} tends to cross the lower hysteresis band at point 1, where upper side IGBT of leg 'a' is switched on. The linearly rising current i_{ca}^{+} then touches the upper band at point 2, where the lower side IGBT of leg 'a' is switched on. The following equations can be written in the respective switching intervals t1 and t2.

$$L\frac{di_a^+}{dt} = (0.5\,V_{dc} - V_s) \tag{27.7}$$

$$L\frac{di_a^-}{dt} = -(0.5\,V_{dc} - V_s) \tag{27.8}$$

Where L is phase inductance, and $i_{ca}^{+}$and i_{ca}^{-} are the respective rising and falling current segments. From the geometry of Fig. 27.4, we can write

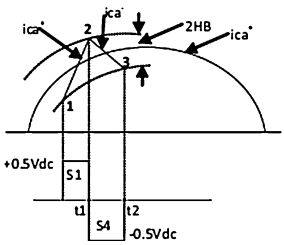


Fig. 27.4 Current and Voltage with AHCC

$$\frac{di_a^+}{dt} + \frac{di_a^-}{dt} = 0 \tag{27.9}$$

$$\frac{di_a^+}{dt} t_1 + \frac{di_a^*}{dt} t_1 = 2 * HB \tag{27.10}$$

$$\frac{di_a^-}{dt} t_2 + \frac{di_a^*}{dt} t_1 = -2 * HB \tag{27.11}$$

$$t_1 + t_2 = T_c = \frac{1}{f_c} \tag{27.12}$$

Where t1 and t2 are the switching intervals and fc is the switching frequency. Adding (27.10) and (27.11) and substituting in (27.12) again subtracting (27.11) from (27.10), we get,

$$t_1 \frac{di_a^+}{dt} + t_2 \frac{di_a^*}{dt} - \frac{1}{f_c} \frac{di_{ca}^*}{dt} = 0 \tag{27.13}$$

$$4HB = \frac{di_a^+}{dt} t_1 - \frac{di_a^-}{dt} t_2 - (t_1 - t_2) \frac{di_a^*}{dt} \tag{27.14}$$

Substituting (27.9) in (27.13) and (27.14) and simplifying

$$4HB = \frac{di_a^+}{dt} t_1 - \frac{di_a^-}{dt} t_2 - (t_1 - t_2) \frac{di_a^*}{dt} \tag{27.15}$$

$$(t_1 - t_2) = \left(\frac{di_{ca}^*}{dt}\right) \Big/ f_c \left(\frac{di_a^+}{dt}\right) \tag{27.16}$$

$$HB = \left\{ \frac{0.125 V_{dc}}{f_c L} \left[1 - \frac{4L^2}{V_{dc}^2} \left(\frac{V_s}{L} + m \right)^2 \right] \right\} \tag{27.17}$$

Where m is the slope of command current wave. Hysteresis band (HB) can be modulated at different points of fundamental frequency to control the switching patterns of the inverter. For symmetrical operation of all three phases, it is expected that the hysteresis band width (HB) profiles HBa, HBb and HBc will be

the same, but have phase difference. The variable HCC is used to produce gate control pulses that operate the voltage source inverter.

27.4 Simulation Result

A typical distribution feeder of 11 kV originating from a sub-station to Walchand College of Engineering Sangli campus load centers has been modeled and considered for simulation. A three-phase 11 kV/433 V, Dy11 transformer is employed in the Institute for catering to the loads locally. Therefore the values for source impedance come out to be 0.02871 Ω, 0.2047 mH. Source voltage is considered as 440 V and of 50 Hz frequency. Filter parameters are selected as Lf = 4 mH, Vdc = 850 V, Cdc = 1400 μF.

The performances of ramp comparison method, HCC and AHCC based shunt active power filter were evaluated through simulation using MATLAB/SIMULINK environment. A thyristor converter with R-L load is taken as t = 0 to t = 0.1 with resistance of 100 Ω and inductance of 50 mH under steady state and for remaining transient period R-L with 50 Ω and inductance of 25 mH.

The simulation results in transient operation using AHCC are presented in Fig. 27.5 The Waveform of the source current without APF and its THD are shown in Fig. 27.5a and b respectively. These current waveforms are for a particular phase (phase a). Other phases are not shown as they are only phase shifted by 120^0 and we have considered only a balanced load. Also the Waveform of the source current with APF and its THD are shown in Fig. 27.5c and d respectively. The APF supplies the compensating current to PCC, which is shown in Fig. 27.5e. The time domain response of the p-q theory controller is shown in Fig. 27.5f which clearly indicates that, the controller output settles after a few cycles. The capacitor voltage superimposed to its reference is shown in Fig. 27.5f. In order evaluate the good performance of the control, the total harmonic distortion (THD) is measured for the source current before and after compensation. It shows that THD improves from 27.78 % to 1.49 % using Adaptive hysteresis current controller than other two methods.

The system is simulated at different operating conditions such as thyristor convertor with firing angle of 0^0 and 30^0. The final values of THD of source current before and after compensation for all three methods are listed in Table 27.1. The source current is giving better result using AHCC than other controllers and their THD's are below the specifications prescribed by IEEE 519 standard recommendations on harmonics levels. As shown in Table 27.1 reactive power is also reduced substantially which shows improvement in power factor to nearly unity.

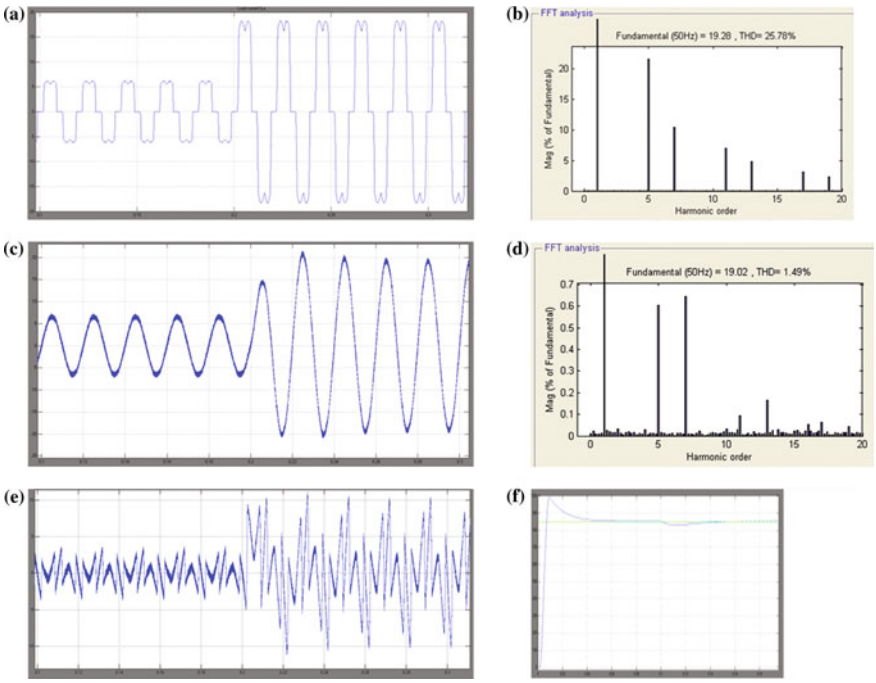


Fig. 27.5 **a** Source current without APF **b** THD of I_{sa} without APF **c** Source current without APF **d** THD of I_{sa} with APF **e** Compensating current **f** V_{dc} with its reference

Table 27.1 Comparison of %THD of source and load currents and powers

Type of controller	Load THD %			Source THD %			PQ before	PQ after
$\alpha = 0^0$								
SPWM	25.79	25.8	25.79	1.6	1.6	1.61	10.22 KW, 1.56 KVAR	10.27 KW, 4.97 VAR
HCC (h = 0.9)	25.76	25.76	25.76	1.57	1.59	1.64	10.23 KW, 1.55 KVAR	10.27 KW, 7.1 VAR
AHCC	25.78	25.78	25.78	1.49	1.51	1.48	10.22 KW, 1.55 KVAR	10.27 KW, 1.29 VAR
$\alpha = 30^0$								
SPWM	29.99	30.02	30.01	2.92	2.95	2.96	8.51 KW, 4.01 KVAR	8.54 KW, 7.9 VAR
HCC (h = 0.9)	30.00	30.02	30.01	2.71	2.86	2.84	8.504 KW, 4.00 KVAR	8.54 KW, 9.33 VAR
AHCC	30.00	30.03	30.01	2.79	2.83	2.77	8.5 KW, 4.00 KVAR	8.54 KW, 3.4 VAR

27.5 Conclusion

An AHCC has been implemented for three phase shunt active power filter. The instantaneous power theory is used to extract the reference currents from the distorted line currents. This facilitates enhancement of power quality through reactive power compensation and harmonics suppression due to nonlinear load. The results obtained indicate that DC-capacitor voltage and the harmonic current can be controlled easily for various load conditions. The performance of the AHCC, fixed HCC and ramp controller technique shunt active power filter are verified and compared with the simulation results. The THD of the source current after compensation is 1.49 % which is less than 5 %, the harmonic limit imposed by the IEEE-519 standard.

References

1. Akagi H (1996) New trends in active filters for power conditioning. IEEE Trans Ind Appl 32(6):1312–1322
2. Emadi A, Nasiri A, Bekiarov S (2004) A book, "on uninterruptible power supplies and active filters". CRC Press, Boca Raton
3. Akagi H (2006) Modern active filters and traditional passive filters. Bull Pol Acad Sci 54(3):255–256
4. Akagi H, Kanazawa Y, Nabae A (1984) Instantaneous reactive power compensators comprising switching devices without energy storage components. IEEE Trans Ind Appl IA-20(3):625–630
5. Bhattacharya S, Divan DM, Banerjee B (1991) Synchronous reference frame harmonic isolator using series active filter. In: 4th European power electronic conference, Florence, vol 3, pp 30–35
6. Singh B, Al-Haddad K, Chandra A (1997) Active power filter with sliding mode control. In: Proceedings of institute electrical engineer, generation transmission and distribution, vol 144, pp 564–568
7. Bhattacharya S, Veliman A, Divan A, Lorenz R (1995) Flux based active power filter controller. In: Proceedings of IEEE-IAS Annual Meeting Record, pp 2483–2491
8. Jou H (1995) Performance compression of the three-phase active power filters algorithms. In: Proceedings of Institute Electrical Engineer, Generation Transmission and Distribution, vol 142, pp 646–652
9. Dixon J, Garcia J, Moran I (1995) Control system for three-phase active power filters which simultaneously compensates power factor and unbalanced loads. IEEE Trans Ind Electron 42(6):636–641
10. Cui Y-l, Liu H, Wang J-q, Sun S-g () Simulation and reliability analysis of shunt active power filter based on instantaneous reactive power theory. J Zhejiang Univ Sci an ISSN
11. Akagi H, Nabae A, Atoh S (1986) Control strategy of active power filters using multiple voltage source PWM converters. IEEE Trans Ind Appl IA-22(3):460–465
12. Kale M, Ozdemir E (2003) A novel adaptive hysteresis band current controller for shunt active power filter. In: Proceedings IEEE Conference on Control Applications
13. Bose BK (1990) An adaptive hysteresis band current control technique of a voltage feed PWM inverter for machine drive system. IEEE Trans Ion Ind Electron 37(5):402–406

Chapter 28
Optimum LQR Switching Approach for the Improvement of STATCOM Performance

L. Yathisha and S. Patil Kulkarni

Abstract Static Synchronous Compensator (STATCOM) is a device capable of solving the power quality problems in the power system. These problems happen in milliseconds and because of the time limitation; it requires the STATCOM that has continuous reactive power control with fast response. In this paper, an optimum Linear Quadratic Regulator (LQR) switching approach for STATCOM control is introduced to improve the performance by achieving the optimum performance between peak overshoot and settling time. Results are compared with the earlier conventional LQR approaches.

Keywords STATCOM · LQR · MIMO · FACTS

28.1 Introduction

Reactive power control is critical consideration in improving the quality of power systems. Reactive power increases transmission losses, degrades power transmission capability and decreases the voltage regulation at the load end [1]. In the past, Thyristor-Controlled Reactors (TCR) and Thyristor-Switched Capacitors were applied for reactive power compensation. However, with the increasing power rating achieved by solid-state devices, STATCOM is taking place as one of the new generation flexible AC transmission systems (FACTS) devices. The

L. Yathisha (✉) · S. P. Kulkarni
Departrment of E&C, SJ College of Engineering, Mysore, India
e-mail: yathisha_171@yahoo.co.in

S. P. Kulkarni
e-mail: pk.sudarshan@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_28,

STATCOM is normally designed to provide fast voltage control and to enhance damping of inter-area oscillations. A typical method to meet these requirements is to superimpose a supplementary damping controller up on the automatic voltage control loop [2]. Many of the methods focus on decoupling the system variables and designing PI controllers. However, STATCOM based power system is a Multiple Input Multiple Output (MIMO) system, where it is not possible to totally decouple the system variables.

The earlier conventional control methods, apply state feedback control techniques [3, 4]. These methods demonstrate the improvement in current control response compared with simple LQR and pole placement methods. In the LQR control method of [3] the weighting matrices Q and R are selected by a trial and error method. Specifically, for LQR1, Q = diag (7.88, 0.024); R = diag (1500.6); for LQR2 Q = diag (19.1, 0.007); R = diag (1105.4) and for LQR3, Q = diag (7.88, 0.024); R = diag (1500.6).

In the current paper, a switching strategy is suggested to switch between controllers LQR1 & LQR2 or between LQR1 & LQR3. A new set of LQR controllers is also proposed that are obtained by selecting more simpler weighting matrices Q = diag(1;1); R = diag(1;1) for *LQR11* and Q = diag(1;1); R = diag(0.01;1) for *LQR'11* to switch between them such that the performance of the STATCOM is improved.

The remainder of the paper is organized as follows. Section 28.2 describes the system configuration and modelling for a STATCOM connected in a distribution system. The design of the LQR control algorithm is detailed in Sect. 28.3. Section 28.4 describes the proposed optimum LQR switching approach for the existing STATCOM model with optimum LQR controllers along with the proposed switching rule. Results and analysis follow in the concluding section.

28.2 System Configuration and Modelling

In this section, a simplified IGBT based STATCOM system is described [5]. By firing the three-phase IGBT Bridge appropriately, a requested bridge side voltage can be generated and current through line impedance R and L is controlled. Eqs (28.1)–(28.4) give a mathematical expression of STATCOM. The variable $\omega$is the angular power frequency, and subscripts d, q represent variables in rotating d-q coordinate system.

$$\frac{di_d}{dt} = -\frac{R}{L} + \omega i_q + \frac{1}{L}(V_{td} - V_{sd}) \tag{28.1}$$

$$\frac{di_q}{dt} = -\omega i_d - \frac{R}{L} i_q + \frac{1}{L}(V_{tq} - V_{sq}) \tag{28.2}$$

$$\frac{dV_{dc}}{dt} = -3\frac{(V_{td} + V_{tq}i_d)}{2CV_{dc}} - \frac{i_L}{C} \tag{28.3}$$

$$Q = \frac{3}{2}(V_{sq}i_{sd} - V_{sd}i_{sq}) \tag{28.4}$$

Consider a linear system,

$$\dot{X} = AX + BU \tag{28.5}$$

$$Y = CX \tag{28.6}$$

Writing (28.1), (28.2) in the state space format as (28.5), (28.6) corresponding matrices can be found as

$$A = \begin{bmatrix} -\frac{R}{L} & \omega \\ -\omega & -\frac{R}{L} \end{bmatrix} B = \begin{bmatrix} \frac{1}{L} & 0 \\ 0 & \frac{1}{L} \end{bmatrix} C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} X = \begin{bmatrix} i_d \\ i_q \end{bmatrix} U = \begin{bmatrix} V_{td} - V_{sd} \\ V_{tq} - V_{sq} \end{bmatrix}$$

28.3 LQR Control Algorithm

The conventional MIMO design approach for STATCOM used in the previous works is the linear quadratic regulator (LQR) optimal control method. This method determines the feedback gain matrix that minimizes in order to achieve some compromise between the use of control effort and the speed of response that will guarantee a stable system. For a given linear system in (28.5).

Determine the matrix K of the LQR vector: $U(t) = -KX(t)$

In order to minimize the performance index: $J = \frac{1}{2}\int_0^\infty (X^T QX + U^T RU)dt$

Where Q and R are the positive-definite Hermitian or real symmetric matrixes, the matrix Q *and* R determine the relative importance of the error and the expenditure of this energy. From the above Equations: $K = -R^{-1}B^T P$

And hence the control law is: $U(t) = -KX(t) = -R^{-1}B^T PX(t)$

In which P must satisfy the reduced Riccati Eq: $PA + A^T P - PBR^{-1} + B^T P + Q = 0$.

The LQR design selects the weight matrix Q and R such that the performances of the closed loop system can satisfy the desired requirements mentioned earlier.

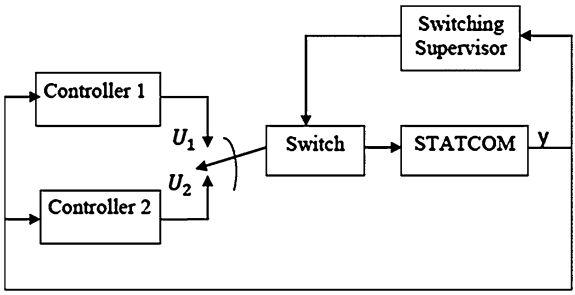


Fig. 28.1 General implementation of switched linear systems

28.4 Proposed Optimum LQR Switching Approach

In this section, mathematical modeling of statcom system with switched linear systems and the proposed switching algorithm will be explained

28.4.1 Switched Linear Systems

Switched systems are composed of a group of sub-systems guided by a switching law that governs the change among these subsystems. Use of appropriate switching in control has proved to give better performance when compared to the performance of a system without switching control. A switched linear system model (refer Fig. 28.1) for the current problem is as follows:

$$\bar{X}(t) = A_{\theta(t)}X(t) \tag{28.7}$$

$$Y(t) = CX(t) \tag{28.8}$$

The switching strategy $\theta(t)$ shown in (28.7) takes values 1 and 2 based on switching rule decided by supervisor leading to closed loop $A_1 = A - BK_1$ and $A_2 = A - BK_2$. Here the two controller gains K_1 and K_2 model the LQR1 & LQR2 of [3] respectively in one experiment and LQR1 & LQR3 of [3] respectively in another experiment. Similarly K_1 and K_2 also model the proposed new controllers *LQR11 & LQR'11* respectively that are based on trivial weighting matrices as mentioned in Sect. 28.1.

28.4.2 Performance Based Switching Strategy

Following theorem from [7] is useful for the purpose of this paper.

Theorem 1 The two primary and secondary state variable feedbacks K_1 and K_2 where $A_1 = A - BK_1$ is known to be asymptotically stable. For an initial state $X = x$, the use of secondary strategy K_2 for some time $T > 0$ will be beneficent if

$$\sigma(x,t) = \prec X, e^{A_2^T T}(T(t) - T_0)e^{A_2 T}x \succ \; > 0$$

The above theorem shows that if the secondary feedback is applied for t seconds, the performance of the system is improved. From the above theorem it was difficult to establish a time of switching. Here one can note that since $T(0) = T_0$, the function $\sigma(x,t)$ is zero for $t = 0$ is positive, then we can say that there will be a time interval of length $t > 0$ where the function σ will be positive. Thus one can infer that if for an initial state x, the following conditions hold

$$\frac{d\sigma(x,t)}{dt}|_{t=0} = \prec x, \left(A_2^T T_0 A_2 + C^T C\right)x \succ \; > 0,$$

Then for small enough values of t functionσ will be positive and secondary control is beneficial. This knowledge can be formulated into following switching algorithm.

28.4.3 Switching Algorithm

The switching control algorithm based on [6, 7] can be explained in following steps:

1. Define K_1as the primary controller and K_2as the secondary controller. Where $A_1 = A - BK_1$ asymptotically stable and $A_2 = A - BK_2$ not necessarily stable.
2. Determine T_0 by solving algebraic Lyapunov Eq: $A_1^T T_0 + T_0 A_1 = -C^T C$.
3. Using, $A_2 = A - BK_2$ define switching matrix: $S = -(A_2^T T_0 + T_0 A_2 + C^T C)$
4. Now, the switching rule is, use secondary controller K_2 with: $\theta(t) = 2$ if $\prec x, Sx \succ \; > 0 = 1$ Otherwise

28.5 Results and Conclusion

The experimental set-up to test the proposed algorithm consists of LQR controller gains K_1(LQR1), K_2 (LQR2) &K_3 (LQR3) described by A and B matrices along with the switching matrices are given below.

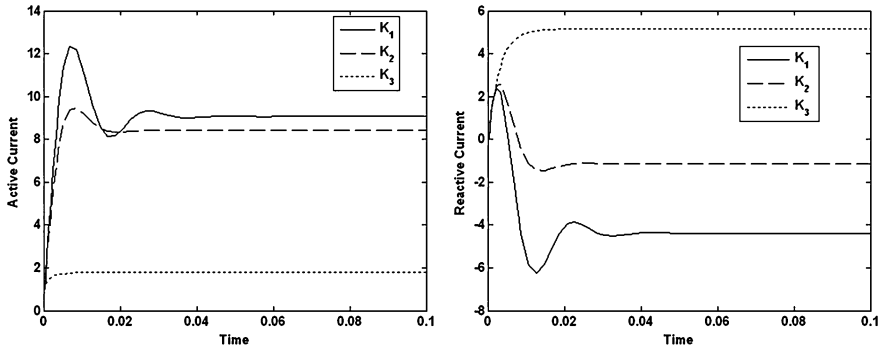


Fig. 28.2 Statcom response with LQR controllers

$$\mathrm{A} = \begin{bmatrix} -5 & 314.5 \\ -314.5 & -5 \end{bmatrix} S = -\left(A_U^T \Gamma_0 + \Gamma_0 A_U + C^T C\right) \mathrm{B} = \begin{bmatrix} 2500 & 0 \\ 0 & 2500 \end{bmatrix}$$

$$\mathrm{K}_1 = \begin{bmatrix} 0.0520 & 0.0090 \\ 0.0090 & 0.0448 \end{bmatrix} S = -\left(A_U^T \Gamma_0 + \Gamma_0 A_U + C^T C\right) \mathrm{K}_2 = \begin{bmatrix} 0.1028 & 0.0229 \\ 0.0229 & 0.0705 \end{bmatrix} \mathrm{K}_3 = \begin{bmatrix} 0.6663 & 0.0874 \\ 0.0874 & 0.1181 \end{bmatrix}$$

$$\mathrm{S}(\mathrm{K}_1 \& \mathrm{K}_2) = \begin{bmatrix} 0.4579 & 0.3308 \\ 0.3308 & 0.9790 \end{bmatrix} \mathrm{S}(\mathrm{K}_1 \& \mathrm{K}_3) = \begin{bmatrix} 11.0013 & 0.5355 \\ 0.5355 & 1.7925 \end{bmatrix}$$

The new LQR controllers K_{11}(*LQR11)* & K'_{11} (*LQR'11*) along with new switching matrix is also given below.

$$\mathrm{K}_{11} \begin{bmatrix} 0.9980 & 0 \\ 0 & 0.9980 \end{bmatrix} S = -\left(A_U^T \Gamma_0 + \Gamma_0 A_U + C^T C\right) \mathrm{K}'_{11} \begin{bmatrix} 10.0103 & -1.017 \\ -0.0102 & 0.9915 \end{bmatrix} \mathrm{S}(\mathrm{K}_{11} \& \mathrm{K}'_{11}) \begin{bmatrix} 9.3801 & -0.9419 \\ -0.2265 & -0.8708 \end{bmatrix}$$

The dynamic response curves for the two state space variables active current $(\mathrm{i_d})$, and reactive current $(\mathrm{i_q})$ are plotted as shown in Figs. 28.2, 28.3, 28.4 with the legend $\mathrm{K}_1, \mathrm{K}_2, \mathrm{K}_3, \mathrm{K}_{11}, \mathrm{K}'_{11}$, Switch between $\mathrm{K}_1 \& \mathrm{K}_2$, Switch between $\mathrm{K}_1 \& \mathrm{K}_3$ and Switch between $\mathrm{K}_{11} \& \mathrm{K}'_{11}$ In order to show the effectiveness of the proposed system peak overshoot and settling time response specifications is also tabulated.

From Fig. 28.2 with the different controllers, LQR2 gives better results compared to LQR1 and LQR3 gives better results compared to LQR2. The proposed system (switch between LQR1 and LQR2) in Fig. 28.3 improves the performance in the reactive current (i_q) state variable and the other proposed system (switch between LQR1 and LQR3) in Fig. 28.3 improves the performance in both active current (i_d) and the reactive current (i_q) state variables. The proposed new LQR switching controllers *LQR11* and *LQR'11* in Fig. 28.4 also improves performance of two state variables.

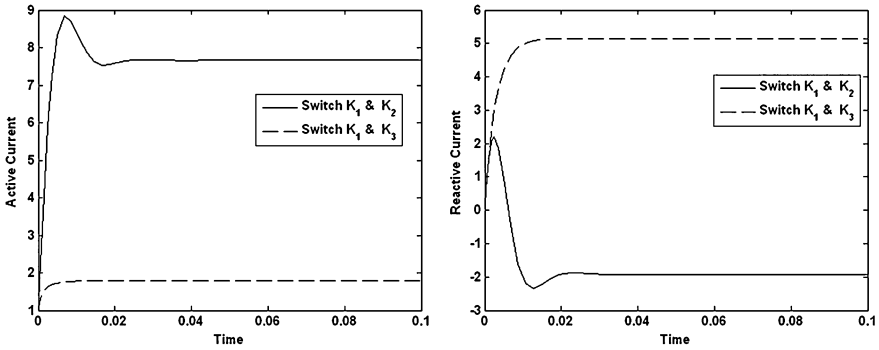


Fig. 28.3 Statcom response with proposed optimum LQR switching controllers

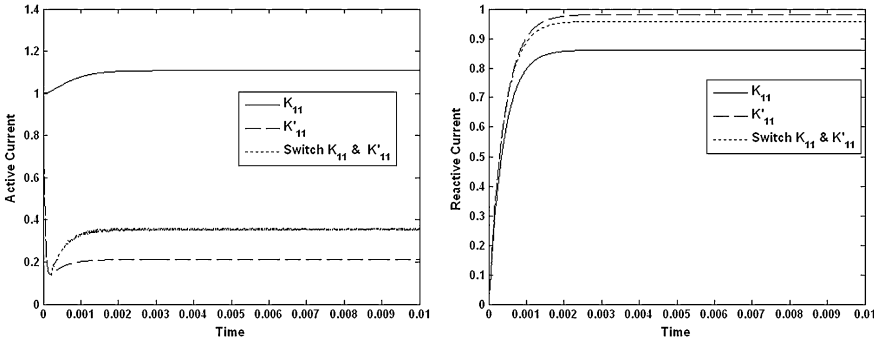


Fig. 28.4 Statcom response with LQR and proposed optimum LQR switching controllers

Table 28.1 System response for active and reactive current state variables

Response(i_d)	K_1	K_2	K_3	Switch K_1&K_2	Switch K_1&K_3	K_{11}	K'_{11}	Switch K_{11}&K'_{11}
Peak overshoot in %	250	100	0	100	0	0	80	50
Settling Time in ms	60	20	7	25	8	0.15	0.15	0.1
Response (i_q)	K_1	K_2	K_3	Switch K_1&K_2	Switch K_1&K_3	K_{11}	K'_{11}	Switch K_{11}&K'_{11}
Peak overshoot in %	250	120	0	30	0	0	0	0
Settling time in ms	30	22	15	22	15	20	22	17

From Figs. 28.2, 28.3, 28.4 and Table 28.1, one can conclude that the proposed optimum LQR switching approach provides optimum performance between peak

overshoot and settling time in the two state space variables of active and reactive currents simultaneously compared to system response with optimally controlled individual controllers with respect to K_1, K_2, K_3 and also with the new controllers K_{11}& K'_{11}..

References

1. Xing L (2003) A comparison of pole assignment and LQR design methods for multivariable control for STATCOM, M.Sc. dissertation, Florida State University
2. Rao P, Crow ML, Yang Z (2000) STATCOM control for power system voltage control applications. IEEE T Power Deliver 15(4):1311–1317
3. Eshtehardiha S, Shahgholian GH (2007) Improvement of STATCOM performance with optimum LQR and pole placement controller based on genetic algorithm. First joint congress on fuzzy and intelligent systems, Ferdowsi University of Mashhad, Iran, 29–31 Aug 2007
4. Eshtehardiha S, Shahgholian GH (2007) Coordinating the multivariable state-feedback controller on static synchronous compensator with genetic algorithm. International conference on intelligent and advanced systems, IEEE, Ferdowsi University of Mashhad, pp 864–869
5. Ajami A, Taheri N (2011) A hybrid fuzzy/LQR based oscillations damping controller using 3-level STATCOM. Int J Comp Electr Eng 3(2):184–189
6. Lalitha SD (2005) Performance based switching control for single input linear time invariant systems. M.S thesis, Department of Electrical and Electronics Engineering, Louisiana State University
7. Yathisha L, PatilKulkarni S, Ananda Murthy RS (2010) Hybrid modelling and switching algorithm for power system with FACTS based controllers. In: Proceedings of international conference on system dynamics and control, pp 367–372

Chapter 29
Squirrel Cage Rotor Design for Safety and Reliability Improvement of a Three Phase Induction Machine

Lokesh Varshney, Vikas Varshney, Albert Newwel and R. K. Saket

Abstract This paper presents design of a squirrel cage rotor for safety and reliability improvement a three phase induction machine. Induction motor with designed rotor is capable for use in overloading conditions. Presented rotor design reduces problems related to temperature rise, tooth pulsation loss, cooling, vibration, crawling, cogging and voltage ripples in three phase induction machine. This design reduces pulsation loss, unbalanced magnetic pull and noise in the machine with increase in the air gap length. Overload capacity and reliability improves with increase the air gap length of the machine.

Keywords Three phase induction machine · Air gap length · Rotor slots · End ring · Rotor core

29.1 Introduction

Engineering design is application of science, technology and invention to produce machines to perform specified tasks with optimum economy and efficiency. The problem of design and manufacture of electric machinery is to built, as economically as possible, a machine which fulfils a certain set of specifications and guarantees. The major considerations to evolve a good design are cost, durability

L. Varshney (✉) · A. Newwel · R. K. Saket
Electrical Engineering Department, Institute of Technology, Banaras Hindu University, Varanasi, Uttar Pradesh, India
e-mail: lokesh_varshney@rediffmail.com

V. Varshney
Maintenance Engineer Division M33, NDMC, New Delhi, India

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_29,

and compliance with performance criteria. It is impossible to design a machine which is cheap and is also durable at the same time. This is because a machine which is to have a long life span must use high quality materials and advanced manufacturing techniques which obviously make it costly. A good design is one where the machine has reasonable operating life, reliability, durability and has a low initial cost. A design sheet of a squirrel cage rotor of given stator of a three phase induction machine has been developed by using Matlab program. Program is designed with keeping some important features in mind like: overload capacity, iron loss, copper loss, temperature rise, tooth pulsation loss, leakage reactance, cooling, vibration, noise, crawling, cogging, voltage ripples, etc. Rotor is designed in such a way to avoid synchronous cusps and magnetic locking etc. Complete care of every aspect of an induction machine related to efficiency, power factor and overload capacity has been taken. In the countries like India where there is sudden increase or decrease in the load, it is really necessary to consider the overload capacity. To increase the overload capacity, the air gap has been increased up to a limit without compromising efficiency, reliability and durability of a three phase induction machine. Increasing the air gap not only increases overload capacity and more cooling it also reduces Pulsation losses, unbalanced magnetic pull, and Noises.

An design algorithm has been developed to give suitable number of rotor slots such that it avoid synchronous cusps, harmonic induction torque which causes Crawling and harmonic synchronous torque which cause Cogging. Selection of suitable number of rotor slots decreases vibrations as well as noise in machine; decreases the voltage ripples, which cause additional iron losses and inductive interference to the communication lines.

29.2 Considerations of Rotor Design

29.2.1 Choice of Average Flux Density in Air Gap

For choice of Bav consider some factor Such as Power factor, Iron loss, Overload capacity.

29.2.2 Choice of Ampere Conductor Per Meter

Its depend on Copper loss, temperature rise, Voltage and Overload capacity.

Table 29.1 Type of rotor slots

Slots	Machines	Magnetizing current	Noise	Leakage reactance	Starting current	Overload capacity
Closed	Small	Low	Quieter	Large	Less	Less
Semi-enclosed	Medium	More	Less quieter	Smaller	More	Better

29.2.3 Length of Air Gap

Length of the air gap is depend on the Power factor, Overload capacity, Pulsation loss, Unbalance magnetic pull, Cooling, Noise .

29.2.4 Number of Rotor Slots

To avoid crawling, cogging, synchronous cusps, vibration, noise and voltage ripples, when Ss (number of stator slots)-Sr (number of rotor slots) should not be equal to 0, +p, +2p, +3p, +5p, +1, +2, +(p + 1), +(p + 2). And number of rotor slot is 15–30 % larger or smaller than the number of stator slots.

29.2.4.1 Choice of Rotor Slots

There are two types of slots preferred in squirrel cage rotor either closed or semi-enclosed. Generally preferred circular shapes for smooth starting (Table 29.1).

29.2.4.2 Rotor Slot Insulation

No insulation is used between bars and rotor core. Near about, 0.15–0.40 mm clearance between rotor bar and core depending upon whether slots are skewed or not should be used.

29.2.5 Area of Rotor Bars

A machine designed with a higher value of rotor resistance has the advantage that it has high starting torque but, a rotor with a high resistance has a disadvantage that I^2R loss is greater and therefore its efficiency is lower under running condition. Higher current density, lower is the conductor area and greater resistance [2].

29.2.6 Design of End Ring

The value of current density chosen for the end rings should be such that the desired value of rotor resistance is obtained. The ventilation is generally better for end rings so a slightly higher value of current density that obtaining in rotor bars can be taken.

29.2.7 Rotor Teeth

The width of rotor slot should be such that the flux density in the rotor teeth does not exceed about 1.7 Wb/m2.

29.2.8 Rotor Core

The flux density in rotor teeth and core can be taken slightly higher than those in the stator teeth and core. This is because the iron losses in the rotor are very small owing to small value of frequency of rotor current.

29.3 Design Calculations

For rotor design calculations various design parameters like: full load output, line voltage, number of phases, frequency, synchronous speed, bore diameter, length of the core number of stator slots, angle of chording or number of slots by which winding is chorded and stator core depth have been used in this paper. Average flux density, ampere conductor per meter in rotor teeth, ampere conductor per meter in rotor core, current density have been considered in the rotor design. First of all, length of the air gap has been calculated. Compensation in the overload capacity, pulsation loss, unbalanced magnetic pull, cooling, noise etc. has been kept in mind. Other effects and power losses have also been considered in the design of a very efficient motor. So, greater is the length of air gap, greater is the overload capacity and cooling but smaller pulsation loss, unbalanced magnetic pull and noise. Now to select the rotor slots all consideration to affect the motor performance from using rotor slots have been considered. Harmonic induction torque (crawling), harmonic synchronous torque (cogging), vibration, noise and voltage ripples affect the performance of a three phase induction motor. Calculated the end ring design (Table 29.2).

Table 29.2 Given data of the machine

Description	Variable	Values
Power output	P	2.2 Kw
Line voltage	V	400 V
Number of phases	ph	3
Frequency	F	50 Hz
Synchronous speed	ns	1500 rpm
Number of stator slots	Ss	24
Stator core depth	d_{cs}	17 mm
Stator winding factor	K_{ws}	0.9330
Turns per phase	Ts	44
Specific magnetic loading	B_{av}	0.48 Wb/m^2
Specific electric loading	ac	28,000 A/m
Current density in rotor bar	δ_{rb}	6 A/mm^2
Current density in end rings	δe	6 A/mm^2

Important points related to rotor design have been summarized as follows:

- Calculate the air gap length and diameter of the rotor D_r $l_{ag} = 0.2 + 2\sqrt{(D_S.L_S)}$
- Rotor slots Sr calculate from para. 2. 4 and rotor slot pitch at air gap $y_r = \pi$ *Dr/S_r
- Rotor bar current I_{rb} = 2 *ph*Kws*Ts*Is*cosΦ/S_r and chose the value of current density
- Obtain the area of each rotor bar a_{rb} Area of each rotor bar $A_{rb} = I_{rb}/\delta_{rb}$ mm^2 Current density in rotor bars may be taken between 4–7 A/mm^2.
- Width of the rotor slot, depth of the rotor slot, slot pitch at the bottom of slots, tooth width at the root, flux density at the root of rotor teeth.
- Calculate the final length of each bar (remember the bar length increase beyond the core on both side and also increase the length for skewing).
- Find the resistance of each bar and copper loss in bars.
- Now, calculate design of end ring, First find end ring current Ie = (Sr*I_{rb})/(π*p)
- Taking the suitable current density in end ring δe, then find the area of end ring a_e
- Chose the suitable value of the depth and the thickness of the end ring.
- And find the outer diameter Doe, inner diameter Die and mean diameter Dme.
- Calculate the resistance of each ring r_e. And obtained the value of copper loss in both end rings.
- Check the full load slip is in under limit or not by using = (Rotor copper loss/rotor output) = s/(1 − s)
- Generally the value of depth of rotor core dcr is taken equal to that of stator core. Find the value of inner diameter of rotor laminations

Table 29.3 Design sheet of rotor obtained

Description	VariabIe	Values
Length of air gap in mm	I_{ag}	0.3885
Diameter of rotor in mm	D_r	105.5601
No. of rotor slots	S_r	22
Slots per pole per phase	q_r	1.8333
Conductor per slot	Z	1
Winding factor	KW_g	1
Slot pitch at the air gap in mm	y_g	15.0739
Width of the rotor slot	W_r	6.8
Slot pitch at the bottom of slots in mm	y_{rb}	11.6835
Tooth width at the root in mm	Wt	4.8835
Angular displacement between adjacent slots in electrical degree in radian	β	0.5236
Distribution factor	K_d	0.9659
Rotor bar current in A	I_{rb}	328.6103
Rotor bar Cross section area in mm^2	a_{rb}	58.5000
Rotor bar Current density in A/mm^2	$\delta_{rb\ b}$	5.6173
length of each bar in mm	l_{rb}	123.5169
Resistance of each bar in ohm	R_{rb}	44.3394×10^{-6}
Copper loss in bars in W	H_{rb}	105.3355
End ring current in A	I_e	575.2995
Outer diameter of end ring in mm	Doe	81.8172
Inner diameter of end ring in mm	Die	57.8172
End Ring Cross section area in mm^2	a_e	95.8833
Mean diameter in mm	Dme	69.8172
Current density in A/mm^2	r_e	5.9927
Resistance of each ring in micro-ohm	dcr	47.9800
Depth of rotor core in mm	A_{cr}	17
Rotor core area in m^2	a_{cr}	0.0013

29.4 Squirrel Cage Rotor Design

Various design data have been illustrated in the following Table 29.3:

Copper loss in 2 end ring in W, He = 31.7599

Total copper loss in W, Ht = 137.0954

$\frac{\text{total rotor copper loss}}{\text{rotor output}} = \frac{s}{1-s}$, Full load slip S = 5.8 %

29.5 Conclusion

Squirrel cage rotor design of a three phase induction machine has been described in this paper successfully. For rotor design calculations various design parameters like: full load output, line voltage, number of phases, frequency, synchronous

speed, bore diameter, length of the core number of stator slots, angle of chording or number of slots by which winding is chorded and stator core depth have been described in this paper. Average flux density, ampere conductor per meter in rotor teeth, ampere conductor per meter in rotor core, current density have been considered in the rotor design. Designed rotor based on suggested data has reduced various problems related to temperature rise, tooth pulsation loss, cooling, vibration, crawling, cogging, voltage ripples, pulsation losses, unbalanced magnetic pull and noise.

References

1. Vincent DT (2005) Electric machines and power system. Prentice-Hall of India Private Limited, New Delhi (India), ISBN-0-87692-544-1
2. Faiz J, Dadgari AA, Horning S, Keyhani A (1995) Design of a three-phase self excited induction generator. IEEE Trans Energy Conversion 10(3):516–523
3. Saket RK (2008) Design, development and reliability evaluation of micro hydro power generation system based on municipal waste water. In: IEEE electrical power and energy international conference 2008 (EPEC 2008), pp 01–08, Canada
4. Sawhney AK (2008) A course in electrical machine design. Dhanpat Rai and co., New Delhi, India. Web-site: http://www.dhanpatraico.in

Chapter 30
Experimental Validation and Performance Comparison of Multi-Loop MPPT Controlled PV Systems on Low to High End Controllers

Atul Gupta, Venu Uppuluri Srinivasa and Ankit Soni

Abstract Maximum power point tracking (MPPT) is used in photovoltaic (PV) systems to maximize the PV array output power, irrespective of the temperature, irradiation conditions and load electrical characteristics. A new PV MPPT control system is developed, consisting of a flyback topology based DC–DC converter controlled by a DSP. In this proposed system a multi-loop control scheme is implemented to control the flyback converter. The system is highly effective for low power applications and can be easily modified to handle various energy sources (e.g., wind-generators). The proposed multi-loop control system is implemented on low to high-end controllers and their performances are compared. Experimental results describe the performance of the proposed design prototype are in agreement with the simulation results. 8-bit microcontroller NXP89V51RD2, 16 bit DSP TMS320LF2401A, 32 bit DSP TMS320F28027 and 32-bit DSP with floating point unit TMS320F28069 are used to realized this proposed design.

Keywords MPPT · Flyback converters · PV system · DSP · String inverter

A. Gupta (✉) · V. U. Srinivasa
Santerno India Design Center, Viman Nagar, Pune, Maharashtra, India
e-mail: atulgupta2006@gmail.com

V. U. Srinivasa
e-mail: venuuppuluri@gmail.com

A. Soni
School of Instrumentation, DAVV, Khandwa Road Campus, Indore, Madhya Pradesh, India
e-mail: soniankit15@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_30,

30.1 Introduction

Solar energy is one of the most important renewable energy sources. With an increasing world-wide interest in renewable energy production and use, there is renewed focus on the power electronic converter interface to these DC energy sources. To increase the performance of the PV systems, it is important to note that the output characteristic of a photovoltaic array is nonlinear and changes with solar irradiation and cell's temperature. Therefore, an efficient modeling of the panel is required for the experimental validation and MPPT technique is needed to maximize the produced energy.

A new approach "converter-per-module" approach offers many advantages including individual module MPPT which gives great flexibility in module layout, replacement, insensitivity to shading, better protection of PV sources, redundancy in the case of source or converter failure, easier and safer installation, maintenance and better data gathering [1, 2]. Placing a DC–DC converter on each sub string and then connecting these strings in series avoids many of the problems such as shading and loss of system-efficiency due to difference in individual characteristics of the panel. Some of the advantages like: better utilization of MPPT on per module basis [3], better protection of module power sources, better data gathering, greater safety during installation and maintenance are equally important.

For controlling the overall system a multi-loop control scheme was proposed which consists of a fast inner current loop and slower outer voltage loop corresponding to input PV current and voltage respectively. A comparatively slow MPPT loop was also running setting the reference point for the voltage loop.

In particular, without lack of generality a stand-alone photovoltaic system was constructed by connecting the DC–DC converter between the solar panel and a DC load and the Perturb & Observe (P&O) MPPT algorithm was implemented. The DC–DC converter was controlled by different controller cards based on low end 8-bit to high end 16-bit & 32-bit DSPs and the performance of the system was compared.

30.2 PV Modeling/Characteristics

Solar cells consist of a p-n junction fabricated in a thin wafer or layer of semiconductor. The simplest equivalent circuit of a solar cell is an ideal current source in parallel with a diode. The current source represents the current generated by photons (often denoted as I_{ph}) and its output is constant under constant temperature and constant incident radiation of light [4]. The current from a PV cell depends on the external voltage applied and the amount of sunlight on the cell. The equation which describes the I–V characteristics of PV cell is:

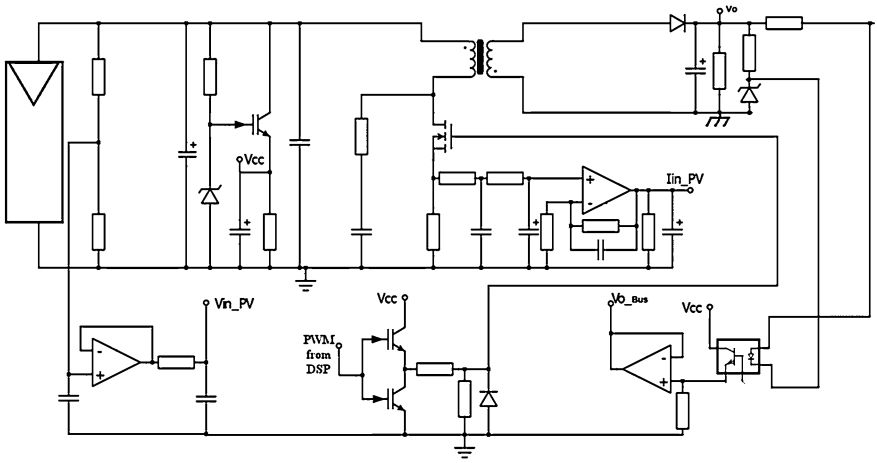


Fig. 30.1 Power board implemented in the system

$$I = I_{ph} - I_o\left[\exp\left(\frac{q \cdot (V + IR_s)}{AkT}\right) - 1\right] \tag{30.1}$$

Where, 'I' is the PV array output current, 'V' is the PV array output voltage, 'q' is the charge of an electron, 'k' is Boltzmann's constant, 'A' is the p-n junction quality/ideality factor, 'T' is the cell temperature and 'Io' is the cell reverse saturation current. The factor 'A' in (30.1) determines the cell deviation from the ideal p-n junction characteristics.

30.3 System Description

To test the performance of the proposed system with different controller cards, the power board shown in Fig. 30.1 was implemented. The experimental setup consisted of a solar panel, DC–DC converter and control board (based on 8-bit, 16 & 32-bit DSP).

The power board consisted of various sensing and driver circuits. The panel voltage was sensed by a voltage divider followed by a buffer circuit, a low valued current sensing resistor was placed in series with the power switch to sense the current. An optocoupler was placed to couple the feedback from the output DC bus voltage. PWM signals from the controller were fed to the driver circuit which finally derived the gate of the power switch.

Table 30.1 Electrical characteristics of PV panel at Standard Test Conditions (STC)

Symbol	Quantity	Value
P_{MPP}	Maximum Power	15 W
V_{MPP}	Voltage at PMPP	18 V
I_{MPP}	Current at PMPP	0.81 A
V_{OC}	Open Circuit Voltage	28 V
I_{SC}	Short Circuit Current	1.05 A
NOCT	Nominal Operating Cell Temperature	44.5

30.3.1 Solar Panel

The solar panel used for the experimental purpose was PAE SE-15. They are made using the latest Copper Indium Gallium diSelenide (CIGS) solar cells. CIGS perform well over a range of light—levels and climatic conditions, providing more kWh per day compared to conventional silicon technology. The technical specifications of the panel are given in Table 30.1.

30.3.2 DC–DC Converter

The DC–DC flyback converter was implemented for working at a switching frequency of 21.59 kHz. The isolation was achieved through a coupled inductor. The ideal gain in CCM is given by:

$$V_o = n \cdot \frac{d}{(1-d)} \cdot V_{in} \tag{30.2}$$

Where, n is the turn ratio, V_{in} and V_o are the input and output voltage respectively. The output ripple frequency is the same as switching frequency. This circuit employs the minimum number of components among all the DC–DC converters (one active switch, one passive switch, one magnetic element and one capacitor) and hence the preferred circuit for low power (up to 200 W). The switching controls algorithms, which are simple and fast, provides a significant improvement in the system's dynamic performance compared to usual analog control techniques [5]. The parameters of the converter designed are given in Table 30.2.

30.3.3 DSP Control Board

The MPPT algorithm and the control of the DC–DC converter were implemented on NXP89V51RD2, TI's TMS320LF2401, TMS320F28027 and TMS320F28069

Table 30.2 DC-DC flyback converter parameters designed in the setup

Parameter	Value	Part No.
Transformer (Inductance) L	9.9 uH	Coiltronics VP5-0155
Capacitance C	220 uF/50 V	
Mosfet Switch	57A/100 V	IRF3710
Diode	40A/45 V	40CTQ045S
Input voltage range	18–25 V	
Output voltage range	23 V	
Switching frequency	21.59 kHz.	

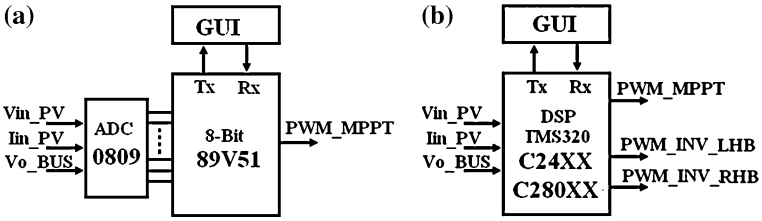


Fig. 30.2 Scheme of controller boards. **a** 8-Bit microcontroller board. **b** DSP controller board

DSPs. The DSP measured the input current and input voltage though the A/D module and calculated the power obtained from the panel. The scheme of the power board designed in the experiment is shown in Fig. 30.2.

The duty cycle is used as the control variable in order to simplify the control structure of the system [6]. The update rate of the MPPT algorithms with different controllers were found experimentally by setting it as fast as possible without causing instability to the system, too fast update rate may cause the system to become instable due to the relatively long time constant of the power stage [7].

30.3.4 MPPT: Perturb and Observe Method

The P&O method is the most popular MPPT algorithm due to its simplicity. After one perturb operation the current power is calculated and compared with previous value to determine the change of power ΔP. If $\Delta P > 0$, then the operation continues in the same direction of perturbation, otherwise the operation reveres the perturbation direction [8]. A common problem in P&O algorithms is that the array terminal voltage is perturbed every MPPT cycle; therefore when the MPP is reached, the output power oscillates around the maximum, reducing the generable power by the PV system.

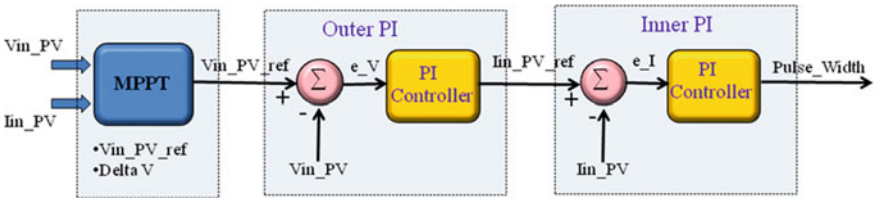


Fig. 30.3 Proposed multi-loop control scheme

30.4 Control Scheme

The proposed control consists of three loops, the outer most MPPT loop set the V_{ref} corresponding to the panel input voltage, the slower outer proportional integral (PI) loop set the I_{ref} for the faster inner PI loop which finally controls the pulse width of the PWM signal which drives the MOSFET gate, the fast inner PI loop regulates the PV array output voltage with respect to I_{ref} which is set in the outer PI loop [9]. The block diagram of the MPPT control and the compensator for digital implementation are shown in Fig. 30.3.

The functions of all the three loops are performed by controller/DSPs. The controller senses the solar panel current and voltage to calculate the solar array output power, V_{ref}, I_{ref} and the pulse width signal for maximum power control. The performance of the control algorithm can further be improved by implementing it on RTOS [10].

30.5 Experimental Results

The above-described scheme was used to test the MPPT algorithm. Fig. 30.4a shows the output waveform of the experiment, the PWM signal, the input current and voltage. The oscillations of the P&O algorithm around the maximum power point can easily be seen from the oscillation of the current waveform in Fig. 30.4b. Table 30.3 tabulates the execution time required by various function loops controlling the system from low to high end controllers. Figure 30.5a shows the waveforms at the fly-back converter. Figure 30.5b shows the timing for various subroutines of the multi-loop control implemented in the microcontroller.

30.6 Conclusion

This paper discussed the implementation of a multi-loop controlled PV system, based on flyback topology and the performance of the proposed PV system was compared over different DSP controllers. A simple method which combined the

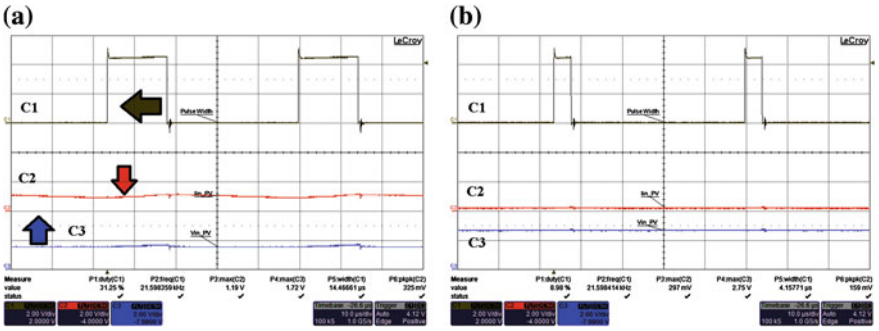


Fig. 30.4 Waveforms at two different operating points (**a**) & (**b**) C_1: Pulse Width, C_2: Iin_PV, C_3: Vin_PV

Table 30.3 Timing of the control loops with various controllers

Controller/ DSP	ADC read (msec)	PI loop (msec)	MPPT loop (msec)	PWM frequency (kHz)	ADC clock (kHz)	MPPT update time (msec)
NXP89V51RD2	3.74	0.901	0.650	21.59	21.59	12.8
TMS320LF2401	0.002	0.003	0.012	21.59	5,000	0.024
TMS320F28027	0.001	0.002	0.010	21.59	10,000	0.017
TMS320F28069	0.0005	0.001	0.008	21.59	20,000	0.011

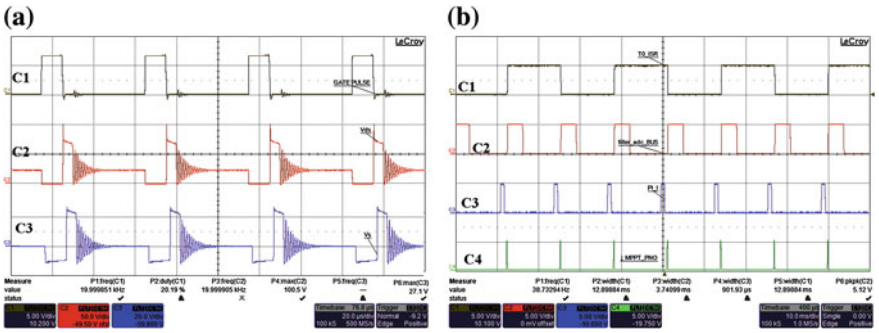


Fig. 30.5 **a** various test points at flyback converter.C_1: Gate Pulse, C_2: V_{DS} of MOSFET, C_3: Output diode. **b** Timings of subroutines at NXP89V51RD2. C_1: T_0 Interrupt, C_2: ADC Filtering, C_3: PI Control loop, C_4: MPPT_PNO loop

basic MPPT algorithm (P&O) and the PI controllers, which were required to track the voltage (Vin_PV), current (Iin_PV) of PV module/string with its references (Vref, Iref) and to controlled the output voltage by setting an appropriate pulse width signal to power switch. The proposed multi-loop control system was constructed and the experimental tests were carried out against changing

environmental conditions (dust on panels, partial shading and impact of wind) and the tracking efficiency was confirmed by simulations and experimental results.

Simulation based numerical results were provided and validated by experimental datas and it was concluded that the performance of the low end controllers are comparable with their high end counterparts, without any significant improvements, if the production cost, the architectural complexity and the development time was considered. However, it was observed that high end controllers still had enough bandwidth or system resources to implement other control function like, incorporating DC-AC full bridge topology based sine wave inverter etc. The proposed approach provided a feasible way to manage low-cost PV power systems design.

References

1. Walker GR, Sernia PC (2002) Cascaded DC–DC converter connection of photovoltaic modules. IEEE press, pp 24–29
2. Calais M, Myrzik JMA, Agelidis. VG (2001) Inverters for single phase grid connected photovoltaic systems—overview and prospects. In: 17th PV solar energy conference and exhibition, Munich, Oct 2001
3. Noguchi T, Togashi S, Nakamoto R (2002) Short-circuit pulse-based maximum-power-point tracking method for multiple photovoltaic-and-converter module system. IEEE T Ind Electron 49:217–223
4. Gupta A, Soni A, Srinivasa UV (2011) Generalized accurate modeling of various PV cells/module using MATLAB Simulink. Int J Recent Trends Eng Technol 6(2):65–67
5. Gupta A, Chandra A (2010) Fully digital controlled frontend converter based on boost topology. In: CCPE, ACEEE 2010 pp 171–176
6. Xiao W, Dunford WG (2004) A modified adaptive hill climbing MPPT method for photovoltaic power systems. In: IEEE power electronics specialists conference (PESC), pp 1957–1963
7. Hohm P, Ropp ME (2003) Comparative study of maximum power point tracking algorithms. Progress in photovoltaics: research and applications, vol 11, pp 47–49
8. Dolara A, Faranda R, Leva S (2009) Energy comparison of seven MPPT techniques for PV systems. J Electromagn Anal Appl 3:152–162
9. Chihchiang H, Chihming S (1998) Study of maximum power tracking techniques and control of DC/DC converters for photovoltaic power Ssystem. IEEE Press, pp 86–91
10. Gupta A, Srinivasa UV (2011) RTOS: a new approach in design and organization of high-speed power control applications. Springer, Verlag

Chapter 31
Effect of Temperature on Si-Ge Hetero-Gate Raised Buried Oxide Drain Tunnel FET Electrical Parameters

Monalisa das and Brinda Bhowmick

Abstract The effect of temperature on SiGe hetero-gate raised buried oxide drain Tunnel FET electrical parameters like tunnelling bandgap, threshold voltage, subthreshold swing, etc. are discussed in this paper. A modified SOI based Silicon hetero-gate TFET structure has been used. The proposed device is almost free from short channel effects. The simulation is performed using Synopsys 2D TCAD tools where non local band-to-band tunnelling is applied.

Keywords Non local band-to-band tunnelling · Hetero-gate · Raised buried oxide · Tunnel FET · SiGe

31.1 Introduction

The constant downscaling of MOSFET devices have lead to fundamental performance limitations like increased leakage current, threshold voltage roll-off, short channel effects, reduction of subthreshold swing being limited to 60 mV/dec [1]. Power dissipation is the greatest hurdle in modern day electronics. The TFET has emerged as a promising candidate for low power applications to overcome the difficulties faced by the conventional devices and some other novel devices like I-MOS

M. das (✉) · B. Bhowmick
Department of Electronics and Communication Engineering,
National Institute of Technology, Silchar, 788010 Assam, India
e-mail: monalisdas@gmail.com

B. Bhowmick
e-mail: brinda_bh@yahoo.co.in

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_31,

[2], nanoelectromechanical FETs [3], suspended gate MOSFETs [4], etc. because of its better switching characteristics.

The TFET is a gated reverse-biased p-i-n diode that works under the principle of band to band tunnelling controlled by the gate. In band-to-band tunnelling (BTBT) the electrons tunnel from the conduction band to the valence band through the band-to band tunnelling width and vice versa and this band-to-band tunnelling width is controlled by the gate. Hence in case of TFET's the tunnelling current which is the main source of current in the device is controlled by the gate and so short channel effects are prevented. TFET's with subthreshold swing (SS) lower than 60 mV/dec have been discussed [5, 6]. Such devices have low OFF-current (*Ioff*), and hence they can be qualified for an ultimately scaled quasi-ideal switch. The On current (*Ion*) of TFET is much low and is far behind the ITRS requirement [7]. In the proposed device SOI technology is used so as to prevent the problems like latch-up, parasitic source and drain capacitances, etc. faced in bulk technology.

The effect of temperature on the electrical parameters like tunnelling bandgap, subthreshold swing, threshold voltages, On-current, etc. for the proposed device is discussed here. Here a hetero-gate dielectric is used so as to increase the on-current, to reduce parasitic ambipolar current and to obtain a improved sub-threshold swing. In order to obtain high on-current, high transconductance and low electric field, fully depleted (FD) SOI is used in the proposed device. A raised buried oxide (BOX) has been used so as to enhance mobility, to reduce electric field and off current at the drain end and to minimize the gate to drain capacitance. A film of SiGe layer is introduced as the source to reduce the tunnelling gap by modulating the Germanium mole fraction and thereby increasing the tunnelling current. The inclusion of the SiGe layer also helps in obtaining an abrupt doping profile. Two main approaches have been used for the fabrication of SiGe heterojunctions i.e. differential epitaxy and selective epitaxy. The advantage of selective epitaxy is that it eliminates the need for both deep and shallow trench isolation. Sometimes combination of selective and non selective epitaxy are found to be useful [8].

31.2 Device Design and Operation

The proposed device (in Fig. 31.1) is a lateral hetero-gate SOI tunnel FET with raised buried oxide and SiGe source. The length of high-k material used is 10 nm and of low-k material is 30 nm. A raised buried oxide is used at the drain end so as to reduce the silicon film thickness and is of a thickness of 10 nm. The channel length is 40 nm. The silicon film thickness at the drain end is 20 nm and the source end has a SiGe film thickness of 20 nm. The SiGe film in the source region in the proposed device has been used in place of silicon film of conventional TFET's so as to reduce the tunnelling bandgap because reduction in tunnelling bandgap leads to increase in tunnelling of electrons thereby giving rise to increase in tunnelling

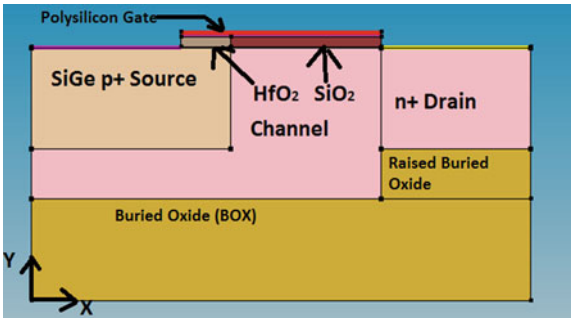


Fig. 31.1 Proposed device

current. The dielectric constant of Si-Ge is a function of mole fraction of Germanium. The gate material is N + polysilicon with a work function of 4.5 eV. The high-k material that is used is hafnium oxide (HfO_2) which has a dielectric constant of 25 and bandgap energy of 6 eV and the low-k material used is silicon-dioxide (SiO2) with a dielectric constant of 3.9 and bandgap energy of 9 eV. A constant oxide thickness of 2 nm is used and that of polysilicon gate is 1 nm. The dopings for p + source, intrinsic and n + drain are $10^{21}, 10^{16}$and 5×10^{19} cm^{-3} respectively. The operation is based on non local band -to-band tunnelling [9].

When a positive gate voltage above a minimum voltage Va, which is required to align the valence band in the source and the conduction band in the channel in the same level, is applied the bands in the lowly doped intrinsic region are pulled downwards and the tunnelling barrier is reduced thereby facilitating tunnelling of electrons from the valence band of the p+ source to the conduction band of the channel region. As a result the device turns on.

31.3 Result and Discussion

For simulation of the proposed device Synopsys 2D TCAD tools where non local band-to-band tunnelling is applied [10]. The high field saturation doping dependent Canali model is used.

31.3.1 Effect of Temperature on Tunneling Bandgap of Proposed Device

The energy bandgap or the tunnelling bandgap is reduced at elevated temperature. This can be understood from the following equation

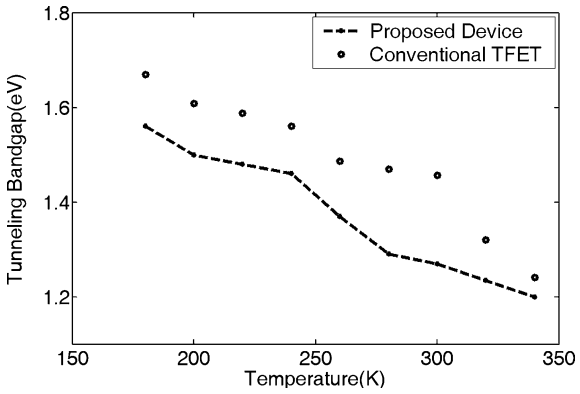


Fig. 31.2 Variations of tunnelling bandgap with temperature

$$E_g(T) = E_g(0) - \frac{\alpha T^2}{T + \beta} \tag{31.1}$$

where α and β are material dependent constant and $E_g(0)$ is the limiting value of E_g at 0 K. Hence with increasing temperature the tunnelling bandgap should be reduced. Figure 31.2 shows the variation of the tunnelling bandgap with temperature for the conventional TFET and proposed device. The Si-Ge have smaller bandgap than silicon because of larger lattice constant and strain.

From Fig. 31.2 it is evident that the proposed device is in accordance with theoretical calculations and the variation in tunnelling bandgap with temperature for the proposed device is much lesser compared to that of the conventional devices. It is observed that the reduction in tunnelling gap with temperature is more for the conventional TFET device and hence proposed device is less dependent on temperature. Since raised buried oxide in drain with Si-Ge layer in the source side provides less electric field, high mobility of carriers, and less channel resistance.

31.3.2 Effect of Temperature on Subthreshold Swing of the Proposed Device

Subthreshold swing (SS) can be defined as the amount of V_{GS} required to change the I_{DS} by 1 decade.

$$\mathrm{SS} = 2.3\left(\frac{E + A'}{E^2}\frac{dE}{dV_{GS}}\right)^{-1} \tag{31.2}$$

$$A' = DEg^{\frac{3}{2}}$$

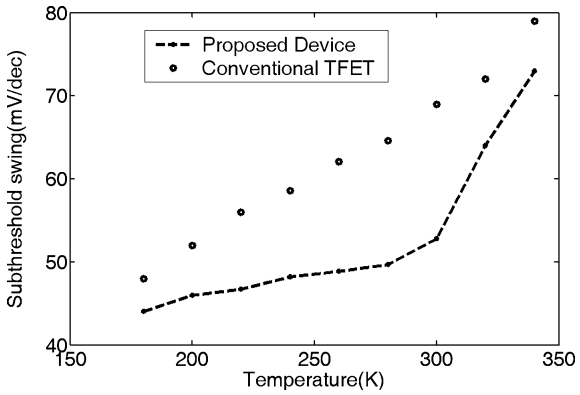


Fig. 31.3 Subthreshold swing increases with the decrease in tunnel bandgap

where E is the magnitude of the electric field, V_{GS} is the gate to source voltage, D is the parameter which depends on the effective mass of valence band and conduction band and E_g is the energy bandgap. From (31.2), it is observed that subthreshold swing (SS) is a function of A', which itself is a function of energy bandgap (E_g). With increasing temperature the tunnelling bandgap reduces, hence A' decreases too. As a result the subthreshold swing (SS) increases. Fig. 31.3 shows the variation in subthreshold swing with temperature for the proposed device.

In Fig. 31.3 it is observed that the SS values are decreased with increasing tunnel bandgap. Comparing conventional device and proposed device it is seen that change in SS value with temperature is more in conventional device. Moreover the proposed device has a subthreshold swing (SS) much lower than the conventional device at room temperature.

31.3.3 Effect of Temperature on Threshold Voltage of the Proposed Device

Threshold voltage is one of the most vital electrical parameter of tunnel FET. It is already established that unlike MOSFET There are two threshold voltages for tunnel FET i.e., gate threshold voltage and drain threshold voltage. In MOSFET's the threshold voltage is defined at the onset of strong inversion but in tunnel FETs the definition of threshold voltage is the voltage at which drain current changes from quasi-exponential to linear. Moreover the energy barrier narrowing in tunnelling current is a complex functions of both gate and drain voltage. Gate and drain threshold voltages are analytically derived considering conductance change method [11, 12].

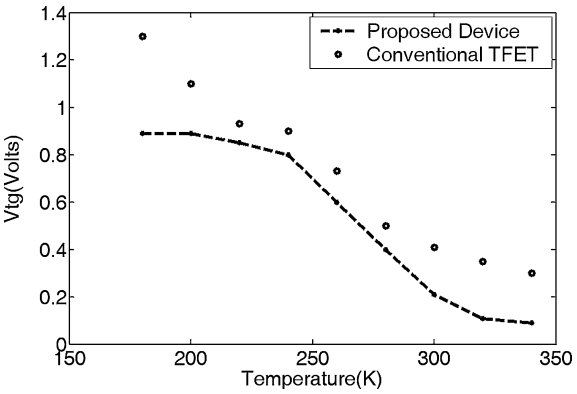


Fig. 31.4 Gate threshold voltage variation with temperature at gate length of 40 nm. This threshold is determined using conductance change method

31.3.3.1 Effect of Temperature on Gate Threshold Voltage

According to the conductance change method the gate threshold voltage can be defined as the gate voltage when the derivative of the transconductance reaches maximum. The gate threshold voltage is given by [13],

$$V_{TG} = V_{\mathrm{FB}} + \psi_{\mathrm{s}} + \frac{\varepsilon^2\tau^2}{2\varepsilon*^2\left(1+\frac{4y}{t_S}\right)}\left[\begin{array}{l}\frac{\varepsilon*}{\varepsilon\tau}y\left(2-\frac{y}{t_S}\right)\psi_S - \frac{3A'^2E}{A'^2-2A'E} \\ +\frac{A'E^2+3A'}{A'^2-2A'E} \\ -\frac{(E^2+A')E^2}{A'^2-2A'E}\left\{\frac{8\varepsilon*^2\,y}{\varepsilon^2\tau^2 t_S}\right\}\end{array}\right] \tag{31.3}$$

where flat band voltage isVFB, $\varepsilon*$ is the gate oxide dielectric constant, τ is the oxide thickness, and t_s is the body layer thickness, E is the magnitude of electric field and A and D are parameters which depend on the effective mass of valence and conduction band electrons determined from device dimensions and material parameters. With increasing temperature the gate threshold voltage tends to decrease. Figure 31.4 shows the variation of the gate threshold voltage with temperature for the proposed device as well as the conventional device.

From Fig. 31.4 we can conclude that the variations in the gate threshold voltage with temperature for the proposed device is lesser compared to the conventional device and the gate threshold voltage is much lesser for the proposed device.

31.3.3.2 Effect of Drain Threshold Voltage with Temperature for the Proposed Device

According to the conductance change method the drain threshold voltage can be defined as the drain voltage where the derivative of output conductance reaches maximum. The drain threshold voltage is given by [12],

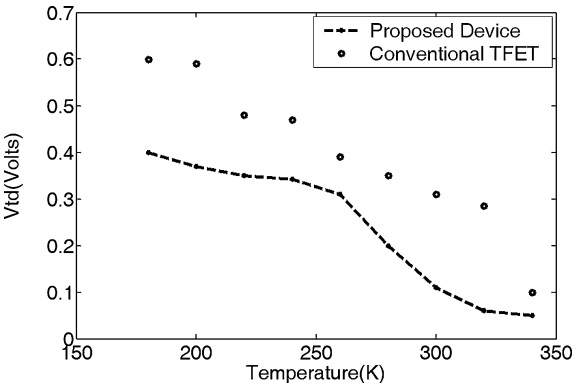


Fig. 31.5 Variation of drain threshold voltage with temperature. The drain threshold voltage is dependent on the tunnelling bandgap which is a complex function of gate and drain voltage

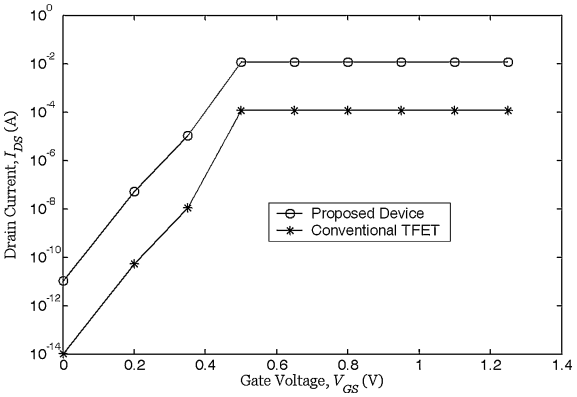


Fig. 31.6 Comparison of I_D-V_{GS} characteristics at $V_{DS} = 1$ V

$$V_{\mathrm{TD}} = \frac{0.066eV}{q} - V_{gs} - V_0 + V_{\mathrm{t}} \ln\left[\left(e^{N/N_{\mathrm{C}}} - 1\right)\left(e^{P/N_{\mathrm{V}}} - 1\right)\right] \tag{31.4}$$

N_{C} and N_{V} are effective density of states located at conduction and valence band edge, N and $P$are electron hole concentrations, $V_o$is built in potential for zero biased junction.

As the temperature is increased the drain threshold voltage tends to decrease. Figure 31.5 shows the variation of the drain threshold voltage with temperature for the proposed device as well as the conventional device.

From Fig. 31.5 it is clear that the variations in drain threshold voltage with temperature for proposed device is lesser than that for conventional device. The drain threshold voltage is much lesser for the proposed device as compared to the conventional devices. This is due to the fact that the tunnel band gap in the proposed device is reduced by inclusion of SiGe layer in the source.

31.3.4 Comparison of DC Characteristics

The I_D-V_{GS} characteristics of the proposed device are compared with conventional TFET. The proposed device has high on current and on–off current ratio. Due to the use of raised buried oxide mobility can be improved and the off current can also be reduced due to the reduced electric field at the drain side (Fig. 31.6).

31.4 Conclusion

The on current of the proposed device is found in the range of 10^{-2} amp which is much higher than that of a conventional SOI TFET. The off current of the device has been reduced to a great extent too. The variation of threshold voltages, subthreshold swing and tunneling bandgap with temperature for conventional TFET is more compared to the proposed device. Hence the proposed device can be qualified as a better successor of the conventional TFET.

Acknowledgments This work was supported by ALL INDIA COUNCIL FOR TECHNICAL EDUCATION (AICTE), under Grant 8023/BOR/RID/RPS-253/2008-09.

References

1. Vandamme EP, Jansen P, Deferm L (1997) Modelling the subthreshold swing in MOSFET's. IEEE Electr Dev Lett 18:369–371
2. Gopalakrishnan K, Woo R, Jungemann C, Griffin PB(2005) JD Plummer, impact ionization MOS (I-MOS) -part II: experimental results. IEEE T Electron Dev 52:77–84
3. Kam H, Lee DT, Howe RT, King T-J (2005) A new nano electromechanical-field effect transistor (NEMFET) design for low power electronics. In: IEDM Technical Digest pp 463–466
4. Abele N, Fritschi N, Boucart K, Casset F, Ancey P, Ionescu AM (2005) Suspended-gate MOSFET, bringing new MEMS functionality into solid-state MOS transistor. In: IEDM Technical Digest pp 1075–1077
5. Choi WY, Park BG, Lee JD, Liu TJK (2007) Tunneling field effect transistors (TFETs) with subthreshold swing (SS) less than 60 mV/dec. , IEEE Electr Dev Lett 28:743–745
6. Qin Z, Wei Z, Seabaugh A (2006) Low-subthreshold-swing tunnel transistors. IEEE Electr Dev Lett 27(4):297–300
7. Semiconductor Industry Association (SIA), International Technology Roadmap for Semiconductors (ITRS). available at www.itrs.net
8. Ashburn P (2004) SiGe heterojunction bipolar transistor. John Wiley & Sons, Ltd, Chichester
9. Zhang, Q., Suta, S., Kosel, T., Seabaugh, A.: Fully-depleted Ge interband tunnel transistor modeling and junction formation. Solid State Electron, vol. 53, pp. 30-35 (2009).
10. Synopsys TCAD sentaurus device manual (2010).
11. SOI Technology, Materials to VLSI, 3rd edn Jean-Pierre Colinge, Springer International, Heidelberg

12. Tsividis Y (1999) Operation and modelling of the MOS transistor. 2nd edn, McGraw Hill, New York
13. Bhowmick B, Baishya S (2012) A physics–based model for electrical parameters of double gate hetero-material nano scale tunnel FET. Int J Appl Inform Syst 3:28–32

Chapter 32
A Novel Inverter Topology for Low Power Drives

G. Nageswara Rao, K. Chandra Sekhar and P. Sangameswara Raju

Abstract This paper presents a novel topology to a low cost converter which drives a spindle motor at high speed with high starting torque utilizing the bipolar starting and unipolar running algorithm. This topology is simple and developed with only eight switches in the converter to drive the spindle motor at high speed with high starting torque. The proposed scheme has been simulated on MATLAB/SIMULINK platform, the results are presented and discussed.

Keywords Bipolar starting and unipolar running drive · HDD spindle motor · Unipolar drive

32.1 Introduction

The Hard Disk Drive (HDD) was made in the year 1956, since then it has grown to be the most effective mass data storage device for computers [1]. Each HDD consists of a spindle motor to turn one or multi platters with the storage media, where the data is stored. So the spindle motor is one of the most important components of a HDD and its performance has a direct impact on HDD performance especially on data access speed and capacity. The spindle motors used in computer hard disk drives are to possess high speed characteristics for fast data

G. Nageswara Rao (✉)
Vijaya Institute of Technology For Women, Vijayawada, India

K. Chandra Sekhar
RVR & JC College of Engineering, Guntur, India

P. Sangameswara Raju
S.V.University, Tirupathi, India

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_32,

access. Spindle motor is a brushless DC (BLDC) motor which has been used in high speed applications due to its high efficiency, high power density and wide range speed controllability. Several methods were proposed in literature [2–5] to achieve high speed operation. High speed operation could be achieved by any of the four schemes. The first method is to design the electromagnets in such a way that, the machine possess low back emf which offers high speed for a given system voltage [4]. But the disadvantage is that a low back emf constant results in low starting torque. The second method is winding method i.e., a series winding start and parallel winding run to run the motor at high speed with large starting torque [5], this requires additional switching devices and more complex control logic. The third method is to use a higher dc bus voltage, where a high starting torque with high speed operation can be achieved. But this scheme has a problem where the switch voltage rating is to be enhanced and it requires a current protection to limit the current during low speeds. This also adds cost and safety hazard to the system.

Apart from the above three methods, the forth method is to use a converter which can provide high speed with high starting torque. To get high torque BLDC motor needs to be operated in bipolar mode and to get high speed BLDC motor should be operated in unipolar mode. In [2], a new converter is proposed which can achieve high starting torque with high speed by using 14 switches of same rating. Drawback of this circuit is, using numerous gate drives and switches.

In this paper a novel inverter topology is proposed which uses bipolar operation to achieve high torque during starting and unipolar operation later to achieve high speed using only 8 switches. The proposed inverter is similar to conventional 3 leg inverter with one additional leg. The model has been verified with an inverter-motor model developed using Matlab/Simulink. Finally a novel converter is proposed to further increase the speed of spindle motor.

32.2 Proposed Inverter Operation

In conventional BLDC motor during bipolar operation, at any time across DC bus, two phases come in series. Only half of the DC bus voltage is applied to each phase, resulting in addition of torque constant on both phases there by achieving high starting torque. But speed will be limited. To get higher speed, full DC bus voltage is to be applied to each phase. This can be achieved in unipolar operation, where each phase conducts only in one direction which in turn reduces the starting torque. Thus in order to get high torque, motor should operate in bipolar mode and to get high speed motor should operate in unipolar mode. Shifting of modes between unipolar and bipolar operation is achieved based on speed requirement. The proposed inverter consists of 4 legs. The 3 phases of BLDC motor is connected to first 3 legs and neutral point is connected to the fourth leg as shown in Fig. 32.1. In bipolar operation first 3 legs are active and the 4th leg is inactive.

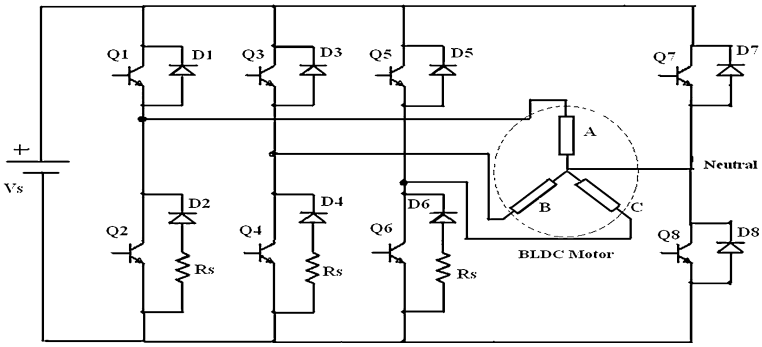


Fig. 32.1 Proposed inverter circuit

32.3 Simulation Results

The mathematical model of BLDC is given appendix and is used for simulation purpose. Macon EC6 215550, a 6 V 1.2 W motor was considered for validating the proposed converter. A MatLab/Simulink based model was developed for the motor and the inverter. Figure 32.2 shows the Simulink model of BLDC motor. It consists of four main blocks. The first one is Torque speed block. In this block electromagnetic torque and load torque are compared speed and position output is generated. Second block is trapezoidal back EMF block, in this block based on position information trapezoidal back EMF's are generated. Third block is converter block. Back EMF, Voltage and position information are inputs for converter block and Current is output. Based on speed information the converter block selects bipolar switching logic or unipolar switching logic. Fourth block is Electromagnetic torque generator block. From the current and position information torque block generates electromagnetic torque. Figure 32.8 shows the speed versus time plot of the motor at 0.5 m N-m of load torque. It is clearly shows that higher speed is achieved as soon as the system is switched from bipolar mode to unipolar (Fig. 32.3). Figure 32.4 shows the Back EMF and corresponding phase current of one phase in bipolar and unipolar mode.

Figures 32.5 and 32.6 show the extended torque versus speed characteristics of the motor with a changeover speed from bipolar mode to unipolar mode as 20,000 rpm with and without current limitations respectively. The changeover speed can be taken as any desired value. Unipolar mode operation without current limitation requires higher rating devices in the topology and is not preferred. This converter topology can work in three modes of operation.

(1) Continuously in bipolar mode (2) Continuously in unipolar mode and (3) bipolar starting and unipolar running mode with current limitation.

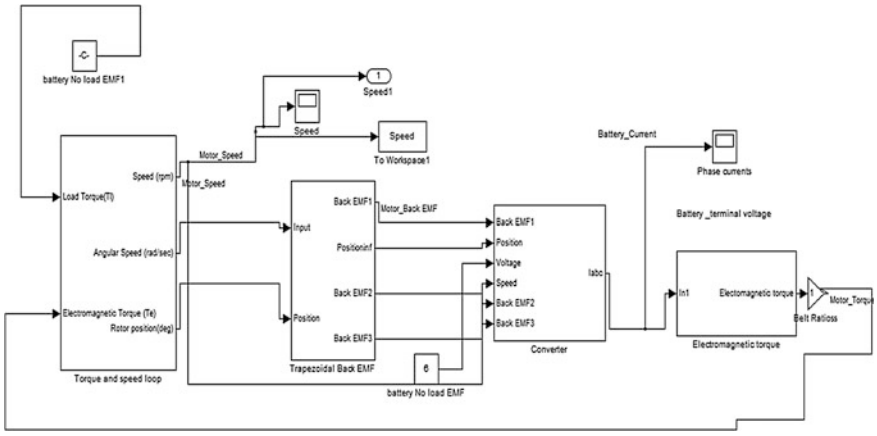


Fig. 32.2 MatLab/Simulink model of BLDC motor

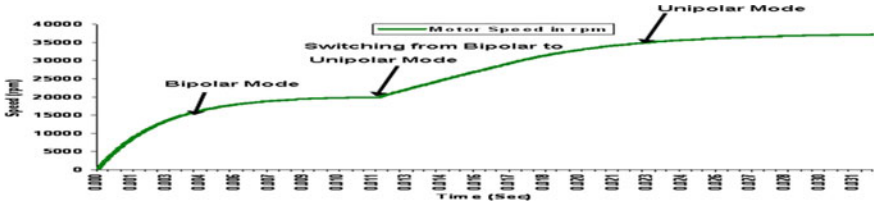


Fig. 32.3 Speed versus time of motor

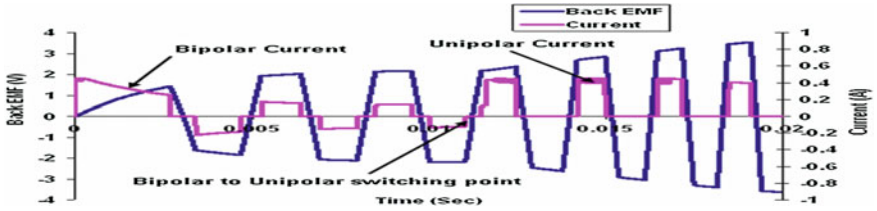


Fig. 32.4 Back EMF and phase current

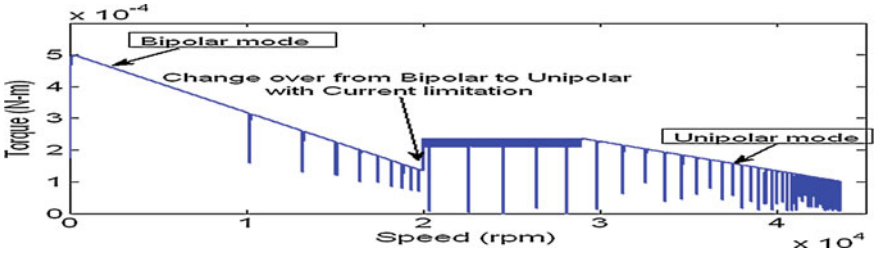


Fig. 32.5 Extended speed torque characteristic with current limitation in unipolar mode

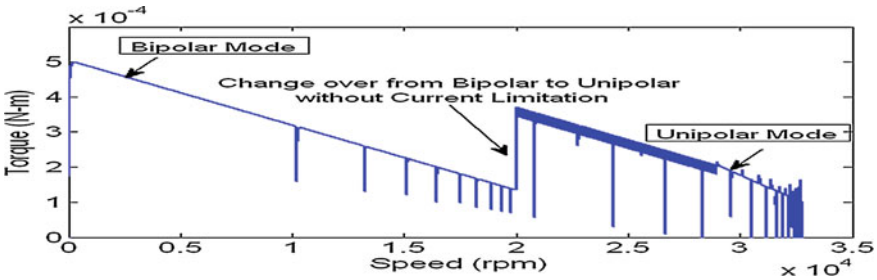


Fig. 32.6 Extended speed torque characteristic without current limitation in unipolar mode

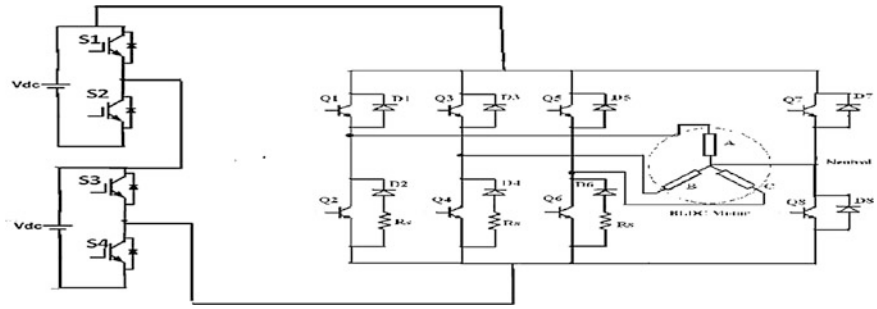


Fig. 32.7 Proposed Novel Converter

32.3.1 Proposed Novel Converter

The proposed Novel converter is shown in Fig. 32.7 It consists of two cascaded half Bridges. By closing switches S1 and S4 we can apply Vdc voltage to ISG. By closing switches S2 and S4 we can apply 2Vdc voltage to ISG. Ths switching table is shown in Table I. The proposed inverter consists of 8 switches, which is a far less count as compared to some of the existing inverter topologies for driving a Spindle motor. With the proposed inverter Spindle motor can achieve high starting torque with high speed using simple control algorithm. The Matlab based Motor model and inverter model are developed for verifying the proposed topology (Fig. 32.7).

Brushless DC (BLDC) Motors are used in various drive applications because of its high efficiency, power density and controllability over wide speed ranges. The speed requirement for the Spindle Motor especially those used in computer hard disks is increasing day by day, to facilitate faster data access. Such high speed operation could be achieved by three methods. The first method is to design the machine with low back EMF constant which offers high speed for a given system voltage. But the disadvantage is that a low back EMF constant results in low starting torque. To get high torque BLDC motor needs to be operated in bipolar mode and to get high speed BLDC motor should be operated in unipolar mode. In

Table 32.1 Switching table

Switches turn ON	Voltage level
S1,S4	Vdc
S2,S4	2Vdc
S2,S3	Vdc

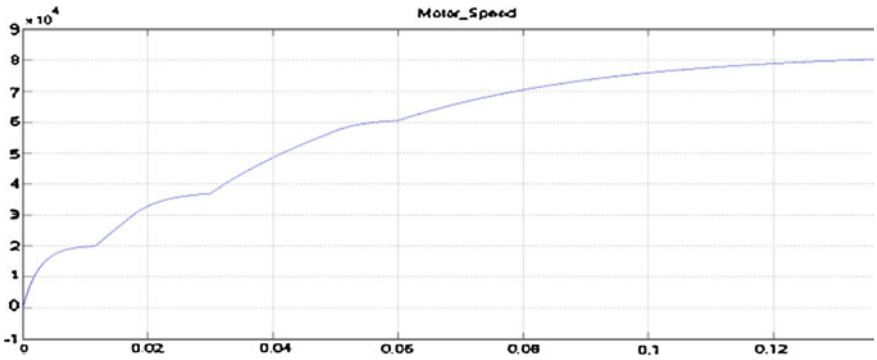


Fig. 32.8 Speed response

[2], a new converter is proposed which can achieve high starting torque with high speed by using 14 switches of same rating. Drawback of this circuit is, using more gate drives and more switches. This adds extra cost and space to the system. In this paper a new inverter topology is proposed which uses bipolar operation to achieve high torque during starting and unipolar operation to achieve high speed using only 8 switches. The model was verified with an inverter-motor mode developed using mat lab/Simulink (Table 32.1)

32.3.2 Simulation of Proposed Converter

Figure 32.8 shows the speed response of BLDC motor with proposed converter. From this figure it is clear that there four modes, in mode one supply is VDC and motor is operating is bipolar mode, in mode two supply is Vdc and motor is operating in unipolar mode. In mode three supply is 2Vdc and motor operating in bipolar mode. In mode four supply is 2Vdc and motor is operating in unipolar mode. With conventional technique motor final speed is 20,000 rpm with the proposed converter motor speed is 780,000 rpm due to Bipolar and unipolar converter and proposed novel converter.

32.4 Conclusion

The proposed cost effective inverter topology uses only eight switches. It drives a BLDC motor at high speeds with high starting torque using bipolar-starting and unipolar—running algorithm. The torque—speed characteristics of spindle motor can be extended to get high speeds with this low cost topology.

References

1. Civilian R, Stupak D (1995) Disk drive employing multi mode spindle drive system. US Patent 5471353, 3 Oct 1995
2. Jang GH, Kim MG (2005) A bipolar-starting and unipolar-running method to drive an HDD spindle motor at high speed with large starting torque. IEEE Trans Magn 41(2):750–755
3. Grochowski E, Hyot RF (1996) Future trends in hard disk drives. IEEE Trans Magn 32(3):1850–1854
4. Ede JD Zhu ZQ, Howe D (2001) Optimal split ratio control for high speed permanent magnet brushless DC motors. IN: Proceedings of the .5th international conference on electrical machines and systems, vol 2. pp 909–912
5. Chen SX, Jabbar MA Zhang OD, Lie ZJ (1996) New challenge: electromagnetic design of BLDC motors for high speed fluid film bearing spindles used in hard disk drives. IEEE Trans Magn 32(5):3854–3856

Chapter 33
Enhancement of ATC in Presence of SSSC Using Linear and Reactive Methods

Y. Chittemma, S. Lalitha kumari and A. Varaprasad Rao

Abstract Fast, accurate algorithms to compute network capabilities are indispensable for transfer-based electricity markets. Available Transfer Capability (ATC) is a measure of the remaining power transfer capability of the transmission network for further transactions. Transmission System Operators (TSOs) are encouraged to use the existing facilities more efficiently. One of the limitations of reactive ATC is the error produced by neglecting the effect of reactive power flows in line loading. This paper presents the determination of shunt reactive power compensation with Flexible AC Transmission System (FACTS) devices, the Static Synchronous Series Compensator (SSSC) to improve the transfer capability of a power system incorporating the Linear and Reactive power flows in ATC calculations. By redistributing the power flow, the ATC is improved. Studies on a sample 5-bus power system model are presented to illustrate the effect of shunt compensation along with line flow control.

Keywords Reactive and linear method ATC · PTDF · SSSC

Y. Chittemma (✉) · S. Lalitha kumari
Department of Electrical Engineering, GMR Institute of Technology,
Rajam, Rajam 532127, AP, India
e-mail: chittemma.y@gmrit.org

S. Lalitha kumari
e-mail: lalithakumari.s@gmrit.org

A. Varaprasad Rao
National Aerospace Laboratories, Department of Electrical Division, Bangalore, India
e-mail: avaraprasadrao@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_33,

33.1 Introduction

AVAILABLE transfer capability (ATC) determines the size of the largest transfer that can be implemented in a certain direction across the power grid without violating security constraints [1, 2]. The determination of ATC requires the continuation version of power flow, steady-state stability, voltage stability, and transient stability simulations. The ATC study starts with a base case that corresponds to an initial operating point computed from a power flow solution using usual data specifications at PQ buses, PV buses, and a slack bus *s*. The transfer direction is then specified by means of participation factors of source and sinks buses.

The traditional method to improve the transfer capability is to install phase shifter transformer and SVC. A key concept in the restructuring of the electric power industry is the ability to accurately and rapidly quantify the Available Transfer Capability (ATC) of the transmission system. This paper presents the application of one type of Flexible AC Transmission System (FACTS) device, the Synchronous Series Compensator (SSSC) to improve the transfer capability of a power system incorporating the Linear and Reactive power flows in ATC calculations.

33.2 Power Flow Control with SSSC

Better utilization of existing power system capacities by installing new power electronic controllers such as FACTS has become imperative. FACTS controllers are able to change, in a fast and effective way, the network parameters in order to achieve better system performance. FACTS controllers, such as phase shifter, shunt, or series compensation and the most recent developed converter-based power electronic controllers, make it possible to control circuit impedance, voltage angle, and power flow for optimal operation performance of power systems, facilitate the development of competitive electric energy markets.

33.3 Multicontrol Functional Model of the SSSC

33.3.1 Operation Principles of the SSSC

A SSSC usually consists of a coupling transformer, an inverter, and a capacitor. As shown in Fig. 33.1, the SSSC is series connected with a transmission line through the coupling transformer. It is assumed here that the transmission line is series connected with the SSSC via its bus. The active and reactive power flows of the SSSC branch entering the bus are equal to the sending end active and reactive power flows of the transmission line, respectively. In principle, the SSSC can

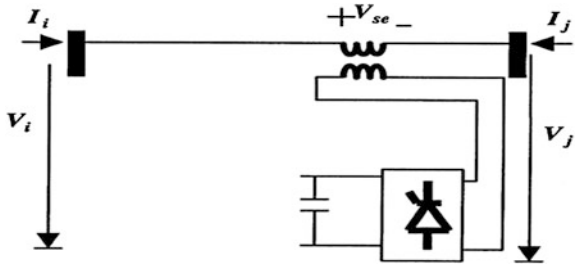


Fig. 33.1 SSSC operating principles

generate and insert a series voltage, which can be regulated to change the impedance (more precisely reactance) of the transmission line. In this way, the power flow of the transmission line or the voltage of the bus, which the SSSC is connected with, can be controlled.

33.3.2 Equivalent Circuit and Power Flow Constraints of the SSSC

An equivalent circuit of the SSSC as shown in Fig. 33.2 can be derived based on the operation principle of the SSSC. In the practical operation of the SSSC, V_{se} can be regulated to control the power flow of line i–j.

It is proposed to improve the performance of the system by in presence of SSSC using all of its advantages. The SSSC equivalent circuit for steady state model is shown in Fig. 33.1.

In the equivalent circuit, $V_{se} = V_{se}\angle\theta_{se}$, $V_i = V_i\angle\theta_i$, $V_j = V_j\angle\theta_j$, then the power flow constraints of the SSSC are:

$$\begin{aligned} P_{ij} = {} & V_i^2 g_{ii} - V_i V_j (g_{ij}\cos(\theta_i - \theta_j) + b_{ij}\sin(\theta_i - \theta_j)) \\ & - V_i V_{se} (g_{ij}\cos(\theta_i - \theta_{se}) + b_{ij}\sin(\theta_i - \theta_{se})) \end{aligned} \tag{33.1}$$

$$\begin{aligned} Q_{ij} = {} & - V_i^2 b_{ii} - V_i V_j (g_{ij}\sin(\theta_i - \theta_j) - b_{ij}\cos(\theta_i - \theta_j)) \\ & - V_i V_{se} (g_{ij}\sin(\theta_i - \theta_{se}) - b_{ij}\cos(\theta_i - \theta_{se})) \end{aligned} \tag{33.2}$$

$$\begin{aligned} P_{ji} = {} & V_j^2 g_{jj} - V_i V_j (g_{ij}\cos(\theta_j - \theta_i) + b_{ij}\sin(\theta_j - \theta_i)) \\ & + V_j V_{se} (g_{ij}\cos(\theta_j - \theta_{se}) + b_{ij}\sin(\theta_j - \theta_{se})) \end{aligned} \tag{33.3}$$

$$\begin{aligned} Q_{ji} = {} & - V_j^2 b_{jj} - V_i V_j (g_{ij}\sin(\theta_j - \theta_i) - b_{ij}\cos(\theta_j - \theta_i)) \\ & + V_j V_{se} (g_{ij}\sin(\theta_j - \theta_{se}) - b_{ij}\cos(\theta_j - \theta_{se})) \end{aligned} \tag{33.4}$$

Where $g_{ij} + jb_{ij} = 1/Z_{se} = g_{ii} = g_{ij} = b_{ii} = b_{ij} = g_{jj}g_{ij}$, and $b_{jj} = b_{ij}$

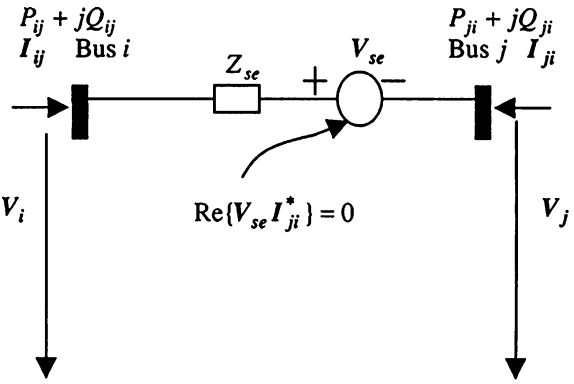


Fig. 33.2 SSSC equivalent circuit

33.4 Available Transfer Capability

For illustration, we assume a point-to-point transfer from the slack bus s, to any bus i, and we would like to maximize this transfer without exceeding the flow limits of any line or transformer.

$$\rho_{jk,i} = \frac{\partial P_{jk}}{\partial P_s} = -\frac{\partial P_{jk}}{\partial P_i} \tag{33.5}$$

These PTDFs are essentially current dividers in linear circuit theory. As such, they are large-change sensitivities and can be used to predict the change in the line flow (line $j-k$) due to a transfer (bus $s$to bus i) as

$$\Delta P_{jk} = \rho_{jk,i} \Delta P_s = -\rho_{jk,i} \Delta P_i \tag{33.6}$$

Note that $\Delta P_s = -\Delta P_i$ is the amount of transferred power from s to i. For a given positive line flow limit $P_{jk}^{\max}$, assumed equal to the line MVA rating, and an initial positive line flow P_{jk}^{0}, the size of the transfer that drives the line to its limit is equal to

$$\Delta P_s^{jk} = \begin{cases} \frac{P_{jk}^{\max} - P_{jk}^{0}}{\rho_{jk,i}} & \rho_{jk,i} \rangle 0 \\ \frac{-P_{jk}^{\max} - P_{jk}^{0}}{\rho_{jk,i}} & \rho_{jk,i} \langle 0 \end{cases} \tag{33.7}$$

In order to determine ATC, the minimum value of ΔP_s^{jk} among all lines in the system is determined

$$ATC_{s \to i} = \min\{\Delta P_s^{jk} : \mathit{alllines} \quad jk\} \tag{33.8}$$

Note that it is the linear relation between the transfer and the line flows that makes linear ATC the fastest algorithm for transfer studies. Since, in general, the complex flow at the sending and receiving ends of a transmission line are different, there is a corresponding k -end circle given by the equation

Table 33.1 Atc results for The 14-bus system with SSSC

Transfer direction	Limiting line	NR PTDF	NR reactive ATC Δp_{ij}	NR linear ATC Δp_{ij}	SSSC PTDF	SSSC reactive ATC Δp_{ij}	SSSC linear ATC Δp_{ij}
1–2	2–1	0.88638	1.0851	1.2496	0.8124	1.0646	1.1273
1–8	8–1	1.4872	1.9186	2.1864	0.7086	1.4727	1.4979
2–3	3–2	0.6380	1.5022	1.6323	0.3771	1.3167	1.6113
2–6	6–2	0.5915	1.5564	1.8244	1.0131	1.0000	1.0679
2–8	8–2	0.5371	1.7127	2.0106	0.7744	1.3218	1.3935
3–6	6–3	0.6431	1.5236	1.5859	0.5382	1.0797	1.5818
4–11	11–4	0.7543	1.1315	1.5197	0.7317	1.0797	1.5168
4–12	12–4	0.6626	1.3367	1.6817	0.6623	1.0342	1.6812
4–13	13–4	0.7398	1.0379	1.6654	0.7358	1.2406	1.6744
7–5	5–7	1.0000	1.0000	1.0000	1.0000	0.8951	1.0000
7–9	9–7	0.8270	1.2090	1.2091	0.8157	1.2241	1.2259
9–10	10–9	0.8962	1.1158	1.1158	0.8866	1.0279	1.1278
9–14	14–9	0.7222	1.3645	1.4845	0.7097	1.0486	1.4091
6–9	9–6	0.2951	1.4737	1.1737	0.2760	1.1742	1.1942
8–4	4–8	0.7797	1.2824	1.2824	0.7226	1.2615	1.3839

$$(P_{kj} - Pk_{kj\Theta})^2 + (Q_{kj} - Q_{kj\Theta})^2 = (S_{kj\Theta})^2 \tag{33.9}$$

Where in general $P_{jk\Theta} \neq P_{kj\Theta}$ and $Q_{jk\Theta} \neq Q_{kj\Theta}$.

As the transfer increases, the flow in the line varies but all feasible operating points in the $P_{jk} - Q_{jk}$ plane lie on the *operating circle* given by (10). On the other hand, in this plane, the MVA rating of the line can be represented by a circle with center at the origin and radius equal to the thermal limit $S_{jk}^{\max}$. This is referred to here as the *limiting circle*. Given an initial power system state without overloaded lines, the ATC calculation must determine the maximum amount ΔP_s for a transfer s to i such that the flows lie inside the limiting circle (i.e., $|S_{jk}| \leq S_{jk}^{\max}$ for all $j - k$ lines (ends j and k) in the system.

33.5 Simulation Results

33.5.1 Simulation of a 14-Bus System

Another sample 14-bus system is considered to illustrate the implementation of SSSC for ATC enhancement. The system consists of 4 generator buses, 11 load buses, and 16 transmission lines. The SSSC is located in the line connected between 3 and 4. The real and reactive power settings of the SSSC are 40 MW and 2 MVAR. The results of the 14-bus system are given in Table 33.1

From the Table, it can be seen that control of line flows by SSSC the ATC capability of from the system is improved and bus voltages are also improved significantly. The magnitudes of real power line flows are also redistributed significantly from high values to low values.

33.6 Conclusions

Improving of ATC is an important issue in the current deregulation environment of power systems. ATC can be limited usually by heavily loaded circuits and buses with relatively low voltages. It is well known that FACTS technology can control voltage magnitude. Using these device may redistribute the load flow regulating bus voltage. Therefore, it is worthwhile to investigate the effect of FACTS controller on the ATC. The paper has presented the application of a new generation FACTS device, the SSSC in determination of ATC along with the line flow control using the real and reactive methods.

References

1. (1996) NERC transmission transfer capability task force. Available transfer capability definitions and determination. North American Electric Reliability Council, Princeton
2. Ejebe GC, Waight JG, Santos-Nieto M, Tinney WF (2000) Fast calculation of linear available transfer capability. IEEE Trans Power Sys 15:1112–1116
3. Pavella M, Ruiz-Vega D, Giri J, Avila-Rosales R (1999) An integrated scheme for on-line static and transient stability constrained ATC calculations. In IEEE Power Eng Soc Summer Meet 1:273–276
4. Repo S (1998) Real-time transmission capacity calculation in voltage stability limited power systems. In: Proceedings of the bulk power system dynamics and control IV-restructuring, Santorini, Greece, 24–28 Aug 1998
5. Gravener MH, Nwankpa C (1999) Available transfer capability and first order sensitivity. IEEE Trans Power Sys 14:512–518
6. Ajjarapu V, Christy C (1992) The continuation power flow: A tool for steady state voltage stability analysis. IEEE Trans Power Sys 7:416–423
7. Stott B, Marinho JL (1979) Linear programming for power system network security applications. IEEE Trans Power Apparat Sys PAS-98:837–848
8. Landgren GL, Anderson SW (1973) Simultaneous power interchange capability analysis. IEEE Trans Power Apparat Sys PAS-92:1973–1986
9. Schauder CD, Grenhardt M, Stacey E, Lemak T, Gyugyi L (1995) Development of a ± 100 Mvar static condenser for voltage control of transmission systems. IEEE Trans Power Deliv 10(3):1486–1493
10. Gyugyi L, Shauder CD, Sen KK (1997) Static synchronous series compensation of transmission lines. IEEE Trans Power Deliv 12(1)406–413
11. Gyugyi L, Shauder CD, Williams SL, Reitman TR, Torgerson DR, Edris A (1995) The unified power flow controller a new approach to power transmission control. IEEE Trans Power Deliv 10(2):1085–1093

12. Gyugyi L, Sen KK, Schauder CD (1999) The interline power flow management in transmission system. IEEE Trans Power Deliv 4(3)1115–1123
13. North American Reliabilty Council (1996) Available transfer capability definition and determination
14. Srinu Naik R, Vaisakh K, Anand Kumar K (2009) Identification of overloading line using ATC linear methods with PTDF. Int J Electron Electr Eng 04:6, SUMMER-2009
15. Srinu Naik R, Vaisakh K, Anand Kumar K (2009) ATC enhancement with TCSC using linear and reactive methods. Int J Electron Electr Eng 03:4, SUMMER-2009
16. Srinu Naik R, Vaisakh K, Kiran Chandra P (2010) Effect of shunt reactive power compensation on ATC using linear methods with FACTS. In: International conference on recent advancements in electrical sciences (ICRACE-2010, IEEE) 8th and 9th Jan 2010 at K.S.R College of Engineering, Tiruchengode-637215
17. Srinu Naik R, Vaisakh K, Anand Kumar K (2010) Determination of ATC with PTDF using linear methods in presence of TCSC. In: International conference on electrical energy systems (ICEES-2010, IEEE), on 26th–28th Feb 2010 at Singapore

Chapter 34
Improvement of Power Quality and Performance Analysis of PV Fed UPQC in Utility Connected System

S. Balasubramaniyan, T. S. Sivakumaran, Thulasidharan and D. Balamurugan

Abstract This study presents a combined operation of the Unified Power Quality Conditioner (UPQC) with Photovoltaic cell system. The proposed system consists of a series inverter, a shunt inverter and a battery connected Photovoltaic array which is connected with the DC link of UPQC through a boost converter. The proposed system compensates the voltage sag, voltage swell, voltage interruption, harmonics, real and reactive power compensation. PV fed UPQC system is simulated in single phase 14-bus and three phase single bus system. The proposed system is validated with the results of computer simulation and hardware implementation. Sag is created by applying heavy load and swell occurs during light load conditions. These power quality problems are compensated with the help of UPQC fed Photovoltaic arrays.

Keywords UPQC · PV

S. Balasubramaniyan · Thulasidharan · D. Balamurugan
Department of Electrical and Electronics Engineering, Mailam Engineering College, Mailam 604 304, India
e-mail: balasipi@yahoo.co.in

Thulasidharan
e-mail: sthulasimails@gmail.com

D. Balamurugan
e-mail: balamrgnd4@gmail.com

T. S. Sivakumaran (✉)
Department of Electrical and Electronics Engineering, Arunai College of Engineering, Tiruvannanmalai 606 603, India
e-mail: sivakumaranphd@gmail.com

T. S. Sivakumaran
Anna University Chennai, Chennai, Tamilnadu, India

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_34,

34.1 Introduction

One of the most interesting structures of energy conditioner is two back-to-back connected DC/AC fully controlled converters. In this case, depending on the control scheme, the converters have different compensation techniques. For example, they can function as active series and shunt filters to compensate simultaneously both the load current harmonics and supply voltage fluctuations [1]. Increase in applications of electronic equipments has heightened the power quality problems [2]. An active shunt filter is a suitable device for current-based compensation. It can compensate current harmonics and reactive power [3]. The active series filter is normally used for voltage harmonics and voltage sag compensation. The two inverters of UPQC share one DC link capacitor for compensating the voltage sag and swell. The harmonic current and voltage affects the power flow and voltage stability [4]. Nevertheless, UPQC cannot compensate the voltage interruption due to lack of energy source in its DC link. Numerous studies are available on operation of UPQC and distributed generation [5]. Combined operation of UPQC and photovoltaic is proposed, in which the battery is connected to UPQC DC link through an uncontrolled rectifier [6, 7]. The VA rating of series and shunt inverters of UPQC are estimated for proposed system [8].

34.2 Block Diagram of UPQC

The block diagram of the proposed system is shown in Fig. 34.1. Here the battery energy is stored from the PV cell which is fed to the capacitor in the UPQC system. A PWM control scheme is used as controller [4, 9, 10]. Then using the series and shunt compensation the power quality problems such as voltage sag, voltage swell, real and reactive power are compensated [11].

34.3 Proposed System

The UPQC has a combination of series and shunt converter which is connected to a dc link capacitor. In the proposed system DC link capacitor is connected to a Battery Energy Storage System (BESS). The DC voltage for BESS is fed by PV Array [12–14] using MISO DC–DC converter. Here the BESS stores energy from PV array during the day time. To get a single DC output, the DC outputs from several PV arrays are combined using MISO DC–DC converter (Boost Converter) [15]. Thus the proposed system effectively compensates the power quality problems using the excess energy fed by PV Arrays (Fig. 34.2).

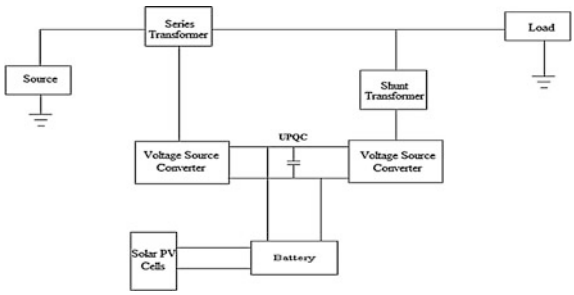


Fig. 34.1 Block diagram of PV fed UPQC

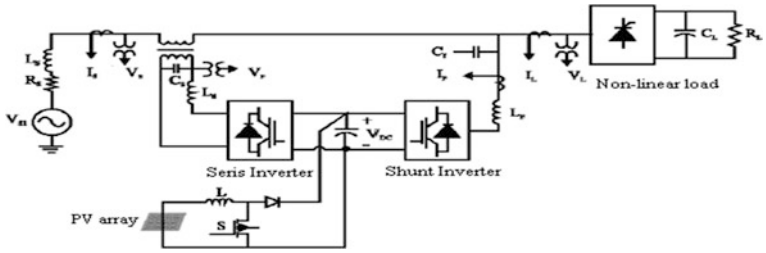


Fig. 34.2 Circuit diagram of PV fed UPQC system

34.4 Simulation Results

In the proposed system MATLAB software is used for the line compensation. Figure 34.3 represents the simulation of PV fed UPQC [16] with 14-bus systems. The corresponding voltage sag, voltage swell, real power and reactive power output is shown. Table 34.1 represents the comparison of the real and reactive power, with and without compensation at different busses.

Figure 34.4 represents the effective compensation of voltage sag across load 1. At 0.2 s load 2 (Unbalanced Load) is added which causes voltage sag, at 0.3 s the voltage is compensated across load 1.

Figure 34.5 shows the compensation of voltage swell. Here at 0.3 s the voltage swell across load-1 occurs due to light load condition and at 0.35 s voltage swell is compensated with the proposed system.

Figure 34.6 shows the effective compensation of real and reactive power across bus 3 under voltage sag condition

Figure 34.7 shows the three phase implementation of the proposed system

Figure 34.8 shows the occurrence of voltage sag at 0.3 s and compensation of the sag at 0.4 s by implementing the proposed technique.

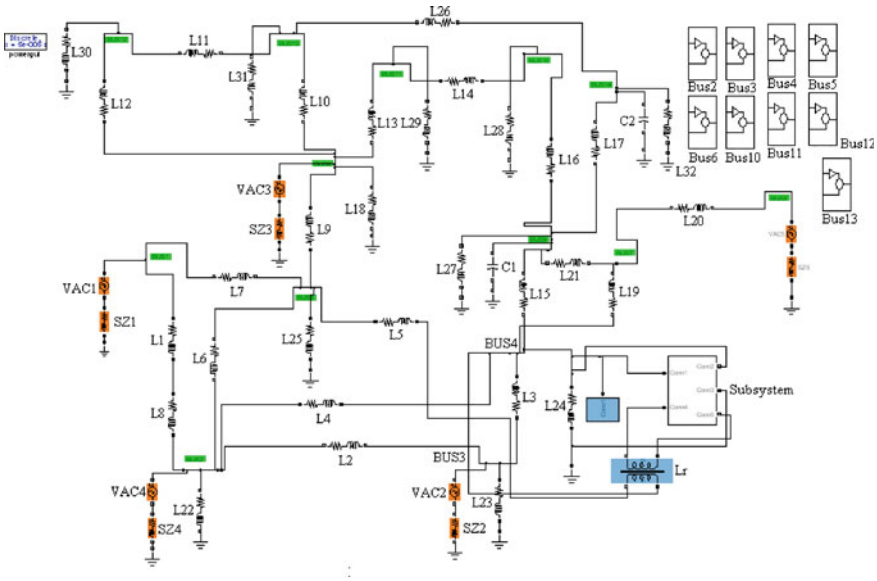


Fig. 34.3 14-bus PV fed UPQC

Table 34.1 Comparison of real and reactive power at different buses in 14-bus proposed system

Bus	Real power (MW) without compensation	Real power (MW) with compensation	Reactive power (MVA) without compensation	Reactive power (MVA) with compensation
7	0.135	0.136	0.032	0.33
1	0.267	0.267	0.279	0.28
3	0.21	0.356	0.693	1.09
4	0.992	1.12	1.03	1.13

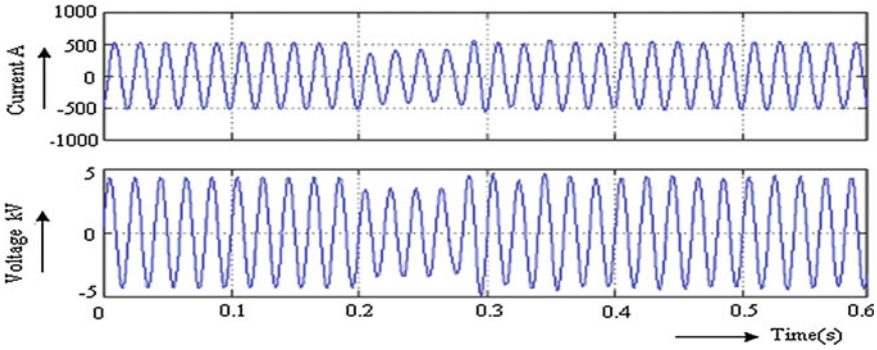


Fig. 34.4 Current and voltage sag compensation across load-1

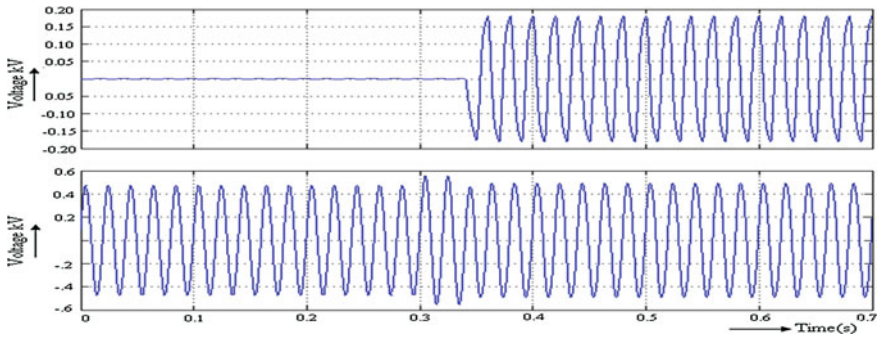


Fig. 34.5 Current and voltage swell compensation across load-1

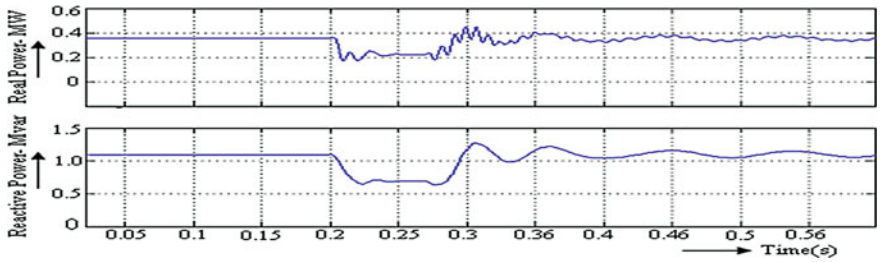


Fig. 34.6 Real power and reactive power across bus 3

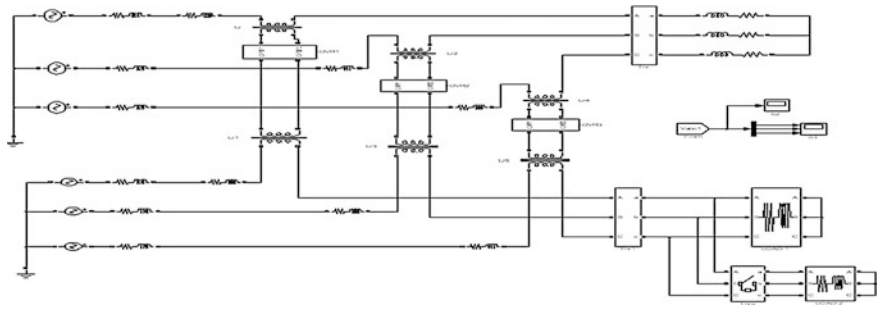


Fig. 34.7 Three phase circuit model with UPQC

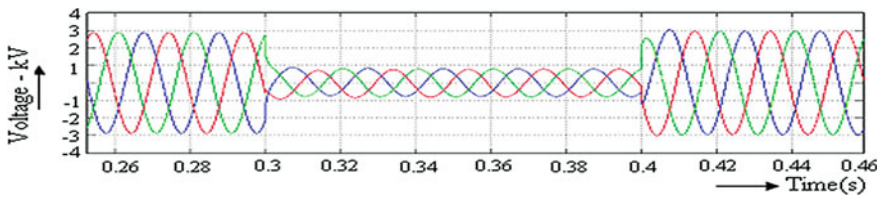


Fig. 34.8 Three phase voltage output with compensation

34.5 Conclusions

This study describes a combined operation of the unified power quality conditioner with photovoltaic generation system. The proposed system can compensate voltage sag, voltage interruption, harmonic generation and real, reactive power compensation. The VA rating of series and shunt inverters of UPQC are estimated for proposed system. The economic saving due to use of proposed system is estimated nearly 20 %. Series converter draws supply from the main source and acts as a controlled rectifier for controlling the terminal voltage and shunt converter controls the power flow. Thus both the converters control the voltage sag and power flow. The circuit with series converter and shunt inverter section is simulated. The circuit is also simulated with photo volatile system, series and shunt inverter. The performances of both the model are compared. The combined circuit gives better performance in power quality problems.

References

1. Davari M, Graduate Student Member, IEEE, Ale-Emran SM, Yazdanpanahi H, Gharehpetian GB, Senior Member (2011) IEEE: modeling the combination of UPQC and photovoltaic arrays with multi-input single-output DC–DC converter. IEEE Trans
2. (1995) IEEE Std 1159: IEEE recommended practice for monitoring eclectic power quality
3. Fujita H, Akagi H (1998) The unified power quality conditioner: the integration of Series- and shunt-active filters. IEEE Trans Power Electron 13
4. Kesler M (2010) Synchronous reference frame based application design and analysis of unified power quality conditioner. Ph.D. Dissertation, Kocaeli University of Kocaeli, Turkey
5. Basu M, Das SP, Dubey GK (2007) Comparative evaluation of two models of UPQC for suitable interface to enhance power quality. Electr Power sys Res 77:821–830
6. Cavalcanti MC, Azevedo GMS, Amaral BA, Neves FAS (2005) A photovoltaic generation system with unified power quality conditioner function. In: 31st annual conference of IEEE industrial electronics society, 6–10, Brazil, pp. 750–755
7. Han B, Bae B, Kim H, Baek S (2006) Combined operation of unified power quality conditioner with distributed generation. IEEE Trans Power Del 21:330–338
8. Kumar GS, Kumar BK, Kumar MM (2010) Optimal VA loading of UPQC during mitigation of unbalanced voltage sags with phase jumps in three-phase four wire distribution system. In: International conference on power system technology, Beijing, China
9. Khadkikar V, Chandra A (2008) A new control philosophy for a unified power quality conditioner (UPQC) to coordinate load-reactive power demand between shunt and series inverters. IEEE Trans Power Del 23(4):2522–2534
10. Forghani M, Afsharnia S (2007) Online wavelet transform-based control strategy for UPQC control system. IEEE Trans 22:481–491
11. Khadkikar V, Chandra A (2011) A novel concept of simultaneous voltage sag/swell and load reactive power compensations utilizing series inverter of UPQC. IEEE Trans Power Electron 26(9)
12. Lorenzo E (1994) Solar electricity engineering of photovoltaic systems. Internationally recognized expert engineers and scientists of IES Solar Energy Institute

13. Altas IH, Sharaf AM (2007) A photovoltaic array simulation model for MATLAB-simulink GUI environment. In: International conference on clean electrical power. ICCEP'07, pp 341–345
14. Esram T, Chapman PL (2007) Comparison of photovoltaic array maximum power point tracking techniques. IEEE Trans Energy Convers 22
15. Matsuo H, Lin W, Kurokawa F, Shigemizu T, Watanabe N (2004) Characteristics of the multiple- input DC–DC converter. IEEE Trans Ind Electron 51
16. Noroozian R, Abedi M, Gharehpetian GB, Hosseini SH (2007) On grid and off-grid operation of multi-input single-output DC–DC converter based fuel cell generation system. In: ACEMP'07 and ELECTROMOTION'07 joint meeting, Bodrum, Turkey
17. Khadkikar V (2012) Enhancing electric power quality using UPQC: a comprehensive overview. IEEE Trans Power Electron 27(5)

Chapter 35
Design of Adaptive FLANN Based Model for Non-Linear Channel Equalization

Sidhartha Dash, Santanu Kumar Sahoo and Mihir Narayan Mohanty

Abstract Wireless Communication systems require the most efficient techniques for reception of error-less data with high data rate. The channels introduce both linear and non-linear distortions. ISI plays a major role in this field. Also these channels contaminate the received sequence with random fluctuation. In this paper, an adaptive algorithm based on FLANN has been developed for channel equalization with analysis of MSE. The FLANN is developed with LMS technique as well as sign regressor based LMS technique and the results are compared. Also the result is compared with the standard adaptive LMS algorithm. The signed FLANN based model shows better performance as compared to LMS based FLANN model.

Keywords Channel equalization · Sign regressor based LMS · Functional link artificial neural network · Signed regressor FLANN

35.1 Introduction

High quality and high-speed is the greatest demand in wireless communication. One of the major limiting factors is inter-symbol interference (ISI). ISI may be due to one or more of the factors like: frequency selective characteristics of the

S. Dash (✉) · S. K. Sahoo · M. N. Mohanty
ITER, Siksha 'O' Anusandhan University, Jagamara, Bhubaneswar, Odisha,
e-mail: sidharthadashiter@gmail.com

S. K. Sahoo
e-mail: sahoosantanu@gmail.com

M. N. Mohanty
e-mail: mihir.n.mohanty@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_35,

channel, time varying multipath propagation that is prominent in mobile communication. Equalizers are usually used to compensate the received signals which are corrupted by the inevitable noise, interference and signal power attenuation introduced by communication channels during transmission [1]. Traditionally linear transversal filters (LTF) [2] are commonly used in the design of channel equalizers. The linear equalizers, however, fail to work well when transmitted signals have encountered severe nonlinear distortion. The use of large constellations provides bandwidth efficient modulation. Quadrature Phase Shift Keying (QPSK) type modulation techniques have constellations, in which signal points are uniformly spread. Information is carried by both signal amplitude and phase; hence they are not constant envelopes. Thus, efficient nonlinear power amplifiers cannot be utilized in the transmitter, without equalization in the receiver [3]. Since Wiener's classical work on adaptive filters [1], the mean-square-error (MSE) criterion has been the workhorse of function approximation and optimal filtering. A variety of approaches employing the MSE criterion have been taken towards solving this nonlinear channel equalization problem.

In this paper the use of signed Regressor LMS algorithm and signed Regressor FLANN algorithm for non-linear channel equalizers have been investigated. Both the algorithms are applied for the design of the channel equalizers in the complex environment.

The paper is organized as follows: Sect. 35.2 outlines the system model and establishes the algorithms under consideration to determine the optimum filter coefficients with respect to a minimum mean square (MSE) design criteria. Sect. 35.3 describes the conventional LMS and signed regressor LMS algorithm and Sect. 35.4 outlines the FLANN as well as signed regressor FLANN. Sect. 35.5 discusses the extensive simulation results in non-linear channel equalization in terms of their effectiveness and Sect. 35.6 concludes the work.

35.2 Model for Channel Equalization

Adaptive filtering techniques are necessary consideration when a specific signal output is desired but the coefficients of filter cannot be determined at the outset. Adaptive channel equalizers have played an important role in digital communication systems. Adaptive equalization at the receiver removes the effects of ISI [4]. In an adaptive equalizer the current and past values of the received signal are linearly weighted by equalizer coefficients and summed to produce the output [5]. An adaptive equalizer is an equalization filter that automatically adapts to time-varying properties of the communication channel. It is frequently used with coherent modulations such as phase shift keying, mitigating the effects of multipath propagation and Doppler spreading. The block diagram of a communication system based on equalizer is shown in Fig. 35.1.

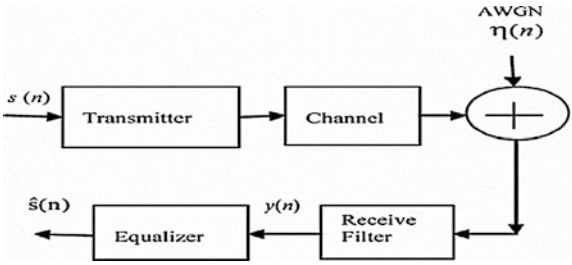


Fig. 35.1 Block diagram of communication system

Here, $s(n)$ is a transmitted symbol sequence, $\eta(n)$ is additive white Gaussian noise, $y(n)$ is a received signal sequence sampled at the rate of the symbol interval T, and $\hat{s}(n)$ is an estimate of the transmitted sequence $s(n)$. The received signal is represented as

$$r(n) = s(n) * h(n) + \eta(n) \tag{35.1}$$

where $h(n)$ is the impulse response of channel .

The non linear channel equation is defined as

$$y(n) = a_0 x(n) + a_1 \tanh^2(x(n-1)) \tag{35.2}$$

where the nonlinearity values of $a_0 = 1$ and $a_1 = 0.8$.

35.3 Algorithm for MSE

Least mean squares (LMS) algorithms are used in adaptive systems to find the coefficients that relate to producing the least mean squares of the error signal [6]. In LMS Algorithm the simplified cost function, $\xi_{LMS}(n)$ is given by

$$\xi_{LMS}(n) = 1/2e^2(n) \tag{35.3}$$

where $e(n)$ is defined as the error function which is the difference between the original and estimated weight of the filter. The cost function in (3) can be thought of as an instantaneous estimate of the MSE (mean square error) that is most useful in practical applications. The update LMS equation is given by

$$W(n+1) = W(n) + \mu e(n) X(n) \tag{35.4}$$

Equation (35.4) requires only multipliers and adders to implement. The performance of LMS algorithm is affected severely in presence of outliers [7]. A simple and efficient normalized Signed Regressor LMS algorithm is suitable for applications requiring large signal to noise ratios with less computational complexity [8]. The convergence speed of sign LMS is faster than the conventional

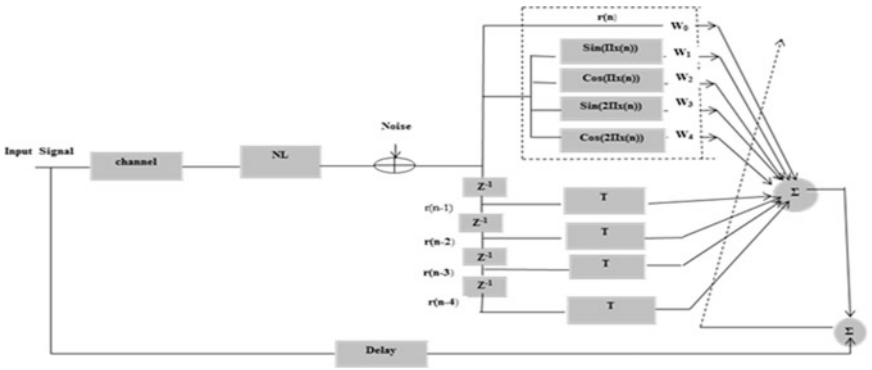


Fig. 35.2 Structure of FLANN

LMS based realizations. The weight update equation of Signed Regressor LMS is given by

$$W(n+1) = W(n) + \mu e(n) sign(X(n)) \tag{35.5}$$

35.4 FLANN and Signed Regressor FLAN

The basic principle of an FLANN is to expand the dimensionality of the input signal space by using a set of linearly independent functions. The expansion can produce complicated decision boundaries at the output space, so the FLANN is capable of dealing with linear inseparable problems [9]. The structure of FLANN used in channel equalization problem is shown in Fig. 35.2. It requires more no of sine and cosine functions that implies the computational complexity is more. The FLANN is a single layer network, where need of hidden layers is removed. Here the functional link acts on an element of a pattern or on the entire pattern itself by generating a set of linearly independent functions, and then evaluating these functions with the pattern as the argument. Thus, separability of input patterns is possible in the enhanced space [1]. However, the FLANN structure offers less computational complexity and higher convergence speed than MLP because of its single layer structure. The expanded function make use of a functional model comprising of a subset of orthogonal sine and cosine basis functions and the original pattern along with its outer products.

The BP algorithm, is used to train the network, becomes very simple because of absence of any hidden layer. The learning ability and the justification for the use of trigonometric polynomial are described in [3]. Inspiring from the fact that signed regressor LMS performs faster as compared to conventional LMS, the proposed Signed regressor FLANN technique uses the signed inputs for the weight update equation.

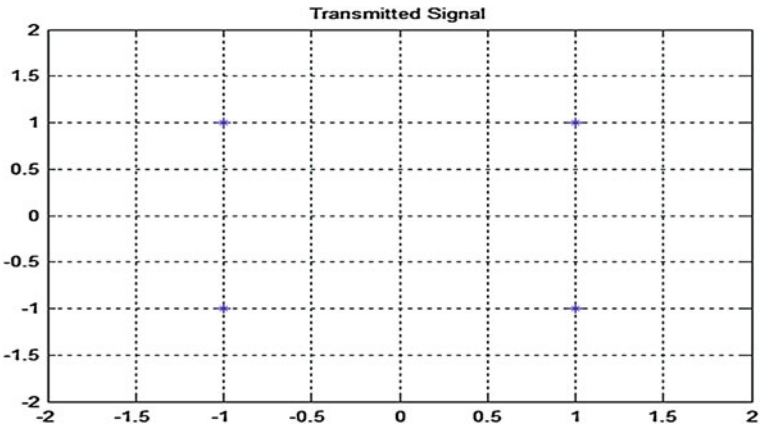


Fig. 35.3 Transmitted signal in QPSK in real valued environment

Fig. 35.4 Received signal in QPSK through non-linear channel

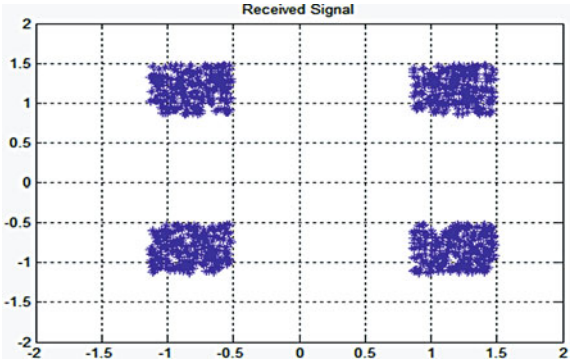
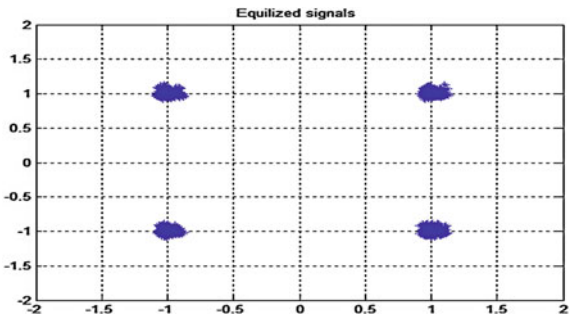


Fig. 35.5 Equalized signal in QPSK by signed FLANN



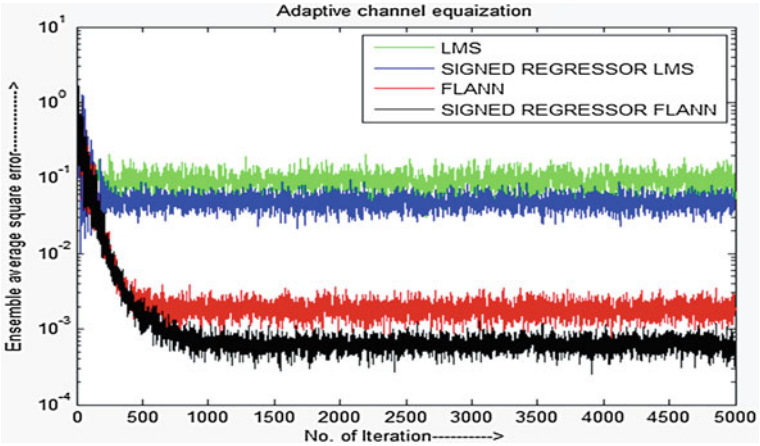


Fig. 35.6 MSE using LMS and FLANN in real valued environment

35.5 Result and Discussion

It has been simulated by taking learning rate parameter μ as 0.01 and SNR as 15db. Here 10,000 random samples have been taken between [−0.5 0.5]. The weight update equation is used for non-linear channel equalization. The non-linear channel used in this proposed model is taken from Eq. (35.2). The result related to QPSK is shown using FLANN (Fig. 35. 3–35. 5) and the error has been analyzed for a non-linear channel using adaptive FLANN (Fig. 35. 6, 35. 7). Also it shows the comparative analysis between FLANN and adaptive FLANN.

35.5.1 Channel Equalization in QPSK Using Adaptive FLANN

Signals in QPSK are shown in Figs. 35.3, 35.4 and 35.5.

35.5.2 MSE Analysis for Non-Linear Channel

MSE using LMS and FLANN in real and complex valued environments is shown in Figs. 35.6 and 35.7 respectively.

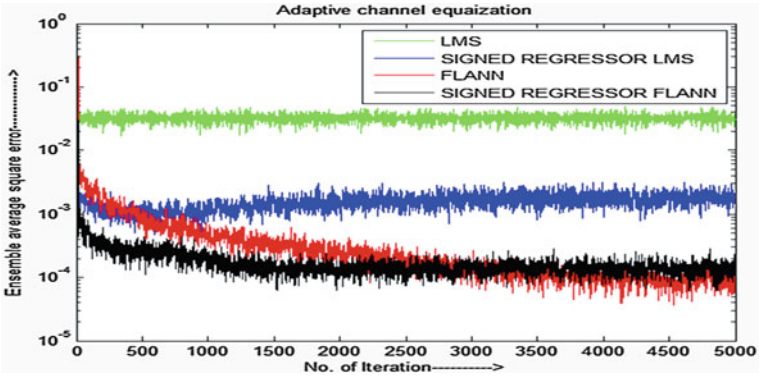


Fig. 35.7 MSE using LMS and FLANN in complex valued environment

35.6 Conclusion

In this paper, Signed Regressor FLANN based non-linear channel equalizer has been presented over real-valued and complex-valued non-linear channels. The simulation results show that such algorithm is capable of constructing a simple network whose performance is close to the optimal solution. There is no need to estimate in advance the number and locations of the equalizer input states. Thus, we can set conditions for increasing the demand for hybrid systems, which also makes the architecture of the constructed FLANN fairly simple.

References

1. Proakis JG (2001) Digital communications, 4th edn. McGraw-Hill, New York
2. Qureshi SUH (1985) Adaptive equalization. Proc IEEE 73:1349–1387
3. Erdogmus1 D, Rende D, Principel JC, Wong TF (2012) Nonlinear channel equalization using multilayer perceptrons with information-theoretic criterion. In: IEEE International Conference on Neural networks for Signal Processing XI [ICNNSP], pp 443–451
4. Guha DR, Patra SR (2009) Channel equalization for lSI channels using Wilcoxon generalized RBF. In: Fourth international conference on industrial and information systems, ICIIS, Sri Lanka, 28–31 Dec 20M
5. Patra JC, Pal RN (1995) Functional link artificial neural network based adaptive channel equalization. EURASIP Signal Process J 43(2):181–195
6. Haykin S (2004) Neural networks. A comprehensive foundation, 2nd edn. Pearson Education, New Delhi
7. Dash S, Mohanty MN (2012) Analysis of outliers in system identification using WLMS algorithm. In: IEEE International conference on computing, electronics and electrical technologies [ICCEET], p 802–806

8. Haykin S (2003) Adaptive filter theory. PHI India
9. Jatoth RK, vaddadi MSBS, Kanoop SSVK (2009) An intelligent functinal link artificial neural network for channel equalization. In: Proceedings of the 8th WSEAS international conference on signal processing, robotics and automation [ISPRA], pp 240–245
10. Mckean JW (2004) Robust analysis of linear models. Stat Sci 19(4):562–570

Chapter 36
A Security Framework for DDoS Detection In MANETs

P. Devi and A. Kannammal

Abstract Mobile Ad-hoc Network (MANET) adopts distributed wireless communication without a centralised control. It is more vulnerable to Denial of Service and Distributed Denial of Service attacks due to dynamic topology, limited physical security and decentralized approach. These attacks may collapse the entire communication networks. The detection of such attacks will improve the network security. This paper produces some clarification and a framework based on the Cluster Analysis to identify and to isolate the attacker from the network for detecting DDoS attack. The traffic is involved for XOR Marking to differentiate legitimate and non-legitimate data packets. Thus origin nodes of DDoS attacks are traced and isolated. Preliminary experiments are done with 2000 DARPA Intrusion Detection Scenario Specific Data Set to evaluate our method. The experimental results show that the proposed system is effective and efficient to identify DDoS attack.

Keywords DDoS attack · DDoS detection · Cluster analysis · XOR marking · Security framework

P. Devi (✉)
Department of Computer Applications, Anna University of Technology, Coimbatore, Tamilnadu, India
e-mail: devipichaimuthu@gmail.com

A. Kannammal
Department of Computer Applications, Coimbatore Institute of Technology, Coimbatore, Tamilnadu, India
e-mail: kanaphd@yahoo.co.in

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_36,

36.1 Introduction

Mobile Ad-hoc Network (MANET) is a distributed network because mobile nodes in the network subject to change and create paths dynamically among themselves to broadcast packets. Each node functions as a router to forward packets if it is not an end node. MANET can be viewed as a random graph because the nodes in the wireless network keep on moving. The nodes can move anywhere and organize themselves into the network. This makes MANET more helpful in establishing communication at the time of natural disasters. Though wireless ad hoc communication is inevitable in our life, it carries security issues along with it. This paper focuses on Distributed Denial of Service attack, which causes massive danger to the society.

The key plan of Distributed denial of Service (DDoS) is to consume system as well as network resources, thus the entire network is disabled. Compromised nodes initiate flooding heavy network traffic which paralyses all genuine nodes. The malicious traffic generated by the non-legitimate nodes will be very high that the victim node cannot afford to it. Also, it is difficult to identify these attacks since most of the IP packets are created in a fake manner.

In this paper, we propose a security framework for DDoS detection which incorporates cluster analysis and XOR marking. This method is used to differentiate legitimate and non-legitimate traffic to isolate compromised nodes. The remainder of the paper is organized as follows. Section 36.2 describes the literature works related to DDoS attacks. Section 36.3 explains the proposed security framework to overcome the DDoS attacks in MANET. The performance evaluation and justification of proposed mechanism over other schemes is discussed in Sect. 36.4. Conclusion of proposed scheme is in Sect. 36.5.

36.2 Related works

DDoS attack is a thriving research area and a lot of research is going on to overcome this attack. In [1], the feasibility of the usage of Management Information Base (MIB) traffic variables for detecting DDoS attacks is being discussed. The variables are extracted from IP-based, TCP-based, ICMP-based and SNMP-based traffic. The traffic variable may contain some value which may tend be a normal one or the one under attack. The Network Management System analyses the MIB variables and concludes whether the system need to be prepared to encounter attacks or not.

Park [2] presents a distributed packet filtering mechanism (DPF) which validates the source address of inbound packets. Routers utilize DPF to find spoofed IP packets and avoid illegitimate traffic. Periphery routers of nodes find attacks and filter non-legitimate traffic so that it can protect the network is explained in terms of D-ward defense solution for DDoS at the source end [3].

The concept of queuing model is introduced in [4]. Queuing model is implemented by intermediate routers for DDos attack detection. The output queue whose rate is higher than the expected threshold is identified to be an attack. Jung [5] States IP addresses of attacking nodes can't be tracked easily in case of DDoS attack. This employs a method based on history of past IP addresses for attack detection. But this method is very difficult to be used since; it can't identify any attacks from a new user.

The probability of an attack with the Bayesian Network based Intrusion Detection Systems are incorporated in [6]. The probability of an attack is determined based on the input parameters. The system involves the usage of three agents which communicate with one another to detect an attack. Often the manual overriding of the attack probability may result in biased and incorrect attack detection.

There is a mechanism where Pushback incorporated with congestion control [7]. In which the nodes assume congestion as a symptom of Distributed Denial of Services attack and performs traffic rate limiting based on the local policy. In case a node fails to control the congestion, but is aware of it, the node sends a request to limit the traffic to the neighborhood nodes. This is practically difficult to be implemented since the request may be considered as a part of diverting the concentration of the node against detection, sometimes.

Size Based Scheduling (SBS) [8] is considered to be a weakness to DDoS attacks. Though this work emphasizes on Least Attained Service (LAS) in association with First in First out (FIFO) in routers, it is stressed that all SBS based methods that support short flows shouldn't diverge much from the performance of LAS. An adaptive cyber security examining system is proposed in [9] which incorporate a number of hybrid techniques such as fusion-based IDS, correlation computation of activity indicators and event classification through network resemblance dimensions.

The K-nearest-neighbor (KNN) Classifier [10] is used to categorise process into normal and intrusive. This method proves to be excellent in attack detection, but is highly expensive when used in real-time environment and when the number of processes is more. The Radial-Basis-Function Neural Network (RBF-NN) is used to identify Distributed Denial-of-Service in [11]. It uses nine packet parameters and the frequencies of these parameters are estimated. Based on the frequencies, the normal and abnormal classes are identified. Errors can be minimized by using K-means centers.

Above schemes show that increase in attack detection rate leads to increase in false alarm rate or to increase in computational overhead. Thus it requires the development of an attack detection system which overcomes the drawbacks in the current systems and for a proactive identification of DDoS attack in MANET.

36.3 Proposed Scheme

In this work, we propose a system which could detect and prevent Distributed Denial-of Service attacks in a Mobile Ad-hoc network. Initially, we use cluster analysis to detect the DDoS attack. Later the Collaborative Peer-to-Peer mechanism is used to encounter the DDoS attack. The detection of DDoS attack is based on detection of each phase of the attack separately. This could be done by extracting the traffic variables, which carry the information required to detect each attack. Thus the network can be made aware of the occurrence of DDoS attack.

DDoS attack involves the following steps

1. Selection of handlers and agents
2. Communication and Compromise
3. Attack

Usually, during the selection of handlers and agents, the attacker sends ICMP echo request packets to find handlers and agents which could help the attack. This is called IPSweep. During this activity, a lot of traffic is transmitted from an attacker to hosts available in the MANET. As a result, the occurrence rate of ICMP packets may be abnormally high compared to the usual network traffic. Thus the occurrence rates of these packets can be used as a measure to detect the initialization of a DDoS attack.

In addition, the distribution of the source IP, destination IP, source port and destination port provides information about the steps in DDoS Attack. During the IPsweep phase, the destination IP address in the network would be distributed randomly. The attack packets have diverse source IP address and are focused on destination IP address at the time of real attack. The degree of divergence can be measured by using the concept of entropy.

According to Shannon and Weaver [12], consider an information source having $n$independent symbols, each with probability of choiceP_i, the entropy$H$is defined as follows

$$H = -\sum_{i=1}^{n} P_i \log_2 P_i$$

When the entropy value is used, the value of occurrence of source IP address will be less compared to that of the destination IP address in the IPsweep phase. In the DDoS attack period, the entropy value of source IP address drastically increases whereas that of destination IP address decreases. The entropy of the source and destination port can also be used to detect DDoS attacks since some of the DDoS attacks use random port numbers during the attack period. In case of ICMP flood attack and UDP flood attack, even the entropy of the packet type are helpful in detecting the attack.

In the final phase, the nodes which act as agents for the DDoS attack generates large quantity of packets towards the victim and the network get jammed. The

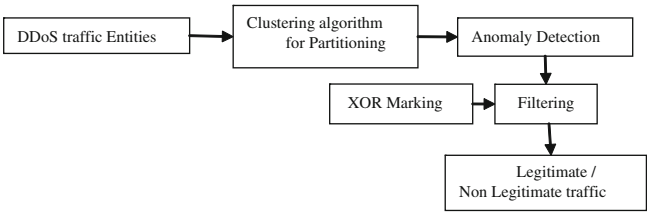


Fig. 36.1 Proposed system framework for DDoS attack detection

number of packets acts as a proof for the occurrence of the attack. Thus the parameters for the detection of DDoS attacks are:

1. Entropy of source IP address and port number
2. Entropy of destination IP address and port number
3. Entropy of packet type
4. Occurrence rate of packet type
5. Number of packets.

The overall proposed system framework is depicted in Figure 36.1.

36.3.1 Attack Detection Using Cluster Analysis

Cluster analysis is one of the most famous techniques which are used to classify a set of similar data into the same group and dissimilar in other groups. In our scheme, we use the hierarchical clustering method. The variables which are used are normalized using

$$z = \frac{x - \bar{x}}{s}$$

to eliminate the effect of difference between scales of the variables. In the above equation$x$is the value of each variable, $\bar{x}$is the mean of sample data,$s$is sample standard deviation. The Euclidean distance is used to compute the dissimilarities. The formula of Euclidean distance is as follows:

$$D(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

where$x$and$y$are two records to be clustered and$n$is the number of variables measured. Figure 36.2 shows the algorithm adopted for clustering technique.

Based on these calculations, the deviations are identified and the DDoS attack is filtered.

```
Input: 2000 DARPA Intrusion Detection Scenario Specific Data Set;
   Define the threshold value;
   For each process P_i in the Data Set:
      Compute the Euclidian distance D_i;
      If D_i < = threshold value then
         Add up P_i into i^th cluster;
      Else then
         Consider the process P_i as a new cluster;
      END
   END
```

Fig. 36.2 Pseudo code for clustering algorithm

36.3.2 XOR Marking

fter detection of the attack, the IP header is XOR marked to differentiate between the normal and non-legitimate traffic. The IP address of the nodes are encoded using two hash functions h_1 and h_2. If the node N_1 decides to mark the identification field of the IP packet, it marks $h_1(N_1)$ in the edge. If the packet is already marked, then the node XORs the$h_2(N_1)$with edge field value and writes the result of XOR in fragmented IP header identification field.

The node uses the last n-bit of IP Address and applies XOR function for the current bit with the previous bit. After this operation is performed, the IP fingerprint is embedded into the identification field. It is impossible for an attacker to fake an IP fingerprint with hop count less than its own [13]. After the encoding style is being changed, the previously marked data is sent along with the newly marked data. At destination, the marking id compared with the marking field dictionary. The marking of the IP packet is performed at the right n bit.

Proposed Algorithm

Based on the combination of cluster analysis and XOR marking, the algorithm is proposed. The steps are as follows.

step.1: Calculate Entropy of packet type
step.2: Calculate the Euclidean distance
step.3: If Euclidean Distance $D(x, y) < E_t$, calculate for next parameter else go to step 4
step.4: (a) Mark the packet with Node N using hash function h_1 as$h_1(N)$ (b) If packet is already marked, then $h_1(N) \oplus h_2(N)$and mark it in IP header identification field.
step.5: At destination, check marking with marking field dictionary If any dissimilarity found, isolate the node initiating the packet Else, allow the node for further transmission

Thus, in our proposed scheme, the Cluster Analysis performs the detection of DDoS attacks and the XOR marking performs the process of marking non-legit-

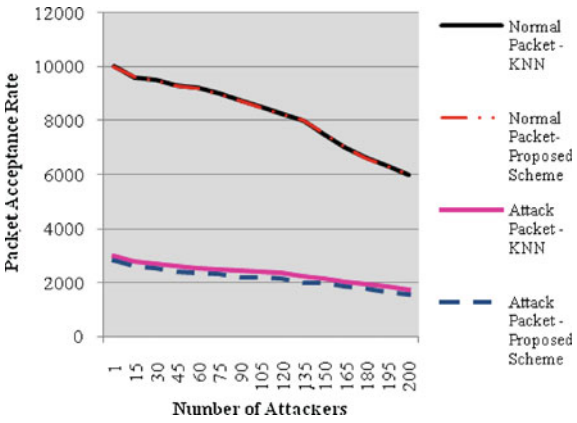


Fig. 36.3 Filtering performance

imate traffic. This greatly helps to avoid DDoS attacks in MANET. The nodes which are found to initiate the DDoS attacks are isolated and thereby avoid further vulnerability for the attack.

36.4 Simulation Results

The experimentation is done using 2000 DARPA Intrusion Detection Scenario Specific Data Set. The proposed scheme is tested for Filtering performance and Attack detection. The filtering performance shows the rate in which the normal traffic is separated from that of attack traffic. In our performance evaluation, we compare the filtering performance of our system with that of the K-nearest neighbor classifier. From the results obtained it could be observed that the packet acceptance rate of the normal packet is almost similar in both cases, whereas the packet acceptance rate is very much decreased in case of our proposed scheme than the KNN classifier scheme. This is clearly visualized in Figure 36.3.

The next experimentation is done over attack detection. The experimentation is done in two parts. The first one is done with respect to High data flood attacks. The other one is done with respect to changes in traffic rates. It has been found that our proposed scheme is more efficient in the attack detection also.

Figure 36.4 shows the attack detection capability of the proposed scheme. It could be found that the proposed scheme proves to be efficient by the reduction of the False Positive rate and the Detection ratio. The previous schemes used provide more false positive responses, whereas, the proposed scheme has greatly reduced the occurrence of false alerts.

The attack detection capability with respect to the traffic is shown in Fig. 36.5. It is evident from the graph that the detection of attackers by our proposed scheme is much efficient. In order to analyze the attack detection, initially the system was subjected to DDoS attack by 12 attackers. The system detected all the 12 attackers.

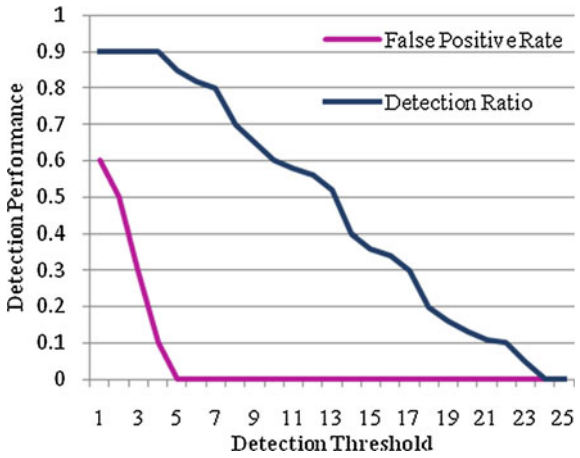


Fig. 36.4 High data flood attack

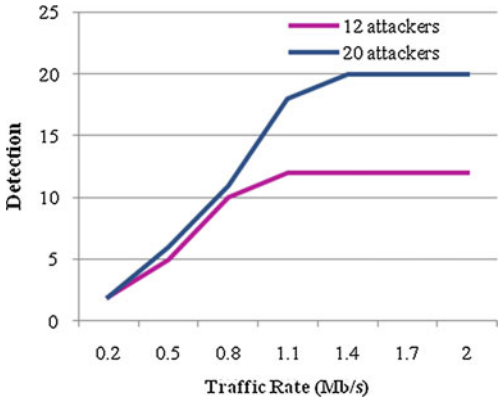


Fig. 36.5 Detection with respect to changes in traffic

The number of attackers was increased to 20. Even in such a scenario, the attackers were identified.

Thus, from the experimental results, it is proved that the proposed scheme is efficient in detecting and preventing DDoS attack from further establishment in the network.

36.5 Conclusion

In this paper, we have studied the impact of DDoS attacks over MANET. Also, we have designed a scheme to overcome the DDoS attack in a Mobile Ad-hoc Network. The newly designed system uses cluster analysis along with XOR marking to detect and prvent the effects of DDoS attacks in the network. The performance analysis was made for the packet acceptance rate and to find the attack detection. From the results obtained, it is evident that the proposed scheme is more efficient

in overcoming most types of DDoS attacks such as SYN flooding, ICMP flooding, Teardrop attack, nuking etc. in a Mobile Ad-hoc Network.

Acknowledgments This work is supported by All India Council for Technical Education under Career Award for Young Teachers Scheme, with File No.1-51/FD/CA/13/2008-09 Dated 29.01.2009.

References

1. Cabrera et al (2001) Proactive detection of distributed denial of service attacks using MIB traffic variables—A feasibility study. In: 7th IFIP/IEEE international symposium on integrated network management, Seattle, pp 1–14
2. Park L(2001) On the effectiveness of route-based packet filtering for distributed DoS attack prevention in power-law internets. SIGCOMM Comp Commun Rev 31:15–26
3. Mirkovic J, Reiher P (2005) D-ward: a source-end defense against flooding denial-of-service attacks. IEEE T Depend Secure Comput 2(3):216–232
4. Jeong WL et al (2006) An effective DDoS attack detection and packet filtering scheme. IEICE T Commun E89-B(7):2033–2042
5. Jung J, Krishnamurthy B (2002) Flash crowds and denial of service attacks: characterization and implications for CDNs and websites. In: ACM conference on Computer and Communications Security, pp 30–41
6. Gowadia V et al (2005) PAID: a probabilistic agent-based intrusion detection system. Comput Security 24 (7):529–545
7. Ioannidis J, Bellovin S (2002) Implementing pushback: router-based defense against DDoS attacks. In: Network and distributed system security symposium, NDSS 2002, San Diego, Feb 2002
8. Serwadda A, Phoha V, Rai A (2010) Size based scheduling: a recipe for DDoS. In: 17th ACM conference on computer and communication security, CCS 10, pp 729–731
9. Wu Q, Ferebee D, Lin Y, Dasgupta D (2009) Monitoring security events using integrated correlation based techniques. In: 5th Annual workshop on cyber security and information intelligence research: cyber security and information intelligence challenges and strategies, CSIIRW 09, pp 47:1–47:4
10. Liao Y, Vemuri VR (2001) Use of K-nearest neighbor classifier for intrusion detection. Comput Security 21(7):439–448
11. Gavrilis D, Dermatas E (2005): Real-time detection of distributed denial-of-service attacks using RBF networks and statistical features. Comput Netw 48(2):235–245
12. Shannon CE, Weaver W (1963) The mathematical theory of communication. University of Illinois Press, Champaign
13. Tariq U, Hong M, Lhee K (2005) PMS an expeditious marking scheme to combat with the DDoS attack. In: 9th International multi-topic conference, IEEE INMIC 2005, pp 1–4

Part II
Short Papers

Chapter 37
Optimal and Robust Framework for Enhancing Network Lifetime Using Power Efficient AODV in Mobile Ad-hoc Network

Bhagyashree Ambore, R. Suma and Jitendranath Mungara

Abstract The proposed system highlights one of the novel approach of enhancing the cumulative lifetime of mobile adhoc network on the backbone of most frequently deployed AODV routing strategies. MANET system is always associated with the design constraint from unwanted power drainage during communication. The proposed system considers the intermediate mobile nodes as vital factor which estimates the current mean power of the cumulative network as comparison threshold in order to evaluate the response of route request message along with impact of the proposed system towards the routing performances. Experimented in NS2, the proposed system shows optimal throughputs which can be definitely used for optimizing the energy on overloaded nodes in MANET and enhance the cumulative network lifetime extensively.

Keywords Mobile adhoc network · Energy · Routing protocol · AODV

37.1 Introduction

The area of mobile adhoc network (MANET) [1] is currently under the scanner of extensive research due to the massive advantages it permits on its application. Using this technology, various users can therefore be connected in networking

B. Ambore (✉) · R. Suma · J. Mungara
Computer Science and Engineering Department,
CMR Institute of Technology, Bangalore, India
e-mail: bhagya.ambore@gmail.com

R. Suma
e-mail: hod.cse@cmrit.ac.in

J. Mungara
e-mail: Jmungara@yahoo.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_37,

areas without any presence of pre-existing networking infrastructure. The mobile nodes can directly communicate with each other within their transmission ranges. In this environment, dual condition can surface where all the mobile nodes which have participated in the transmission involuntarily generate a wireless network, consequently, such types of wireless adhoc network can be visualized as mobile adhoc network. For the purpose of introducing diversified power issues in MANET [2], various power effective routing strategies has already been seen in the review of literature. The failure or degradation of energy in mobile nodes will not only influence the node itself but it will also have impact into its potential to forward the packets on behalf of others and therefore influence the cumulative network lifetime. Hence, majority of the researchers has attempted for designing power aware routing algorithms for specific mobile adhoc network scenario. Unfortunately, it is still in infancy stage as it is still not obvious that which one of the list of routing protocols is best for majority of scenarios as every routing protocols is designed to work for only specific environment. But, it is also highly feasible to unite and incorporate the current solutions in order to facilitate maximum power efficient routing techniques. As power efficiency is also vital issue in many other network layers, considerable efforts has already been given for designing power aware MAC as well as transport protocols. Each layer is believed to function in remoteness in layered network architecture but, as some current research suggested, the cross-layer design is indispensable to exploit the highest power performance. In fact, many routing protocols analyzed in survey also deploy the similar concept, i.e. they utilize lesser layer techniques such as transmission energy control and sleep mode methods in their routing layer protocols. Almost all mobile devices are supported by battery powers, so the power-efficient issue is one of the most important design issues in MANET. Solutions to the energy-efficient issue in MANET can generally be categorized as follows: (1) Low-Power Mode, in which mobile devices can support low-power sleeping mode. The main research challenges in low-power mode are that at what time mobile node can turn to sleeping mode, and at what time it should wake up. Corresponding issues are addressed in [3] and etc.; (2) Transmission Power Control: In wireless communication, transmission power has strong impact on transmission range, bit error rate and inter-radio interference, which are typically contradicting factors. By adjusting its transmission power, mobile node can select its immediate neighbors from others, thus the network topology can be controlled in this way. How to determine transmission power of each node so as to determine the best network topology has been addressed in [4] and etc.; (3) Power-Aware Routing: Other than the common shortest-hop routing protocols, such as DSDV [5], AODV [6], DSR [5], and etc., power-aware routing protocols take various power metrics or cost functions into account in route selection.

The proposed system is designed on the backbone of frequently used AODV routing protocol for enhancing the cumulative lifetime of mobile adhoc network. In Sect. 37.2, we give an overview of related work which identifies all the major research work being done in this area. Proposed system is illustrated in Sect. 37.3.

Section 37.4 discusses about result analysis and finally in Sect. 37.5, we make some concluding remarks.

37.2 Related Work

Arulanandam [7] has shown that energy is the scarcest resource for the operation of the mobile ad hoc networks. Sunsook et al. [8] considered energy constrained routing protocols and workload balancing techniques for improving MANET routing protocols and energy efficiency. Rekha [9] proposes a cost based power aware cross layer design to AODV. Rutvij [10] has discussed some basic routing protocols in MANET like DSDV, DSR, TORA and AODV. Xiangpeng [11] has presented a comprehensive energy optimized routing algorithm and its implementation to AODV. Abdusy Syarif [12] has presented some improvement suggestion to AODV routing protocol. Preeti et al. [13] has tried to remove the existence of misbehaving nodes that may paralyze or slows down the routing operation in MANET. Patil et al. [14] concentrated on emergency search and rescue operations which rely heavily on the availability of the network. Sanjay [15] has presented some improvement suggestion to AODV routing protocol.

37.3 Proposed System

The main aim of the project work is to introduce a novel scheme based on AODV, called Energy Saving Ad-hoc On-demand Distance Vector for routing in MANETs. In this proposed system, a new architecture based on enhancement in AODV is proposed for conducting energy efficient routing in MANETs. The proposed scheme (Fig. 37.1) achieves the energy information exchange among neighboring nodes through already-existed signaling packets in AODV and introduces a new network parameter as the comparison threshold, called current average energy of the network, which can estimated the mean power utilization of the network. In the proposed scheme, each intermediate node determines whether to forward RREQ packet by comparing its remaining energy with current mean power of network. If the energy of the node is larger than the threshold, it will forward the RREQ packet immediately. Otherwise, the node will wait for a while to decide whether the packet should be forwarded or dropped according to the number of the identical RREQ packets received during the waiting period. After that, effects of the proposed routing protocol on network performance are addressed. Both analytical and simulation results shows that the proposed routing scheme is comparatively easier for execution and can facilitate a maximized cumulative network lifetime.

The routing protocols have multiple operations to be performed apart from instituting correct and resourceful routes among the twosome of mobile nodes. The

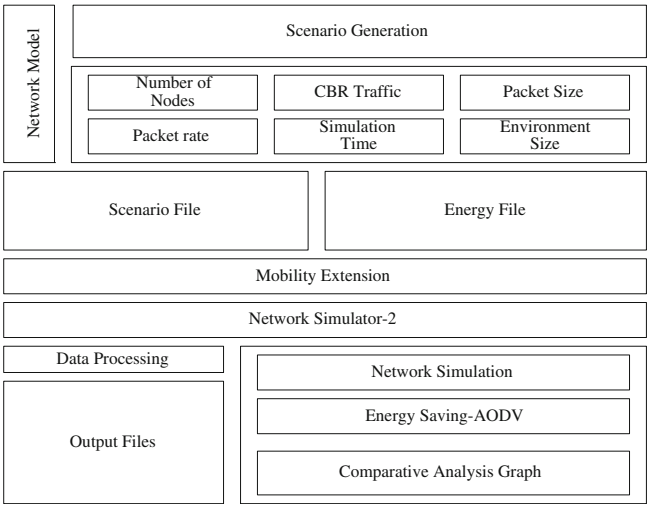


Fig. 37.1 Proposed architecture

Fig. 37.2 Simulation result-I

most prominent aim of the routing protocol is to render the entire networking to operate for as long duration as possible. Such types of parameters are very essential to facilitate the minimum energy path through which the cumulative utilization of power for delivering a packet is reduced. In such experiments, the wireless link is interpreted with the cost of link in terms of transmission power over the link and minimum energy path is another factor which reduces the sum of the cost of link along the same path. But, unfortunately, if such types of routing parameters are selected than it may yield to unbalanced power utilization among the mobile nodes. It was also seen that when certain specific mobile nodes are incorrectly overloaded in order to support majority of packet-relaying operation, such nodes may utilize higher battery power and impede running earlier than other

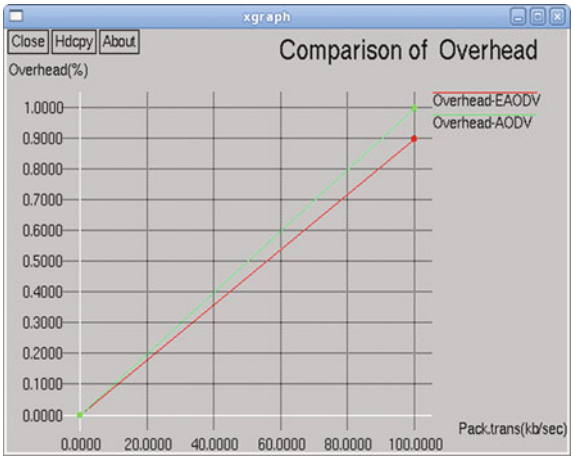


Fig. 37.3 Simulation result-II

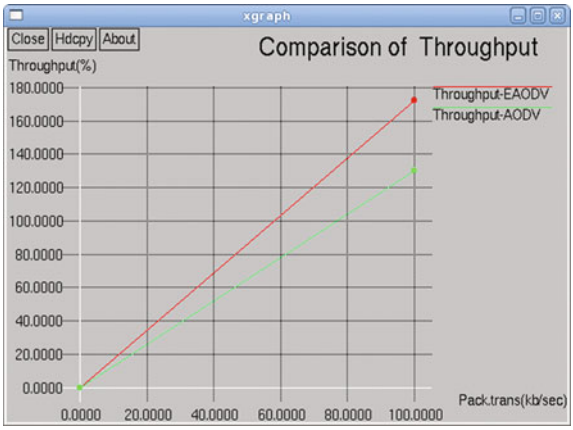


Fig. 37.4 Simulation result-VII

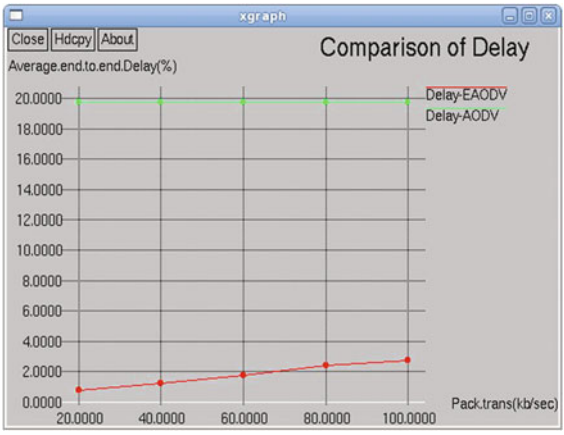


Fig. 37.5 Simulation result-VIII

mobile nodes thereby disturbing the cumulative functionality of the mobile adhoc network. An estimation algorithm to obtain the network average remaining energy is introduced. With such assessment, intermediary mobile nodes can reasonably decide if their residual power is adequate or not. By preventing overused nodes from participating in route discovery processes, the proposed scheme using AODV effectively balances energy consumption around the network. Simulation results will show that the proposed energy saving schema can evidently increase the lifetime of the network.

37.4 Result Analysis

The proposed system is designed on Linux platform using network simulator (NS2). The performance analysis is checked with respect to delivery ratio, network overhead, delay, and cumulative network life time.

The above graph in Fig. 37.2 shows the comparison of packet delivery ratio for proposed energy aware-AODV and AODV with *X*-axis of packet transmission (kb/s) and *Y*-axis of Delivery rate (10^3). The Fig. 37.3 shows the comparison of overhead for proposed energy aware-AODV and AODV Fig. 37.4.

The above graph in Fig. 8 shows the comparison of overall network lifetime for proposed scheme, energy aware AODV, and AODV with *X*-axis of packet transmission (kb/s) and *Y*-axis of network lifetime (s). The Fig. 37.5 shows the comparison of delay for energy aware AODV, and AODV with *X*-axis of packet transmission (kb/s) and *Y*-axis of average end-to-end delay (%)

37.5 Conclusion

In this proposed system, a unique power-aware routing protocol using AODV is presented. In the route discovery process of the proposed scheme, transitional and intermediate nodes estimate the current mean power of the cumulative network as an evaluation threshold to establish how to retort to the received route request packets. An evaluation algorithm to accomplish the network mean residual power is highlighted. With such estimation, intermediate nodes can reasonably judge whether their residual power is adequate or not. By averting overused nodes from participating in route discovery processes, the proposed routing scheme using AODV efficiently stabilize the power expenditure around the cumulative network. For the analysis of the proposed system, with the assistance of graphical representation, the network lifetime of the proposed scheme and AODV with different levels of mobility and network loads is shown in result analysis. It is also shown that graph of data delivery rates with different levels of mobility and the network performance in term of mean end to end delay.

References

1. http://en.wikipedia.org/wiki/Mobile_ad_hoc_network. Accessed 28 Mar 2012
2. Tseng Y-C, Hsu C-S, Hsieh T-Y (2002) Power-saving protocols for IEEE 802.11-Based multi-hop ad hoc networks, INFOCOM
3. Singh S, Raghavendra CS (1998) "Power efficient MAC protocol for multihop radio networks", Proceedings of IEEE international personal, indoor and mobile radio communications conference, pp. 153–157
4. Hu L (1993) "Topology control for multihop packet radio networks", IEEE transactions on communications, vol. 41. pp. 1474–1481
5. Perkins CE, Bhagwat P (1994) "Highly dynamic destination—sequenced distance-vector routing (DSDV) for mobile computers", Proceedings of ACM SIGCOMM, London, pp 234–244
6. Perkins CE, Royer EM (1999) "Ad-hoc on-demand distance vector routing", Proceedings of 2nd IEEE workshop on mobile computing systems and applications, New Orleans, pp 90–100
7. Arulanandam K, Parthasarathy B Dr (2009) "A new energy level efficiency issues in Manet" international journal of reviews in computing 2009 IJRIC. All rights reserved
8. Jung S, Hundewale N, Zelikovsky A (2005) "Energy efficiency of load balancing in MANET routing protocols" SAWN-2005
9. Rekha P, Damodaram A Dr (2008) "Cost based power aware cross layer routing protocol for Manet" IJCSNS international journal of computer science and network security, vol. 8 no. 12. Dec 2008
10. Jhaveri RH, Patel AD, Patel AD (2010) "MANET routing protocols and wormhole attack against AODV" IJCSNS international journal of computer science and network security, vol. 10 no. 4. Apr 2010
11. Jing X, Lee MJ (2004) "Energy-aware algorithms for AODV in ad hoc networks", ICMU-2004
12. Bhati P (2011) "An efficient agent-based AODV routing protocol in MANET", vol. 3 no. 7 July 2011
13. Patil AP (2011) "Design of an energy efficient routing protocol for MANETs based on AODV", IJCSI international journal of computer science issues, vol. 8, issue 4, no 1, July 2011 ISSN (Online), pp 1694–0814
14. Patil AP (2011) "Design of an energy efficient routing protocol for MANETs based on AODV", IJCSI international journal of computer science issues, vol. 8, issue 4, no 1, July 2011 ISSN (Online), pp 1694–0814
15. Sharma S Dr, Patheja PS (2012)"Improving AODV routing protocol with priority and power efficiency in mobile ad hoc WiMAX network", international journal of computer technology and electronics engineering (IJCTEE), vol. 2, Issue 1

Chapter 38
Voice Transformation Using Radial Basis Function

J. H. Nirmal, Suparva Patnaik and Mukesh A. Zaveri

Abstract This paper presents novel technique of voice transformation (VT), which transform the individual acoustic characteristics of the source speaker so that it is perceived as if spoken like target speaker. Using features namely line spectral pairs (LSP) and pitch as spectral and glottal parameters of the source speaker are transformed into target speaker parameters using radial basis function (RBF). The results are evaluated using subjective and objective measures based on voice quality method. The listening tests prove that the proposed algorithm converts speaker individuality while maintaining high speech quality.

Keywords Voice transformation · Line spectral pairs · Long term prediction · Radial basis function

38.1 Introduction

The basic objective of the voice transformation is to mimics the individual acoustic characteristics of the source speaker so it posses the characteristics of the target speaker [1]. It has many applications like personification of text to speech,

J. H. Nirmal (✉)
Department of Electronics Engineering,
K.J.Somaiya College of Engineering, Mumbai, India
e-mail: jhnirmal1975@gmail.com

S. Patnaik
Department of Electronics Engineering, S V National Institute of Technology, Surat, India
e-mail: ssp@eced.svnit.ac.in

M. A. Zaveri
Department of Computer Engineering, S V National Institute of Technology, Surat, India
e-mail: mazaveri@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_38,

audio dubbing, audio based learning tool, audition test [2, 3], it is an alternative method for building the synthetic voices. The VT is carried out in two modes. In the first training mode, it analyses and extracts the spectral and glottal features of the source and target speaker and appropriate mapping function is developed. In the transformation mode, the mapping function developed in the training mode transform the acoustic characteristics of the source speaker into target speaker so it posses the characteristics of the target speaker [4]. Initially, Abe et al. [5] have used vector quantization but the problem of the discontinuities are produced due to the hard partition of the acoustic cues which causes degradation of the quality of speech. Fuzzy based vector quantization and mapping techniques have been proposed in [6].Different types of codebook mapping methods such as STASC also studied in [7, 8]. Valbret et al. [9] used dynamic frequency warping (DFW).The further improvements related to frequency-warping have been presented in [11].

Stylianou et al. [12] used Gaussian mixture models (GMMs) as a soft partitions. Kain et al. modified the GMM approach in [13] but it results into over smoothing. The strong vocoding technique, namely, STRAIGHT [14] and phase reconstruction based method [15]. Researchers have also provided the solution for over smoothing problem like hybridization of GMM with DFW [16] and GMM with codebook mapping [17]. Desai et al. [18] compared the performance of the ANN and GMM and it is reported that the ANN performs better than GMM. ANN used for VC to exploit the nonlinear relationship between the vocal tract shape of source and target speaker [19, 20]. The approach proposed in this paper differs from various methods reported in the literature as below:

a. LSP and fundamental frequency (F0) are used to extract the source and target features of parallel set of data. Using these features the RBF based neural network is trained for an appropriate mapping function that transforms the vocal spectral and glottal excitation cues of the source speaker into target speaker's acoustic space.
b. This proposed VT system has been evaluated using both the subjective evaluation and the objective measures like mean square error, pitch and formants.

The outline of the paper is as follows, Sect. 38.2 explains the LSPs based methodologies of VT. Section 38.3 and 38.4 describe the results and evaluation using subjective and objective measures respectively. Finally, conclusions are discussed in section.

38.2 Proposed Algorithm for Voice Transformation

The proposed algorithm consists of two phases. In first phase, it extracted the LSP and pitch based features from source and target speaker data. It is followed by the second phase where it used RBF neural network which is trained to learn the nonlinear mapping function for source to target speech transformation using the features extracted in the first phase.

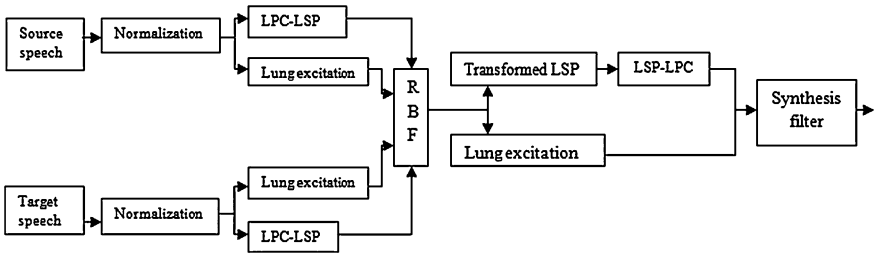


Fig. 38.1 LSP, pitch based VT using neural network

LSP are an alternative to LPC as the LPC has many problems such as stability check, quantization and interpolation [21]. LSP are popular due to its excellent quantization characteristics and consequent efficiency in representation.When the LSP are in ascending order in the range [0, 1], the resulting filter is guaranteed to be stable. The proposed algorithm translates the vocal frequencies (LSP and pitch) of source speaker into target speaker using neural network in two modes. In the training mode the source and target speech is normalized to some predetermined amplitude range and the pitch information is extracted to produce a residual, residual contains vocal tract information which is modeled by LPC. Applying the LPC analysis filter to the residual will result in the vocal tract information being removed leaving lung excitation signal. LPC produces the results in an unstable synthesis filter. So we convert LPC parameters into LSP. We have mapped the lung and LSP parameters of source speaker into target speaker using RBF neural network with spread factor of 0.009 [22] and error threshold of 0.02. In the transformation phase the LSP and lung parameters of test speech samples are transformed as a target speech. The analysis process is essentially reverse of the synthesis process and results in reconstructed speech as shown in Fig. 38.1.

38.3 Simulation Result

The proposed algorithm is evaluated using standard databases as well as our own databases consisting of Gujarati and Marathi regional Indian languages. Our own database consists of 48 sentences recorded with 16 kHz sampling frequency using a high quality microphone (Sony V_120).Our database is parallel database; These sentences are collected from five males and five female's speakers.The standard databases, namely, CMU ARCTIC database consisting of utterances recorded by 7 speakers.The sample results of our LSP based speech conversion algorithm for inter and intra gender speaker are performed.

Table 38.1 MSE for LSP based VT systems

Conversion	LSP based
Male to Female	4.8642e–004
Female to Female	4.3574e–004
Female to Male	4.4977e–004
Male to Male	5.1750e–004

38.4 Objective and Subjective Evaluations

In this section we have evaluated our algorithm based on objective and subjective parameters. We used two objective based evaluation parameters (i) mean square error (MSE) and (ii) similarity measure using pitch and formant. This similarity measure allows us to differentiate between male to female and female to male VT.

38.4.1 MSE Based Objective Evaluation

In this section we provide the objective evaluation for LSP based VT systems to measure the differences between the target and transformed speech signals. Since many perceived sound differences can be interpreted in terms of differences of spectral features and mean squared error (MSE) are considered to be a reasonable metrics; The MSE between target audio vector p and transformed audio vector s is calculated as per below Eq. 38.1 on a sample by sample basis. The average square difference between two vectors is used to evaluate the objective performance of mapping algorithms shown in Table 38.1.

$$E = \frac{1}{N}\sum_{i=0}^{N-1}[s(i) - p(i)]^2 \tag{38.1}$$

38.4.2 Pitch and Formant Based Similarity Measures

This similarity measure is used to find out how close the pitch and formants of speech produced by the application with the source and target speaker. In order to evaluate the performance of the LSP based transformation we have done comparison based on the fundamental frequency (f0) and spectral formant frequencies, the fundamental frequency (f0) of the vocal fold vibration determines the perceived pitch of a voice, most often it is higher in females than in males. The spectral formants determine the characteristic timbre of each speaker's voice, so that the listener can recognize familiar voices and discriminate the gender from familiar or unfamiliar voices [23]. We have calculated the pitch and the formant

Table 38. 2 Pitch and formants of source, target, male to female transformed speech, 1-Male to Female

	Source signal Male	Target signal Female	Transformed signal LSP
Average pitch	216.71	317.56	320.758
First Formant	545.6	659.7	961.888
Second Formant	1826.1	1924.8	2058.4
Third Formant	2797.62	2926	3093.90

frequencies of source, target and transformed speech and shown in following Tables 38.2, 38.3.

As per the Tables 38.2 and 38.3 the average pitch of the target speaker (male) is less than the source speaker (female).Transformed speech signal's average pitch is less than pitch of source signal and closed to the average pitch of the target speech and vice versa, The fundamental frequencies (pitch period) of the women are higher than the men, as can be observed from Tables 38.2 and 38.3 the target speech is similar like the transformed speech and it is clearly transformed from men to women and vice versa.

38.4.3 Subjective Evaluation

To evaluate the overall accuracy of the conversion, a Mean Opinion Score (MOS) test is carried out for evaluating the similarity between the converted voice and the target voice to find the performance of LSP based transformation. 5 listeners give scores between 1 and 5 for measuring the similarity between the output of the two VT systems and the target speaker's natural utterances. The results of this average similarity test are provided in Fig. 38.2. It shows that the proposed algorithm performs better for intra-gender voice transformation than inter-gender voice transformation.

38.5 Conclusion

This paper proposed a novel technique using LSP as spectral features and pitch as glottal features. These features are transformed from source to target using RBF network. Subjective and objective evaluations are performed on speech quality. The fundamental frequency range of female speech is higher than male speech is shown. The inter gender and intra gender speech conversion are also performed. Objective and subjective analysis shown that LSP Pairs method produce very convincing output with high quality.

Table 38.3 Pitch and formants of source, target, female to male transformed speech, 1-Female to Male

	Source signal Female	Target signal Male	Transformed signal LSP
Average pitch	372.5	192.71	189.6
First formant	698.16	512.450	1021.8
Second formant	2037.5	1746.36	2002.6
Third formant	3008.0	2787.4	3217.7

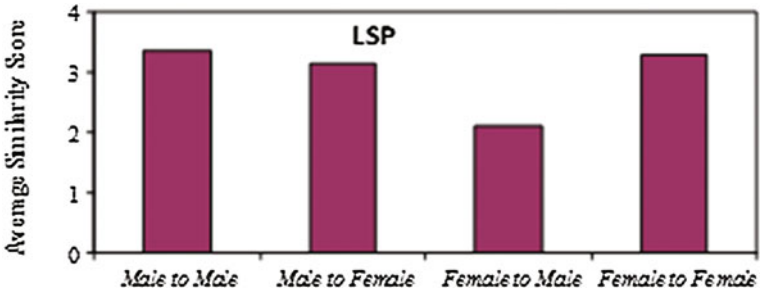


Fig. 38.2 Average similarity scores between transformed and target speakers voice

References

1. Kain A (2001) 'High resolution voice transformation', PhD dissertation, Oregon Health and Science University
2. Daniel, erro. eslava (2008) Intra-lingual and cross-lingual voice conversion using Harmonic plus stochastic models. PhD dissertation universitat politècnica decatalunya
3. Turk O (2007) Cross-lingual voice conversion. PhD dissertation Bogazii University
4. Sreenivasa Rao K (2010) Voice conversion by mapping the speaker-specific features using pitch synchronous approach. Computer speech and language, Elsevier, vol. 24, pp 474–494
5. Abe M, Nakamura S, Shikano K, Kuwabara H (1988) Voice conversion through vector quantization. International conference on acoustics, speech, and signal processing, ICASSP. p 655
6. Abe MA (1999) Segment-based approach to voice conversion. International conference acoustics, speech, and signal processing, ICASSP. p 765
7. Arslan LM, Talkin D (1999) Voice conversion by codebook mapping of line spectra frequencies and excitation spectrum. International proceedings Eurospeech. Rhodes, vol. 3, pp 1347–1350
8. Shikano K, Nakamura S, Abe M (1999) Speaker adaptation and voice conversion by codebook mapping. IEEE international symposium on circuits and systems, vol. 1, pp 594–597
9. Arslan LM (1999) Speaker transformation algorithm using segmental codebooks. STASC Speech Commun 28(3):211–226, 469–471
10. Valbret H, Moulines E, Tubach JP (1992) Voice transformation using PSOLA technique. Acoustics, speech, and signal processing, ICASSP pp I145–I148
11. Shuang ZW, Bakis R, Shechtman S, Chazan D, Qin Y (2006) Frequency warping based on mapping formant parameters. In: Proceedings of international conference spoken language process

12. Stylianou Y, Cappa O (1998) A system for voice conversion based on probabilistic classification and harmonic plus noise model. International conference acoustics, speech and signal processing, Proceedings pp 281–285
13. Kain A, Macon MW (1998) Spectral voice conversion for text-to-speech synthesis. Proceedings ICASSP, Seattle, pp 285–288
14. Toda T, Saruwatari H, Shikano K (2001) Voice conversion algorithm based on Gaussian mixture model with dynamic frequency warping of STRAIGHT spectrum. International conference on acoustics, speech, and signal processing, Proceedings. ICASSP. pp 841–844
15. Ye H, Young S (2006) Quality-enhanced voice morphing using maximum likelihood transformations. IEEE transactions audio, speech, language process, vol. 14, no. 4, pp 1301–1312
16. Ohtani Y, Toda T, Saruwatari H, Shikano K (2006) 'Maximum likelihood voice conversion based on GMM with straight mixed excitation'. In: Proceedings Interspeech
17. Desai S, Raghavendra EV, Yegnanarayana B, Black AW, Prahallad K (2009) Voice conversion using artificial neural networks. In: Proceedings of IEEE international conference acoust, speech, and signal processing, pp 3893–3897
18. Chen W-Q, Zhang JL, Xiuguo B (2010) An improved method for voice conversion based on Gaussian mixture model. International conference on computer application and system modelling, PP V4-404-408
19. Narendranath H, Murthy A, Rajendran S, Yegnanarayana B (1995)'Transformation of formants for voice conversion using artificial neural networks', Speech communication, vol. 16, pp 207–216
20. Chen Z, Zhang LH (2010) A ANN base high quality method for voice conversion'. International conference on wireless communications networking and mobile computing (WiCOM)
21. Grassi S (1997) Dufaux, Ansorge; Pellandini, 'Efficient algorithm to compute LSP parameters From 10th-order lpc coefficients'. International conference on acoustics, speech, and signal processing, vol. 3, pp 1707–1710
22. Lan Vince McLoughlin (2008) Line spectral pairs. Elesevier signal processing, pp 448–467
23. Lan Mcloughlin (2009) Applied speech and audio processing with matlab examples (1st edn). Cambridge Publication, Cambridge
24. Vergin R, Azarshid F, Shahguansy D (2006) Robust gender dependent acoustic phonetic modeling in continous speech recognition based on new automatic Male Female classification. International conference spoken language processing, pp 1–4
25. Pawan K, Jakhanwal N, Bhowmick A, Chandra M (2011) Gender classification using pitch and formant. International conference on communication, computing & security pp 319–324

Chapter 39
IPTC Based Ontological Representation of Educational News RSS Feeds

Shikha Agarwal, Archana Singhal and Punam Bedi

Abstract In order to make online news retrieval more appropriate in respect to end user, it is necessary to make machines aware of a greater part of the underlying semantics. Ontology will help to realize this future of web. In this paper, we have worked upon how to represent Educational News RSS Feeds by using an ontological structure. We are also incorporating International Press Telecommunication Council (IPTC) standards in our design since IPTC has proposed various standards for news industry to make the system more interoperable. Our main objective is to make a structure which can satisfy the specific demands of various categories of end users of the Educational domain. Designed ontology is then tested to meet various criteria mentioned in the paper.

Keywords Semantic web · Ontology · News RSS feeds · IPTC · OWL · Protégé

39.1 Introduction

In today's busy world online news industry needs to represent news in a manner, using which end user can access specific news of their interest and need. This can be achieved using semantic web technologies [1]. In this paper, we have presented how to design ontology for Educational News domain. Motivation is that Education can

S. Agarwal (✉) · A. Singhal · P. Bedi
University of Delhi, Delhi, India
e-mail: shikha_8june@rediffmail.com

A. Singhal
e-mail: singhal_archana@yahoo.in

P. Bedi
e-mail: punambedi@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_39,

enhance competitiveness of a country in Global economy. We have taken a corpus of nearly 500 headline news in RSS format from reliable source, for initial analysis. We are emphasizing on news representation by classifying keywords of the domain into concepts and sub concepts. For the first level of categorization in ontology, we are incorporating standards given by IPTC [2]. We have carefully further refined categories into subcategories and identified their properties. We have implemented our design with OWL ontology language using Protégé tool. Here, we are giving criteria's to test the designed ontology as testing the designed structure or knowledge base makes it logically sound. Testing and Evaluation of the proposed Ontology shows that most of the educational news can be properly categorized and implicit knowledge can be mapped to the properties of the concepts. In this paper our focus is on giving proper semantic structure to educational news.

39.2 Related Works

To represent concepts, properties and relations in news domain, various Ontologies have been developed. Papyrus [3], Mesh [4], Neptuno [5] all projects are approaches to develop news domain ontology. In these projects main focus is on the creation and maintenance of digital achieve of newspaper. In our work, we are focusing on educational domain, giving proper structure to the news headlines, using Ontology.

It has been noticed that many different approaches are used to design Ontology [6]. On-To-Knowledge methodology [7] builds an ontology-based tool environment to improve knowledge management. They used OIL, we recommend OWL. Ontoedit methodology [8] focuses on three main steps for ontology construction. These are requirements specification, refinement, and evaluation. We have further refined the steps. METHONTOLOGY [9], framework enables the construction of Ontologies at the conceptual level, as opposed to the implementation level. They describe process to build ontology for centralized ontology based systems. They do not provide guidance for decentralized ontology development and do not focus on post development processes. However, most of the approaches found till date are not able to present proper structure of the educational news. In [10], it is found that OWL ontology has been developed for basic IPTC news architecture, and linked it with other multimedia metadata standards. We have further classified the basic classification given by IPTC to enrich news metadata. Model has been implemented using OWL_DL [11] in Protégé [12].

39.3 Background

Knowledge Representation (KR) and Information retrieval (IR) using taxonomic structure of Ontology [1] gives more accurate results. To develop Ontology of Educational News Domain we propose to use OWL [11] as KR language as it is

W3C recommendation since 2004. To make our Ontology abreast with news industry standards [2] we further propose to use IPTC NewsCodes [2] standard for basic first level categorization of concepts.

News agencies supply lots and lots of news reports to news organizations that deliver news to the general public via print media, broadcast news and more recently the Internet. Internet news can be represented using formats like RSS [13] and NewsML [2]. Our aim is to improve knowledge management and IR of online news.

39.4 Ontology of Educational News

It is observed that in educational news represented by various news sources, it is not easy to find relevant news according to end users specific needs. For instance, the search service of the online newspaper Hindustan Times [14], Times of India [15], IndianExpress.com [16] (with most readers in India), gives search results just on the basis of keywords, having no semantic relatedness. They lack proper semantic description of structure of the Educational news. A properly designed ontology will give better classification and IR results. Thus in this paper we have emphasized on the proper design of ontology, having semantic relatedness in its entities. Phases of Ontology design, we adopted are: Planning, Design, Refinement and Evaluation.

Planning. Determine domain, Scope and intended use: In our work domain is Educational news headlines and intended use of the ontology is better classification of news items of RSS feeds for more precise news retrieval. Scope of the model is to make it clear that in what type of situations ontology is going to give required results. End user may enquire about an event held at a university or about entrance exams to be held at any one or two geographic locations, in specific period of time.

Source of data: RSS feed educational news items from portal 'Times of India'.

Analysis: We have analyzed news headlines to identify terms, properties, and relations. Using our domain knowledge we have made a glossary (enumerated terminology) of the domain, which is unstructured at this step.

Design. In this phase we first design an abstract conceptual model based on outcome of phase I. Then we implement the model to give a formal representation.

Conceptual model: Basically this step is about identifying Concepts. In our model, initial concepts i.e., main categorization is basically based upon standard IPTC news codes subject codes [2]. We have further refined concepts into sub concepts, based on thorough study of the domain and analysis of corpus taken. Few are given in Table 39.1.

We have created three levels in ontology for proper news classification. We have identified relations among concepts and sub concepts, which will make a super class-subclass hierarchy (refer Table 39.1). Properties or Attributes can be of three types [11]: Data Type, Object and Annotation. These attributes fall under two

Table 39.1 Identified sub-concepts of some basic IPTC concepts

S. No.	Concept	Sub concepts	Type of relation
1.	School	Primary, secondary, higher secondary	'Whole-part'
2.	Organization	Private, government	'Is-a'

categories: Generic and Specific. In Table 39.2 we have explained these points, by taking some examples.

Formal model: This step is about formalizing above designed model. Figure 39.1 shows the snapshot of our designed Ontology created in Protégé. Our ontology contains total 31 Classes, 30 Object properties and 80 Data Type properties.

Once ontology structure is build next step is to populate it with assertions [11] in news items. To check the design and integrity of the structure created in Protégé, reasoners are provided as pug-ins. Reasoner used to check design, shows no inconsistency in our Ontology.

Refinement. It is advisable to go through all the steps again to further improve knowledge representation to unearth any improvement if required.

We have experienced that manual classification is expensive and hard to scale. Moreover, hand crafted rules require expertise which is hard to find. Thus automated classification will help. We need to create a system to place novel knowledge into ontology with high Precision and Recall [17]. For automatic ontology population Natural Language Processing will be required. We will pre-process feeds. Pre-processed feeds or assertions will be represented as vectors using normalized TF_IDF [18]. Traditional TF_IDF approaches consider the full text of news items but in our approach we will consider only the concepts that appear in knowledge base.

Evaluation. We have evaluated the Ontology based on Generic and Specific criteria. We have designed Ontology following Tom Gruber's five principles [19] (Generic criteria) to ensure knowledge sharing: clarity, coherence, extendibility, minimal encoding bias and minimal ontological commitment. Specific criteria are based on predefined domain and scope of work. Evaluation on these criteria shows that final structure of our designed ontology meets the mentioned domain and scope.

39.5 Experimental Study

Along with above mentioned testings, we also check that any other news taken for testing purpose can be properly categorized to one of the categories decided in the structure. Then it is checked that the inherent information in news can be mapped to the properties of that category.

Table 39.3 shows proper mapping of information in feeds to the concepts of Ontology. First column of table shows mapping of inherent news information to

Table 39.2 Identified attributes in news feed items

S.No.	News headline 'title'	Subject (Domain)	Predicate (Property)	Object (Range)	General Attribute (few)	Specific Attribute (few)
1.	"Tamil Nadu: MBBS second phase counselling from July 28"	Course MBBS	Counsel-ling	Date	Category_ belongs_to:univ, Subcategory: course Date: 19-july-2011 Source: TOI (Above three are Data type properties with values)	Counselli-ng_of (Object Property) Course: MBBS Univ Name, Address, Website, VC
2.	"Law says No to raising MPs quota in Kendriya Vidyalya' admission"	School admission	Check_on_quota	Quota	Category_ belongs_to: School Date: 20-july-2011, Source:TOI	Check_on_quota (Object Property), School name, geographic location.

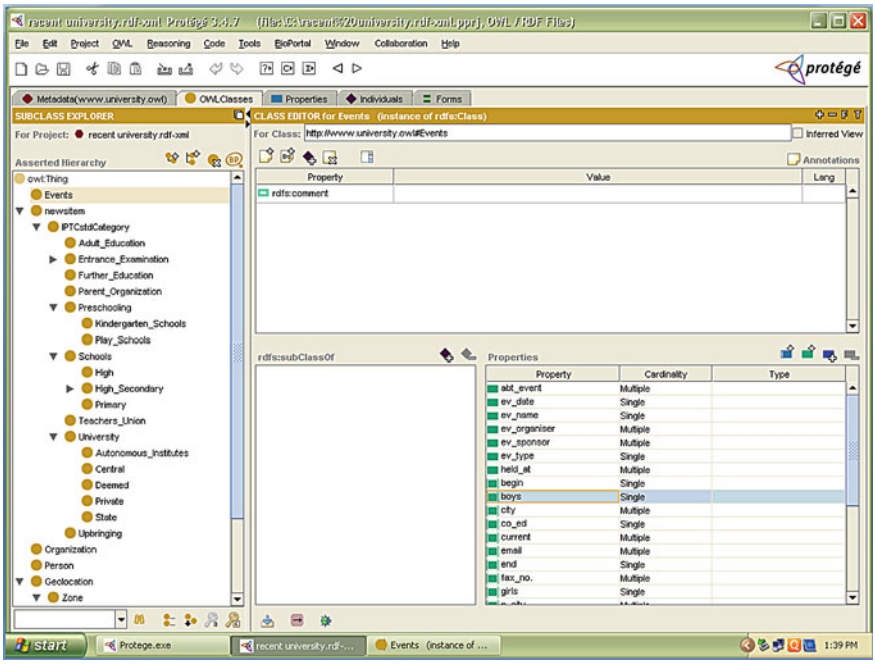


Fig. 39.1 Snapshot of ontology created in protégé

Table 39.3 Mapping values in feeds into concepts and properties of ontology

News feed 1: "Narayan murthy graces first AU annual convocation"	
Generic attribute: value	Specific attribute: Value
Source: "TOI"	Sub Category: "State University"
Date: 14-July-2011	Univ_name: "AU"
Category: "University"	Geographic Location="Allahabad"
	Website: "www/ddd/….."
	Event: "Convocation"
News feed 2: "HSC and SSC supplementary exams in Maharashtra from Sept 29"	
Generic attribute: value	Specific attribute: Value
Source: "TOI"	Sub category: "Higher secondary school"
Date: 09-July-2011	Board: "……"
Category: "School"	Geographic location="Maharashtra"
	Website: "www/xxx/….."
	Event: "Examination"

generic attributes and second column shows the mapping to specific attributes. Category of RSS feed item is identified from metadata or tag title which are present in our Ontology. Similarly we have tested our structure by picking some more news headlines and found successful mappings.

39.6 Conclusion and Future Work

Appropriate semantic representation of entities of a domain is of prime prerequisite for IR. Therefore, in this paper a modest attempt has been made to design an ontological structure using IPTC standards for Educational News Headlines.

Proper classification of news items will give more precise and accurate results to the end user. Thus, we are working on improving the classification of news items using statistical methods by assigning weights to identified concepts in news items. Efforts will be made to design a system of representation and retrieval which can assist the online newspaper industry to meet specific requirements of end users of the domain.

References

1. Fensel D, Musen MA (2001) The semantic web: a brain for humankind, Vrije University of Amsterdam, Stanford University. In: 1094-7167/01/$10.00 IEEE
2. IPTC Standards (2009) Document revision 1, International press telecommunications council, Specification version 2.4
3. Payrus News, www.ict-papyrus.eu.
4. Mesh Ontology, mklab.iti.gr/project/Mesh
5. Neptuno Project, ir.ii.uam.es/neptuno
6. Subhashini R, Akilandeswari J Dr (2011) A survey on ontology construction methodologies. In: IJECBS (Online http://www.ijecbs.com)
7. York sure, Rudi studer (2002) On-to-knowledge methodology. In: CiteSeer
8. Sure Y, Erdmann M, Angele J, Staab S, Studer R, Wenke D (2002) Onto edit: collaborative ontology development for the semantic web. In: Springer, Berlin, Heidelberg
9. Fernandez M, Gomez-Perez A, Juristo N (1997) Methontology: from ontological art towards ontological engineering. In: Proceedings of the AAAI97 Spring symposium series on ontological engineering
10. Troncy R (2008) Bringing the IPC news architecture into the semantic web. In: Sheth A et al. (eds.): ISWC 2008, LNCS 5318, pp 483–498
11. A practical guide to building owl ontologies using protégé 4 and CO-ODE tools Edition 1.1
12. Protege OWL Tutorial Available. http://www.co-ode.org/resources/tutorials/Ontology Development 101: a guide to creating your first ontology
13. About RSS, http://rss.softwaregarden.com/aboutrss.html.
14. Hindustan Times e-news, www.hindustantimes.com.
15. Times of India e-news, timesofindia.indiatimes.com.
16. Indian Express e-news, www.indianexpress.com.
17. Precision and Recall, en.wikipedia.org/wiki/Precision_and_recall.
18. Daiki Nagao (2008) Web contents recommender system on RSS using weighted TFIDF. In: University of Aizu, Graduation Thesis
19. Gruber TR (1993) Toward principles for the design of ontologies used for knowledge sharing. Revision: 23 Aug 1993

Chapter 40
Design of Optimized Modular Multiplier Using Montgomery Algorithm for RSA Cryptosystem

Sandip Kakde, Pravin Zode and Pradnya Zode

Abstract Modular multiplication plays a vital role in RSA Cryptography and Elliptical Curve Cryptography. We have implemented a 256-bit Modular multiplier using Montgomery Reduction Algorithm in VHDL. The output of the Montgomery multiplier is $Z = X*Y\ R^{-1}$ mod M. Our main aim is to calculate the area required for the modular multiplier using Montgomery reduction algorithm. It is a full-featured circuit including Carry save Adders, shift registers, multiplexers, parallel registers component and are too big to fit into a single Altera Stratix Device on the Field Programmable platform, so that we are unable to test them in real hardware. However, each sub-component was simulated in Model-Sim SE 6.0 and Altera Quartus II 8.0 and proved functionally correct.

Keywords Montgomery algorithm · RSA · Cryptography · Modular arithmetic

40.1 Introduction

Since long era, network security is a very big issue and number of systems and algorithms are developed in order to secure the data. Advanced Encryption System (AES), Data Encryption System (DES), RSA, Elliptical Curve Cryptography,

S. Kakde (✉) · P. Zode · P. Zode
Yeshwantrao Chavan College of Engineering, Nagpur, India
e-mail: Sandip.Kakde@gmail.com

P. Zode
e-mail: PravinZode@yahoo.com

P. Zode
e-mail: Pradnya.U@rediffmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_40,

Digital Signature, Diffie Hellman Key exchange are the main algorithms for the data security. However RSA (public key encryption algorithm by Rivest, Shamir, Adelman), Digital Signature, Diffie Hellman key exchange are some examples of algorithms that use a series of modular multiplications to compute modular multiplication and exponentiation. In this age of universal electronic connectivity, it is of utmost importance to store information securely. This led to a heightened awareness to protect the data from disclosure, to guarantee the authenticity of data and messages, and to protect systems from network-based attacks. Cryptography plays a major role in mobile phone communications, sending private emails, transmitting financial information, security, computer passwords and electronic commerce digital signature and so on. Modular exponentiation a^b mod m, and implicitly modular multiplication a*c mod m are the operations intensively used, underlying many cryptographic schemes. Nowadays, more and more reconfigurable hardware devices are used in network applications due to their low cost, high performance and flexibility. Such applications include extensible network routers, firewalls and Internet enable sensors, etc. These reconfigurable hardware devices are usually distributed in a large geographic area and operated over public networks, making on-site configuration inconvenient or infeasible.

Therefore, robust security mechanisms for remote control and configuration are highly needed.

40.2 Mathematical Background

Since large number of bit lengths provides greater security to RSA Cryptography, computing complexity will be increased. The integers used in number theoretic cryptographic algorithms are hundreds or thousands of bits long. Thus efficient implementations of multi-precision integer operations are required. Since multiplication and division are more costly operations as far as hardware part is concerned but still are used more frequently than addition and substraction, a lot of efforts are made to optimize them. In particular there are many studies on modular multiplication method, for example, Karatsuba Method, Montgomery Method and Barrett's Reduction method. The Montgomery modular multiplication algorithm has been widely implemented in both software and hardware. Compared to the software implementations, the hardware implementations are faster as a dedicated datapath is used. However, they are fixed in functions and are not able to respond to new algorithms. The Montgomery modular multiplication algorithm was designed to avoid division in modular multiplications. Given two n-bit inputs, X and Y, this algorithm gives $Z = X*Y\ R^{-1}$ mod M, where R equals to 2^n and M is the n-bit modulo. Algorithm 1 shows the Montgomery modular multiplication algorithm in detail.

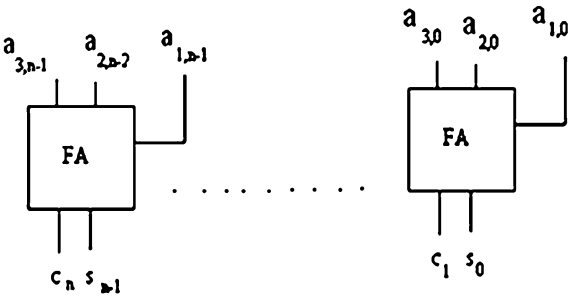


Fig. 40.1 Schematic of full adders

40.2.1 Modular Multiplication

The Montgomery Algorithm [9] is described below. The adders, we constructed for modular multiplication are using shift-add multiplication algorithm. Let X and Y are two k-bit positive integers, respectively. Let X_i and Y_i are the ith bit of X and Y, respectively. The algorithm 1 is stated as follows:

```
Algorithm 1:
Input: X, Y, M, n
Output: Z = X*Y*R^-1 mod M
p: = 0;
for i in 0.. n-1 loop
a: = p + x(i)*y;
if (a mod 2) = 0
then p: = a/2;
else
p: = (a + m)/2;
end if;
end loop;
if p >=m
then z: = p-m; else
z: = p; end if;
```

40.3 Implementation

The carry-save adder (CSA) avoids carry propagation by treating the intermediate carries as outputs instead of advancing them to the next higher bit position, thus saving the carries for later propagation. The sum is a (redundant) n-digit carry-save number, consisting of the two binary numbers S (sum bits) and C (carry bits). A carry save adder accepts three binary input operands or, alternatively, one binary

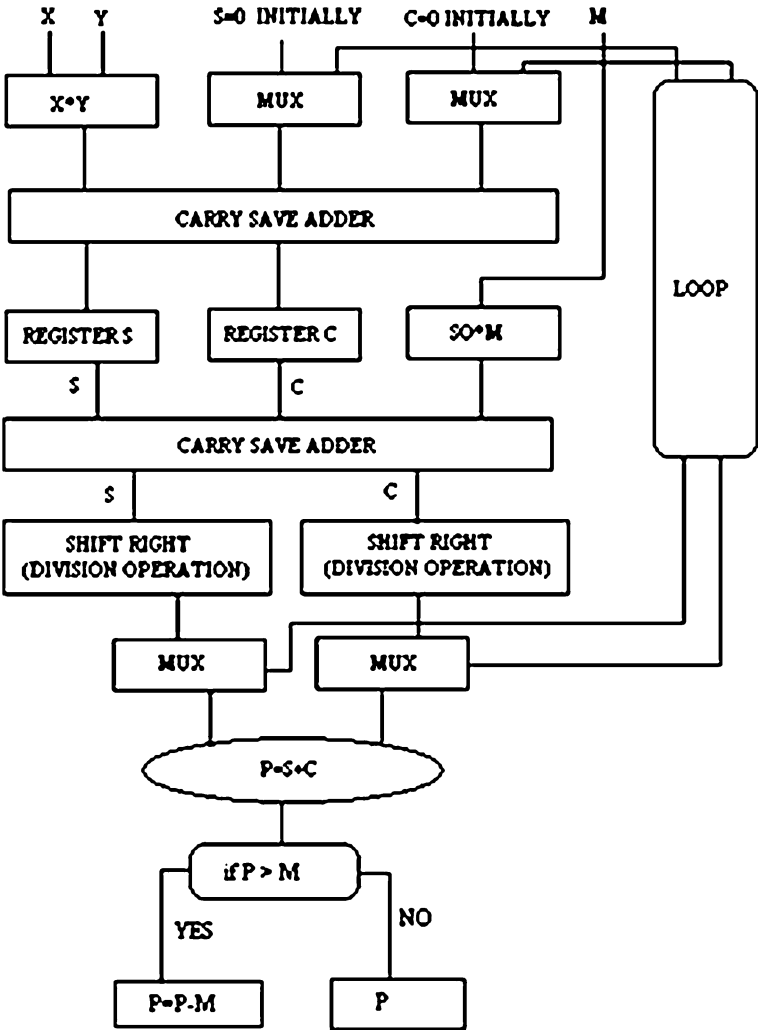


Fig. 40.2 Architecture for algorithm 1

and one carry-save operand. A carry save adder is an important block which adds the intermediate results. A carry save adder consists of a ladder of stand alone full adders and carries out a number of partial additions as shown in Fig. 40.1.

Architecture (Fig. 40.2)

The top level entity indicates only the input and output ports. The description of the hardware culminates in a top level module named Montgomery Multiplier and it has five inputs namely input operands X, Y and clock, reset and start. The outputs are Z and done. For the purpose of our development, we kept the modulo value M, as constant. Once the architecture was specified, the components of the

Table 40.1 Results of Montgomery multiplier on various logic families

Montgomery multiplier (bit length = 8)			
Family	ACEX1 K	FLEX10 K	Cyclone
Total logic elements	104/576 (18 %)	104/3,744 (3 %)	80/5,980 (1 %)
Total pins	28/66 (42 %)	28/189 (15 %)	28/185 (15 %)
Device	EP1K10TC100-1	EPF10K70RC240-4	EP1C6Q240C8

Table 40.2 Results of 256-bits Montgomery multiplier on STRATIX II family

Serial number	Family	Montgomery multiplier (bit length = 256) STRATIX II
	Device	EP2S90F1508C3
1	Area (combinational LUTs)	1808/72768 (2 %)
2	Dedicated logic registers	778/72768 (1 %)
3	Total pins	772/903 (85 %)

design were described in structural VHDL code style. This approach makes the performance of the design less dependent on the synthesis features of the VHDL compiler suite. Only the components containing a finite state machine were expressed in a mixture of structural and algorithmic description and state machine encoding was left to the compiler. All the subcomponents are simulated and tested the functionality by writing test benches for each sub-component.

40.4 Results

The primary objective of this paper is to provide a fair comparison of the efficiency of architecture for the modular multiplication in hardware. The design statistics parameter is area that is the number of LUTs (Look Up Tables) utilized for implementation of the Multiplier. Results of 8-bit Montgomery multiplier on various logic families are shown in Table 40.1 and the results of 256-bit Montgomery multiplier on Stratix II logic family has been tabulated in Table 40.2. The area versus bit lengths graph is as shown in Fig. 40.3. The Modular Multiplier using Montgomery Reduction Algorithm is designed and simulated using ModelSim.

40.5 Conclusions

In this paper, Modular Multiplier design with the Montgomery's reduction algorithm has been implemented using VHDL language and simulated by ModelSim and Altera Quartus II simulators. We have presented a new implementation of

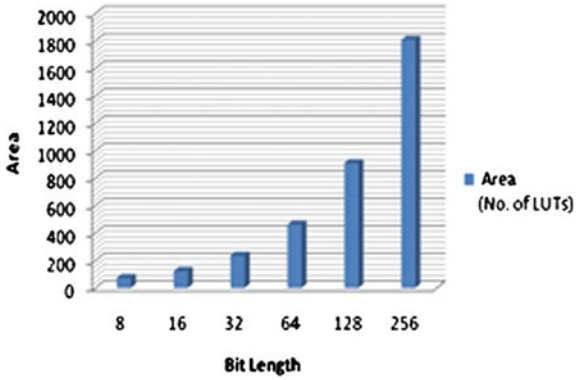


Fig. 40.3 Area versus bit length plot

modules which are the cores of many public key cryptography algorithms. Our multiplier is the improved Montgomery multiplier design with CSA stages and parallel register. The area required will be less as compared to classical modular multiplication.

References

1. Eldridge SE, Walter CD (1993) Hardware implementation of Montgomery's modular multiplication algorithm. IEEE Trans Conzpzrt 42:693–699
2. Clerk Maxwell J (1892) A treatise on electricity and magnetism, vol 2, 3rd edn. Clarendon, Oxford, pp 68–73
3. Koc CK, Acar T, Kaliski B (1996) Analyzing and comparing Montgomery multiplication algorithms. IEEE Micro 16(3):26–33
4. Blum T, Paar C (1999) Montgomery modular exponentiation on reconfigurable hardware. Proceedings of 14th symposium on computer arithmetic, pp 70–77
5. Bunimov V, Schimmler M (2004) Area-time optimal modular multiplication. Embedded Cryptographic Hardware: Methodologies and Architectures, 2004, ISBN: 1-59454-012-8
6. Lu J, Quan W. Implementing a 1024 bit RSA on FPGA. Reconfigurable Network Group
7. Takagi N, Yajima S. Modular multiplication hardware algorithms with a redundant representation and their application to RSA cryptosystem
8. Bernal A, Guyot A (1998) Design of a modular multiplier based on Montgomery's algorithm. Proceedings of the 13th international conference design of circuits and integrated systems (DCIS'98), Novemb 1998
9. Sutter GD, Deschamps J-P, Imaña J L (2010) Modular multiplication and exponentiation architectures for fast RSA cryptosystem based on digit serial computation. 0278-0046/$26.00 © 2010 IEEE

Chapter 41
Automatic Generation of P2P Botnet Network Attack Graph

K. Muthumanickam and E. Ilavarasan

Abstract Attack Graph is a useful representation to reflect attack route existing in network, because it reflects the life path of attack vulnerabilities. As P2P (peer to peer) roBotNetwork (Botnet) has a unique distributed and coordinated attacking behavior, it is difficult to detect its traces. In order to detect and mitigate P2P botnet attack, it is necessary to consider (i) all the hosts in the network as victim of attackers (ii) network-level information which carries malicious programs and commands. Traditional attack graph generation techniques generally have limitations such as time complexity, high space requirement and scalability of network attack graphs. This paper propose an idea to build an efficient botnet detection system model for automatically generating and analyzing network attack graph for P2P botnet using host-level and network-level information.

Keywords **A**ttack graph · Attack route · Host analysis · Network analysis · P2P botnet

41.1 Introduction

A Botnet is an interconnected collection of compromised computers under remotely controlled by BotMaster. Botnet can be used for many illegal activities

K. Muthumanickam (✉) · E. Ilavarasan
Department of Computer Science and Engineering, Pondicherry Engineering College, 605 014, Puducherry, India
e-mail: kmuthoo@yahoo.com

E. Ilavarasan
e-mail: eilaravasan@pec.edu

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_41,

such as distributed-denial-of-service (DDoS) attacks, mass spam mailings, key-logging, and compromising computers to prepare them for infection by near future attacks. The P2P botnet is the one which has unique character as well as distributed attacking coordinated behavior. So P2P Botnet is a serious threat and difficult to detect.

Traditionally attack graph is a useful tool for analysis of network security [2]. In Internet world, attack graphs are manually generated by Red Teams. However, their works were error-prone and tedious for larger network such as Botnet. Various approaches have been proposed to automatically generate attack graphs for analyzing network security [3–8]. Most of previous works on attack graph generation only encounter scalability problem. But suffer from efficiency and space requirement [9].

In this article, a method is proposed to automatically generate attack graph for P2P botnet attack. The organization of this article is as follows. Section 41.2 describes the model for generating network attack graph. Section 41.3 presents the architecture and algorithm for generating P2P attack graph. Section 41.4 talks about analysis of the proposed algorithm. Finally, Sect. 41.5 arrives at the conclusion.

41.2 Model for Attack Graph Generation

The model composed of three parts: Host Vulnerability List, Network Link Path, and Attack store.

41.2.1 Host Vulnerability List

Host vulnerability list consists of different vulnerabilities that can cause the system to compromise by the bot. The resultant data from the host analysis is taken since it monitors the intrusion activity of the node and if the analysis value confirms the bot activity in the host then this will be given as the input and then the links of the particular node is calculated.

41.2.2 Network Link Path

Assume the network that consists of many computers is taken as nodes. These nodes are connected with P2P architecture. The link path will give a list of IP addresses that is connected with the current node. This also includes the network analysis report, which specifically identifies the bot infected node in network

Table 41.1 Attack rules used for graph generation

Rule_Index	Rule_info	Src_criterion	Dest_changes	Sensitivity_Value
1	Registry changes	Registry keys and values	Operating system failure	8
2	File system changes	File system with .exe and .dll	Functions of .dll are modified	5
3	Infector Comparison	Internet connection	Transfer of malicious programs	2
4	P2P node detection	Network connection	Irregular data transfer	7

analysis. The link path is also a vulnerability list that is identified in the network so that it is also given with an ID as shown in Table 41.1.

41.2.3 Attack Store

This is the database that is connected to the attack graph simulator. The attack store consists of existing attack rules for various attack scenarios. The attack rules are used to specify the type of an attack that occurred in the system by the predefined table which is stored in the database. Table 41.1 shows the attack rules used for graph generation.

The attack rule consists of Rule_Index, Rule_info, Src_criterion, Dest_changes and sensitivity_value. Rule_Index is the unique number given to each attack which identifies an attack rule, Rule_info is the description of attack rule, Src_criterion describes the essential conditions in the attackers' host which is needed to launch an attack, Dest_changes represent the effects caused by an attack to the victim and Sensitivity_value shows the difficulty of an attack.

41.2.3.1 Attack Scenario

We have three different scenarios which are given as input to the attack graph simulator. They are: (1) Fully Secured—It is also called as the Strict Protection Mode. This scenario is selected in case where all the hosts in the network are secured. (2)Partially Secured—This scenario is selected when only few nodes in the network are secured. Usually this will be default since there will only be few nodes that are to be very carefully maintained in the network. (3) Singly Secured—This scenario is selected when one host in the network is most secured. If the attack occurs in the network then the attack graph simulator will generate the possible attack path from the infected node to the secured node.

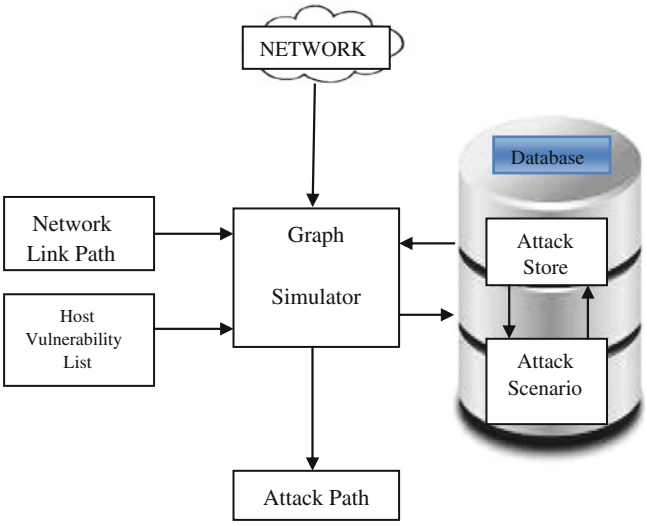


Fig. 41.1 Architecture of attack graph simulator

41.3 Generation of Attack Graph

41.3.1 Architecture of the Graph Simulator

After collecting the information from host and network, it is possible to generate the attack graph for the network. Architecture of the attack graph simulator is given in Fig. 41.1. The Vulnerability of the each and every host is collected by the host analysis part of the architecture. In addition, the link path of the nodes is also collected from the network analysis part. With these values the attack scenario and attack store is also given as the input to the attack graph simulator.

41.3.2 Algorithm to Simulate Attack Graph

In order to launch an attack, the botmaster make use of vulnerabilities in victims and try to attack a host randomly that possesses vulnerabilities in a network. Success of this attempt will freely enable the attacker to attack another host by using privileges and resources of the victim. This process continues until attackers reach their goal.

Based on the information and analysis arrived above, this paper describes a way to find out all the victims in a network. The working principle of the algorithm is as follows:

(1) Select the scenario type, if it is fully secured then the alert will be produced for the network if any one of the nodes is attacked. This is attacker's entry point in the network and this process continues.
(2) If the scenario selected is partial then the secured computers list are taken and the compromised computer is matched with the list. If it is matched, the attack graph will be marked. Else the linked path is obtained for the current node and then the connected node is compared with the secured list. If matched then the graph is marked in second iteration and so on.
(3) Third case considers single system that is considered as secured. First it is compared with the secured node; if it is matched the node is marked in first iteration of the attack graph in the first branch. Else the linked path is obtained and the consecutive nodes are compared with the secured system. If it matches then the graph is marked in second iteration and so on.

```
Algorithm: Attack-graph_Simulation (Input)
Input: Host Vulnerability List, Network Link path and
Attack store
Output: Attack graph
Get Host analysis report
(1) Select scenario
(2) If (scenario == fully secured)
(3) {any node- > attacked
(4) Anynode.generateattackgraph();}
(5) elseif (scenario ==partial)
(6) {get secured node list;
(7) If (node == secured node list)
(8) {node.generateattackgraph();}
(9) Else
(10) {node.anynode- > attacked?
(11) Anynode.generate.attackgraph()}}
(12) Get secured node
(13) Else
(14) {If (node == secured node)
(15) Alert();
(16) Else
(17) {attackgraph();}}
(18) Function alert()
(19) {node.last.display(attacked);}
(20) Function attackgraph()
(21) {node.registers(secured)
(22) For(i = f.n;i <=lastnode;i ++)
(23) {j = i- > nextnode; (24) If(j == lastnode)
(25) {display.j;
(26) Make an edge from previous & current j values
(27) Alert();}}}
```

The nodes in attack graph generated based on the above algorithm are represented as hosts in network. The attack routes in the attack graph represent the shortest route from victim (first compromised host) to secured system (attacker's target).

41.4 Analysis

Suppose the number of the hosts in the network is n, the number of attack rules in the attack store is r, and information list for the nodes is l, the time complexity of the algorithm is n2+nl+r. A queue is maintained to store the nodes generated in the attack graph. Thus compared to other algorithms, the proposed algorithm has better time and space complexity. Since n nodes are included in the attack graph, it additionally solves the scalability problem.

41.5 Conclusion

The previous approaches for attack graph generation generally suffer from scalability, different attack methods, time complexity and high space requirement. As an alternative, this paper proposes an automatic attack graph simulation model based on Dijikstra's algorithm and breadth-first search algorithm. This method solves P2P botnet detection from global angle which is additionally efficient and scalable. Although the proposed method is effective to determine P2P botnet traces, still it has to be improved to overcome problems such as duplicate node reduction and generating the attack rule store automatically for predicting attack scenario which are the directions for our future work.

References

1. Zhong S, Yan D, Liu C (2008) Automatic generation of host-based network attack graph. 2009 IEEE world congress on computing science and information engineering, 2008
2. Swiler L, Phillips C, Ellis D, Chakerian S (2001) Computer attack graph generation tool. In: Proceedings of DARPA information survivability conference and exposition II9DISCX'010, 2001
3. Phillips CA, Swiler LP (1998) A graph-based system for network-vulnerability analysis. In: Workshop on new security paradigms, 1998, pp 71–79
4. Sheyner O, Haines JW, Jha S, Lippmann R, Wins JM (2002) Automated generation and analysis of attack graphs. In: IEEE symposium on security and privacy, 2002, pp 273–284
5. Jajodia S, Noel S, OBerry B (2003) Topological analysis of network attack vulnerability. In: Kumar V, Srivastava J, Lazarevic A (eds) Managing cyber threats: Issues, approaches and challenges. Kluwer Academic publisher, Dordrecht

6. Lippmann R, Ingols K, Scott C, Piwowarski K, Kratkiewicz K, Artz M, Cunningham R (2006) Validating and restoring defense in depth using attack graphs. In: Proceedings of Milcom 2006, Washington, DC, 2006, pp 1–10
7. Man D, Zhang B, Yang W, Jin W, Yang Y (2008) A method for global attack generation. Networking sensing and control, 2008
8. Tang Li Z, Lei J, Wang L, Li D (2007) A data mining approach to generating network attack graphs for intrusion prediction. Fuzzy systems and knowledge Discovery, 2007
9. Xie A, Zhang L, Hu J, Chen Z (2009) A probability-based approach to attack graph generation. In: Second international symposium on electronic commerce and security, 2009, pp 343–347
10. Xie A, Chen G, Wang Y, Chen Z, Hu J (2009) SSIRI 2009 short paper: a new method to generate attack graphs. In: Third IEEE international conference on secure software integration and reliability improvement, 2009
11. Man D, Zhang B, Yang W, Jin W, Yang Y (2007) A method for global attack graph generation, 2007
12. Ou X, Boyer WF, McQueen MA (2006) A scalable approach to attack graph generation. In: Jules A, Wright RN, Di Vimercati SDC (eds) ACM conference on computer and communications security, ACM, 2006, pp 336–345

Chapter 42
Parallelization of Fractal Image Compression Over CUDA

Shazeb Nawaz Khan and Nadeem Akhtar

Abstract Fractal Image Compression has been a promising scheme to allow for substantially high compression ratios for image compression. Several algorithms have been proposed based on this scheme. But the enormous amount of computing and therefore the long runtime needed to encode the image makes these algorithms impractical for commercial purposes especially on personal computers. Recently most of the personal computers & laptops are being shipped in with dedicated graphic processing units (GPU). In this paper we present an approach to parallelize the Fractal Encoding Scheme over GPU using CUDA. This makes fractal image compression feasible for personal computers.

Keywords Fractal image compression · Graphic processor · Parallel computing · CUDA

42.1 Introduction

Fractal Image compression (FIC) [1, 2] involves finding out self similar patterns in parts of the input image. Jacquin [3] proposed one such scheme in which the image is partitioned into square blocks of pixels. The best matching block is found for each and the corresponding indices are stored as a mapping function. This mapping function is later referred to iteratively recover the compressed image. A large

S. N. Khan (✉) · N. Akhtar
Aligarh Muslim University, Aligarh, India
e-mail: shazebnawazkhan@gmail.com

N. Akhtar
e-mail: nadeemalakhtar@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_42,

number of comparisons are needed to find the best matching blocks. To obtain a high level of detail more number of blocks is needed to be considered. Thus the number of comparisons needed increases yet more. In FIC the encoding takes far more time than decoding does. The decoding of compressed image can be done in real time, but encoding cannot. There have been several efforts [4, 5] to speed up the encoding time. One dimension to be explored in order to speed up the encoding process is to parallelize the procedure. An approach towards parallelizing over OpenMP is discussed in [6]. But with OpenMP we cannot involve a large number of processing cores on a personal computer. In this paper we present an approach to parallelize the encoding algorithm over graphic processing units (GPU). GPU serves better due to decrease in communication overhead, and due to the intrinsic data parallel nature of the sequential algorithm.

The structure of this paper is as follows. In Sect. 42.2 the sequential encoding scheme is discussed. Section 42.3 introduces the CUDA programming model. The parallelization approach and the algorithm used for implementation over CUDA are introduced in Sect. 42.4. We finally conclude with the results in Sect. 42.5.

42.2 Sequential Algorithm

The Encoding Algorithm [1] for FIC involves the input image to be partitioned into square blocks (of pixels) of equal size. The input image Ω_{orig} is termed as the Range image. The range image is resized to half along each dimension. The resized image is termed as Domain image. The blocks in range image and domain image are referred to as range blocks and domain blocks respectively. The size of range blocks is equal to the size of domain blocks. The blocks assumed in the range image are non-overlapping. While the domain blocks are overlapping and begin one pixel apart. Thus for example an 512×512 image will have $(512/n) \times (512/n)$ range blocks, the domain image will be 256×256 pixels with $(256 - n + 1) \times (256 - n + 1)$ domain blocks, considering the block size to be n x n pixels. Each domain block is indexed as (i, j) while each range block is indexed as (k, l) represented in (row, column) format.

The Algorithm aims at finding the closest matching domain block for every range block. For a 512×512 image, and a block size of 16×16 pixels, the range image has $32 \times 32 = 1{,}024$ range blocks and the domain image has $241 \times 241 = 58{,}081$ domain blocks. Thus in order to find the best matching domain block 58,081 block comparisons are needed to be made per range block; consequently $1{,}024 \times 58{,}081$ comparisons are needed for the entire range image. Moreover these comparisons are not scalar, they are block comparisons. The following affine transformation is applied to each domain block to arrive at the best approximation of each range block

$$\left(D_{i,j}\right) = \alpha D_{i,j} + t_0 \tag{42.1}$$

where $\alpha = [0,1]$, α is a real number and $t_0 \in [-255, 255]$, $t_0 \in Z$.

Each domain block is transformed and compared to each range block $R_{i,j}$. The exact transformation on each domain block that is the determination of α and t_0 is found minimizing

$$\min \sum_{m,n} \left(R_{k,l}\right)_{m,n} - \left(\Gamma\left(D_{p,q}\right)\right)_{m,n}$$

with respect to α and t_0

$$\alpha = \left(N_s^2 \sum_{m,n} \left(D_{i,j}\right)_{m,n} \left(R_{k,l}\right)_{m,n} - \left(\sum_{m,n} \left(D_{i,j}\right)_{m,n}\right)\left(\sum_{m,n} \left(R_{k,l}\right)_{m,n}\right)\right) / \left(N_s^2 \sum_{m,n} \left(\left(D_{i,j}\right)_{m,n}^2\right) - \left(\sum_{m,n} \left(D_{i,j}\right)_{m,n}\right)^2\right) \tag{42.2}$$

$$t_0 = \left(\left(\sum_{m,n} \left(R_{k,l}\right)_{m,n}\right) - \alpha\left(\sum_{m,n} \left(D_{i,j}\right)_{m,n}\right)\right) / \left(N_s\right) \tag{42.3}$$

where m, n are the block indices, and N_s = size of a block in pixels.

Each transformed domain block $\Gamma(D_{i,j})$ is compared to each range block $R_{k,l}$ in order to find the closest domain block to each range block. This comparison is performed using the following distortion measure

$$d_{l2}\left(\Gamma\left(D_{p,q}\right), R_{k,l}\right) = \sum_{m,n} \left(\left(\Gamma\left(D_{p,q}\right)\right) - \left(R_{k,l}\right)_{m,n}\right)^2 \tag{42.4}$$

Each distortion is stored and the minimum is chosen. The transformed domain block which is found to be the best approximation for the current range block is assigned to that range block, i.e. the coordinates of the domain block along with its α and t_0 are saved into the file describing the transformation. This file containing the transformation mapping is the representation of the image in compressed form.

$$\Gamma\left(D_{i,j}\right)_{best} \rightarrow R_{k,l}$$

The reconstruction process of the original image consists of the applications of the transformations described in the fractal code book iteratively to some initial image Ω_{init} until the encoded image is retrieved back. This initial image can be any image. The obtained image Ω_n is independent of Ω_{init}. The transformation over the whole initial image can be described as follows:

$$\begin{aligned} \Omega_1 &= \eta(\Omega_{init}) \\ \Omega_2 &= \eta(\Omega_1) \\ \Omega_3 &= \eta(\Omega_2) \\ \ldots &= \ldots \\ \Omega_n &= \eta(\Omega_{n-1}) \end{aligned} \tag{42.5}$$

where η can be expressed as two distinct transformations $\eta = \Gamma(\Omega)\Psi(\Omega)$.

$\Gamma(\Omega)$ represents the down sampling and low-pass filtering of an image Ω to create a domain image e.g. reducing a 512 × 512 image to a 256 × 256 image as we described previously. $\Psi(\Omega)$ represents the transformations defined by our mappings from the domain blocks in the domain image to the range blocks in the range image as recorded in the compressed file. Ω_n will converge to a good approximation of Ω_{orig} in less than 7 iterations.

42.3 The CUDA Programming Model

NVIDIA introduced the Compute Unified Device Architecture [7] in the year 2006 with the objective to extend the applicability of GPU beyond Graphic domain and to utilize the parallelism for general purpose computing. CUDA comes with a software environment that allows programmers to use C/C ++ as a high level programming language.

The CUDA parallelization model comprises of a ***Grid***. A grid is a 1 (or maybe 2) dimensional array of blocks. A ***block*** is further a 1, 2 or 3 dimensional array of threads. A ***thread*** is the basic unit of processing in CUDA. Several threads are capable of execution in parallel. The blocks are expected to be independent of each other in execution that means there should be no data dependency between the blocks in a grid.

Each thread has private local memory. Each thread block has shared memory visible to all threads of the block and with the same lifetime as the block. All threads have access to the same global memory. The number of blocks executing in parallel depends on the actual GPU hardware. On feeble GPUs CUDA uses the SIMT [8] scheme under which it splits a block into warps of 16 or 32 threads, and then executes all threads in a warp in parallel (Fig. 42.1).

42.4 The Parallelization Approach

The core sequential algorithm for the Fractal Encoding scheme serves as a good candidate for parallelization because of the enormous amount of computation involved in finding the best matching Domain block for each Range block. The C like representation of the sequential encoding algorithm is presented below:

42.4.1 Sequential Encoding Algorithm

```
Rn : The number of Range blocks across one dimension
Dn : The number of Domain blocks across one dimension
min: a threshold value initialized high enough, may be ∞
```

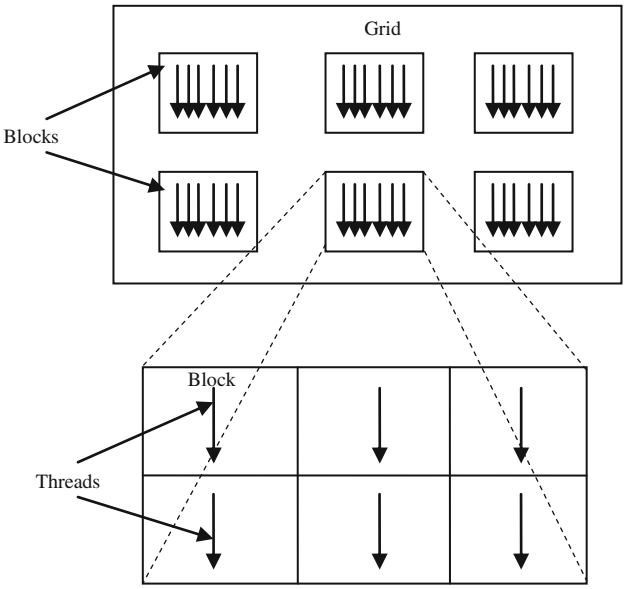


Fig. 42.1 CUDA parallelization model

```
for each Range block (k,l) : 0 <= k,l <= Rn
      for each Domain block (i,j) : 0 <=i,j <= Dn
            (α, t0) ← getAlpha(i,j,k,l)
            match ← getMatchMeasure(i,j,k,l)
            if(match < min)
            {min ← match
            (p,q) ← (i,j)
            }
            store the mapping (k,l) → (p,q)x(α, t0) in list
      }
}
```

In this algorithm there are two loops in nested order. The outer loop is for each Range block. The inner loop is for each Domain block. We pick up each range block at once and for each range block we try to find the best matching domain block by iterating throughout all the Domain blocks.

The function getAlpha() computes the affine transformation parameters α and t_0 for each domain block with respect to the current range block. Once the parameters are obtained then the getMatchMeasure() function finds the distortion measure among the two blocks (range and domain blocks) as the root mean square value of the difference in the corresponding pixel intensity levels. This process is repeated for all domain blocks and the indices of the best matched domain block corresponding to each range block are stored in order along with the transformation parameters as a list. This list serves as the mapping function to regenerate the

original image while decoding. This list is preserved as the compressed form of the image in a file.

The sequential algorithm shows scope for parallelization at two levels:

1. Parallelism at the level of Range Blocks: Each Range block can be processed in parallel to find the best matching domain block in an independent manner.
2. Parallelism at intra block level: All pixel values in a block can be processed in parallel while finding α and t_0 coefficients.

However there are data dependencies occurring at two levels:

1. Updating the 'min' value for match measure: CUDA is a Concurrent Read Exclusive Write platform. Therefore explicit measures need be taken to keep this instruction atomic.
2. Intra Domain level dependency: This dependency occurs inside the 'getMatchMeasure()' function, where the root mean square error value for all the pixels is computed.

We consider each range/domain block as corresponding to a block of threads on the CUDA model. Thus a single thread handles the processing of one pixel value from the domain block along with the corresponding single pixel value in the range block.

42.4.2 Parallel Encoding Algorithm

```
for each Range block (k,l): 0 <=k,l <=Rn in parallel
      for each Domain block (i,j): 0 <=i,j <= Dn
            (α, t0) ← getAlpha(i,j,k,l)
            match ← getMatchMeasure(i,j,k,l)
            if(match < min)
            {min ← match
            (p,q) ← (i,j)
            }
            store the mapping (k,l) → (p,q)x(α, t0) in list
      }
}
```

As illustrated in the parallelization scheme introduced here the outer for loop is executed in parallel. The scheme avoids the first data dependency (discussed above) by serializing the inner for loop. This makes the blocks independent of each other as is needed on the CUDA architecture. The second dependency is handled by using the parallel reduction algorithm. The computations inside the inner for loop are executed in parallel for each pixel value in a range/domain block, and the reduction instructions like summation and averaging are done using parallel reduction [8] in $\log_2(N_s^2)$ steps.

Fig. 42.2 Original Lena image

Fig. 42.3 Decoded image after compression with parallel compression algorithm

The transformation mapping between the Range and Domain blocks along with α and t_0 are stored onto a list and later written into a file, representing the compressed image. The storage of transformation mapping involves three integer values 'p', 'q' and 't_0' (that is at most 6 bytes) and a floating point value α (that is 6 bytes) amounting to consumption of 12 bytes per range block. The range block indices may be implied as the sequential ordering of the records hence need not be

stored explicitly. Thus a 512×512 grey scale image requires just $12 \times 1{,}024 = 12$ kB (for 1024 range blocks) of memory for storage.

42.5 Results and Conclusion

The illustrated parallel algorithm was coded in C and executed under Mac OS X Snow Leopard (Darwin) platform running on Intel Core 2 Duo CPU along with an NVIDIA GeForce 320 m GPU with 6 multiprocessors having 8 cores each (total 48 cores). The sequential algorithm was run on the CPU of same machine. The image taken is a 512×512 grey scale bitmap image. The block size chosen is 16×16 pixels i.e. $N_s = 16$. The average runtime for the sequential algorithm on CPU was noted to be 1021.5 s, while the average runtime for the parallel algorithm was noted to be 95 s. The speedup therefore is more than ten times.

The parallelized algorithm involves no amendments in the mathematical approach towards generation of transformation mapping list, hence the evaluation and comparison of measures of distortion (SNR) and compression ratio shall be redundant. This algorithm computes the same transformation mapping as the sequential version does. The original Lena image and the decoded form of Lena image compressed with the proposed parallel algorithm are depicted below. The initial image Ω_{init} taken here is a 512×512 image with a central white 256×256 square block surrounded by black background region. The algorithm yields a speed up of more than 10 times with the similar compression performance as was observed with the sequential version (Figs. 42.2, 42.3).

Acknowledgments I am thankful towards my supervisor Mr. Nadeem Akhtar for his guidance throughout the work. I am also thankful to my family and friends for their encouragement throughout.

References

1. Fisher Y (1994) Fractal image compression theory and application. Springer, New York
2. Barnsley MF, Hurd LY (1993) Fractal image compression. A.K. Ltd., Peters
3. Jacquin AE (1992) Image coding based on a fractal theory of iterated contractive image transformations. IEEE Trans Image Process 1(1):18–30
4. Zhuang W, Bixi Y (2010) An effective fractal image compression algorithm. ICCASM, 2010
5. Kung CM, Yang WS, Ku CC, Wang CY (2008) Fast fractal image compression base on block property. ICACTE, 2008
6. Hua C, Xi-jin G (2010) OpenMP parallelization of Jacquin fractal image encoding. IEEE, 2010
7. David BK, Wen-mei WH (2010) Programming massively parallel processors. Morgan Kaufman, Burlington
8. NVIDIA CUDA C Programming Guide, NVIDIA CUDA Version 4.0 (2011)

Chapter 43
Distributed Shared Files Management

Saurabh Malgaonkar and Sakshi Surve

Abstract Most often file sharing is the common and basic requirement of any domain. Users can use a system that connects to all the peers in the network to access the shared files. Files of interest can then be downloaded directly from the users in the network.

Keywords P2P File Access · Network Sharing · Distributed Shared File System · Large File Sharing

43.1 Introduction

In a computer system a file is a named object that comes into existence by explicit creation, is immune to temporary failures in the system and persists until explicitly destroyed. The two main purposes of using files are as follows:

1. Permanent storage of information.
2. Sharing of information.

Common methods of storage, transmission, and distribution used in file sharing include manual sharing using removable media, centralized server on computer networks, World Wide Web-based hyperlinked documents and the use of distributed peer-to-peer networking. The most common and feasible approach is to

S. Malgaonkar (✉) · S. Surve
Thadomal Shahani College of Engineering, Bandra (West), Mumbai 400 050, India
e-mail: saurabhmalgaonkar@gmail.com

S. Surve
e-mail: geetams24@rediffmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_43,

use peer to peer file sharing for implementing a distributed shared files management system. Peer to peer file sharing is economically efficient, when the user wants to find specific information, searching for the same would require a lot of human efforts and time. If the upcoming technologies are clubbed with the existing ones it can help better understand the whole system. Thus extending the idea of peer to peer in the file sharing environment helps better built the whole system. In [1] the key features in P2P file transfer are highlighted. In [2] a reliable and simple P2P file sharing system is described which avoids unnecessary data redundancy and connectivity issues among peers by maintaining an adapter which optimizes the working of the entire file sharing system. The disadvantages of a client server file system which do not scale with respect to the number of users and exhibit a single point failure are further highlighted in [3] and also insists on the utilization of a fault tolerance mechanism.

43.1.1 User Interaction Scenario

When a user interacts with the system when joined to a particular network, the user adds the file entry that needs to be shared among the clients in the network. The user mentions the category of the file and also adds its description so the other users are aware about the contents of the file. A user can also search for a particular file entry from the required parameters (name, category or description) as the users in a network can share hundreds of file entries and it is impossible to look for a particular entry manually. Once the user finds the required file entry, the user with its help can access the file by receiving it from the client who is sharing that particular file (Fig. 43.1).

43.2 DSFM Design

It is necessary to keep the system in a constant flow and achieve the targeted goals of the proposed system at the same time. The three most important components of the system are:

1. *Client* The client system will allow the clients to share the files that are required as well as provide an interface that will allow the client to access the files shared by all the nodes in the system (centralized view of shared files). When a client joins the network its shared files will be added to the system and when the client leaves the network all his shared files entries are discarded from the system. After a client receives the updated shared files list, it can access the file from the respective clients.
2. *Controller* The main controller will store the address and shared files information of the client peers and will be responsible for distributing them to all the

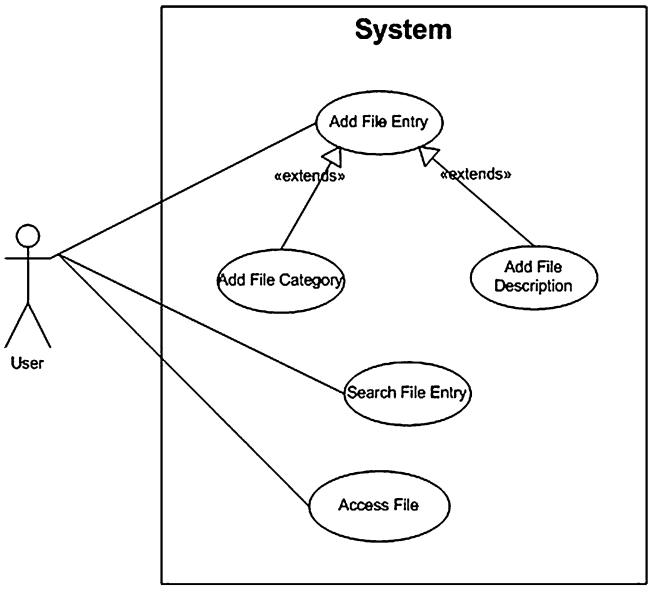


Fig. 43.1 User interaction with the system

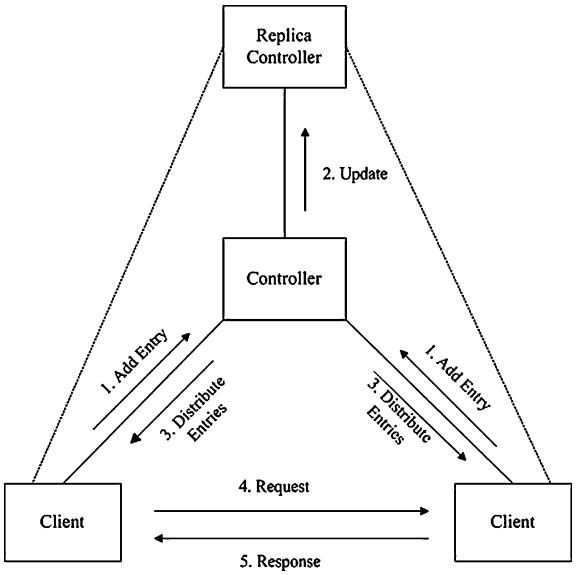


Fig. 43.2 Overall system workflow

clients in the network. Its basic task is just to index the file entries from all the clients and distribute them accordingly. Also when a client joins or leaves the network it will update its shared files entries accordingly and inform the remaining clients. So through the controller we will be able to achieve the

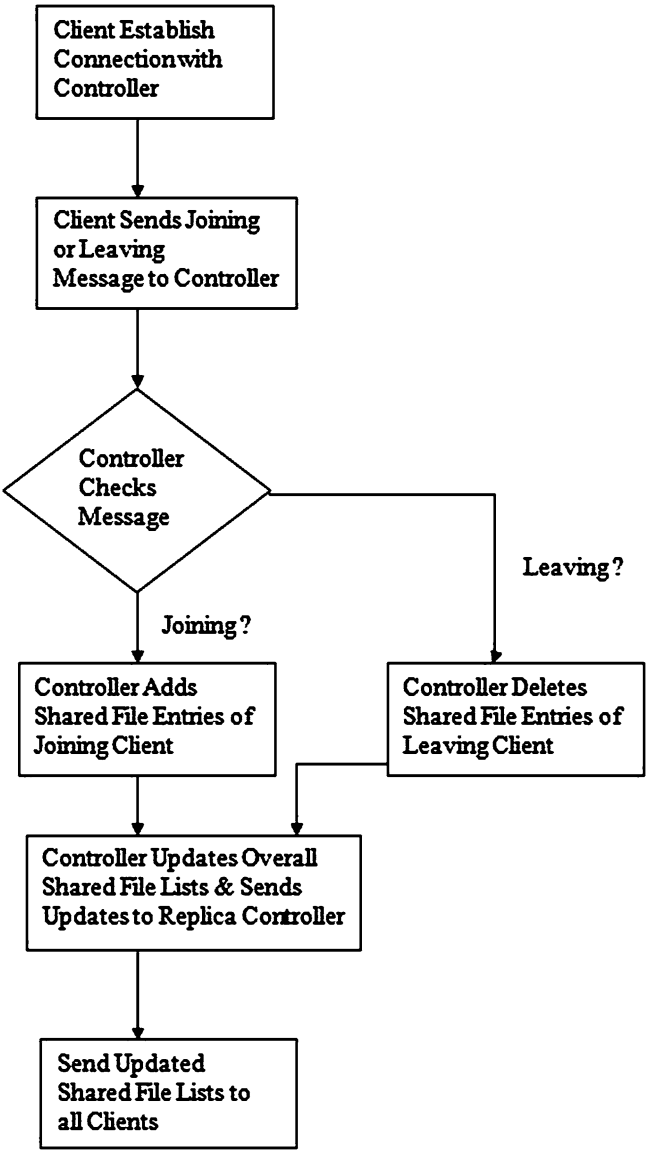


Fig. 43.3 General flow chart

scenario of generating a centralized view of the shared file entries of all the clients in the network and then presenting this view to all the clients in the network.

3. *Replica Controller* The project plans to replicate the main controller so even when the main controller fails the system does not comes to a standstill, the

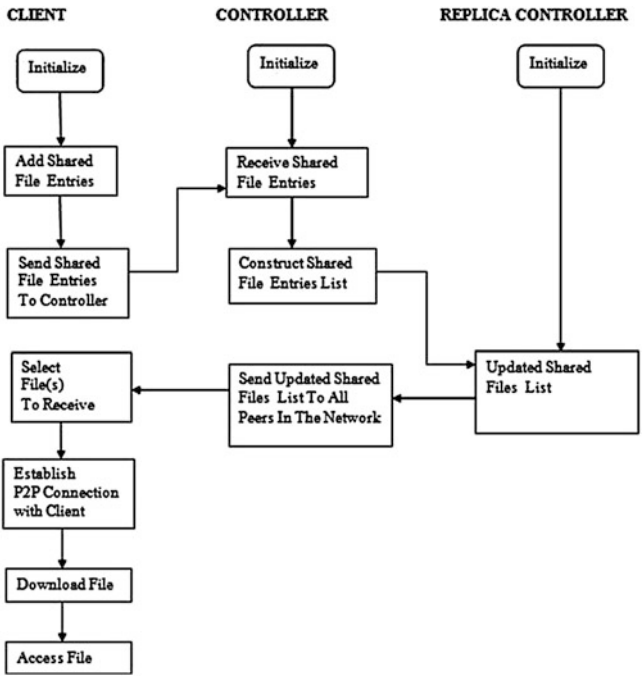
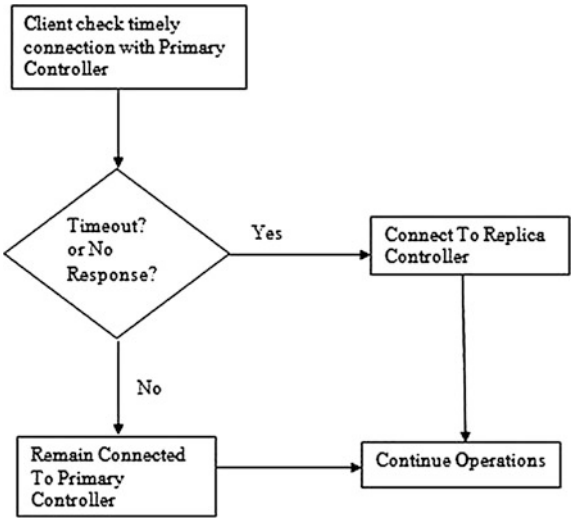


Fig. 43.4 Basic file sharing process

Fig. 43.5 Handling primary controller failure



operations can be handled from the replica controller whose functionality is same as that of the primary controller and the system continues to operate in a continuous flow (Fig. 43.2).

File	Category	Description
conf.txt	Network	Server Configuration Settings
image.jpg	Sports	Jeff Thomson Image
KEVIN1.JPG	Game	Kevin Rayman Resident Evil
DSFM.pdf	Project	Rough Project Documentation

Fig. 43.6 Overall shared files interface

43.2.1 General Basic Functionality

The following diagram denotes the basic interaction among the modules and functionality of each module. This is the normal scenario highlighted when the system is working with the primary controller when fully functional (Fig. 43.3).

43.2.2 File Sharing Process

The following diagram illustrates the scenario that enables to achieve the basic file sharing process among the various clients in the network. It is necessary to initialize the controller and replica controller first and then the clients can join the network. The controller only performs the task of indexing all the file entries of all the clients in the network and the replica controller updates itself regarding this information. Once the overall shared files list is constructed, the controller distributes it to all the clients in the network. When a client receives the overall shared files list, it can then establish a direct P2P connection with the client hosting that particular file and download from it directly and then access it. The updating of the replica controller will ensure no data loss when the controller fails and when the clients will connect to the replica controller. So the previous entries will not get lost and the same functionality as that of the controller will be provided by the replica controller (Fig. 43.4).

43.2.3 Fault Tolerance Scenario

The following diagram illustrates the scenario about how the clients detect the temporary no response or failure of the primary controller and automatically redirect to the replica controller and the system successfully stays in operation without any loss (Fig. 43.5).

43.3 Implementation

The overall networking framework is designed and being developed for the project utilizing the open source NETTY networking library [4] support provided by JDK 1.6 [5] network application development kit and a reliable mySQL database

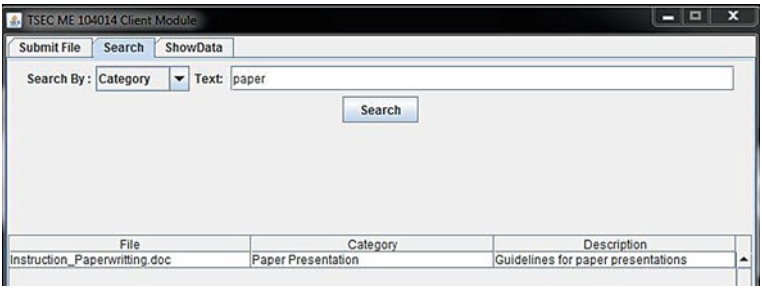
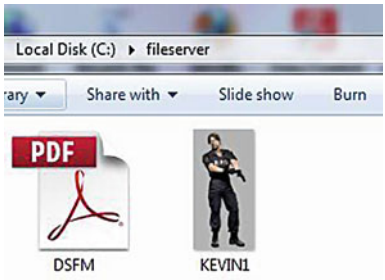


Fig. 43.7 Search results retrieved by a client

Fig. 43.8 P2P file transfer



support provided by the WAMP [6] server application. The whole system has been successfully deployed and tested on an actual network. Five nodes were required on which controller, replica controller, database server and two clients were installed and made to run.

43.3.1 Centralized View of Files

The clients have shared their file entries but they receive such centralized view of shared file entries (Fig. 43.6).

43.3.2 File Entry Search

A new file entry was added by one of the client whose entry was searched using the category criteria and that particular entry was successfully retrieved and accessed by the other client (Fig. 43.7).

43.3.3 P2P File Transfer

One client in the network has successfully received the shared files of the other client. One file sharing scenario was tested when the main controller was operational and the other when the replica controller was operational (Fig. 43.8).

References

1. Ying G, Yao LG, Jian Cong H (2011) A new method of file transfer in computational grid using P2P technique. IEEE international conference on network computing and information security, pp 332–336
2. Zhao R, Liu R, Fu G (2011) P2P file sharing software in IPv4/IPv6. IEEE international conference on software and networks, pp 367–370
3. Chakravarthy S, Hota C (2010) Secure resilient high performance file system for distributed systems. IEEE international conference on computer & communication technology, pp 87–92
4. Netty the Java NIO Client Server Socket Framework. http://www.jboss.org/netty. Accessed 24 Dec 2011
5. Java. http://www.java.com/en/. Accessed 18 Dec 2011
6. WAMP Server. www.wampserver.com/en/. Accessed 20 Dec 2011

Chapter 44
Trusted Computing Architecture for Wi-Fi Protected Access 2 (WPA2) Optimization

Swati Sukhija and Shilpi Gupta

Abstract The Wi-Fi Protected Access 2 (WPA2) is the most secured and recommended protocol for wireless networks today. WPA2 addressed the vulnerabilities of previous protocols wired equivalent privacy (WEP) and Wi-Fi protected access (WPA). WPA2 implemented block cipher AES to provide stronger encryption but it is still vulnerable to various attacks due to transmission of unencrypted management and control frames and group key sharing among peers connected to wireless network. With the rapid popularity of wireless networks, secure transmission of data is extremely essential. The solution for WPA2 shortcomings has been proposed and implemented in this paper and thus, provides protection to wireless networks from several attacks.

Keywords Advanced encryption standard (AES) · Extensible authentication protocol (EAP) · Robust security network (RSN) · Wi-Fi protected access 2 (WPA2)

44.1 Introduction

Solutions for WPA2 vulnerabilities have been discussed in this paper by incorporating trusted computing in order to provide security for wireless networks. Trusted computing aims at addressing the workstation security issues by some

S. Sukhija (✉) · S. Gupta
Department of Computer Science and Engineering, Amity University, Noida, India
e-mail: sukhija.swati@gmail.com

S. Gupta
e-mail: sgupta5@amity.edu

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_44,

software amendments and thus, establishing a trust relationship between clients connected to network. Trusted computing enables binding of data to applications, users and workstations [1].

44.2 Wi-Fi Protected Access 2 (WPA2)/IEEE 802.11i

IEEE 802.11i was proposed in 2004 as solution for IEEE 802.11 and was completely implemented by Wi-Fi Protected Access 2 (WPA2) thus, providing enhancement over Wi-Fi Protected Access (WPA). Counter mode with Cipher Block Chaining Message Authentication Code Protocol (CCMP) was introduced in WPA2 to provide data encryption with Advanced Encryption Standard (AES) block cipher. The encryption standard Temporal Key Integrity Protocol (TKIP) is also available for legacy WPA supported devices. WPA2 suffers from various vulnerabilities which are as follows.

44.2.1 Unencrypted Control Frames

Control frames are unencrypted and are thus, prone to Denial of Service (DoS) attacks [2]. Control frames aid in data frames delivery and used for acknowledgement of received data and acquisition of channel. Various control frames are Request to Send (RTS), Clear to Send (CTS), Acknowledgement and Power Save Poll [3].

44.2.2 Unencrypted Management Frames

Management frames are unencrypted, thus providing the attacker the means to analyze network layout leading to possibility of a DoS attack [2]. These frames aid in initial communication establishment between access point and client stations, thus providing authentication and association services [3, 4].

44.2.3 Hole 196 Vulnerability

WPA2 is prone to the Hole 196 vulnerability. Group Temporal Key (GTK) exchange/distribution occurs during 4-way handshake or group key handshake process. GTK is shared among client stations associated with the authenticator. Thus, an authorized user can sniff and decrypt data of other authorized users and may install malware and compromise other user's devices [5], [6]. A malicious

authorized user can spoof authenticator's MAC address and transmit GTK encrypted packets to launch ARP poisoning attack [2].

44.3 Proposed Solution to WPA2 Vulnerabilities

An encryption algorithm based on stream cipher which can be extended to unencrypted control and management frames in wireless network has been proposed in this paper. Pseudorandom keystream and substitution box (S-box) values are evaluated and XOR operation is then applied on subsequent values to generate cipher text in the proposed algorithm. The steps for encryption algorithm are as follows.

44.3.1 Calculate Passkey Numeral for Encryption

- Random number is generated between 1,024 and 999,999.
- Length of number is calculated.
- Sum of ASCII value of digits in number are calculated.

Thus, Passkey numeral = Numeral length + Sum of ASCII value of digits.

44.3.2 Calculate a0, a1, a2 and a3 Parameters

- a0 = Sum of digits at even positions of passkey numeral
- a1 = Sum of digits at odd positions of passkey numeral
- a2 = Product of digits of passkey numeral
- a3 = (Passkey numeral) mod (256)

44.3.3 Calculate b0, b1, b2 and b3 Parameters

In order to compute b0, b1, b2 and b3 values, encryption parameters EP1, EP2, EP3 and EP4 are required which are computed using Table 44.1:

$$b0 = EP1[0] + EP2[0] + EP3[0] + EP4[0] \quad (1)$$

$$b1 = EP1[1] + EP2[1] + EP3[1] + EP4[1] \quad (2)$$

$$b2 = EP1[2] + EP2[2] + EP3[2] + EP4[2] \quad (3)$$

Table 44.1 Encryption parameters

	EP1 parameter	EP2 parameter	EP3 parameter	EP4 parameter
0	a0 XOR a1	EP1 + 15	a2 XOR a3	EP3 + 55
1	a0 XOR a2	EP1 + 25	a1 XOR a3	EP3 + 65
2	a0 XOR a3	EP1 + 35	a1 XOR a2	EP3 + 75
3	a2 XOR a3	EP1 + 45	a1 XOR a3	EP3 + 85

$$b3 = EP1[3] + EP2[3] + EP3[3] + EP4[3] \tag{4}$$

44.3.4 Calculate c0, c1, c2 and c3 Parameters

$$c0 = ((EP1[b2] \text{ XOR } EP2[b2]) * a0) + b2 \tag{5}$$

$$c1 = ((EP1[b1] \text{ XOR } EP3[b1]) * a1) + b1 \tag{6}$$

$$c2 = ((EP1[b0] \text{ XOR } EP4[b0]) * a2) + b0 \tag{7}$$

$$c3 = ((EP2[b3] \text{ XOR } EP3[b3]) * a3) + b3 \tag{8}$$

44.3.5 Calculate Substitution box (S-box) Values

The calculated substitution box (S-box) values are shown in Table 44.2

44.3.6 Calculate Message Parameter

Message parameter = Passkey numeral (obtained in step 3.1) + Randomly generated key between 1,024 and 9,999 + Average of a0, a1, a2 and a3 parameters (obtained in step 3.2) + Average of b0, b1, b2 and b3 parameters (obtained in step 3.3) + Average of c0, c1, c2 and c3 parameters (obtained in step 3.4).

44.3.7 Message Encryption

- Reverse the plaintext to be encrypted to obtain Partial Message Encryption 1 (PME1).

Table 44.2 S-box values

	0	1	2	3
0	(EP1[b0] XOR c0) * c0	(EP1[b1] XOR c1) * c0	(EP1[b2] XOR c2) * c0	(EP1[b3] XOR c3) * c0
1	(EP2[b0] XOR c0) * c1	(EP2[b1] XOR c1) * c1	(EP2[b2] XOR c2) * c1	(EP2[b3] XOR c3) * c1
2	(EP3[b0] XOR c0) * c2	(EP3[b1] XOR c1) * c2	(EP3[b2] XOR c2) * c2	(EP3[b3] XOR c3) * c2
3	(EP4[b0] XOR c0) * c3	(EP4[b1] XOR c1) * c3	(EP4[b2] XOR c2) * c3	(EP4[b3] XOR c3) * c3

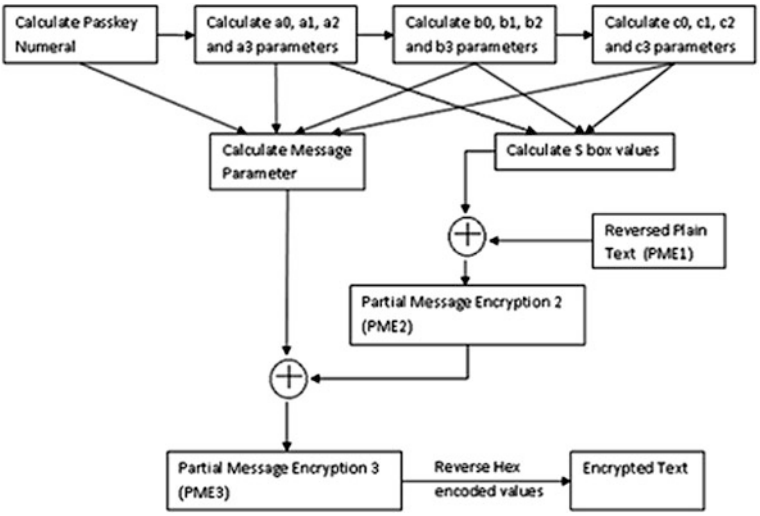


Fig. 44.1 Encryption process

- Perform PME1 XOR S-box [Row-index, Column-index] (obtained in step 3.5) operation to obtain Partial Message Encryption 2 (PME2) where, $0 \leq$ Row-index ≤ 3 and $0 \leq$ Column-index ≤ 3.
- Perform PME2 XOR Message parameter (obtained in step 3.6) operation to compute Partial Message Encryption 3 (PME3).
- Reverse hex encoded value of PME3 to compute Partial Message Encryption 4 (PME4) which is the resultant encrypted text.

The process of message decryption is identical as the above encryption process. The above algorithm is based on trusted computing where trust relationship is established between a pair of communicating devices. The passkey numeral used for encryption is being generated randomly in the first step of encryption algorithm. This algorithm has been implemented at application level of OSI model under the WPA2 umbrella (Fig. 44.1).

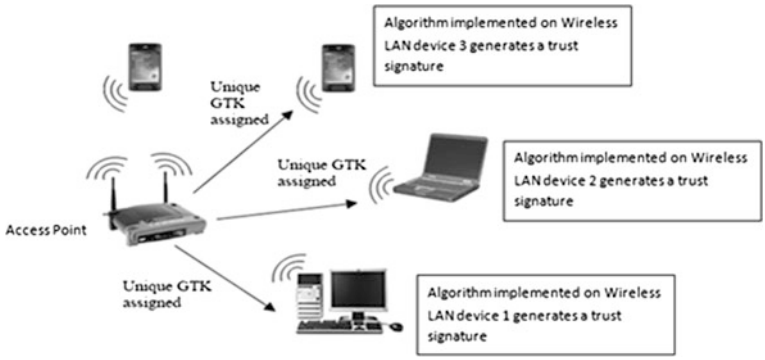


Fig. 44.2 Unique GTK assigned to clients by access point

Identical GTK is shared among all peers connected to the network in WPA2 thus, leading to Hole 196 vulnerability. In order to address the above issue, authenticator can assign random and unique GTK to every peer in network [6]. Thus, in the proposed solution a trust signature is generated for devices which have the above algorithm implemented. The access point then generates a unique GTK to these devices on the basis of trust signature as depicted in Fig. 44.2. Thus, during a multicast or broadcast communication, sender device sends encrypted text using its key to access point. The access point then transmits the encrypted text along with the sender station's GTK to the recipient stations. The recipient stations then decrypt the text at their end using this GTK. At the end of the session, a new GTK is assigned to the sender station. The Hole 196 vulnerability has been overcome in this procedure as each peer connected to network is unaware about GTK of rest of the peers.

44.4 Algorithm Efficiency Experimental Results

The algorithm efficiency has been measured on the basis of memory requirements during the encryption and decryption process. The algorithm has been implemented in Java with JDK version 1.7.0 under Windows 7 Professional 64-bit on Intel® Core™ i5- CPU 2.50 GHz and 8 GB RAM. The used heap size for encryption/decryption process was 180.72 MB and available heap size was 340.68 MB using Netbeans profiler. The above implemented algorithm has been integrated with a mobile emulator as depicted in Fig. 44.3. Sun Java Wireless toolkit has been used for developing the application thus, enabling it to run on wireless devices.

Fig. 44.3 Mobile emulator: **a** cipher text after encryption **b** plain text after decryption

44.5 Conclusion

This paper presents optimizations for WPA2 by incorporating the concept of trusted computing. WPA2 is prone to several attacks due to its vulnerabilities of unprotected control and management frames and sharing of GTK among the peers. Thus, the proposed solution addresses the WPA2 susceptibility and thereby, provides secured transmission of data over wireless networks.

References

1. Blight DC, Trusted computing, voyager systems.http://www.blackhat.com/presentations/win-usa-04/bh-win-04-blight/bh-win-04-blight.pdf
2. Arana P (2006) Benefits and vulnerabilities of Wi-Fi protected access 2 (WPA2), INFS 612-Fall 2006
3. Gast MS (2005) 802.11 Wireless networks: the definitive guide, 2nd edn. O'Reilly Media, Inc. ISBN:978-0-596-10052-0
4. WLANs–WPANs management frames, http://www.wireless-center.net/WLANs-WPANs/1449.html
5. Sukhija S, Gupta S (2011) Wireless network security protocols: a comparative study. Int J Emerg Technol Adv Eng, 2(1), ISSN:2250–2459
6. WPA2 Hole196 vulnerability: exploits and remediation strategies. A whitepaper by AirTight Networks, Inc., www.airtightnetworks.com

Chapter 45
Parallel Pseudo-Exhaustive and Low Power Delay Testing of VLSI Systems

Deepa Jose and P. Nirmal Kumar

Abstract The aim of the paper is to conduct parallel delay testing of modules with different input capacities in a SOC, using mutual BIST pattern generator; especially iterative system realisations well suited for VLSI fabrication technologies. The quality of timing optimised and high performance digital VLSI systems is assured only through delay testing. A unique accumulator based Iterative Pseudo-Exhaustive Two-Pattern (IPET) generator for parallel delay BIST is presented. Generally, the accumulator belongs to the data-path of the SOC. Hence, IPET test can be generated using micro-code self-test strategy. Reduced hardware overhead due to accumulator based design and test time due to parallelism is found to be beneficial. A CMOS implementation of Low Power Architecture for delay testing is carried out, which reduces test power and test time. These architectures can be used as efficient chip-level designs for high speed and low power BIST of SOCs.

Keywords Delay testing · Parallel BIST · Low power · Digital VLSI · Pseudo-exhaustive test

D. Jose (✉) · P. N. Kumar
Department of Electronics and communication Engineering,
Anna University, Chennai, India
e-mail: deepajose11@gmail.com

P. Nirmal Kumar
e-mail: nirmal@annauniv.edu

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_45,

45.1 Introduction

The advent of nanotechnology has massively increased the density and operating frequency of the SOC. Iterative system realisations, which consist of interconnected modules, currently gain more importance in the modern high-speed digital systems. Iterative systems are well suited for VLSI fabrication technologies and offer advantages like ease of bypassing faulty cells, high flexibility in design, function and performance, and its close resemblance with Field Programmable Gate Array (FPGA).

Pseudo-Exhaustive testing of repetitive structures like Multipliers, Adders, FFT processors, Bit-sliced microprocessors, Iterative Logic Arrays (ILAs), Digital Signal Processing systems, Data-path architectures and Embedded Memories require special set of test patterns in the BIST environment. For such complex VLSI circuits with large number of inputs (n), exhaustive testing requires 2^n test patterns and the test time increases exponentially with n. For a (n,m,k)-CUT with n-inputs, m-outputs and cone-size k, the Pseudo-exhaustive testing approach involves applying exhaustive test to the m-output cones. In such cases Pseudo-exhaustive testing objectives can be formulated so that the entire n-bit space will be exhaustively covered, if for all $n-k+1$ contiguous k-bit subspaces, each of the 2^k patterns occur at least once [1, 2]. This modified scheme is called as Recursive Pseudo-Exhaustive (RPE) testing. Common failure mechanisms that appear in high speed digital CMOS VLSI circuits like gate oxide shorts, bridging lines, trapped carriers in the gate oxide, electromagnetic interference cannot be modelled as stuck-at faults but only as delay faults. For delay faults, detection of CMOS stuck-open faults and transition faults Two-pattern test is required. Delay testing involves checking for accurate temporal behaviour of circuits.

45.2 Proposed Work and Problem Approach

In most of the discussed methods for BIST, very few efficient testing schemes for parallel testing of modules with different cone sizes, on a VLSI chip with delay testing capability using the same BIST pattern generator, are proposed. If parallel testing is made possible, using the same BIST pattern generator the test time and cost can be considerably reduced [3]. Accumulator based Recursive Pseudo-Exhaustive Two-pattern schemes is a solution to this problem. Most of the Pseudo-Exhaustive One-pattern generators discussed earlier need exclusive hardware. Nevertheless these methods have drawbacks like increased area overhead, performance degradation and dynamic power due to large number of switching transitions of the test patterns entering the CUT. Accumulator based Recursive Pseudo-Exhaustive One-pattern generators have been hardly proposed. Moreover, no scheme for accumulator based Recursive Pseudo-Exhaustive Two-pattern test generation exists except for [3]. Hence, a unique accumulator based Iterative

Pseudo-Exhaustive Two-Pattern (IPET) test generator for parallel delay testing is implemented. This leads to lesser hardware overhead and at-speed testing. Generally, the accumulator belongs to the data-path, hence the IPET test can be generated using micro-code self-test strategy, thereby no hardware overhead. A CMOS implementation of a Low Power Architecture (LPA) for delay testing as well as combinational fault testing for BIST is implemented [4].

45.3 Iterative Pseudo-Exhaustive Two-Pattern Generator

The IPET generator shown in Fig. 45.1, is a module with n-inputs E[n:1] and n-outputs A[n:1]. The circuit can generate a Two-Pattern (n,k)- Pseudo-Exhaustive test at the accumulators output for any value of k, ($1 \leq k \leq n$) [3]. At a time, only one E[k] signal is enabled. Depending on the values of E[k], the n-stage selective counter and accumulator are reconfigured to work as sub-stage counters and sub-stage accumulators respectively. The increment value of selective counter is unique for different values of E[k]. This architecture generates either One-pattern exhaustive test, when E[1] is enabled, or One-pattern(n,k) -Pseudo-Exhaustive test, when E[k + 1] is enabled. The n-stage generic accumulator belonging to the data-path architecture continuously accumulates the output from selective counter. The carry generator module in Fig. 45.1 is designed to implement modulo division. The input of carry generator circuit is obtained from the E[k] signals and the carry outputs of the n-stage generic accumulator. When E[k] is enabled, the corresponding cout[k] signal of the n-stage accumulator is passed as carry input(cin) to the accumulator. The carry input of each sub-accumulator is driven by the carry output of the preceding sub-accumulator. By repeatedly enabling E[k] for all values of k,$2 \leq k \leq n$, we can generate (n,k)- IPET test patterns for all $k \leq n$. The E[k] signal can be iterated to generate the IPET test by adding a $y = [\log_2 n]$ stage counter to drive the inputs of a y:n decoder.

45.3.1 Program Code Self-Test Strategy for IPET Test

A Two-pattern test algorithm proposed in [5] is applied to IPET architecture via a control module to generate IPET test. The algorithm was proved to generate all n-bit Two-Pattern tests within $2^{k \times 2^{(k-1)}} + 1$ clock cycles. The algorithm consists of three steps: (a) Generation of n-bit S-CIRCLE ($2^n - 3$), (b) Generation of n-bit CIRCLE (2^n-2), (c) Generation of Zero-Transitions, where, ‘n’ is the input size.

STEP(i)—Generates patterns originating from number A and takes ‘i’ jumps to reach number B.

CIRCLE(i)—Generates patterns originating from number A and performs consecutive Steps(i) until returning to number A.

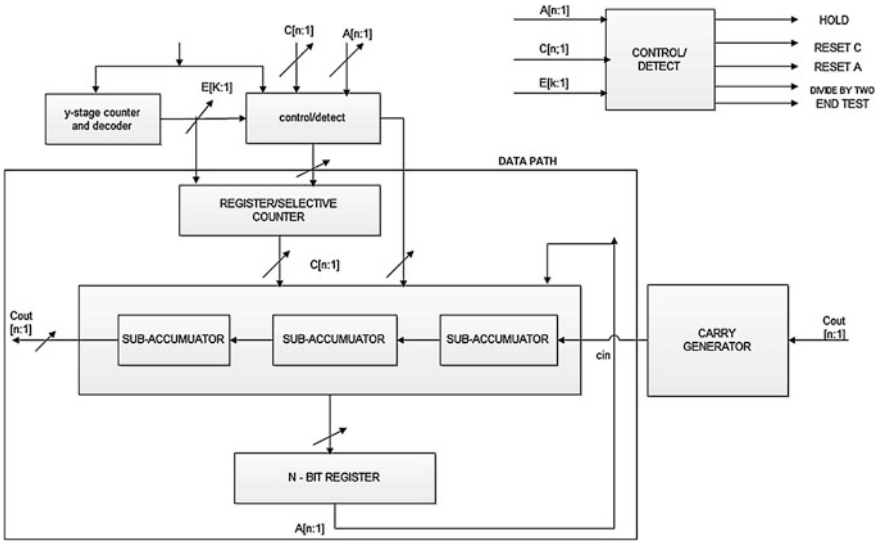


Fig. 45.1 Iterative pseudo-exhaustive two-pattern generator based on data-path architecture

SEQUENCE(k)—Starts from number A and generates consecutive Steps(i) such that i = 1,2,3….k.

S-CIRCLE(i)—Generates consecutive SEQUENCE(i) till it reaches number A again.

S-CIRCLE changes its step value for each jump.

The control module detects specific states of the selective counter and the generic accumulator. Based on these states, the control module generates the required control signals to control the IPET generator to ensure generation of (n,k)-IPET test. The accumulator based Two-pattern test algorithm is implemented using C-program. This code is modified to obtain the IPET test. A modified assembly code using minimum instruction set for implementing the IPET algorithm has been executed using 8051 microcontroller within 0.1425 ns.This code calculates the appropriate increment value of the selective counter and mathematical manipulations are done using modulo division operation to generate the IPET test. A part of the IPET results are shown in Table 45.1. If the accumulator belongs to the data path of the processor, the IPET test can be generated with accurate results using program code stored in memory, thereby replacing the hardware overhead. The mathematically formulated methods in the code are transposed to the IPET architecture.

45.4 Low Power Architecture for Scan Based Delay Testing

Generally, SOC consists of register chains. In such conditions, the register chains can be utilised as scan-chains for delay testing. Unlike the conventional scan cells, the proposed architecture shown in Fig. 45.2 allows the application of arbitrary

Table 45.1 A part of (6,2) iterative pseudo exhaustive two-pattern test results

Output(n = 6, k = 2)	Count	Output	Count	Output	Count
(11) (11) (11)	1	(010)(010)	11	(010)(010)	35
(01) (01) (01)	2	(011)(011)	12	(111)(111)	36
(10) (10) (10)	3	(101)(101)	13	(110)(110)	37
(11) (11) (11)	4	(001)(001)	14	(101)(101)	38
(10) (10) (10)	5	(101)(101)	15	(100)(100)	39
(01) (01) (01)	6	(001)(001)	16	(001)(001)	40
(11) (11) (11)	7	(100)(100)	17	(010)(010)	41
(00) (00) (00)	8	(110)(110)	18	(001)(001)	42
(01) (01) (01)	9	(010)(010)	19	(111)(111)	43

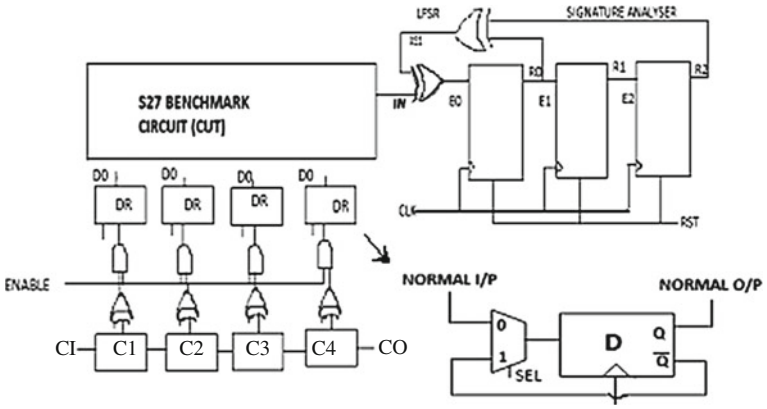


Fig. 45.2 LPA output is applied to the CUT and CUTs output is applied to signature analyzer

vector pairs to the CUT in two consecutive clocks, resulting in testing of delay faults. The Low Power Architecture (LPA) in Fig. 45.2 is implemented. The LPA reduces the number of switching transitions 'f' from propagating into the CUT subsequently reducing the dynamic power. This is done by modifying the scan chain using a separate C-chain of flip-flops and a trigger logic consisting of array of XOR gates and AND gates [4].

Assume that the test generated for a delay fault includes the vector pair (P1, P2), To obtain a 5- bit test vector (n-bit pattern) P1 from previous test vector P2 in the D-flip-flops of the D-register chain, the reformatted difference data between P1 and P2 is fed as a serial input (CI) to the C-chain. This shift mode requires five clock cycles (n-clock cycles) and the upload mode requires only one clock pulse. The C-chain outputs are fed to the AND–XOR array. If the changes in data are detected, the output of the XOR gate will be high. This output is fed to a two input AND gate with Enable signal which generates the Clock enable signal for the D-flip-flops of the D-register chain. If data change is detected, the clock of only the corresponding flip-flops is enabled, inverting the values in these enabled flip-flops alone. The Enable signal for the upload mode is made high after 5-clock cycles

Table 45.2 Power and frequency comparisons obtained

Parameters	Normal BIST	LPA BIST
Test power	3865 mW	2994 mW
Test speed	139 MHz	327 MHz

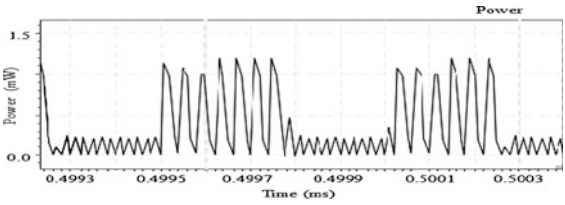


Fig. 45.3 Power results obtained for CMOS implementation of low power architecture

(n-clock cycles). After the shift operation, in the consecutive clock edge, the test vector P1 is changed to test vector P2. Retaining data causes reduced number of transitions at the D-register outputs, which are the inputs of the CUT. This will reduce the number of transitions during shift operations from propagating into the CUT resulting in reduction in dynamic power.

As the Enable signal can be generated internally, no additional pin is required. This signal can be easily generated through a pulse after receiving n-clock cycles. The same SELECT signal in conventional scan chains can be used to determine the Test and Normal select mode signal of the multiplexer [4]. In the normal mode, the circuit does the normal operation. Therefore, the proposed LPA requires no additional test pin compared to conventional scan structures. The proposed Low Power Architecture does not increase the delay during normal operation when testing for delay faults. The generated test vectors are applied to s27 benchmark circuit and the output of CUT is applied to a signature analyser. From the generated signature, any stuck-at-faults in the CUT can be identified. This architecture generates Two-Pattern test for delay testing as successive patterns appear in consecutive clocks at the input of the CUT.

$$P_{Dynamic} = fCV_{cc}^2 \quad (1)$$

The power, time and fault coverage obtained for this modified BIST architecture is compared with the normal LFSR based BIST in Table 45.2. The results indicate that even with the scan chain architecture, the dynamic power during testing process is minimised due to the modification in the scan chain. The disadvantage of the LPA is the increased area-overhead due to the AND–XOR logic. Built-in data registers of SOC can be configured to work in test mode for the LPA to reduce the hardware overhead. The circuit level CMOS implementation of LPA is executed in 0.25 μm tech and signal waveforms and power results obtained (Figs. 45.2, 45.3).

45.5 Conclusion

By generating the IPET patterns, this paper reiterates the fact that accumulator based Recursive Pseudo-Exhaustive Two-pattern test generation schemes for parallel BIST, are superior to all other existing methods as far as hardware overhead and delay testing is concerned. Also test time is reduced due to parallelism. If the accumulator belongs to the data path of the processor, the IPET test patterns can be generated using micro-code self-test strategy, thereby no hardware overhead. The results from the implementation of Low Power Architecture for delay testing as well as combinational testing prove its improved performance in terms of test-time and test-power compared to LFSR based BIST. At the core-level, these architectures can be used for different high speed and low power BIST strategies in VLSI systems.

References

1. Rajski J, Tyszer J (1993) Recursive pseudoexhaustive test pattern generation. IEEE Trans Comput 42:1517–1521
2. Dasgupta P, Chattopadhyay S, Chaudhuri PP, Sengupta I (2001) Cellular automata-based recursive pseudo-exhaustive test pattern generation. IEEE Trans Comput 50:177–185
3. Voyiatzis I, Gizopoulos D, Paschalis A (2010) Recursive pseudo-exhaustive two-pattern generation. IEEE Trans Very Large Scale Integr Syst 18:142–152
4. Hosseinabady Mohammad, Sharifi Shervin, Lombardi Fabrizio, Navabi Zainalabedin (2008) A selective trigger scan architecture for VLSI testing. IEEE Trans Comput 57:316–328
5. Voyiatzis (2007) Accumulator-based pseudo-exhaustive two-pattern generation. J Syst Arch 35:846–860
6. Voyiatzis I (2004) A counter-based pseudo-exhaustive pattern generator for BIST applications. Microelectron J 35:927–935

Chapter 46
An Algorithm for Traffic Analysis Using RFID Technology

RamaKrishna Kothamasu, Rajesh Madugula and Priti Kumari

Abstract In this paper, we presented a method for analyzing traffic on roads by using recent sophisticated technology, Radio Frequency Identification (RFID). First, we gave a brief look at what is RF technology and how to setup readers on roadside for reading the tags. Initially each vehicle is tagged and when it is passed through the RF reader range, the tag data is read by reader. We proposed Traffic Analysis algorithm works at centralized system based on all reader's data. Algorithm itself had three modules. First module for reading data, second module is about analyzing traffic at a particular reader and third module is for analyzing traffic in between readers.

Keywords Radio frequency identification (RFID) · RFID tag or transporter · RF reader · Traffic analysis algorithm · Traffic analysis at reader · Traffic analysis in between readers · RF middleware

46.1 Introduction

RFID technology research started at MIT. It is not complete replacement for current existing barcode systems but with its special advantages, it is ahead of cutting edge of all other automatic reading technologies (Barcode technology,

R. Kothamasu · P. Kumari (✉)
NIT, Patna, India
e-mail: priti_kaushik2003@yahoo.com

R. Kothamasu
e-mail: krkbpk@gmail.com

R. Madugula
Vignan University, Guntur, India
e-mail: rajesh.chinna20@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_46,

Contact memory buttons etc.) [1–3]. Briefly RF technology contains two modules, first module is, reading data from RF tags through antennas by using Radio frequency and it is done by the coupling between tag and antenna. RF tag can be either Active tag or Passive tag [4]. Active tag contains itself a battery to run the tag and, it can operate up to kilometers and it gives good Return on Investment (ROI) [5, 6] if we use in product manufacturing units as well as these are expensive. Passive tags are very cheap and these operate by taking antenna radio frequency signal. Second module is sending data to centralized host or particular host, for this, we have RF middleware.

46.1.1 Traffic Analysis Using RF Technology

Now every country benefitting from RF technology, one of them is Automatic Toll Billing system. In this human intervention is not needed, when RF tagged vehicle is reached, Toll gate it automatically charges according to vehicle type. Almost all vehicles in developed countries are RF tag enabled, If not so, we assumed all vehicles are RF tagged. From the Fig. 46.1, RF antennas broadcasting the signals to read RF tagged vehicles are placed on road side and are connected to RF readers. The computer works on Traffic Analysis algorithm based on RF readers data and gives the traffic analyzed results to WebBrowser or PDA or to a dedicated Monitor. Before placing the antennas, we need to find ambient electromagnetic noise (AEN) by conducting full faraday cycle analysis. Because if we operate RFID with unlicensed band it may interfere with other devices, which are, operate in same frequency.

Perfect coupling [7] is needed while reading the RF, if we place tag on windshield it gives better read results instead of placing on oil tank. We had better not placing tag on oil tank and not tagging the tankers vehicle because RF waves behave differently with liquids and metals. By using the Traffic Analysis algorithm, the computer analyses at what reader traffic existed and checking in between readers as well.

46.1.2 Traffic Analysis Algorithm

```
1 (a, b, NoOfReaders) {
1 /*module 1*/
1. CT: = CurrentTime();
2. for i : = 0 to NoOfReaders step i ++ do {
3. for j : = 0 to 30 step j ++ do {//Taking values up to 30 min
4. for k : = 0 to n step k ++ do {
5. a[i][j][k] : = ReadEPC();
6. b[i][j][k] : = RFTagSpeed();
7. NoOfTags ++;}
```

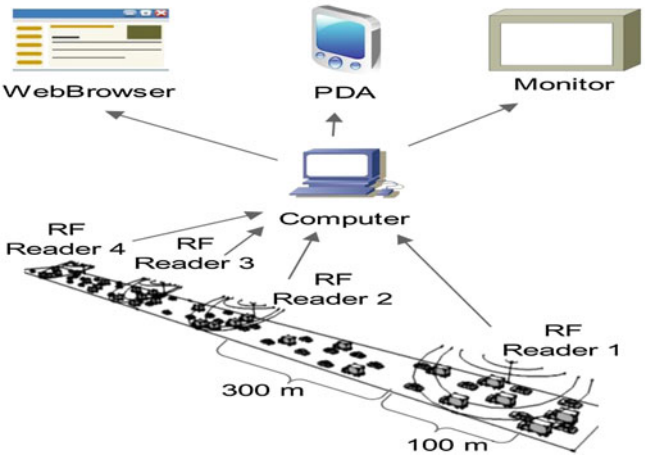


Fig. 46.1 Traffic analysis using RF technology

```
8. C[i][j] = NoOfTags;}}
   /*module 2*/
9. for i : = 0 to NoOfReaders step i ++ do {
10. for j : = 0 to 30 step j ++do {
11. if (((min(c[i][j],c[i][++j])/max
   (c[i][j],c[i][++j])*100) > 75) {
12. if (c[i][j] < c[i][++j]) {/*Line 12-17 doesn't carry
   significant role but while
13. x : = max (c[i][j],c[i][++j]); developing as a program-
   ming code it plays vital role*/
14. y : = min(c[i][j],c[i][++j]);}
15. else {
16. x : = min (c[i][j],c[i][++j]);
17. y : = max(c[i][j],c[i][++j]);}
18. for l : = 0 to x step l ++ do {
19. for k : = 0 to k < y step k ++ do {
20. if(a[i][j][k] ==a[i][++j][l])
21. hit ++;}}
22. Value1 : = hit/min(c[i][j],c[i][++j])*100;
23. if Value1 > 75 then
24. print ''traffic from the moment CT + j to CT ++j at the
   reader i'';}}}
   /*module 3*/
25. for i : = 0 to NoOfReaders step i ++ do {
26. for k : = 0 to 30 step k ++ do {
27. for j : = 0 to 30 step j ++ do {
```

```
28. if (((min(c[i][k],c[++i][j])/
    max(c[i][k],c[++i][j])*100) > 70) {
29. if (c[i][k] < c[++i][j]) {
30. x = max(c[i][k],c[++i][j]);
31. y = min(c[i][k],c[++i][j]);}
32. else {
33. x = min(c[i][k],c[++i][j]);
34. y = max(c[i][k],c[++i][j]);}
35. for p : = 0 to x step p ++ do {
36. for m : = 0 to y step m ++ do {
37. if (a[i][k][m] = = a[++i][j][p]) {/*Line 37, Comparing
    all ReadEPC of current reader
38. hit ++;
39. with all ReadEPC of next reader*/
40. temp1 = temp1 + b[i][j][m];
41. temp2 = temp2 + b[i][j][m];}}}
42. if (((hit/min(c[i][k],c[++i][j]))*100) > 80) {
43. val = ((temp1/hit) -- (temp2/hit));/*avg difference
    from current reader to next reader*/
44. if val > 0 thenprint ''each vehicle Average speed
    increased by val from reader i to ++i '';
45. elseprint ''Average speed of each vehicle decreased by
    val from reader i to ++i '';}}}}}
```

Module 1: All the readers read all the tags and their Tag Speed [8] which are passed through their range based on Anti-Collision protocol [9], in algorithm we assigned each EPC (Electronic License plate) and their respective speed of the tag are assigned to 3D arrays(In program code from line 1 to line 8).

Module 2: From Fig. 46.1, we analyzed the traffic at a RF Reader 1, read the tags for every minute. Analyzing purpose, we took following tabulated values for Reader 1. From line 1, we assumed current time is 9:45.

Analyzing traffic from 9:45 to 9:46, From Line 11, we compared No of Reads (total no of vehicles) at 9:45 and at 9:46. Instead of comparing each ReadEPC, first we compared Total Reads to minimize the computation at computer as shown in the Fig. 46.1. It was 13/15 = 86.66 %.Further we comparedeach TagEPC at time 9:45 with each TagEPC at time 9:46 by Reader 1, it is 11/13 = 84 %.(Line 18 to Line 21 compares each ReadEPC if Line 20 is true it increments hit value by 1). From the Fig. 46.1, Reader 1 covers the distance of 100 m. Speed = 100 m/60 s1.667 m/s→6.0012 kmph ∼ 6 kmph, means 11 vehicles are moving not more than 6 kmph at the Reeder1from time 9:45 to 9:46, it means traffic is existed at this time.From Line 18 to Line 24, print the traffic results if there are75 % of same vehicles from current minute to next minute at the same reader. So from Table 46.1 we can conclude the following results.

From Time 9:46 to Time 9:47, It is 8(matched ReadEPC (Line 22))/10(min (10, 13)) →80 %. Therefore, we can say traffic is existed at this time (line 24). From

Table 46.1 Read values at reader 1 from 9:45 to 9:48 ("–" indicates read nothing) at RF reader 1

Time 9:45		Time 9:46		Time 9:47		Time 9:48	
(Total reads 15)		(Total reads 13)		(Total reads 10)		(Total reads 12)	
TagEPC	TagSpeed	TagEPC	TagSpeed	TagEPC	TagSpeed	TagEPC	TagSpeed
602	9	101	5	591	15	191	12
095	3	045	13	124	17	392	3
423	11	546	12	592	14	045	6
249	8	591	5	091	5	224	9
592	4	124	13	546	13	105	5
091	12	592	3	015	17	646	8
325	9	091	11	423	13	897	11
991	5	991	6	065	18	198	6
124	11	249	10	101	3	499	5
591	5	423	9	045	9	091	9
101	3	325	6	–	–	101	4
085	7	065	8	–	–	603	9
546	13	007	11	–	–	–	–
065	8	–	–	–	–	–	–
045	5	–	–	–	–	–	–

Line 5: From Table 46.1, first column is a[1][time 9:45][] and Line 6: From Table 46.1, second column is b[1][time 9:45][]

Time 9:47 to Time 9:48, It is 3 (matched ReadEPC)/10(min (10, 12)) →80 %. Therefore, it did not execute the line 23 and line 24. So finally, we concluded that, there is traffic from 9:45 to 9:47and no traffic from 9:47 to 9:48. For the sake of understanding, we took up to 4 min only. For getting better results, you can takehours of values. In this way we can find the traffic with in antennas range of all Readers. It very difficult to cover entire city with antennas, for this we developed module 3 to know traffic from one reader to next reader.

Module 3: Here we are analyzed the traffic by using data from RF Reader 1 and RF Reader 2 (From Fig. 46.1). Let us say Reader 2 had the Table 46.2 values.

We applied module 1 on Table 46.2 values, it was given "No traffic from 10:15 to 10:18".By comparing Table 46.1 values with Table 46.2 values, from Line 28, and first it compares the no of vehicles after that comparing each ReadEPCof vehicles. Comparing Table 46.1 first column with Table 46.2 first column, it is 15/15→ 100 %, so we can compare ReadEPC of both columns (Line 35 to Line 40: line 38 got 12 matched tags). Average of TagSpeed of Reader 1 of matched tags (12 tags) is (9 + 11 + 8+4 + 12 + 5+11 + 5+3 + 7+13 + 8)/12→ 8 kmph and average of TagSpeed of Reader 2 of matched tags (12 tags) is

(19 + 11 + 28 + 34 + 12 + 15 + 21 + 25 + 20 + 17 + 13 + 18)/12→ 20 kmph. Traffic decreased as we moved from Reader 1 to Reader2 and average of each vehicle's speed increased from 8 to 20 kmph. From Line 37, if i = 0, k = 0, j = 1, Table 46.1 column 1 compares with Table 46.2 column 2, as the j value increases, comparison is done with Table 46.2 next columns(Analyzing Traffic from Reader 1 to Reader 2). From Line 25 to Line 27, If i = 2, k = 3,

Table 46.2 Read values at reader 2 from 10:15 to 10:18 ("–"indicates read nothing) at RF reader 2

Time 10:15		Time 10:16		Time 10:17		Time 10:18	
(Total Reads 15)		(Total Reads 5)		(Total Reads 10)		(Total Reads 12)	
TagEPC	TagSpeed	TagEPC	TagSpeed	TagEPC	TagSpeed	TagEPC	TagSpeed
602	19	101	35	591	25	191	32
995	23	045	33	124	27	392	23
423	18	546	12	592	34	045	36
249	28	991	25	091	35	224	29
592	34	924	23	546	23	105	45
091	12	–	–	015	37	646	38
523	21	–	–	423	23	897	31
991	15	–	–	065	18	198	25
124	21	–	–	101	23	499	35
591	25	–	–	045	39	091	19
101	20	–	–	–	–	101	24
085	17	–	–	–	–	603	29
546	13	–	–	–	–	–	–
065	18	–	–	–	–	–	–
123	15	–	–	–	–	–	–

From line 5: From Table 46.2, first column is a [2] [time 10:15] [] and from line 6: From Table 46.2, first column is b [2] [time 10:16] []

j = 4, RF Reader3 (i.e. Table 3 column 4) compares with RF Reader4 (i.e. Table 4 column 4), as the j value increases, comparison is done with Table 4 next columns(Analyzing Traffic from Reader 3 to Reader 4). Here Tables 3, 4 is not there but they are corresponding to Reader 3, Reader 4. From Line 25, asi value increases, comparison is made between next two consecutive readers. Therefore, in this way computer can analyze the traffic in between the Readers.

46.2 Conclusion

In this paper, we presented Traffic Analysis algorithm using RFID technology. This algorithm works at centralized host or computer, which is connected, to all the readers. The algorithm itself had three modules, first module is for reading the data, and second module is for finding the traffic at all the readers. Third module is for finding the traffic in between all two consecutive readers. The robustness of RFID technology in various manners has made possible to a reliable, error-free and productive service. With the use of RFID, flexibility of system is maintained with the option of modification of information at any time.

References

1. Wolfram G I, Gampl B I, Gabriel P (eds) (2008) The RFID roadmap the next steps for Europe. Springer, Germany, ISBN 978-3-540-71018-9, pp 127–134, 91–107
2. Fourikis N Advanced array systems, applications and RF technologies: California, USA-A Harcourt science and technology company, ISBN 0-12-262942-6, pp 9–19, 31–82
3. Shepard S (2005) RFID radio frequency identification. Mcgraw-hill, United States of America, ISBN 0-07-144299-5, pp 42–52
4. http://www.rfidtags.com/?gclid=CIHTo8zQk7ACFQp76wodnQGfkg
5. Wessel R (2006) Clothing manufacturer invests its ROI in RFID. RFID J. http://www.rfidjournal.com/blog/entry/2547, Case Study
6. Gambon J (2006) RFID frees up patient beds. RFID J. http://www.rfidjournal.com/article/view/2549
7. Lee E-K, Yoo YM, Park CG (2009) Installation and evaluation of RFID readers on moving vehicles. In: Proceedings of the sixth ACM international workshop on VehiculAr InterNETworking, pp 99–108. ISBN: 978-1-60558-737-0
8. Manikonda P, Yerrapragada AK, Annasamudram SS (2011) Intelligent traffic management system. In: IEEE conference on sustainable utilization and development in engineering and technology (STUDENT), The University of Nottingham, Semenyih, Selangor, Malaysia, 20–21 October 2011
9. Klair DK, Chin K-W, Raad R (2010) A survey and tutorial of RFID anti-collision protocols. IEEE Commun Surv Tutor 12(3):400–421, third quarter
10. http://articles.economictimes.indiatimes.com/2007-09-25/news/27681008_1_rfid-tagsrfidenabled-traffic-congestion

Chapter 47
Trust-Based Grid Resource Management

Damandeep Kaur and Jyotsna SenGupta

Abstract Grid Resource Management and security issues have become critically important with the fast expansion of grid systems. The present day research is moving towards achieving a secured architecture for resource management in Grid System. We hereby, through this paper, wish to present a secured grid resource management for global grids by addition of a trust-based layer whereby allowing grid resource to enter the commercial area, wherein it shall help the grid consumer in Decision Making, as the system will offer only those gird resources which assure of a high degree of trust relationship of grid resource provider. For the successful deployment of a Grid infrastructure, it is essential to access and make maximum use of the resources that are available on the Grid and this is possible only if the secured resources can be tracked effectively and efficiently. It is achieved by presenting a secured Grid Resource Management System based on Trust-Management System.

Keywords P2P · Trust · Reputation · Grid resource management · Grid services · Virtual organization

D. Kaur (✉)
University College of Engineering, Punjabi University, Patiala, India
e-mail: daman_811@yahoo.com

J. SenGupta
Department of Computer Science, Punjabi University, Patiala, India
e-mail: jyotsna.sengupta@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_47,

47.1 Introduction

Grid computing [1] is a coordinated resource sharing and problem solving in any dynamic environment. Grid resource management can be utilized to improve the quality of service offered by the grid so that various heterogeneous resources present in the grid can be managed properly. Resource management is a complex task involving security, fault tolerance along with scheduling. Grid applications compete for resources that are very different in nature, including processors, data, scientific instruments and other services. This paper focus on implementation of security in grid resource management in the form how much trust can be done on a grid resource depending on various parameters like nature of job, performance, availability etc.

47.2 Related Work

47.2.1 Common Security Threats in Grid Environment

1. Individual malicious entities always provide bad services when selected as service providers.
2. Malicious collectives:-Malicious entities form a malicious collective by assigning the maximum trust value to other malicious entities in the network.
3. Malicious collectives with camouflage:-Malicious entities provide bad services in p % of all cases when selected as service providers. Malicious entities form a malicious collective by assigning the maximum trust value to other malicious entities in the network.
4. Malicious spies:-Some malicious entities, known as malicious spies, always provide good services when selected as service providers, but they also give the maximum rating values to those malicious entities that always provide bad services.
5. Sybil attack: In this type of attack, an entity is selected as a service provider, it provides a bad service, after which it is disconnected and replaced with a new entity identity.
6. Man in the middle attack:-Malicious entity can intercept the messages from a benevolent service provider entity to the requestor and rewrite them with bad services, making therefore the reputation of the benevolent entity to decrease.
7. Driving down the reputation of a reliable grid entity:-In such a situation an interaction is never performed with an entity which is actually benevolent but whose reputation has been driven down by malicious participants and entity will not probably be chosen as the entity to have an interaction with.
8. Partially malicious collectives:-When malicious entities, for certain services behave properly, while for other specific services, they act maliciously.

Malicious entities form a malicious collective by assigning the maximum trust value to other malicious entities in the network.

9. Malicious pre-trusted entities:- Some or all the pre-trusted benevolent entities become malicious ones, maybe by always providing bad services when selected as service providers or by rating with maximum trust value other malicious entities who always provide bad services when selected as service providers.

47.2.2 P2P Trust and Reputation Models in Grid

(i) In GridEigenTrust [4], the author exploits the beneficial properties of Eigen Trust, extending the model to allow its usage in grids and integrate the trust management system as part of the QoS management framework, proposing to probabilistically pre-select the resources based on their likelihood to deliver the requested capability and capacity. The global trust for an organization with regard to another organization is built from the direct trust that can be acquired during time from transactions that happened between members of these organizations and by considering also trust information acquired from 3rd party sources. The same trust aggregation scheme can be employed at the level of organization members, each of them storing the trust values for its transaction partners. GridEigenTrust allows obtaining the trust value for an organization by aggregating the trust values of its members.

(ii) Path Trust is a reputation system proposed for member selection in the formation phase of a Virtual Organization. To enter the Virtual Organization formation process, a member must register with an Enterprise Network (EN) infrastructure by presenting some credentials. Besides user management, EN supplies with a centralized reputation service. At the dissolution of the VO, each member leaves feedback ratings to the reputation server for other members with whom they experienced transactions. The system requires each transaction to be rated by the participants. Path Trust is one of the first attempts to apply reputation methods to grids by approaching VO management phases. They approached only partner selection and did not tackled organizational aspects. Their model still lacks dynamicity, as the feedback is collected only at the dissolution of the VO.

47.3 Secured Resource Management

47.3.1 Trust Parameters

1. Satisfaction parameter deals with number of desired features fulfilled by the resource provider. Here it means the amount of satisfaction a grid resource provider can provide in terms of resource consumer's requirement [9, 10].

$$\textbf{Satisfaction} = \frac{\sum_{\mathbf{i}=\mathbf{0}}^{\mathbf{i}=\mathbf{n}} \textbf{Rank Obtained value C.R } (\mathbf{j}, \mathbf{i})}{\textbf{Total Rank value}}$$

where

CR Customer Requirement
i Name of each C.R
j Grid Resource Provider

2. Feedback is the rating given by the consumer after every transaction to the service of grid resource provider.

$$\textbf{Feed Back} = \textbf{Current feedback} + \sum_{\mathbf{i}=\mathbf{0}}^{\mathbf{i}=\mathbf{n}-\mathbf{1}} \textbf{Previous feedback}$$

Rating of Feedback is [0, 1]

If the feedback is to be given by the consumer which is not part of the grid, then the feedback is stored in the recommendation table of a random grid resource selected randomly by grid portal in encrypted format and the key is stored at grid resource where the task was performed.

3. Feedback Decay Function is based on the concept that grid is dynamic so is the credibility of the feedback as resource provider's behaviour can change with time. The decay function is calculated using credibility of function and number of transactions. The feedback decay is important to know to what extent feedback given by the various grid entities about a grid resource provider is important.

   To calculate the feedback decay,
   Feedback Decay = 0 -if last successful transaction is within set time
   .1–if there is no transaction is within set time
   .5–if last unsuccessful transaction is within set time

4. Trust context factor aaggregates the feedback from each transaction as transactions may differ from one another.

   Some of the issues included in Trust context Factor:

(a) Consistency = No of Successful Transaction/Total Number of Transactions
(b) Type of job
(c) Non-repudiation = No of Incomplete Transaction/Total Number of Transactions
(d) Size of job = Large/Medium/Small
(e) Type of Service Provider = Excellent/Good/Average/Poor

Equation 47.1 is a general trust metric derived from the Grid Peer Trust algorithm with few modifications to overcome the drawbacks of PeerTrust model.

$$\textbf{Trust\_final} = \textbf{Satisfaction} + \alpha * \textbf{F} - \gamma + \beta * \textbf{TCF} \quad (47.1)$$

Where

Satisfaction = is the amount of desired features fulfilled by the resource provider.

F = Feedback; Nt = Total Number of Transaction

Feedback Value can be from -1 to 1

FeedBack Decay (γ) is calculated by accessing the last successful transaction time.

TCF = Trust context factor

α, β denote the normalized weight factors for the feedback factor and the trust context factor

47.3.2 Calculating the Trust of Entities

There are few assumptions with respect to grid environment:

- All resource providers must be part of any VO of global grid.
- The basic information of resource provider is stored in trust data of VO.
- Resource Consumer may or may not be part of grid
- Feedback is sent in encrypted form.
- Each grid entity has a trust manager that is responsible for feedback submission and trust evaluation, a small database that stores a portion of the global trust data, and a data locator for placement and location of trust data over the network.

The following is a listing of steps for calculating Trust using GridPeerTrust Algorithm

1. Local Index Table
2. Global Index Table
3. Local Feedback Table
4. Global FeedBack Table

Trust Data contains the following information about any resource provider.

Features; Number of Transaction; Current Feedback (u,i); Previous Feedback

Inputs: Client's requirement, Local Index Table, Global Index Table, Local Feedback Table, Global FeedBack Table

1. The Trust Manager receives Grid Resource Provider List; J [n] created in the Grid, where n is the number of resource Providers.
2. Sort the resources as per client's (Resource Consumer) requirement.
3. Calculate Satisfaction value for these sorted resources.
4. While trust calculated for ten resource provider from Resource Database.
5. Get the feedback (local &global) for the resource provider using

Table 47.1 Security analysis

Trust models security threats	GridPeer trust	GridEigen trust	Path trust
Malicious individual grid entity	Y	Y	Y
Malicious collectives	Y	Y	X
Malicious collectives with camouflage	Y	Y	X
Driving down the reputation of a reliable entity	Y	Y	Y
Malicious spies	Y	X	X
Sybil Attack	Y	Y	Y
Man in the middle attack	Y	Y	Y
Partially malicious collectives	Y	Y	X
Malicious pre-trusted entity	#	X	#

Note Y: Handled
N Not handled
Not applicable

$$\textbf{Feed Back} = \textbf{Current feedback} + \sum_{\mathbf{i=0}}^{\mathbf{i=n-1}} \textbf{Previous feedback}$$

6. Get the Number of transaction from Resource Database
7. Calculate the Decay function

   Feedback Decay = 0-if last successful transaction is within set time
   1–if there is no transaction is within set time
   5–if last unsuccessful transaction is within set time

8. Calculate Trust

   $\textbf{Trust\_final} = \textbf{Satisfaction} + \alpha * F - \gamma + \beta * TCF$

9. End While

Output: Sorted Trust based Resource Provider List, Service Provider List and Recommender's List

47.4 Comparative Analysis

Trust Computation Accuracy [7, 11, 12] (Tables 47.1, 47.2).

To compare the performance of GridPeerTrust with other existing trust models we use a parameter named, Transaction Consistency Rate (TCR) described as the ratio of the number of successful transactions to the total number of transactions. We determine TCR against the variation of (malicious individuals) *m_ind* and Malicious Collectives *m_group*. All experimental results are averaged over 50 runs.

Table 47.2 Security analysis

System	GridEigen trust	Path trust	GridPeer trust
Classification	Probalistic	Probalistic	Probalistic
Centralized data	Yes	Yes	No
Trust metric	[0,1]	[0,1]	[0,1]
Trust aggregation	Yes	Yes	Yes
Type of feedback	Continuos	Negative & positive	Continuos
Storage cost	No	No	No
Scalability	Not applicable	Medium	High

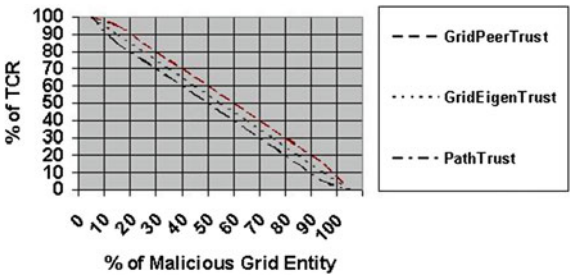


Fig. 47.1 Comparing GridPeerTrust with other models in terms of TCR against *m_ind*

1. Malicious Individuals: In Fig. 47.1, we see that both GridPeerTrust shows superiority over the remaining trust models as the amount of malicious entities in the network increase beyond 40 %.In this model as a list of type service provider is maintained and a grid resource provider giving TCR less than 40 % is considered a bad service provider so its easy to discard malicious entities even their percentage increases. Due to the ease of accessibility, networks today are home to a significantly large number of malicious entities, in other words, threats and risks are implicitly increasing as grid expands. So, in such environment GridPeerTrust would be the best option.
2. Malicious Collectives: In Fig. 47.2, we see that in GridEigen Trust, forming a malicious collective does not increase the global trust values of malicious entities enough in order for them to have impact on the grid due to the presence of pre-trusted entities. A user will always have the opportunity to perform a transaction with one of those pre-trusted entities and if an interaction is performed with a malicious entities (which occurs again around 10 % of the times), it will be identified as malicious by the whole system where as in GridPeer Trust, the accurate management of the credibility of an entity as a recommender, as well as the context factor allows this model to effectively overcome this threat. In path trust, these attacks are actually not handled but avoided by collecting fees for every transaction that are supposed to capture the additional profit gained by the fake transactions, but the more vulnerable a reputation system is to this attack, the higher the fees have to be.

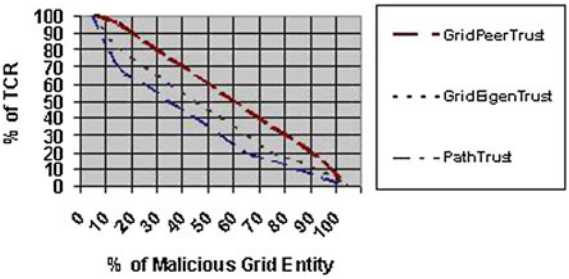


Fig. 47.2 Comparing GridPeerTrust with other models in terms TCR against Malicious Collectives(*m_group*

47.5 Conclusion and Future work

In this paper, we have described a framework for calculating trust in Grid environment. The paper mostly focused on issues related to implementation of security in grid resource management in the form how much trust can be done on a grid resource depending on various parameters like nature of job, performance, availability etc. We have identified several of these issues. Second we have experimented with an architecture and algorithm to gain experience with this new area of research for the Grid community. We have identified a framework and algorithm that is a combination of other research efforts. The underlying algorithm is based on introducing decay function that is updated with feedback based trust calculation algorithm. At present we are enhancing and evaluating our framework by introducing a variety of reputation measurements that are controlled through adaptive parameters.

References

1. Foster I, Kesselman C, Tuecke S (2001) The anatomy of the grid: enabling scalable virtual organizations. Int J Supercomput 15(3):200–222
2. Marmol FG, Perez GM (2009) Security threats scenarios in trust and reputation models for distributed systems. Elsevier Comput Secur 28(7):545–556
3. Welch V, Siebenlist F, Foster I, Bresnahan J, Czajkowski K, Gawor J, Kesselman C, Meder S, Pearlman L, Tuecke S (2003) Security for grid services. In: Proceedings of the HPDC-12
4. Kamvar SD, Schlosser MT, Garcia-Molina H (2003) The EigenTrust algorithm for reputation management in P2P networks. In: Twelfth international world wide web conference, Budapest, Hungary, 20–24 May 2003, ACM Press
5. Alunkal BK (2003) Grid eigen trust: a framework for computing reputation in grids. Thesis Master of Science in Computer Science in the Graduate College of the Illinois Institute of Technology. http://www.iit.edu/∼alunbeu/thesis/thesis-final.pdf
6. Kerschbaum F, Haller J, Karabulut Y, Robinson P (2006) Pathtrust a trust-based reputation service for virtual organization formation. In: iTrust2006: Proceedings of the 4th international conference on trust management, volume 3986 of lecture notes in computer science, pp 193–205, Springer
7. Xiong L, Liu L (2004) PeerTrust: supporting reputation based trust to P2P E-communities. IEEE Trans Knowl Data Eng 16:843–857

8. Tian C, Zou S, Wang W, Cheng S (2006) An efficient attack-resistant trust model for P2P networks. IJCSNS 6(11):251–258
9. Ma B, Sun J, Yu C (2006) Reputation-based trust model in grid security system. J Commun Comput 3(8) (Serial No.21)
10. Papalilo E, Freisleben B (2008) Managing behaviour trust in grid computing environments. J Inf Assur Secur 3(1):27–38
11. Azzedin F, Maheswaran M (2002) Towards trust-aware resource management in grid computing systems. In: CCGRID, 2nd IEEE/ACM international symposium on cluster computing and the grid (CCGRID'02), p 452
12. Vijayakumar V, Wahida Banu RSD (2008) Security for resource selection in grid computing based on trust and reputation responsiveness. IJCSNS Int J Comput Sci Netw Secur 8(11):107–118

Chapter 48
Switch Line Fault Diagnosis in FPGA Interconnects Using Line Tracing Approach

Shilpa Dandoti and V. D. Mytri

Abstract This paper aims at fault detection and location in interconnect of a Field Programmable gate array. We discuss the testing of interconnect types like single lines, double lines, global interconnect, Long length lines, switching matrices, Buffer drivers, Quad lines and direct lines. The proposed testing scheme uses a test manager which defines a part of the chip as the pattern generator and the other half as response analyzer. The chip is reconfigured several times to cover all portions of interconnect. The outcome of each reconfiguration is a bit which provides a pass or fail result. Testing is done in two phases, phase one involves several reconfigurations intended to detect various faults in the interconnect structure. The test manager provides the required test sequence in each configuration. This phase involves extensively testing the complete interconnect structure for all possible faults namely, configurable interconnection points struck on, configurable interconnection points struck off, wire struck-at-1, wire struck-at-0, two adjacent wires short and wires open.

Keywords Fault diagnosis · FPGA interconnects · Switch fault · Line transition

S. Dandoti (✉)
JNTU, Hydrabad, India
e-mail: shilpashrigiri@gmail.com

V. D. Mytri
Shetty Institute of Engineering, Gulbarga, India
e-mail: vdmytri@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_48,

48.1 Introduction

An interconnect is usually analyzed as a set of nets with different faults,such as stuck at, bridge (short) and open. Detection is usually accomplished by comparing the signatures received at the output pins with the ones provided with the input pins, i.e. no internal probing. FPGA test can be substantially more complex than application-specified integrated circuit (ASIC) test,providing motivation for new efficient testing techniques. Information regarding defect location is particularly important in today's test environment since new techniques [1] have been developed that can reconfigure FPGAs to avoid faults. To operate effectively, these approaches require that the specific location of the fault be clearly identified. The reconfigurability of FPGAs plays an important role in reducing on-chip testing hardware relative to ASICs. While ASIC discrete Fourier Transform (DFT) approaches require the modification of circuit functionality to perform test, FPGA test hardware can be swapped out of the device once verification is complete. Reconfigurability does incur other test costs, including increased test generation complexity and increased test application time. Unlike ASICs, which require a single configuration for fault detection, FPGAs require multiple configurations to test an assortment of switch settings. In general, fault coverage is directly related to the number and scope of test configurations that are created. The fault coverage issue has been further complicated in recent years by the introduction of FPGA devices [2, 3] with millions of programmable switch points. This device capacity growth strongly suggests the need for a hierarchical and incremental approach to FPGA test and diagnosis. To support this need, contemporary FPGA devices now allow for rapid partial device reconfiguration [3, 4]. Research in FPGA testing has investigated a wide range of test architectures and techniques. The FPGA test problem has been divided by several researchers into the interconnect test problem [7–10] and FPGA logic test problem [11, 12].

48.2 Line Switch Faults

Data used in FPGA are granulated data which are requested a periodically on demand and consist of few bytes of discrete information. The previous coding schemes were devised for the applications with this kind of I/O patterns. However, a new pattern of data transmission has become a great concern with the widespread use of practical systems. Data are transferred like a stream when used in digital applications. Once an operation is started, it requires transmitting large amounts of data from a few kilobytes to hundreds of megabytes. As streaming becomes one of the major data transfer patterns, we have one more degree of freedom, i.e. the sequence of data that we can exploit to reduce the number of interconnect transitions faults during data transition is of greater importance. A new coding scheme called interconnect switching scheme (ISS) is proposed in this paper. It is different from previous transition-reduction coding schemes in that it is aimed at

applications with the stream-type data transfer pattern. ISS reduces the number of interconnect transitions by rearranging the transmission sequence of data. An algorithm called shuffle algorithm is presented to show the feasibility of ISS. This algorithm reduces around 10 % ofinterconnect transitions in transmission of the benchmark files. For the brevity of description, let us define some terms and notations first. A switched sequence is defined as the sequence of words transmitted according to an ISS algorithm. Let us express the hamming distance between a word, W and a interconnect, B, as H (W; B). Since I/O coding was proposed into reduce transient noises, there have been many efforts to reduce the dynamic fault of interconnects by coding which can minimize the number of interconnect transitions in transmission. BI coding is a general-purpose coding that is suitable for the transmission of uncorrelated data. Some variations of BI such as partial BI (PBI) coding and weight-based BI coding have also been developed by exploiting some prior knowledge on data. For instruction address interconnect and data address interconnects, addresses are highly correlated, localized, and even sequential. Gray code, T0 code, inc-xor, and Working zone encoding can be more efficient than BI for such interconnects.

The waveform of eight-bit interconnects when the eight data are transmitted by the (a) original order and (b) a different order. So far, most of the previous transition-reduction coding schemes have been developed for the granulated data, therefore, they have not considered the sequence of data as an important factor. ISS is the first general-purpose coding scheme that employs the sequence of data in reducing the number of interconnects transitions. ISS needs no prior information on data to be sent, and it can be applied to any application that transmits more than two data sequentially for most operations.

48.3 Switch Coding Algorithm

In this algorithm, only two most recent words are considered in switch operation. The encoder has a register called shuffle register to store a word temporarily. Any word that stayed in this register is tagged as a shuffle. At every cycle, two words, a shuffle and a new incoming word compete for transmission so that the winner is sent and the looser is held as the shuffle for next competition. The coding information is transmitted simultaneously with the word by an auxiliary line called S-line that is added in parallel to the interconnect. The S-line is set to 1 when a latter word is selected while it is reset to 0 whenever a shuffle is transmitted. For example, Fig. 48.1b can be obtained by applying US to Fig. 48.1a, and the S-line becomes 01111010. If a switched sequence is compared with the original sequence, all words displaced from the original positions are shuffles. Furthermore, the original sequence can be recovered from a switched sequence by relocating every shuffle to the position of its preceding shuffle; the first shuffle is placed at the beginning of the sequence. Using this property, decoding can be achieved with two first in–first out (FIFO) buffers and a shift register. The one buffer is reserved for

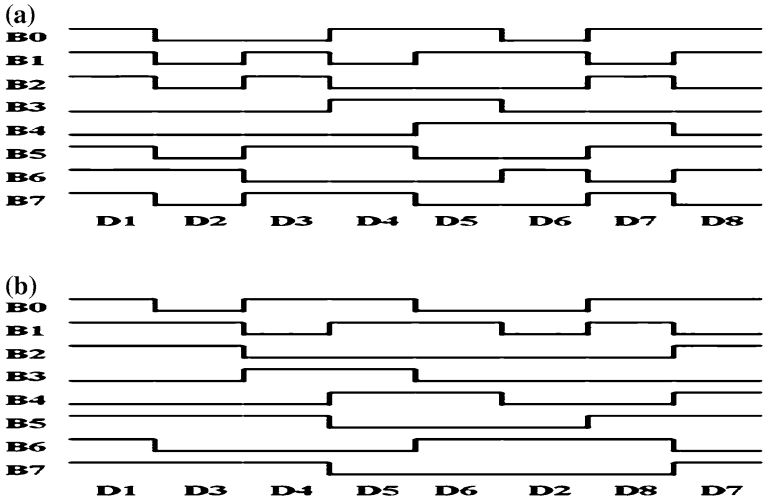
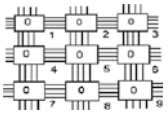


Fig. 48.1 Illustration of the effect of transmission sequence on interconnects transitions

Fig. 48.2 Structure of a 3 × 3 CLB



shuffles and the other is for no shuffles. The SR consists of n bits that are initially reset to 0. Assume that the decoder sends a decoded word at the first-half cycle while it receives a word from interconnects at the second-half cycle. For every cycle, the decoder receives a word with an S-value from interconnect, and stores the word in either of the buffers according to the S-value. SR is shifted from SR1 to SRn, and the S-value goes to SR1. For the first n-1 cycle, i.e., until the first S-value reaches to SRn_1, the decoder only stores the received words. From nth cycle, the decoder starts to send a word to the next stage (client). According to the value of SRn, the decoder fetches a word from either of the buffers and sends it to the client. The latency of the decoder becomes n-1 cycles. Let us define lagging distance of a shuffle as the number of its losses in the competitions, and the maximum lagging distance (l_{max}) of a switched sequence as the biggest lagging distance of all shuffles in the sequence. Then, n should be greater than or equal to $l_{max} + 2$, and the FIFOs should have at least $l_{max} + 1$ slot. The shortcoming of US is that l_{max} is not predictable in advance. Because it can be a very large number, the decoder should prepare a huge number of slots.

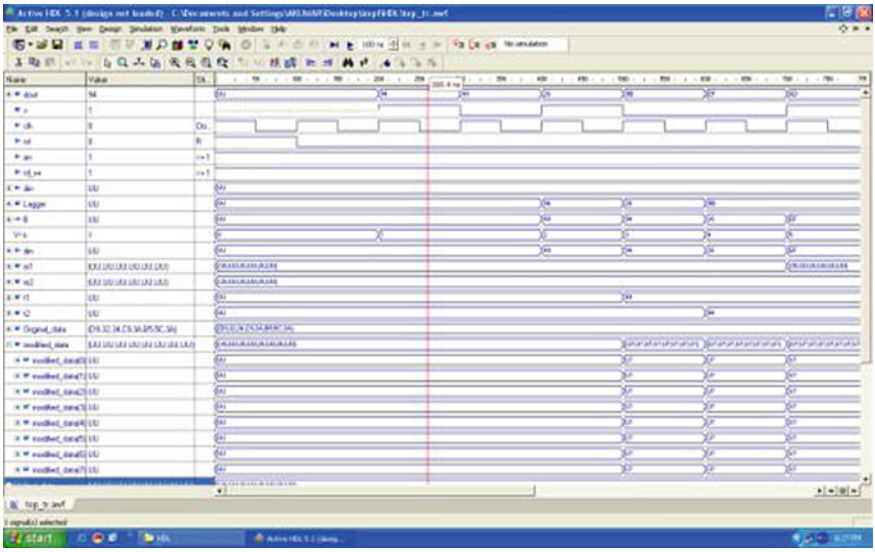


Fig. 48.3 Simulation plot showing the regenerated after coding observations made or the developed the Lagger algorithm

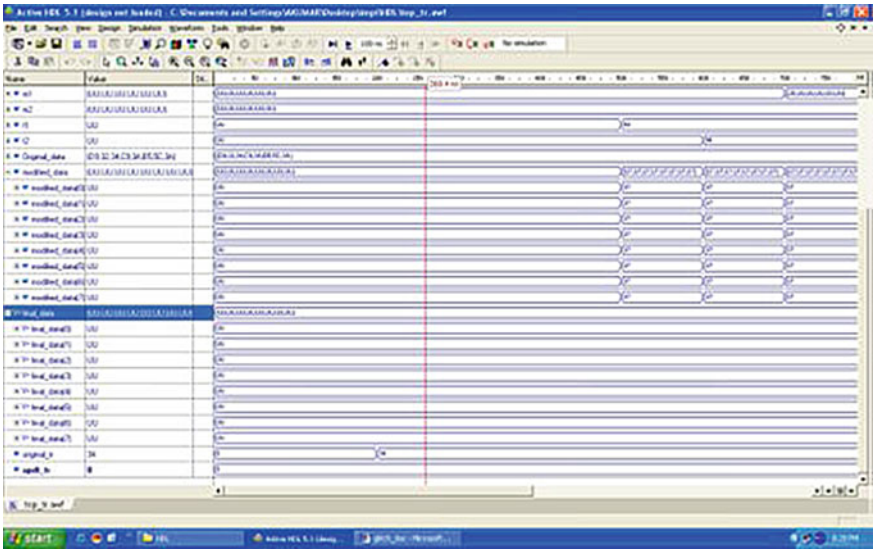


Fig. 48.4 Figure illustrating the original data

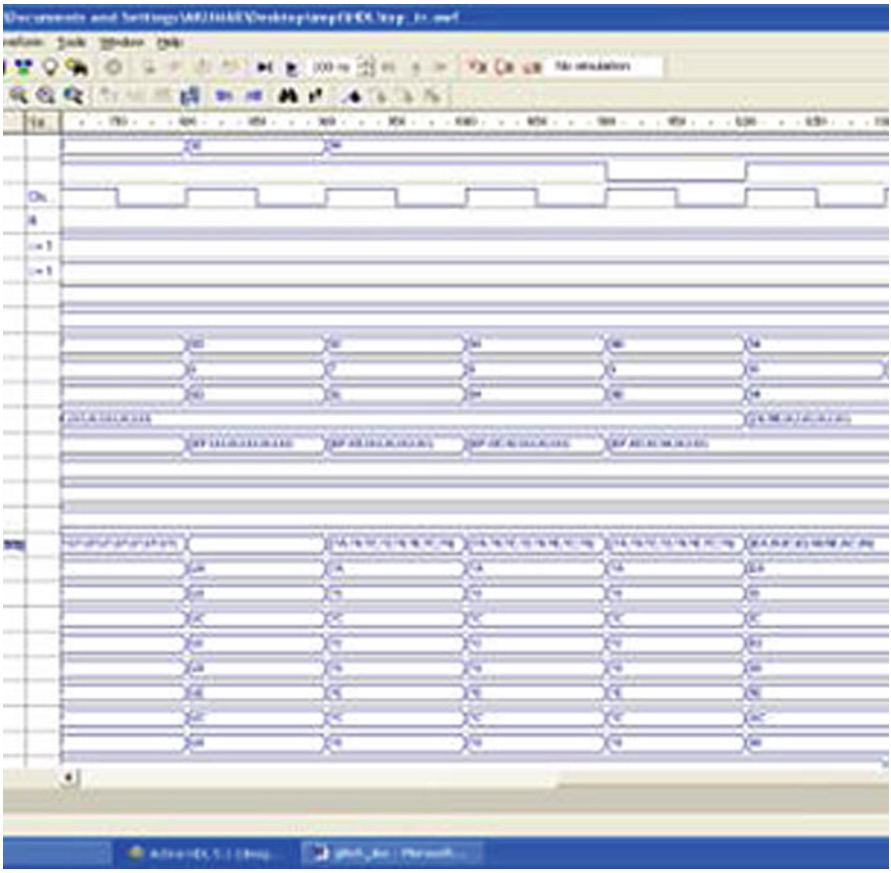


Fig. 48.5 Figure illustrating the bus, Lagger and input data line for the designed encoder and decoder unit

48.4 Fault Tracing Approach

The fault location in FPGAs has been carried out by phase path method. The location of the fault has been carried out in six different phases. As the target FPGA is a XILINX 4000E, which has a symmetrical architecture of CLB's we consider a 3 × 3 CLB structure as shown in figure below for the experimental purpose and this can be extended for a n X n structure. For the location of the fault in a 3 × 3 CLB structure as shown in the figure below (Fig. 48.2).

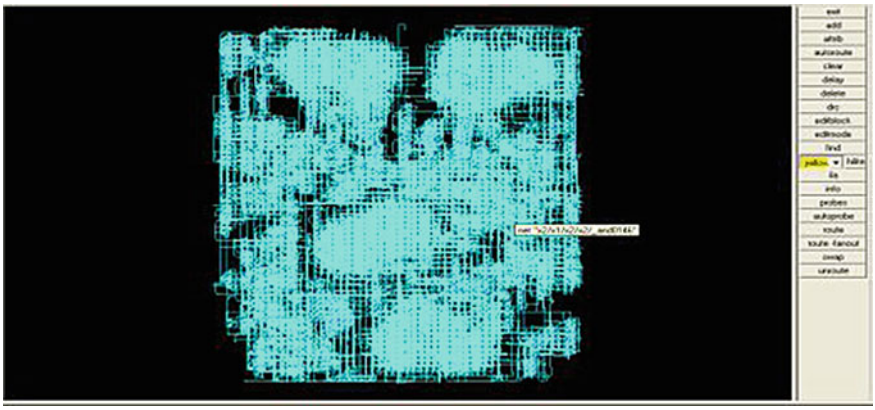


Fig. 48.6 Logical placement of the proposed system

48.5 Analysis of the Steps

If we closely observe all the six phases, we find that we consider all the 9 CLBs for test in each of the six phases with a different set of input output ports. By testing with all these 6 phases we can clearly pass through the faulty switch in all of the phases only once and its flag is incremented by 1 in all the phases. Thus the CLB or switch with the flag count of 6 (the highest possible value as we use 6 phases) is the faulty switch and this need only to be reprogrammed for getting the non faulty FPGA configuration. Thus we observe that six phases are sufficient to test and locate any fault occurred in either CLB or the Interconnects or the I/O lines in the targeted FPGA.

48.6 Simulation Result

The developed system is defined using VHDL definition language and the timing results and the implementation on a targeted FPGA device is observed as, (Figs. 48.3, 48.4, 48.5, 48.6).

Default fault location = ck1 (under asynchronous mode communication)
Testing selection method = round robin
Total number of communicating nodes = 4
Total amount of data generated per node = 3 bytes
Total amount of expected data in processing = 12 bytes
Total transition taken = 5580 (under non synchronous round robin based comm.)
(Total time = processing time + comm. Time

48.7 Conclusion

This paper is aimed to detect and locate the faults in FPGA interconnect during data transition. The testing has been done to achieve minimum number of reconfigurations and maximum fault location capabilities. The results from each reconfiguration are stored in the on chip RAM. By looking at this RAM lookup tables, users will be able to locate the fault. Various kinds of faults are simulated using VHDL programming. The testing phases are successfully tested upon these faults and the simulation results show that the proposed model works as per requirement. This model can be extended to locate multiple faults and the resolution can be made as high as needed. The number of reconfigurations in this case depend on the extend of resolution required, in data transition.

References

1. Lakamraju V, Tessier R (2000) Tolerating operational faults in cluster-based FPGAs. In: 8th international ACM/SIGDA symposium on field programmable gate arrays, New York, USA
2. Altera Web Site [Online] http://www.altera.com/
3. Xilinx Corporation (2000) Virtex data sheet
4. Atmel Corporation (1997) Configurable Logic Design and Application Book
5. Altera Corporation (2000) Altera apex data sheet
6. Betz V, Rose J, Marquardt A (1999) Architecture and CAD for Deep-Submicron FPGAs. Kluwer, New York
7. Renovell M, Portal JM, Figueras J, Zorian Y (1998) Testing the interconnect of RAM-based FPGAs. IEEE Des Test Comput 15(1):45–50
8. Renovell M, Zorian Y (2000) Different experiments in test generation for XILINX FPGAs. In: Proceedings of international test conference, Washington, USA, pp 854–862
9. Zhao L, Walker DMH, Lombardi F (1998) Bridging fault detection in FPGA interconnects using IDDQ. In: International symposium on field programmable gate array, New York, USA, pp 95–104
10. Stroud C, Wijesuriya S, Hamilton C, Abramovici M (1998) Built-in self-test of FPGA interconnect. In: The Proceedings of international test conference, Washington, USA, pp 404–441
11. Renovell M, Portal JM, Figueras J, Zorian Y (1998) "SRAM-based FPGAs: Testing the LUT/RAM modules. In: Proceedings of international test conference, Washington, USA, pp 1102–1111
12. Metra C, Pagano A, Ricco B (2001) On-line testing of transient and crosstalk faults affecting interconnections of FPGA-implemented systems. In: Proceedings of international test conference, New York, USA, pp 939–947
13. Naeimi H, DeHon A (2007) Fault secure encoder and decoder for memory applications. In: Proceedings of 22nd IEEE international symposium on defect fault tolerance VLSI systems, California, pp 409–417
14. Pomeranz I, Reddy SM (2010) On multiple bridging faults. In: 28th VLSI test symposium (VTS), pp 221–226
15. Fourlas G.K (2009) Multiple faults diagnosability of hybrid systems. In: 17th Mediterranean conference on control and automation, pp 365–370

Chapter 49
An Approach to Encryption Using Superior Fractal Sets

Charu Gupta and Manish Kumar

Abstract The voluminous digital data exchange between various computers has introduced large amount of security vulnerabilities. Encryption schemes have been increasingly studied to meet the demand for real-time secure transmission of data over the Internet and through wireless networks. In this paper, we try to study a new cryptographic key exchange protocol based on superior Mandelbrot and Superior Julia sets. In this study we analyze a cryptographic system utilizing fractal theories; this approach uses concept of public key cryptography by taking advantage of the connection of Superior Julia and Superior Mandelbrot sets. This paper exploits the main feature of public key security.

Keywords Superior Mandelbrot and Julia sets · Fractal geometry · Public key

49.1 Introduction

Enhanced security will definitely be a great relief for paranoid people. RSA cipher is most commonly used for public-key encryption depends on the difficulty of factoring large numbers [20]. RSA uses a variable size encryption block and a variable size key [17]. This study proposes a new fractal (based on Superior Mandelbrot and Superior Julia sets) encryption key protocol which secures

C. Gupta (✉)
Computer Science Department, NIMS University, Jaipur, India
e-mail: charuvek@yahoo.co.in

M. Kumar
Sri Ram Murti Smarak International Business School, Kanpur Road, Lucknow, India
e-mail: dr.manish.2000@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_49,

transmission of data between computers. The working of the proposed scheme depends on the strong interconnection between two Superior Julia and Superior Mandelbrot sets which generates the corresponding public and private keys. In general, a security protocol uses public-key cryptosystem to exchange the secret key between communicating nodes and then uses secret-key algorithms with the agreed secret key as the password to ensure confidentiality on the data transferred (Branovic et al. 2003) [1].

Fractal is a geometric pattern that cannot be represented by classical geometry as it is a geometric figure that repeats itself under several levels of magnification; a shape that appears irregular at all scales of length, e.g. a fern. This geometric figure, built up from a simple shape, by generating the same or similar changes on successively smaller scales; it shows **self-similarity**… The backbone of the fractal is iteration method i.e. feedback system.

The Superior Mandelbrot Fractal sets and Superior Julia sets is the set of points on a complex plane. The Fractal image can be *generated* by applying Eq. 49.1 recursively [13].

$$xn = \beta n\, f(xn-1) + (1-\beta n)xn - 1,\ n = 1, 2, \cdots, \tag{49.1}$$

where $0 \leq \beta n \leq 1$ and $\{\beta n\}$ is convergent to β away from 0.

The difference between the Mandelbrot set and the Julia set is that the Mandelbrot set iterates $Z^2_{n-1} + c$ with Z starting at 0 and varying c with every iteration, while Julia set iterates $Z^2_{n-1} + c$ for fixed c and starting with non-zero value of Z [6, 7]. All points, Z_n, must reside on the Mandelbrot set or the Julia set, respectively. In our work, we are depending on the intrinsic connection between both of the Superior Mandelbrot and the Julia Fractal sets. The connection between the Mandelbrot set and the Julia set is that each point c in the Mandelbrot set specifies the geometric structure of the corresponding Julia set.

49.2 Review of Literature

Recently Rani and Kumar [12] have given the concept of Superior Julia set and Superior Mandelbrot set [11, 12] in the study of discrete dynamical system. They can be used effectively to play a significant role in cryptographic security.

49.2.1 Superior Mandelbrot and Superior Julia Sets

Rani and Kumar [12]. The Superior Mandelbrot set (*SM* set) for the polynomial $Qm,c(z) := z_m + c$, where $m > 1$ is a positive integer, is defined as the collection of values of c for which the superior orbit of $z = 0$ does not escape to infinity. The Mandelbrot sets for $Qm,c(z)$ are SM sets with $\beta = 1$. The escape criterion plays a

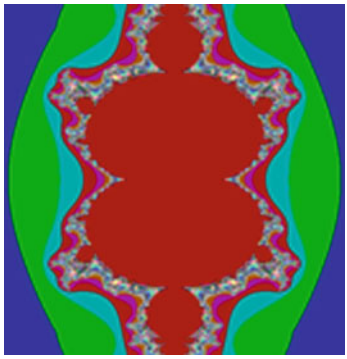


Fig. 49.1 Superior Mandelbrot sets for z = z3 + c

vital role in the generation and analysis of Mandelbrot sets and its variants. Let *X* be a metric space of complex numbers, *D* be nonempty convex subset of *Z* and *T* be a self map of *D*, let z_0 ε *D*. The Mann Iteration is defined by:

$$z_{n+1} = (1 - \omega)z_n + \omega.\theta_c(z) \tag{49.2}$$

Where and $\theta_c(z)$ can be a quadratic, cubic, or biquadratic polynomial. $0 \leq \omega \leq 1$ n ≥ 0. When we apply Mann Iteration on Mandelbrot with complex polynomial equation z_{n+c}, is called Superior Mandelbrot set. The iteration of all values for which the point escapes from the unit circle. The set of all those points is known as escape set.

This paper extends their following result, which gives a general escape criterion for the polynomial *Qm,c(z)*, $m \geq 2$ (Fig. 49.1).

The Superior Julia sets is the set of points whose orbit are bounded under the superior iteration of the function $Q_{a,,b}(z)$, where $\{\beta_n\}$ converges to a non-zero number between 0 & 1. The general superior escape criterion is obtained for the function $Q_{a,b}(z) = z^3 + a\,z + b$ by iterating it in SO and define prisoner set using cubic superior escape criterion.is given by max where $0 < \beta n <= 1$, n = 1,2,…….m > 1 & a positive integer and c is a complex parameter.

49.3 Superior Mandelbrot and Superior Julia Sets Used for Key Generation in RSA

In the proposed protocol sender and receiver will agree and use a public domain value, c. The receiver and sender generates their private key in the form of (e,n) and (k,d) respectively. As these keys are generated both sender and receiver will use their private key values and the value of c as input to the superior Mandelbrot function to produce the public keys z_nd and z_ke. And then both sender and receiver must exchange the public keys. Using key distribution authorities sender will obtain receiver's public key, z_nd and uses this value together with her private key and the plaintext, as inputs to the Superior Julia sets to produce cipher text V

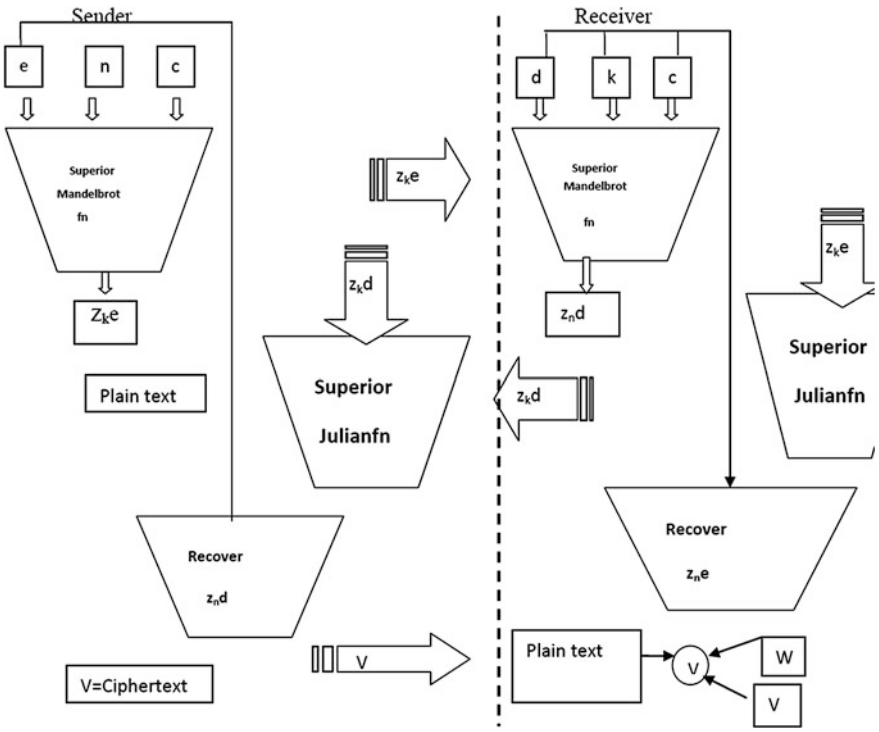


Fig. 49.2 Proposed algoritham

which will be sent to Bob. To retrieve the original text receiver must have sender's public key, $z_k e$ and the cipher text V which will then be used as input values together with his own private key to Superior Julian function for deciphering V as shown in Fig. 49.2 below.

References

1. Menzes A, Oorschot PV, Vanstone S (1996) Handbook of applied cryptography. CRC Press, Boca Raton, pp 425–488
2. Ms A (2007) Public key cryptography: applications algorithms and mathematical explanations. Tata Elxsi, India
3. Preneel B, Mercierlaan K (2004) Cryptanalysis of message authentication codes. Department Electrical Engineering, Katholieke Universiteit Leuven, Belgium
4. Damgard A (1989) Design principle for hash functions. In: Brassard G (ed) Advances in cryptology-crypto 89 proceedings. Lecture notes in computer science, vol 435. Springer-Verlag, Berlin
5. Yang F (2005) Cryptanalysis on an algorithm for efficient digital signatures. In: Cryptology ePrint archive 2005/456

6. Dobbertin H (1996) The Status of MD5 after a recent attack. In: RSA Labs' CryptoBytes, vol 2(2), Summer
7. El-Bolok H, El-Mageed TA, El-Salam NA, Elgtawal IA (2003) Public key cryptosystems and its applications in digital signature. Helwan University, Faculty of Engineering
8. Dugelay JL, Polidori E, Roche S (1996) Iterated function systems for still image processing. In: IWISP-96, Manchester, UK, 2005
9. Barnsley M (1993) Fractals everywhere, 2nd edn. Academic Press Professional Inc., San Diego, ISBN: 10: 0120790610, pp 550
10. Barnsley M, Demko S (1985) Iterated function systems and the global construction of fractals. Proc R Soc Lond 399:243–275. http://adsabs.harvard.edu/abs/1985RSPSA.399..243B
11. Rani M (2002) Iterative procedures in fractals and chaos. Ph D thesis, Department of computer science, Faculty of technology, Gurukula Kangri Vishvavidyalaya, Hardwar
12. Rani M, Kumar V (2004) Superior Mandelbrot set. J Koreans Soc Math Educ Ser D 8(4):279–291
13. Koblitz N (1994) A course in number theory and cryptography, 2nd edn. Springer, Berlin, ISBN: 0387942939, p 235
14. Goldreich O, Goldwasser S, Micali S (1986) How to construct random functions. J ACM 33(4):210–217
15. Atkinson R (1995) Security architecture for the internet protocol. IETF Network Working Group, RFC 1825
16. Atkinson R (1995) IP authentication header. IETF Network Working Group, RFC 1826
17. Rivest R, Shamir A, Adleman L (1978) A Method for Obtaining Digital Signatures and Public-Key Cryptosystems. Commun ACM 21(2):120–126
18. ElGamal T (1985) A public key cryptosystem and a signature scheme based on discrete logarithms. In: IEEE Trans Inf Theory IT 31(4):469–472
19. Diffie W, Hellman M (1976) New directions in cryptography. In: IEEE Trans Inf Theory It-22(6):644–654
20. Stalling W (2004) Cryptography and network security. PHI, New Delhi
21. Hou YC, Tu SF (2005) A visual cryptographic technique for chromatic images using multi-pixel encoding method. J Res Pract Inf Technol 37(2):179–192
22. Fisher Y (1995) Fractal image compression: theory and application. Springer-Verlag, New York, ISBN: 0-387-94211-4, pp 341

Chapter 50
Shape Based Image Retrieval Using Gradient Operators and Block Truncation Coding

Padmashree Desai and Jagadeesh Pujari

Abstract The need of Content Based Image Retrieval (CBIR) arises because of digital era. It is very much required in the field of radiology to find the similar diagnostic images, in advertising to find the relevant stock, for cataloging in the field of geology, art and fashion. In CBIR, the set of image database is stored in terms of features where feature of an image can be calculated based on different criteria like shape, color, texture and spatial locations etc. Among three features shape is the prominent feature and helps to identify the image correctly. In this paper, we are proposing Shape Based Image Retrieval (SBIR) to retrieve shape features extracted using gradient operators and Block Truncation Coding (BTC). BTC improves the edge maps obtained using gradient masks like Robert, Sobel, Prewitt and Canny. The proposed image retrieval techniques are tested on generic image database with 1000 images spread across 10 categories. The average precision and recall of all queries are computed and considered for performance analysis. Among all the considered gradient operators for shape extraction "shape mask with BTC CBIR techniques" give better results. The performance ranking of the masks for proposed image retrieval methods can be listed as Canny (best performance), Prewitt, Sobel and lastly the Robert.

Keywords CBIR · BTC · Shape · Canny · Prewitt · Sobel · Robert

P. Desai (✉)
CSE Department, BVBCET, Hubli, Karnataka, India
e-mail: padmashri@bvb.edu

J. Pujari
SDM College of Engg & Tech, Dharwad, Karnataka, India
e-mail: jaggudp@yahoo.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_50,

50.1 Introduction

The last few years, there is an emerging need to organize and efficiently use large pools of images that have been collected over the last decades and contain information potentially useful to areas such as medicine, journalism, weather prediction, environmental sciences, art, fashion and industry. It is estimated that there are more than 20 million pages containing hundreds of millions of images on world wide web pages alone [1]. Traditionally, images were retrieved by their filename, other technical characteristics such as date and size or through text keywords, in the case of manually annotated images. Manual annotation, except for being a time consuming and not real-time process, can describe only a very small percentage of the information that an image contains. Recently, there is an increasing effort to organize and retrieve images by content based on characteristics such as color, texture, and shape. A number of methods in the literature [1–6] perform indexing and retrieval based on global image characteristics such as color, texture, layout, or their combinations. Color feature of the image is extracted by computing a color histogram for each image that identifies the proportion of pixels within an image holding specific values [3]. Texture measures visual patterns in images and how they are spatially defined. This is measured by relative brightness of pairs of pixels such that degree of contrast, regularity, coarseness and directionality. Shape of the image is obtained either by segmentation or edge detection to an image [2]. Edge detection can be done using many gradient operators like Sobel, Roberts, Prewitt and Canny [6]. Here the paper discusses shape extraction using edge detection with improvement achieved by applying BTC on the edge image obtained using gradient operators. And also the shape feature is calculated using moment invariants which are invariant to translation, rotation and scaling. The results depict the improvements in the edge map and also the performance of retrieval of images.

50.2 Proposed Method

Proposed method use different gradient operators to get the edges of the image. To eliminate the discontinuities in the edges, slope magnitude is applied to the gradient of the images. Then BTC is applied on obtained shape mask to obtain clear boundary of the image. Using canny, Sobel, Prewitt and Robert shape masks with BTC four edge maps are considered. Moment invariants [5] are applied for the mask obtained and shape feature vector is constructed. Similarity measure is done using Canberra distance formula. Flow of proposed method is shown in Fig. 50.1.

We can define an edge as a strong intensity contrast or a jump in an image's intensity within a limited spatial range, i.e. intensity changes between one image pixel to the next. Edge detection [5] is a fundamental tool used in most image processing applications to obtain information from the frames as a precursor step

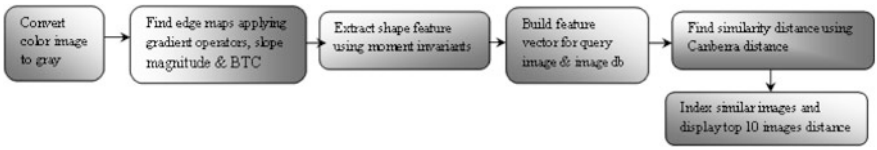


Fig. 50.1 Frame work for proposed work

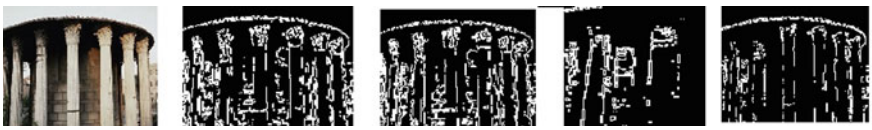


Fig. 50.2 Query image edge maps using Canny, Sobel, Prewitt and Robert

to shape feature extraction and object segmentation. Proposed method use gradient operators such as Sobel, Prewitt, Roberts and Canny for edge detection. All these gradient-based algorithms have kernel operators that calculate the strength of the slope in directions which are orthogonal to each other, commonly vertical and horizontal. Later, the contributions of the different components of the slopes are combined to give the total value of the edge.

50.2.1 Block Truncation Coding

Block Truncation Coding, or BTC [6], is a type of lossy image compression technique for gray scale images. It divides the original images into blocks and then uses a quantiser to reduce the number of gray levels in each block whilst maintaining the same mean and standard deviation. Shape of an image is calculated using moment invariants [5]. Canberra distance formula is used for calculating the distance and is given by equation 1.

$$CD_k = \sum_{i=1}^{n} \frac{|x_i - y_{ik}|}{|x_i| + |y_{ik}|} \tag{1}$$

Where CD is the Canberra distance, x and y are the feature vectors of query and database images respectively, n is the length of the feature vector. And $k = 1$ to m, where m is the number of images in image database. Images are indexed based on the distance between the query image and images in the database. Similar images are displayed in the ranking order. Edge maps obtained using different gradient operators after applying slope magnitude and BTC are shown in the following Fig. 50.2.

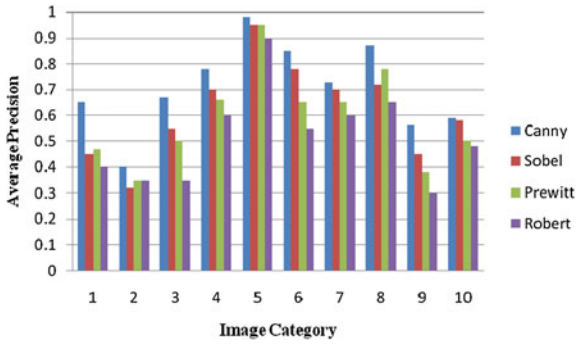


Fig. 50.3 Weighted average precision graph

50.3 Results and Discussions

For the purpose of experimentation, an image database Wang's data set [7] having established ground truth is used. A set of 1000 images assorted into 10 categories with 100 images in each category, forms the dataset. The images are of size either 384 × 256 or 256 × 384. The 10 image categories available are: (i) Dinosaurs (ii) Buses (iii) Buildings (iv) Horses (v) Beaches (vi) Elephants (vii) Flowers (viii) Africa (ix) Food and (x) Mountains. The images in each category are numbered (Category 1: 0 to 99, Category 2: 100 to 199 etc.…).Ten query images from each category are used to check the performance efficiency. Mat lab Ver. 7.0 is used to implement the system.

To assess the retrieval effectiveness, the precision and recall are used as statistical comparison parameters. Recall measures the ability of retrieving all relevant or similar items in the database. It is defined as the ratio between the number of relevant or perceptually similar items retrieved and the total relevant items in the database. Precession measures the retrieval accuracy and is defined as the ratio between the number of relevant or perceptually similar items and the total number of items retrieved. Weighted average precision is also evaluated for the 10 query images among all categories. Retrieved results for top 10 images using different operators are shown in the weighted average precision bar graph Fig. 50.3. This shows that for different categories across different gradient operator canny with BTC technique gives better result.

50.4 Conclusion

CBIR, the problem of searching huge image repositories according to their content, has been the subject of vital amount of research in the last decade. The aim of a CBIR algorithm is to predict the concerned images in a database that are relevant to an arbitrary query. The key contribution of this is to improve the performance of gradient operators and slope magnitude by applying BTC and invariant moments to extract shape invariant to translation, rotation and scaling of an image. Proposed methods achieve the good performance results.

References

1. Datta R, Joshi D, Li J, Wang JZ (2008) Image retrieval: ideas, influences, and trends of the new age. ACM Comput Surv 40(2, Article 5):1–60
2. Wu Y, Wu Y (2009) Shape-based image retrieval using combining global and local shape features. In: IEEE 2nd international conference on image and signal processing CISP 09
3. Ha J-Y, Kim G-Y, Choi H-I (2008) The content-based image retrieval method using multiple features. In: IEEE fourth international conference on networked computing and advanced information management, NCM 2008, vol. 1, pp. 652–657
4. Desai P, Pujari J, Parvitikar S (2011) Image retrieval using shape feature: a study. In: ACEEE CIIT 2011, CCIS 250, pp. 817–821,© Springer
5. Desai P, Pujari J, Goudar RH (2012) Image retrieval using wavelet based shape features. (JISC) J Inform Syst Commun ISSN: 0976-8742, E-ISSN: 0976-8750, 3(1):77–79
6. Kekre HB, Sudeep T, Mukherjee P, Miti K (2010) Image retrieval with shape features extracted using gradient operators and slope magnitude technique with BTC. Int J Comput Appl (0975-8887), 6(8):28–23
7. http://wang.ist.psu.edu/.

Chapter 51
Performance Evaluation of TCP Congestion Control Variants Using Dynamic State Routing In Wireless Ad-hoc Network

Mayank Kumar Goyal, Yatendra Kumar Verma, Paras Bassi and Paurush Kumar Misra

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. TCP/IP protocol plays an important role in developing communication systems and providing better and reliable communication capabilities in almost all networking environment. The paper aims to investigate the performance of the TCP congestion control variants in MANET and its behaviour with respect to Dynamic State Routing. TCP optimization in MANETs is a challenging issue because of some unique characteristics of congestion control. The results presented in this paper clearly indicate that the Dynamic State Routing achieves maximum throughput, higher packet delivery ratio and less average end to end delay when TCP congestion control agent used is TCP Vegas.

Keywords DSR · NS2 · TCP · Tahoe · Reno · NewReno · Vegas · HTTP & MANETS

M. K. Goyal (✉) · Y. K. Verma · P. Bassi · P. K. Misra
Deptt Of CSE/IT, JIIT University, Noida, Uttar Pradesh, India
e-mail: mayankrkgit@gmail.com
URL: http://www.jiit.ac.in

Y. K. Verma
e-mail: yatendra54@gmail.com
URL: http://www.jiit.ac.in

P. Bassi
e-mail: paras123@gmail.com
URL: http://www.jiit.ac.in

P. K. Misra
e-mail: misrapaurush@gmail.com
URL: http://www.jiit.ac.in

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_51,

51.1 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]. Routing is the process of moving a data packet from source to destination. Routing is usually performed by a dedicated device called a router [2, 3]. The routers are free to move randomly and organize themselves arbitrarily thus, the network's wireless topology may change rapidly and unpredictably [4].

This paper is organized as follows. Section 51.2 describes TCP Congestion control mechanism. In Sect. 51.3, Performance metrics and Sect. 51.4 describes the Results.

51.2 TCP Congestion Control Mechanism

TCP is the reliable connection orientated transport layer protocol that provides reliable transfer of data between the nodes. It ensures that the data is delivered to the destination correctly without any loss or damage [2]. The data is transmitted in the form of continuous stream of octets. The most famous implementation of TCP called Tahoe, Reno, New-Reno & Vegas [5]. Following are the different versions of TCP congestion control.

Tahoe: TCP Tahoe is the first TCP variant that incorporates congestion control mechanisms. Indeed its implementation added a number of new algorithms and refinements to earlier implementations.

Reno: TCP Reno adds some intelligence to TCP Tahoe so that the packets which are lost are detected earlier and the pipeline is not vacant every time a packet is lost.

New-Reno: TCP NewReno is able to detect multiple packet losses and thus is much more effective than RENO in the case of multiple packet losses.

Vegas: TCP Vegas detects congestion at an incipient stage based on increasing Round-Trip Time (RTT) values of the packets in the connection unlike other flavours like Reno, NewReno, etc.

51.3 Performance Metrics

Performance metrics is the basic route map to analyze the performance of the routing protocol. Some attributes used are End-to-End Delay, Packet Delivery Fraction, Number of Packets dropped and Throughput.

Table 51.1 Simulation results for 25 nodes

Performance metric	Tahoe	Reno	Newreno	Vegas
Start time(s)	10	10	10	10
Stop time(s)	60	60	60	60
Generated packets	3229	3229	3229	4987
Received packets	394	394	394	409
Packet delivery ratio	12.201	12.201	12.201	14.201
Total dropped packets	15	15	15	6
Avg.end-to-end delay(ms)	0	0	0	0
Avg.throughput(kbps)	1641.1	1641.1	1641.1	2854.3

Fig. 51.1 Xgraph for TCP Tahoe using DSR

51.4 Results

Simulation results are taken for 25 nodes and are shown in Table 51.1. Table 51.1 represents the values for performance metrics considered above for Dynamic Source Routing when implemented with different versions of TCP congestion control mechanism.

Figures 51.1, 51.2 represents identical results for DSR implementation with TCP Tahoe & TCP Reno.

Figures 51.3,51.4 does not represent identical results for DSR implementation. Table 51.1 clearly indicates that DSR implementation with TCP Vegas achieved maximum throughput, higher packet delivery ratio and less average end to end delay.

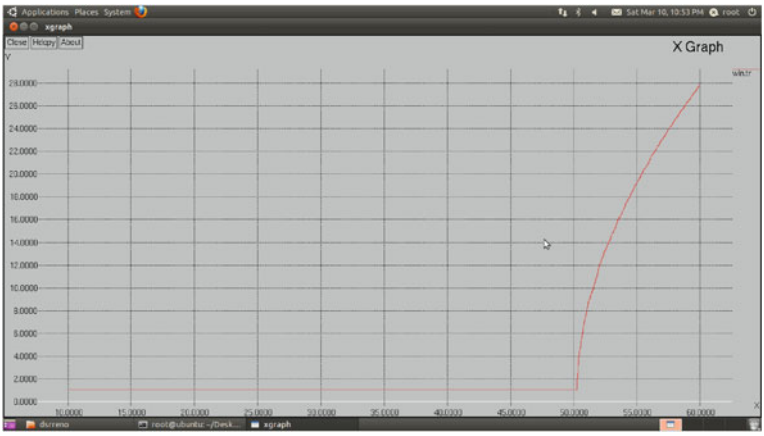


Fig. 51.2 Xgraph for TCP Reno using DSR

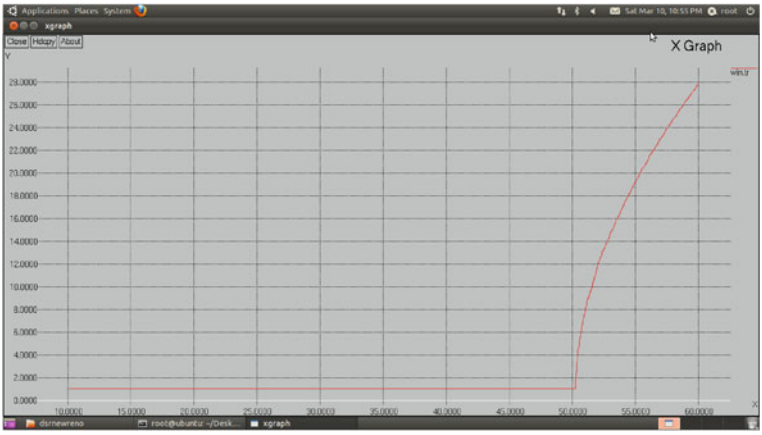


Fig. 51.3 Xgraph for TCP NewReno using DSR

51.5 Conclusion

The implementation of TCP directly impacts the TCP throughput, packet delivery ratio, average end to end delay and the communication efficiency of networks. The average bandwidth available to different versions of TCP congestion control mechanism are not the same. But for this network simulation environment, we prefer to adopt TCP Vegas because the average bandwidth is less than other TCP versions. TCP-Vegas experienced an inaccuracy problem of Base RTT due to frequent path change caused by node mobility causing performance degradation. Therefore, some modification in order to estimate an exact Base RTT over a new

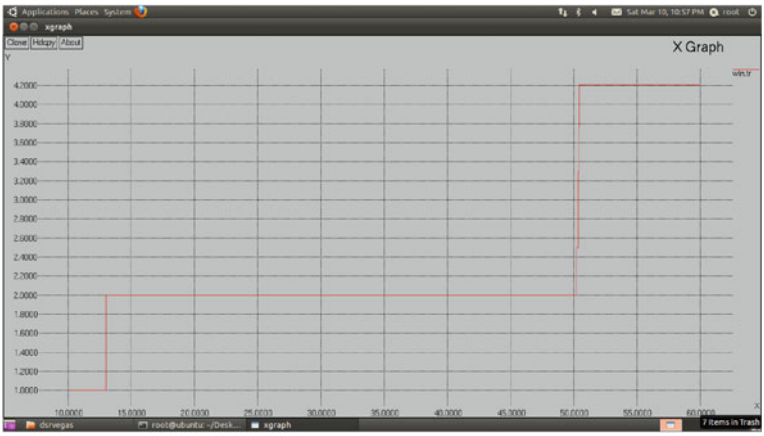


Fig. 51.4 Xgraph for TCP Vegas using DSR

path is required to improve the TCP performance for MANET. Dynamic state routing achieved its best performance metrics i.e. maximum throughput, higher packet delivery ratio and less average end to end delay when TCP congestion control agent used was TCP Vegas. This is due to fine tuning of congestion window size by taking into consideration the RTT of a packet, whereas other reactive protocols like TCP Tahoe, RENO and NewReno continue to increase their window size until packet loss is detected. Therefore, TCP algorithm must study constantly to reliable and adaptive TCP.

References

1. Abdullah A, Ramly N, Abdullah M, Derahman MN (2008) Performance compariso,n study of routing protocols for mobile grid environment. IJCSNS Int J Comput Sci Netw Secur 8(2):82–88
2. Johnson DB, Maltz DA, Broch J (2001) DSR: the dynamic source routing protocol for multi-hop wireless ad hoc networks
3. Johnson D (2003) Dynamic source routing for mobile ad hoc networks. IEFT MANET Draft, April 2003
4. Ch. Routing Protocols for Ad-HOC Wireless Networks, pp. 299 {364 Prentice Hall Communications Engineering and Emerging Technologies Series, New Jersey: Prentice Hall Professional Technical Reference, 2004.
5. Altman E, Jimenez T (2003) NS simulator for beginners. Lecture notes. Univ.de Los Andes, Merida, Venezuela and ESSI. Sophia-Antipolis, France
6. Chakeres ID, Belding-Royer EM (2004) AODV routing protocol implementation design. ICDCS Workshops 2004, pp 698–703
7. Staub T (2004) 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
8. Alexander Z (2003) Performance evaluation of AODV routing protocol: real-life measurements, SCC, June 2003

Chapter 52
Security Based Requirements Engineering for E-Voting System

P. Salini and S. Kanmani

Abstract The election process is in need of secured electronic system that voters can rely on and have trust. Unfortunately, a recent study revealed that various E-Voting Systems show serious specification, design, and implementation flaws. When system is being built, tasks such as Security Requirements Elicitation, Specification and Validation are essential to assure the Quality of the resulting secure E-Voting System. In this paper we propose, to adopt Security Requirements Engineering in the early phases of E-Voting System development and consider the Security requirements as functional requirements. This helps in standardizing the Security Requirements for secure E-Voting System with completeness.

Keywords: E-Voting · Functional Requirements · Threats · Security issues · Security Requirements

52.1 Introduction

Manual voting systems have been deployed for many years with enormous success. If those systems were to be replaced with Electronic Voting Systems, we have to be absolutely sure that they will perform at-least as efficient as the

P. Salini (✉)
Department of Computer Science and Engineering, Pondicherry Engineering College, Puducherry, Puducherry, 605014, India
e-mail: salini@pec.edu

S. Kanmani
Department of Information Technology, Pondicherry Engineering College, Puducherry, Puducherry 605014, India
e-mail: kanmani@pec.edu

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_52,

traditional voting systems without any security issues. Failures or flaws in E-Voting Systems will put at risk to Democracy in the country implementing them. The main reason for this being the worst-case scenario is really catastrophic. The aim of electronic voting schemes is to provide a set of protocols that allow voters to cast ballots while a group of authorities collect votes and output the final tally. The Problems with voting machines extend from the quality of the locks, to the need for a printed audit trail, to the hacking of the communication links.

The importance of Securing E-Voting System is that, E-Voting web application is the only thing standing in the way of an attacker and sensitive information. An attacker may be able to View or manipulate sensitive information, obtain unauthorized access to E-Voting System and able to take control of the whole application. So an E-Voting System should consider the following minimum requirements:

- To ensure eligible voters only able to cast a vote
- To ensure that every vote casted is counted
- To maintain the voter's right and to express his or her opinion in a free manner, without any influence
- To protect the secrecy of the vote at all stages of the voting process
- To guarantee accessibility and availability of the system to as many voters as possible
- To increase voter confidence by maximizing the transparency of information

The reality of E-Voting Security is that the security done at application system is not fully secure. To secure E-Voting web application, means it involves security at three layers: the network layer, host layer, and the application layer. You have to secure infrastructure and environment where the E-Voting System is put on. The reliance on firewall and host defenses are not sufficient when used in isolation. Moreover the requirements must be clear, comprehensive, consistent and unambiguous. This statement has significance for Security Requirements (SR) and if you say application must be secure, it is not Security Requirements. It is hard to construct secure E-Voting System or to make statements about security unless we know what to secure, against whom and at what extent. So the solution to the problems is Security Requirements engineering (SRE), a phase that comes before design and programming, will play a more important role that determines the success of applications. SRE is moving to the forefront of gaining increased significance in software engineering for building secure service oriented applications. In this paper in Sect. 52.2 we find the Security Requirements for E-Voting System, in Sect. 52.4 the results are discussed and Sect. 5 concludes with future works.

52.2 Security Requirements Engineering for E-Voting System

The Security Requirements Engineering is the process of eliciting, specifying, and analyzing the SR for system. In this paper, we will discuss how to develop E-Voting System with secure, robust, accurate and quality. There are many SRE

methods [1] which consider SR as constraints on functional requirements [2] or non functional requirements [3–6] and that can be used to elicit and model the SR for E-Voting System. By considering the SR as functional requirements in the Requirement phases, the SR and domain knowledge for E-Voting System can be captured in a well-defined model. SR integration with the artifacts of other phases can be cost effectively improved and can effect a significant reduction of the problems encountered in the E-Voting System due to poor SRE and Management.

In Our method of SRE, first find the objective of E-Voting System and since the development is based on the multilateral view of the stakeholders find people from voters' community, candidate, security experts, election officers, government representatives, developers and requirements engineering team. Next step is to identify the business assets like voters and candidate's details, votes, voter's credentials, voter's secret etc. The brainstorming technique can be used for elicitation of requirements for E-Voting System. Then draw a rough architecture diagram with high level of abstraction and elicit the Non-Security goals and business requirements. Some of the requirements for the E-Voting System are: Record the selection of individual vote choices for each contest, Before the ballot is cast the voter is allowed to review his choices and, if he desires, to delete or change his choices before the ballot is cast, Prevent the voter from over-voting and Check for total votes for each candidates.

The non-security requirements are categorized as essential and non essential requirements and prioritized according to the Stakeholders preference. Next, for better understanding use case modeling [7] of the applications is developed. Then security goals/security objectives, threats and vulnerabilities can be identified with respect to assets, business goals and organizational principles.The following are some of the identified threats and vulnerabilities for the system and with them we will able to find the security requirements for the system.

Threats

- Password Cracking of users of E-Voting System
- Network eavesdropping between browser and Web server to capture voters credentials
- SQL injection, to execute commands and access or modify data like vote
- Cross-site scripting (XSS) where an attacker injects script code
- Information disclosure secret to whom the voter voted
- Unauthorized access to the election database
- Discovery of encryption keys used to encrypt sensitive data in the database
- Unauthorized access to Web server resources and static files of E-Voting

Vulnerabilities

- Weak or blank passwords, Passwords that contain everyday words
- Lack of password complexity enforcement
- Missing or weak input validation at the server
- Failure to validate cookie input
- Failure to encode output leading to potential cross-site scripting issues

The impact of threats and vulnerabilities are analysed. Risk Assessment is done by using Microsoft method [8] of risk analysis for E-Voting System. The threats and vulnerabilities can be categorized with respect to the security goals and security policies of the organization and prioritized based on the level of security and assets to be secured. The detailed set of misuse case diagram [9] of the applications should be developed that encompass the most significant threats to the system e.g. tamper misuse case, unauthorized users misuse case.

The SR is the counter measure that the application should have, as the functional requirements. Some of SR for E-Voting System identified based on business and system assets are: Election process should not be subject to any manipulation, system should provide accurate time and date settings, and system should not allow improper actions by voters and election officials should not allow voter submissions to be observed or recorded in any way that is traceable to the individual voter, should not allow tampering with audit logs, Use secure authentication, secure communication channels, remote procedure call encryption and Firewall policies that block all traffic except expected communication ports. With SR the Use Cases Diagram, UML Diagrams for the E-Voting System can be generated with high level of abstraction. This process is repeated for iteration based on the level of security needed and we can reach low level requirements which can be used in the design phase.

52.3 Discussions

In the previous section we have identified the list of some Security Requirements and they are based on the business and system assets by applying our SRE method. In this paper we have identified the need for systematically eliciting and producing a complete set of requirements specification. We found that using SRE methods, we will be able to get better set of security requirements for secure E-Voting System and reduce the burden of the developers to develop a vulnerability free system.

52.4 Conclusion and Future Work

Security Requirements have to be considered as functional requirements in the early phase of Requirements Engineering (RE), and SRE Methods can be used to elicit requirements for a secure E-Voting System. It is important to identify potential security aspects in early phase of development for a secure E-Voting System and solve problems in practice through secure design.

As a future work the Security Requirements identified from RE Phase should be carried to Design phase because good design will give Vulnerability free E-Voting. We also intent to do penetration testing and find the results based how far our application is vulnerable.

References

1. Fabian B, Gurses S, Heisel M, Santen T, Schmidt H (2009) A comparison of security requirements engineering methods. Requir Eng: Secur Requir Eng 15:7–40, Springer-Verlag London Limited 2009
2. Haley CB, Laney R, Moffett JD, Nuseibeh B (2008) Security requirements engineering: a framework for representation and analysis. IEEE Trans Softw Eng 34(1):133–152
3. Mead R, Houg ED, Stehney TR (2005) Security quality requirements engineering (square) methodology. Technical report CMU/SEI-2005-TR-009, Software Engineering Institute, Carnegie Mellon University 2005.
4. Wang H, Jia Z, Shen Z (2009) Research in security requirements engineering process. In: IEEE 16th international conference on Industrial Engineering and Engineering management, 2009, pp 1285–1288
5. Apvrille A, Pourzandi M (2005) Secure software development by example. IEEE Secur Priv 3(4):10–17
6. Graham D (2006) Introduction to the CLASP process. Build Security In, 2006. https://buildsecurityin.us-cert.gov/daisy/bsi/articles/best-practices/requirements/ 548.html.
7. Jacobson I (1995) Modeling with use cases: formalizing use case modelling. J Object-Oriented Program8 (2):18–24
8. Meier JD, Mackman A, Dunner M, Vasireddy S Escamilla R, Murukan A (2003) In: Improving web application security :threats and countermeasures. Microsoft Corporation, Published, June 2003
9. Sindre G, Opdah AL (2004) Eliciting security requirements with misuse cases. Requir Eng 10:34–44, Springer-Verlag London Limited 2004

Chapter 53
Analysis of 3 Dimensional Object Watermarking Techniques

Deepika Khare, Sumita Verma, Ravindra Gupta and Gajendra Singh Chandel

Abstract Due to the explosive growth of Internet and the development of digital content designing and processing techniques, many valuable materials can be represented in digital forms for exhibition and access via Internet. Due to the characteristics of easy duplication and modification of digital contents, it is necessary to develop a variety of watermarking techniques for various protection purposes such as ownership claiming and authentication. In this survey paper, we examine 3D model watermarking technologies developed over the last decade. We classify various algorithms into two classes: robust watermarking and fragile watermarking.

Keywords 3D watermarking · 3D mesh · Digital signature · Copyright protection · Information security

53.1 Introduction

In recent years, 3D graphic models have become more accessible to general end users due to the usage of advanced scanning devices and the virtual-reality modeling language (VRML) for graphic description. Moreover, due to the

D. Khare (✉) · S. Verma · R. Gupta · G. S. Chandel
SSSIST, Sehore, Madhya Pradesh, India
e-mail: deepika.united@gmail.com

S. Verma
e-mail: avantika08@gmail.com

R. Gupta
e-mail: ravindra_p84@rediffmail.com

G. S. Chandel
e-mail: hod.cseit@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_53,

explosive growth of Internet and the development of digital content designing and processing techniques, many valuable materials can be represented in digital forms for exhibition and access via Internet.

Due to the characteristics of easy duplication and modification of digital contents, it is necessary to develop a variety of digital signature or watermarking techniques for various protection purposes such as model authentication and ownership claiming. Digital signatures [1, 2] are designed for the receiver of electronic documents to verify the identity of the sender and to check the originality of the documents. The watermarking schemes are usually designed for the sender to check the copyright ownership (robust watermarking) or for the receiver to verify the authentication of the received media (fragile watermarking). The main difference between digital signature and watermarking techniques is that the former attaches a small piece of information (the digital signature) transmitted with the original documents whereas the latter embeds invisible information (the watermarks) in the original media. Encryption techniques [3, 4] can be symmetric or asymmetric [5].

Instead of using digital signature for all kinds of electronic documents, researchers develop various watermarking schemes for various multimedia data types such as audio, images, video, and 3D models. Watermarks can be invisibly/inaudibly embedded in these media by altering some of their lower significant bits. Watermarking schemes usually don't need any complex computation thus they can be performed in a fast and low cost way comparing to the digital signature schemes. According to the application purposes, watermarking techniques can be classified into robust and fragile schemes. Robust watermarking is usually designed for ownership claiming while fragile watermarking is used for digital content authentication and verification. The design goal of robust watermarking is to make the embedded watermarks remain detectable after being attacked. In contrast, the requirements of fragile watermarking are to detect the slightest unauthorized modifications and locate the changed regions.

53.2 Robust Watermarking for 3D Models

Ohbuchi et al. [6] proposed three requirements for 3D robust watermark embedding: unobtrusive, robust, and space efficient. Unobtrusive means the embedding must not interfere with the intended use of a model, such as viewing. Robust means the embedded watermarks should remain detectable after being maliciously attacked. Space efficient means an embedding method should be able to embed sufficient amount of information into models.

We here introducing a triangle similarity quadruple (TSQ) embedding scheme proposed in [6] as an illustration. The TSQ algorithm uses a dimensionless quantity pair such as {b/a, h/c} in Fig. 53.1 to define a set of similar triangles. The algorithm uses a quadruple of adjacent triangles that share edges in the configuration depicted in Fig. 53.2 as a Macro-Embedding-Primitive (MEP). Each MEP stores a quadruple of values {Marker, Subscript, $Data_1$, $Data_2$}. A marker is a

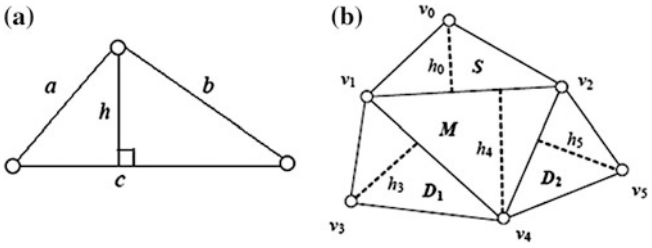


Fig. 53.1 **a** Example of dimensions quantities that define a set of similar triangles. **b** A macro-embedding-primitive

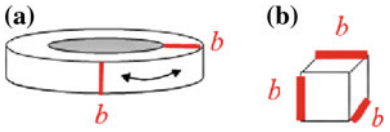


Fig. 53.2 The binary state space for a vertex: **a** the state space formed by cylindrical parameterization; **b** the state space formed by the conversion function for computing value indices

special value (in this case {b/a, h/c}) that identifies MEPs. In Fig. 53.2, the triangle marked M stores a Marker, S stores a Subscript, and D_1 and D_2 store data values $Data_1$ and $Data_2$. While each MEP is formed by topology, a set of MEPs are arranged by the quantity of subscript. The TSQ algorithm embeds a message according to the following steps:

1. Traverse the input triangular mesh to find a set of four triangles to be used as an MEP.
2. Embed the marker value in the center triangle of the MEP by slightly changing the coordinates of vertices $_{v1}$, $_{v2}$, and $_{v4}$ such that the quantity pair $\{_{e14}/_{e24}, {}_{h4}/_{e12}\}$ = {b/a, h/c} (refer to Fig. 53.2) where $_{eij}$ represents the edge between $_{vi}$ and $_{vj}$.
3. Embed a subscript and two data symbols in the remaining three triangles of the MEP by slightly changing the coordinates of vertices v0, v3, and v5.
4. Subscript is embedded in the pair $\{e_{02}/e_{01}, h_0/e_{12}\}$, and two data symbols are embedded in the pairs $\{e_{13}/e_{34}, h_3/e_{14}\}$ and $\{e_{45}/e_{25}, h_5/e_{24}\}$. For each of the three triangles, the algorithm first modifies the ratio h_i/e_{ij} by changing only h_i, and then modifies the ratio e_{ij}/e_{kl} while keeping the height hi constant.
5. Repeat (1) to (3) until all the data symbols of the message are embedded.

The TSQ extraction algorithm is a public scheme since it does not require the original 3D model for extraction. The quantity pair {b/a, h/c} is the key to identify marker triangles. Watermarks embedded by the TSQ algorithm remain detectable against translation, rotation, and uniform-scaling transformations of the marked 3D models. However, the watermarks will be destroyed by randomization of coordinates, or by topological alteration such as re-meshing, smoothing, and simplification operations.

Table 53.1 A comparison of fragile watermarking and digital signature applied in 3D models

Functions	Digital signature	Fragile watermaking
Verify the ID of the sender	Yes	No
Check the integrity of the documents	Yes	Yes
Locate the changed regions	No	Yes
Implementation cost	High	Low
Computation speed	Slow	Fast
Application fields	General	Different methods for different media

53.3 Fragile Watermarking for 3D Models

Fragile watermarking techniques for still images have been widely studied and investigated in recent years [7, 8]. On the other hand, fragile watermarking for 3D models got relatively less notice. There are two major functions in 3D fragile watermarking: integrity checking and changed region locating. Moreover, a good fragile watermarking scheme should be invariant to translation, rotation, and uniformly scaling operations. Possible applications of public fragile watermarking include demonstrating a digital material having not been changed (or having been changed) in an official situation (e.g., in a court) and to confirm the received digital material having not been changed at the receiver end. The functions of the 3D fragile watermarking scheme are similar to that of digital signature. A comparison of these two schemes is summarized in Table 53.1.

The authentication scheme for fragile watermarking consists two modules: computing location indices and computing value indices.

53.3.1 Computing Location Indices

Step 1: Given a vertex coordinate v, the specified parameterization S : R3 → R2 converts the vertex coordinate into a parameter coordinate by cylindrical parameterization [9]. A cylindrical parameterization process can be expressed as:

$$S_{(m,n)}(v) \rightarrow (\alpha, \beta) = \left(\sqrt{\|v-m\|^2 - (n \cdot (v-m))^2}, n \cdot (v-m) \right) \tag{53.1}$$

where (α, β) is the coordinate in the parameter domain, m is 3-D point and n is its orientation. The range for each dimension of the parameter domain is $\alpha \in (0, \infty)$ and $\beta \in (-\infty, \infty)$, respectively.

Step 2: Convert the parameter coordinate formed in Step 1 into the so-called bin coordinate, i.e., the location index (Lx, Ly). This conversion can be accomplished by quantizing the parameter domain. Assume that the size of a 2-dimensional watermark pattern is WM_X_SIZE × WM_Y_SIZE, the quantization formula for a cylindrical parameterization domain is as follows:

$$L = (L_x, L_y) = \left(\left\lfloor \frac{\alpha}{b} \right\rfloor \%WM\_X\_SIZE, \left\lfloor \frac{\beta}{b} \right\rfloor \%WM\_Y\_SIZE \right), \quad (53.2)$$

where b is the quantization step for ordinary numeric values and % represents a modulus operator. A very important feature of the above design is that the quantized parameterization domain and the watermark pattern together form a binary state space. Such a state space is helpful for defining a legal domain of alternation for a given vertex. The state space corresponding to the cylindrical parameterization is illustrated in Fig. 53.2a.

53.3.2 Computing Value Indices

Even though any functions for converting a floating-point number into an integer can be used to calculate value indices, the following conversion function was designed since it is able to form a binary state space. Assuming that the size of each look-up table is LUT SIZE, the conversion function is formulated as

$$p = (p1, p2, p3) = \left(\left\lfloor \frac{v_x}{b} \right\rfloor \%LUT\_SIZE, \left\lfloor \frac{v_y}{b} \right\rfloor \%LUT\_SIZE, \left\lfloor \frac{v_z}{b} \right\rfloor \%LUT\_SIZE \right), \quad (53.3)$$

where b is the same quantization step as used to compute location indices. As we have already mentioned, the quantization step b can be hard-coded into the implementation process. In addition, Fig. 53.2 reveals that the domain of acceptable alternation for a given vertex can be defined as the intersection of the binary state spaces where the values of both hash functions applied to that vertex match each other.

53.4 Performance Evaluation

For robust watermarking schemes, the embedded watermarks should resist against various malicious attacks. The possible attacks include cropping, smoothing, simplification, noising, re-meshing, vertex re-ordering, translation, rotation, and scaling (uniformly or non-uniformly) operations. A good robust watermarking scheme should resist against as many attacks as possible. For fragile watermarking schemes, the extraction algorithm should be a public scheme, and it should detect and locate the changed regions. Both robust and fragile watermarking schemes control the distortion caused by the watermark embedding. Only a few researches discussed about the topic of distortion control. In 3D models, the distance between two surfaces X and Y can be defined by L2 measurement,

$$d(X,Y) = \sqrt{\frac{1}{area(X)} \int\limits_{x \in X} d(x,Y)^2 dx} \tag{53.4}$$

where d(x, Y) is the Euclidean distance from a point x on X to the closest point on Y. We modify the definition of L2 measurement to d(M, M') for representing the average distortion of all vertices in a model,

$$d(M,M^*) = \frac{1}{|M|} \sum_{i=1}^{|M|} \left| v_i - v_i^* \right| \tag{53.5}$$

where M and M' are the original and marked models, respectively. The L_2 measurement should be treated as one measurement of performance evaluation criteria for 3D model watermarking algorithms.

53.5 Conclusion

In this paper, we performed a survey on current 3D model watermarking techniques by classifying major algorithms into classes, describing main ideas behind each algorithm, and comparing their strength and weakness. The major function of robust watermarking is for ownership claiming. The design goal of robust watermarking is to make the embedded watermarks remain detectable after being attacked. There three requirements for 3D robust watermark embedding: unobtrusive, robust, and space efficient. Unobtrusive means the embedding must not interfere with the intended use of a model. Robust means the embedded watermarks should remain detectable after being maliciously attacked. Space efficient means an embedding method should be able to embed sufficient amount of information into models. The major function of fragile watermarking is for digital content authentication and verification. The design goal of fragile watermarking is to detect the slightest unauthorized modifications and locate the changed regions. There are two major functions in 3D fragile watermarking: integrity checking and changed region locating.

References

1. Maurer U (2004) New approaches to digital evidence. Proc IEEE 92(6):933–947
2. Wells TO (2000) Electronic and digital signatures: in search of a standard. IT Prof 2(3):24–30
3. Beth T, Gollman D (1989) Algorithm engineering for public key algorithms. IEEE J Select Areas Commun 7(4):458–466
4. Diffie W, Hellman ME (1979) Privacy and authentication: an introduction to cryptography. Proc IEEE 67(3):397–427

5. Diffie W, Hellman ME (1976) New directions in cryptography. IEEE Trans Inform Theory 22(6):644–654
6. Cayre F, Macq B (2003) Data hiding on 3-D triangle meshes. IEEE Trans Signal Process 51(4):939–949
7. Praun E, Hoppe H, Finkelstein A (1999) Robust mesh watermarking. In: ACM SIGGRAPH, Los Angeles, CA, Aug 8–13, 1999, pp 49–56
8. Hoppe H (1996) Progressive meshes. In: ACM SIGGRAPH, New Orleans, LA, Aug 4–9, 1996, pp. 99–108
9. Johnson AE, Hebert M (1999) Using spin-images for efficient multiple model recognition in cluttered 3-D scenes. IEEE Trans Pattern Anal Mach Intell 21:433–449
10. Xie L, Arce GR (2001) A class of authentication digital watermarks for secure multimedia communication. IEEE Trans Image Process 10(11):1754–1764
11. Barni M, Bartolini F, Piva A (2001) Improved wavelet-based watermarking through pixel-wise masking. IEEE Trans Image Process 10(5):783–791
12. Ohbuchi R, Masuda H, Aono M (1998) Data embedding algorithms for geometrical and non-geometrical targets in three-dimensional polygonal models. Comput Commun 21(15):1344–1354
13. Ohbuchi R, Masuda H, Aono M (1998) Watermarking three-dimensional polygonal models through geometric and topological modifications. IEEE J Select Areas Commun 16(4):551–560
14. Ohbuchi R, Takahashi S, Miyazawa T, Mukaiyama A (2001) Watermarking 3D polygonal meshes in the mesh spectral domain. In: Proceedings of the graphics interface, Ottawa, Ontario, June 7–9, 2001, pp 9–17
15. Ohbuchi R, Mukaiyama A, Takahashi S (2002) A frequency-domain approach to watermarking 3D shapes. Comput Graph Forum 21(3):373–382
16. Benedens O (1999) Geometry-based watermarking of 3-D models. IEEE Comput Graph Appl 19(1):46–55

Chapter 54
Graph Based Approach for Heart Disease Prediction

M. A. Jabbar, B. L. Deekshatulu and Priti Chandra

Abstract The diagnosis of Disease is a significant and tedious task in Medicine. The detection of heart disease from various factors or symptoms is a multilayered issue which is not free from false presumptions often accompanied by unpredictable effects. Thus the attempt to exploit knowledge and experience of several specialists and clinical screening of data of patients collected in data bases to facilitate the diagnosis process is considered a good option. The health care industry collects huge amounts of health care data which unfortunately are not mined, to discover hidden information for effective decision making. Discovery of hidden patterns and relationships often goes unexploited. Advanced data mining techniques can help remedy this situation. Weighted association rule mining is the most useful data mining technique. Weighted association rules are association rules with weights or strength of presence. As data mining techniques are being introduced and widely applied to nontraditional item sets, existing approaches for finding frequent item sets were out of data as they cannot satisfy the requirement of these domains. Hence, an alternative method of modeling the objects in the said data set is graph. Graph based algorithms efficiently solve the problem of mining association rules. In this paper we propose an efficient algorithm which integrates weighted association rule mining and graph based approach for heart disease prediction for Andhra Pradesh population.

Keywords Andhra Pradesh · Heart disease · Maximum weighted clique · Weighted association rule mining

M. A. Jabbar (✉)
JNTU, Hyderabad, India

B. L. Deekshatulu
IDRBT, RBI Govt of India, Hyderabad, India

P. Chandra
Advanced System Laboratory, Hyderabad, India

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_54,

54.1 Introduction

India is undergoing rapid epidemiological transition as a consequence of economic and social change, and cardiovascular disease is becoming an increasingly important cause of death. India's disease pattern has undergone a major shift over the past decade. As per WHO report, at present out of every 10 deaths in India, eight are caused by non communicable diseases such as cardiovascular diseases, and diabetes in urban India. In rural India, six out of every 10 deaths is caused by NCD's [1].

Data from registrar general of India shows that heart attacks are major cause of death in India. Studies to determine the precise cause of death in rural areas of Andhra Pradesh have revealed that cardio vascular disease cause about 30 %in rural areas [2].Hospitals and clinics accumulate a huge amount of patient data over the years. This data provide a basis for analysis of risk factors for many diseases. we can predict the level of heart attack to find patterns associated with heart disease. Data mining is used to discover interesting patterns in medical data.

Data mining is a technology that blends traditional data analysis methods with sophisticated algorithms for processing large volumes of data. Data mining is also known as knowledge discovery in data bases is the process of automatically discovering useful information in large data repositories [3].Association rule mining, one of the most and well researched techniques of data mining was first introduced by Agrawal et al. [4].It aims to extract interesting correlations, frequent patterns, associations among sets of items in transactional data bases or other data repositories. Weighted association rules are association rules with weights or strength of presence. The weights associated with the items signify the importance of items. The items are given different weights in the transaction data bases. The main focus in weighted frequent item sets mining concerns satisfying the downward closure property. The downward closure property is usually broken when different weights are applied to the items according to their significance [5].

Graph mining is the task of finding novel, useful and understandable graph theoretic patterns in graph representation of Data[6].Graph based algorithms can efficiently solve the problem of mining association rules.

In this paper we integrated the concept of weighted association rule mining and graph based approach for heart disease prediction. We applied our new method on heart disease patients of Andhra Pradesh population. The paper is organized as follows. In Sect. 54.2 a brief review of some of the works on heart disease diagnosis is presented. An introduction about heart disease and its effect is given in Sect. 54.3. Heart disease data sets is presented in Sect. 54.4. Proposed method is explained in Sect. 54.5. Section 54.6 deals with results and discussion. We conclude our Remarks in Sect. 54.7.

54.2 Related Work

Numerous works in literature related with heart disease diagnosis using data mining techniques have motivated our work. A model intelligent heart disease prediction system (IHDPS) was proposed by Palaniappan [7]. IHDPS was capable of answering queries that the conventional decision support system were not able to. The problem of identifying constrained association rules for heart disease prediction was studied by Carlos Ordonez [8].Three constraints are introduced to decrease the no. of patterns. Their experimental result shows that constraints reduced the no. of discovered rules remarkably besides decreasing the running time.

Enhanced prediction of heart disease with feature subset selection was proposed by Anbarasi et al. [9].the objective of their work is to predict more accurately the presence of heart disease with reduced no. of attributes.

Patil et al., proposed an efficient approach for the extraction of significant patterns from heart disease ware house for heart attack prediction [10].

Jabbar et al., proposed cluster based association rule mining for heart attack prediction [11].Their method is based on digit sequence and clustering. The entire data base is divided into partitions of equal size, and from each cluster heart disease patterns are mined. Their approach reduces main memory requirement since it considers only a small cluster at a time and it is scalable and efficient.

54.3 Basic Concepts

In this section we will present basic concepts of association rule mining, weighted association Rule mining, graph concepts, and heart disease.

54.3.1 Association Rule Mining

Association rule finds interesting associations and/or correlations among large sets of data items. Association rule shows attribute value conditions that occur frequently together in a given data set. A typical and widely used example of association rule mining is market basket analysis. The problem of association rule mining can be described as below. If $I = \{I_1, I_2, \ldots I_n\}$ is the set of items. Suppose D is data base transaction set and each transaction T contains set of items, such that $T \subseteq I$. Each transaction has identifier called as TID. Suppose A is a set of items and transaction T is said to contain A only if $A \subseteq T$. association rule is an implication like $A \Rightarrow B$ in which A, $B \subset I$ and $A \cap B \neq \emptyset$. There are 2 important basic measures for association rules, support and confidence.

$$\text{Support } (A = > B) = P\ (AUB) \tag{54.1}$$

$$\text{Confidence } (A = > B) = \text{Support } (AUB)/\text{support } (A) \tag{54.2}$$

Association rule mining has been decomposed to the following 2-step process. (1) Finding all frequent item sets (2) generate strong association rules from the frequent item sets. These rules must satisfy minimum support and minimum confidence. The overall performance of mining association rules is determined by first step.

54.3.2 Weighted Association Rule Mining

weight of an item is a non negative real number which is assigned to reflect the importance of each item in the transaction data base [12] for a set of items $I = \{I_1, I_2, \ldots I_n\}$,weight of a pattern $P\{x_1, x_2, \ldots x_n\}$is given as follows

$$Weight(p) = \sum_{q=1}^{length(p)} weight(xq)/length(p) \tag{54.3}$$

A weight support of a pattern is defined as a resultant value of multiplying the pattern support with the weight of pattern. So the weighted support of a pattern P is given as

$$\text{Wsupport } (p) = \text{weight } (P) \times \text{support } (P) \tag{54.4}$$

An item set X is called a large weighted item set if the weighted support of the item set X is greater than or equal to the weighted support threshold. Support(X) $\geq$ Wminsupport

The motivations behind introducing weights are listed below.

1. For unweighted case, whether a rule is interesting or not depends on the count of the item sets in a data base. Weights can represent the knowledge for the items.
2. Weights can provides the users with a convenient way to indicate the importance of the attributes, and obtain more interesting rules
3. Weights can be adjusted according to the human experts. Different knowledge can generate different rules.
4. Suppose we separate the supports and weights, we can find item sets having both sufficient support and weights. However this may ignore some interesting knowledge.
5. Suppose we separate the supports and weights and find the item sets with sufficient support or weights. However it may not allow us to handle the mining process efficiently.

6. Multiplication can provide us a balance between the weight and support. Weights can be "Adjusting factor of support". The multiplication is the easiest way to do this. It will dramatically lower the weighted support, if the weight is very low, and it will keep the original support value if the weight is 1

54.3.3 Graph Concepts

A graph G consists of a pair (V(G),X(G)) where V(G)is a non empty finite set whose elements are called points or vertices and X(G)is a set of unordered pairs of distinct elements of V(G).The elements of X(G) are called lines or edges of the graph G.A graph H = (v1, x1) is called a sub graph of G = (V, X) if vi⊆V and x1⊆X. A sub graph G1 is said to be complete if there is an arc for each pair of vertices. A complete sub graph is also called a clique. A clique is maximal if it is not contained in any other clique. In the maximum clique problem the objective is to find the largest complete sub graph in a graph. If a weight is given to each edge, the sub graph is known as a weighted sub graph, and the weight of the sub graph is given by the sum of the weights in the edges [13].

$$\text{Clique weight} = \sum \text{weight } (i,\ j),\ \forall \text{ edge } (i,\ j)\ \in \text{ clique} \tag{54.5}$$

A clique can represent a common interest group. If a graph with weights in the edges is used, the highest weighted clique corresponds to the common interest group whose elements communicate the most among themselves. The maximum clique problem is an important problem in combinatorial optimization with many applications like market analysis, project selection and signal transmission [14].

54.3.4 Weighted Cosine Measure

In this subsection we introduce a new measure weighted cosine measure for finding weighted association rules. The cosine similarity, also known as uncentered Pearson correlation, has been widely used for mining association patterns, which contain objects strongly related to each other. As a widely used measure, the cosine similarity is suitable for high –dimensional item vectors, has the important null variant property [15].cosine measure often produces high quality results across different domains. In the field o association analysis, cosine similarity is defined as an interesting measure to a specific item set. For a 2 item set X = {i_2, i_1}, the cosine measure of X is defined as

$$\text{Cos(X)} = \text{Support } (I_2,\ I_1)/\sqrt{(\text{support } (i_2) * \text{support } (i_1))} \tag{54.6}$$

For k-item set $X = \{i_k, \ldots i_2, i1\}$ cosine measure is defined as

$$\text{Cos}(X) = \text{support } (X)/K \sqrt{\pi \text{ support } (I_k)} \tag{54.7}$$

Range of cosine measure is [0, 1] if the resulting value of the above equation is between 0 and 1, and then the item sets are correlated to each other.

$$\text{Weighted cosine} = \text{weighted support}(A \cap B)/\text{SQRT } (\text{Support } (A) * \text{WSupport}(B)) \tag{54.8}$$

$$\text{Where Wsupport} = \text{Weight of item} * \text{support count} \tag{54.9}$$

According to Piatetsky and Shapiro [16] a good measure should satisfy the three key properties. Consider the interesting rule A =>B to be evaluated by a measure M, the three properties are

(1) M = 0 if X and Y are statistically independent
(2) M Monotonically Increase with P (AUB) when P (A) and P(B) remain the same.
(3) M Monotonically Decrease with P (A) or P(B) when the rest of parameters P (AUB) and P(B) or P (A) remain unchanged. Our new measure satisfies the above three properties.

54.4 Heart Disease

Heart disease is a type of cardiovascular disease. In addition to heart disease, the term cardiovascular disease encompasses a variety of heart conditions, such as high blood pressure and stroke. Coronary heart disease (CHD) is caused by a narrowing of the coronary arteries, which results in a decreased supply of blood and oxygen to the heart.CHD includes myocardial infarction, commonly referred to as a heart attack, and angina pectoris, or chest pain. A heart attack is caused by the sudden blockage of coronary artery, usually by a blood clot, and chest pain occurs when the heart muscle does not receive enough food. Another type of heart disease is a heart rhythm disorder, which includes rapid heart murmurs, and other unspecified disorders. Congestive heart failure (CHF) which is often the end stage of heart disease is another disease of the heart [17].In our research we mined association among various attributes which leads to heart disease

Table 54.1 Cardiology patient data

Sl.no	Attribute name	Comments	Numeric value
1	Age	Age in years	–
2	Sex	Patient gender	0 or 1
3	Chest pain type	Angina, abnormal angina, no tang, asymptomatic	1-4
4	Blood pressure	Resting blood pressure upon hospital admission	–
5	Cholesterol	Serum cholesterol	–
6	Fasting blood sugar	Is fasting blood sugar less than 120	0,1
7	Resting ECG	Normal, abnormal, hyp	0,1,2
8	Maximum heart rate	Maximum heart rate achieved	–
9	Induced angina	Does the patient experience angina as a result of exercise	0,1
10	Old peak	ST depression induced by exercise relative to rest	–
11	Slope	Up, flat, down	1–3
12	CA	Number of major vessels colored by fluoroscopy	0,1,2,3
13	Thal	Normal, fixed defect, reversible defect	3,6,7

54.5 Methodology

This section describes our proposed algorithm in detail. The algorithm has two parts.

(1) First part calculates weighted cosine values of each edge and constructs the weighted graph.
(2) In the Second part algorithm generates maximum cliques from the constructed weighted graph and builds the relation among the attributes which leads to heart disease.

We have taken 13 attributes [18] and assigned weights to each attribute which will are shown in Tables 54.1, 54.2

54.5.1 Algorithm Description

STEP 1) Assign weights to each attribute $0 < W_a < 1$
STEP 2) Wsup (a) = (W_a)* support count
STEP 3) Generate candidate 2-item sets
STEP 4) Compute Wcosine of 2-item sets

$$\text{Wcosine (A, B)} = \text{Wsup (A} \cap \text{B)}/\sqrt{\text{Wsup (A)*Wsup (B)}}$$

STEP 5) If Wcosine (item set) < minimum weighted cosine, Prune corresponding Item set

Table 54.2 Medical data attributes and their corresponding weights

Attribute and corresponding weight
1) Age > 40 = 0.3, Age 41–65 = 0.7, Age > 65 = 0.8
2) Male = 0.3, Female = 0.8
3) Chest pain type <=3 = 0.5, Chest pain type >=3 = 0.3
4) BP > 120 = 0.8
5) Cholesterol > 240 = 0.9
6) Fbs = True = 0.5, Fbs = False = 0.3
7) Resting ECG > 2 = 0.3
8) Maximum heart rate = 0.9
9) Induced angina = yes 0.8, Induced angina = no = 0.3
10) Old peak > 0 = 0.5
11) Slope >=2 = 0.3
12) CA <=3 = 0.3 CA > 3 = 0.3
13)Thal <=3 = 0.3, Thal > 3=0.5

STEP 6) Obtain the adjacency matrix of the weighted graph (V, E) and generate the
Graph using 2-item set weighted cosine
STEP 7) Using recursive DFS, find maximum K-cliques in a graph. Where k = floor (1/2)*(log n/log2) and n = |v|, that corresponds to association rules to Predict Heart disease.

54.6 Results and Discussion

To assess the performance of our graph based weighted association rule mining, we applied on heart disease data collected from various corporate hospitals in Andhra Pradesh. Association rules generated predicting heart diseases are

(1). (Age 41–65 and BP > 120 and maximum heart rate achieved and CA <=3) => heart disease
(2). (Age 41–65 and gender male and cholesterol > 240 and maximum heart rate achieved and CA <=3) =>heart disease
(3). (gender male and maximum heart rate achieved and old peak > 0 and slope >=2 and CA <=3) => heart disease
(4). (BP > 120 and resting ECG > 2 and maximum heart rate achieved and old peak > 0 and slope >=2 and CA <=3) => heart disease
(5). (chest pain > 3 and BP > 120 and old peak > 0 and slope >=2 and CA <=3) => heart disease

Rule 1 says that if a person‘s age is above 45 and has high blood pressure and when he attains maximum heart rate then he has a high probability of having heart

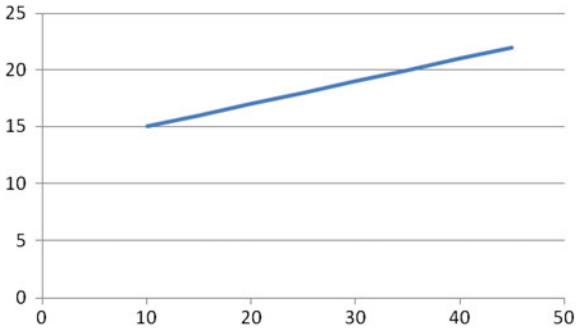


Fig. 54.1 No. of vertices vs. Execution time

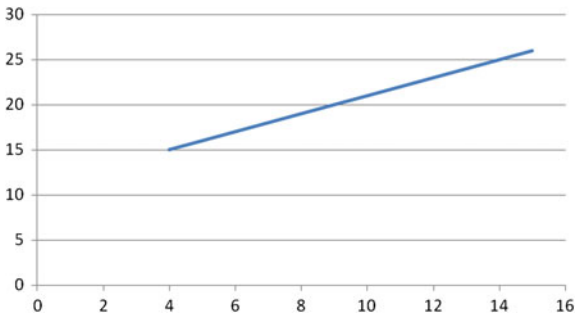


Fig. 54.2 No. of Edges vs. Execution time

disease. Rule 2 states that if a person is male and age is above 40 years with high blood pressure and cholesterol levels and maximum heart rate, then he has high chance of getting heart disease. Rule 3 to 5 confirm heart disease for risk factors like chest pain, blood pressure and with maximum heart rate achieved. These Rules characterize the patient with coronary disease.

54.7 Performance Evaluation

We evaluated the performance of our algorithm on medical data sets. Experiments are carried out using java programming language. In Fig. 54.1 running time is measured upon No. of vertices in a graph

For a certain range of vertices, the running time is moderately increasing in polynomial behavior. However, if the number of vertices exceeds range, the running time increases with an astronomical growth rate. In Fig. 54.2 Running time is measured on different graph density i.e. the number of edges in a graph varies. And the running time increases in a polynomial time.

Time complexity: Time taken for generating frequent item sets and constructing the weighted graph is O(n) and to find maximum weighted clique is $O(N^3)$where n is no. of items.

54.8 Conclusion

The objective of our work is to predict more accurately the presence of heart disease. Rules with risk factors like Age, high BP, and cholesterol were extracted. The Results and Discussion section discussed several important association rules predicting presence of heart disease. Graph based weighted association rule mining efficiently solve the problem of mining association rules. In future work we plan to develop data mining system for predicting heart disease using artificial intelligence.

References

1. The Times of India (August 14, 2011)
2. Gupta R (2010) Recent trends in CHD epidemiology in India. Ind Heart J 132(5):634–642
3. Hint Wint Khaing (2011) Data mining based fragmentation and prediction of medical data. IEEE
4. Agrawal R, Imielinski T, Swami1 A (1993) Mining association rules between Sets of Items in large databases. In: ACM SIGMOD Int'l conference on management of data, Washington, DC, 1993
5. Yun U, Leggett J (2004) Weighted frequent item mining with a weight range and a minimum weight. In: SIAM (2004)
6. Cook DJ, Holder LB et al (2005) Graph based mining of complex data. Advanced methods for knowledge discovery from complex data, pp 75–93
7. Palaniapan S (2008) IHDPS using data mining techniques. IEEE 2008
8. Ordonez C (2004) Improving heart disease prediction using constrained association rules. In: seminar presentation at Tokyo (2004)
9. Anbarasi M et al (2010) Enhanced prediction of heart disease with feature subset selection using genetic algorithm. IJEST 2(10):5370–5376
10. Patil SB et al (2009) Extraction of significant patterns from heart disease warehouses for heart attack prediction. IJCSNS 9(2): (February 2009)
11. Jabbar MA, Deekshatulu BL, Chandra P (2011) Cluster based association rule mining for heart attack prediction. JATIT 32(2):196–201
12. Yun U, Leggett JJ (2005) WFIM-Weighted frequent item set mining with a weight range and minimum weight. In: 5th SIAM international conference on data mining pp. 636–640 (2005)
13. Cavique L (2007) A scalable algorithm for the market basket analysis. J Retaia Consum Serv 14:400–407
14. Berge C (1991) Graphs, 3rd edn. North-Holland mathematical library, North Holland
15. Tan PN, Kumar V, Srivastava J (2002) Selecting the right interesting measure for association patterns. In: KDD, pp. 32–41
16. Piatetsky-Shapiro G (1991) Discovery, analysis and presentation of the strong rule. In KDD, MIT Press, Cambridge, pp 229–248
17. National academy on an Aging society. www.agingsociety.org
18. UCI Machine Learning Repository http://archive.ics.uci.edu/ml/datasets/Heart+Disease

Chapter 55
Elimination of Black Hole and False Data Injection Attacks in Wireless Sensor Networks

R. Tanuja, M. K. Rekha, S. H. Manjula, K. R. Venugopal, S. S. Iyengar and L. M. Patnaik

Abstract Wireless Sensor Networks (WSNs) are currently being used in a wide range of applications that demand high security requirements. Since sensor network is highly resource constrained, providing security becomes a challenging issue. Attacks must be detected and eliminated from the network as early as possible to enhance the rate of successful transactions. In this paper, we propose to eliminate Black Hole and False Data Injection attacks initiated by the compromised inside nodes and outside malicious nodes respectively using a new acknowledge scheme with low overhead. Simulation results show that our scheme can successfully identify and eliminate 100 % black hole nodes and ensures more than 99 % packet delivery with increased network traffic

Keywords Security · Sink acknowledgement · Negative acknowledgement · Black hole attack · Packet delivery rate

R. Tanuja (✉) · M. K. Rekha · S. H. Manjula · K. R. Venugopal
Department of Computer Science and Engineering,
University Visvesvaraya College of Engineering, Bangalore University, Bangalore,
e-mail: r_tanuja@yahoo.com

S. S. Iyengar
Florida International University, USA

L. M. Patnaik
Indian Institute of Science, Bangalore,

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_55,

55.1 Introduction

In Wireless Sensor Network the sensor nodes are usually deployed in harsh, unattended, remote areas and have limited sensing, computation and communication capabilities. The sensor networks are often exposed to various malicious attacks and the conventional defense mechanisms are not suitable because of its highly resource constrained nature. The security mechanisms should be strong enough and undoubtedly energy efficient to prevent attacks by malicious nodes to reduce the wastage of sensor resources and to provide authentication and integrity to sensed data. This paper proposes to detect and eliminate black hole attack which is a simple form of selective forwarding attack, where a malicious node may drop all the packets passing through it without forwarding to the sink node. We consider false data injection attack from outside malicious nodes i.e. where an attacker injects false data reports into the network and depletes the energy of the forwarding nodes.

55.2 Related Works

Zia and Zomaya [1] have analyzed Attacks, countermeasures and threat models in different layers of WSNs. Arif et al. [2] have designed Virtual Energy-Based Encryption and Keying (VEBEK) scheme, resulting in reduced number of overhead messages thereby increasing the lifetime of WSNs. The intermediate nodes can verify the authenticity and integrity of the incoming packets using a predicted value of the key generated by the senders virtual energy, without any need for specific rekeying messages. This work does not address insider attacks and dynamic paths. Misra et al. [3] have proposed an efficient technique, BAMBi, to mitigate the adverse effects of black hole attacks in WSNs. Bysani and Turuk [4] have discussed about selective forwarding attack, its types and some mitigation schemes to defend such attacks. Kaplantzis et al. [5] have developed a centralized Intrusion Detection Scheme (IDS) based on Support Vector Machines and sliding windows. It uses only two features to detect selective forwarding and black hole attacks. Ba et al. [6] have discussed a deterministic key management scheme, DKS-LEACH, to secure LEACH protocol against malicious attacks.

55.3 Problem Definition and Algorithm

Sensor Networks are deployed in harsh, unattended remote areas that are susceptible to various inside compromised node attacks (Black Hole) and outside malicious node attack (False Data Injection attacks). Node based authentication using cryptographic keys is ineffective in addressing insider attacks. The outside

malicious node false data injection attacks needs to be detected and eliminated. The objectives are: (i) To detect and eliminate black hole attack using a new acknowledgement scheme with low overhead. (ii) To ensure authenticity and integrity of transmitted packets by preventing false data injection by outside malicious nodes.

55.3.1 Algorithm

The algorithm for detection and elimination of Black Hole and False Data Injection attack consists of six steps:

(i) Keying process: This process involves dynamic key generation. When a node senses some data, it must authenticate the sensed data before transmitting to the sink node. Here we have used virtual energy-based keying process [2]. The dynamic key is generated as a function of current virtual energy of the sensor node. The key for first packet is generated as a function of initial virtual energy and initial vector of sensor node. Later keys are generated based on current virtual energy and previous key of the sensor. The dynamic key obtained from keying process is fed to RC4 algorithm to get permutation code Pc. The permutation is mapped to a set of actions to be taken on the message. Eg., Simple operations like shift, interleaving, 1's complement etc. The resultant packet format is: {ID, {ID, TYPE, DATA, event ID}Pc}.

(ii) DownStream Process: Downstream represents direction towards sink node. When a source node sense some event, it appends nodeID, type and event ID along with the sensed data and encode the whole data using virtual energy-based encryption mechanism [2]. Then forward the packet along with its plaintextID to next hop and wait for a pre-defined time to receive sink acknowledgement from its downstream neighbor. It stores the event ID in its cache until it receives ACK_ SINK.

(iii) Adressing False Data Injection attacks: When a forwarder node receives a packet, it authenticates the packet by performing virtual energy-based decoding and compares the plaintextID with decodedID. Malicious packets inserted by outsiders (False Data Injection attack) will be dropped immediately. The authentic packets will be forwarded to next downstream node along the path to the sink node after doing encoding operation. After forwarding, it will store the event ID and upstream nodeID in its own cache until it receives ACK SINK. This process continues up to the sink node. Table 55.1 shows the actions taking place in the downstream direction and elimination of False Data Injection.

(iv) Upstream Process: Upstream refers direction towards source node. After verifying the received packet, the sink node will send an acknowledgement back to the source node through intermediate nodes. The acknowledgement, ACK SINK consists of event ID and the upstream nodeID. If a node receives

Table 55.1 Algorithm for downstream process of BHnFDIA

```
Let t= predefined time, msg_pc=message encrypted using pc, FN=Forwarder
node S=source ;
begin
  if ( node v == S and S sense some event) then
       msg = append (event_ID, nodeID, type, sensed_data)
       key = DynamicKey(virtual_energy, plaintextID)
       pc = RC4 (key, plaintextID)
       msg_pc = encode (msg,pc)
       pkt = append ( plaintextID, msg_pc )
       forward pkt to next hop
       wait(t)
       cache(event_ID)
   endif
   if ( node v == FN) then
       receive pkt
       key = DynamicKey(virtual_energy, plaintextID)
       pc = RC4 (key, plaintextID)
       msgID = decode (msg,pc)
       x = compare(msgID,plaintextID)
       if (x == true) then
          reencode and forward pkt to next hop
          wait(t)
          cache(event_ID,upstream_nodeID)
       else
          find key by decrementing virtual_energy threshold times
          if failed drop packet
       endif
   endif
end
```

the acknowledgement from sink within the time interval, it will compare the event ID field in ACK SINK with the one stored in its own cache. If it matches, the corresponding transmission will be considered as successful and removes the corresponding entry from its own cache and forwards ACK SINK to its upstream node. This process will continue up to the source node.

(v) Adressing Communication Errors: When a packet or ACK traverses through the network, they can be lost due to some communication error. A node C will transmit a packet threshold number of times and wait for acknowledgements before considering the downstream node D as malicious. If node C fails to receive ACK SINK, the downstream node D is considered as malicious or black hole.

Table 55.2 Algorithm for upstream process of BHnFDIA

```
(i) Upstream Process
begin
  if ( node v == Sink) then
       verify pkt
       send ACK_SINK in upstream direction
   endif
   if ( node v == FN) then
       receive ACK_SINK
       if ( ACK_SINK event_ID == cache event_ID ) then
            remove corresponding entry form cache
            forward ACK_SINK in upstream direction
       endif
   endif
end
(ii) Elimination of BH attack
 Let j = 0, threshold = 5, SN = suspected node
begin
    if ( node v == predecessor (SN) in downstream direction ) then
       wait for ACK_SINK till time-out
       if time out occurs
          send NACK towards S, increment j
          wait for ACK_SINK for next packet
          if j exceeds threshold then
             mark SN as BH
             redirect successive packets to another route
             broadcast ALERT_INFO among neighbours
          endif
       endif
    endif
end
```

(vi) Addressing Black Hole Attack: When a packet traverses from source to sink through multiple hops, if a malicious node acts as a black hole, it will drop all the incoming packets without forwarding to sink node [4]. No acknowledgement is be sent to upstream node. After timeout, the node just before the attacker in the downstream direction marks the malicious node and sends a negative acknowledgement, NACK towards the source node. The successive packets received at the node just before the black hole in the downstream direction will be re-directed using another route to the sink node. It will broadcast an ALERT INFO message to all its neighbours so that they can avoid this particular node from the routes. Table 55.2 gives the steps involved in downstream process and elimination of Black Hole attack.

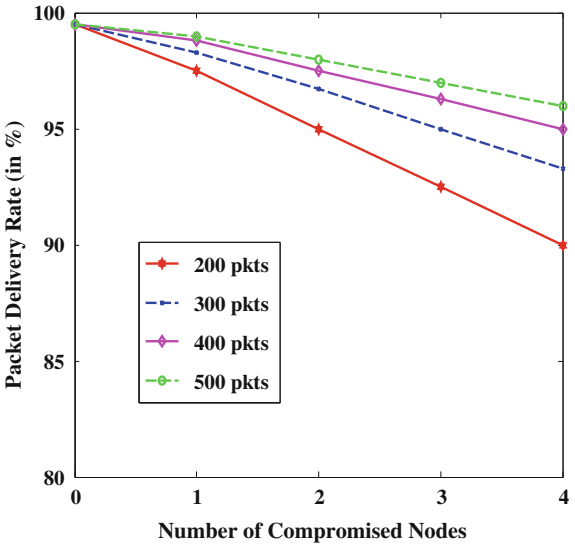


Fig. 55.1 Successful delivery of packets

55.4 Implementation and Performance Evaluation

We evaluate the performance of our scheme by simulation using MATLAB and compare it with other existing schemes in terms of packet delivery rate and filtering efficiency. Nodes are randomly deployed into 100 × 100 m^2. All sensor nodes are assumed to have same communication ranges. The routing algorithm is deployed on unreliable MAC protocol and there may be ACK or packet drops in the network. The network also experience black holes. Outside attackers may have spoofed valid node identifier. The inside attacker may have all the valid cryptographic details of the node.

Packet Delivery Rate: The packet delivery rate is calculated as the ratio between the number of packets that are sent by the source node and the number of packets that are received by the sink node. Figure 55.1 shows the results for successful packet delivery rate of our algorithm without enabling re-transmissions. As can be seen from the Fig. 55.1, packet delivery rate increases with increase in the packet count. This is because only a small threshold number of packets, say 5, need to be dropped in the process of detecting a single black hole. After dropping threshold packets, the upstream node of black hole will re-route the successive packets and informs neighbor nodes to avoid black hole through ALERT_INFO message. The downward slope is obviously due to the increase in black holes. As the number of compromised nodes increase, more packets will be dropped until all the black holes are detected.

Comparison of Packet Delivery Rate: Figure 55.2 compares the packet delivery rate of BHnFDIA with previous work VEBEK schemes in the presence

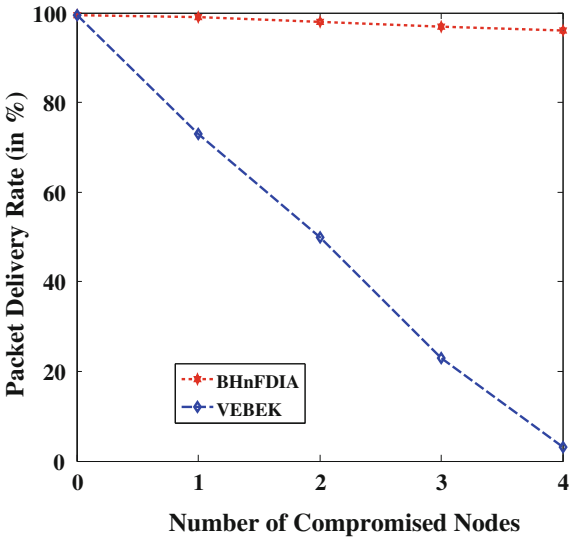


Fig. 55.2 Comparison of packet delivery

and absence of black holes. When there are no black holes, both schemes have almost same packet delivery rate. But when black holes are present, our scheme has 30–95 % more successful packet delivery. The energy consumption for keying process being same for both, but with ACK_SINK packet our scheme provides more security to address insider attacks As authentication is performed at every hop, malicious data inserted by outside attackers will be dropped within one hop itself. Hence the filtering efficiency is almost 100 %, irrespective of the number of malicious packets.

55.5 Conclusions

In WSNs for several applications, security is a major concern. In this paper, we propose algorithm to overcome Black Hole and False Data Injection Attack (BHnFDIA) in WSNs. It provides a new acknowledgement based detection scheme which helps to simplify the elimination of black holes and guarantees successful delivery of packets to destination. Our algorithm can eliminate false data injection by outside malicious nodes. Simulation results show that our algorithm can successfully identify and eliminate 100 % black hole nodes. Since authentication is performed at every hop malicious packets are immediately removed with 100 % filtering efficiency. Our scheme ensures more than 99 % packet delivery with increased network traffic. Our future work will incorporate other insider attacks without adding much communication overheads.

References

1. Tanveer Z, Albert Z (2006) Security issues in wireless sensor networks. In: Proceedings international conference systems and networks communication (ICSNC 06), Oct 2006
2. Uluagac AS, Beyah RA, Li Y, Copeland JA (2010) VEBEK: virtual energy-based encryption and keying for wireless sensor networks. IEEE T Mobile Comput 9(7):994–1007
3. Misra S, Bhattarai K, Guoliang X (2011) BAMBi: blackhole attacks mitigation with multiple base stations in wireless sensor networks. In: Proceedings international conference communications (ICC 2011), July 2011
4. Bysani LK, Turuk AK (2011) A survey on selective forwarding attack in wireless sensor networks. In: Proceedings international conference on devices and communications (ICDeCom), Feb 2011
5. Kaplantzis S, Shilton A, Mani N, Sekercioglu YA (2007) Detecting selective forwarding attacks in wireless sensor networks using support vector machines. In: Proceedings third international conference on intelligent sensors, sensor networks and information, pp 335–340, Dec 2007
6. Ba M, Niang I, Gueye B, Noel T (2010) A deterministic key management scheme for securing cluster-based sensor networks. In: Proceedings 2010 IEEE/IFIP international conference on embedded and ubiquitous computing, pp 422–427, Dec 2010

Chapter 56
An Efficient Algorithm for Finding Frequent Sequential Traversal Patterns from Web Logs Based on Dynamic Weight Constraint

Rahul Moriwal and Vijay Prakash

Abstract Many frequent sequential traversal pattern mining algorithms have been developed which mine the set of frequent subsequences traversal pattern satisfying a minimum support constraint in a session database. However, previous frequent sequential traversal pattern mining algorithms give equal weightage to sequential traversal patterns while the pages in sequential traversal patterns have different importance and have different weightage. Another main problem in most of the frequent sequential traversal pattern mining algorithms is that they produce a large number of sequential traversal patterns when a minimum support is lowered and they do not provide alternative ways to adjust the number of sequential traversal patterns other than increasing the minimum support. In this paper, we propose a frequent sequential traversal pattern mining algorithm with weights constraint. Our main approach is to add the weight constraints into the sequential traversal pattern while maintaining the downward closure property. A weight range is defined to maintain the downward closure property and pages are given different weights and traversal sequences assign a minimum and maximum weight. In scanning a session database, a maximum and minimum weight in the session database is used to prune infrequent sequential traversal subsequence by doing downward closure property can be maintained.

Keywords Sequential traversal pattern mining · Weight constraint · Web usage mining · Data mining

R. Moriwal (✉)
Shri Vaisnav Institute of Technology and Science, Indore, Madhya Pradesh, India
e-mail: rmoriwal@gmail.com

V. Prakash
Shri Vaisnav Institute of Technology and, Science, Indore, Madhya Pradesh, India
e-mail: vijayprakash15@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_56,

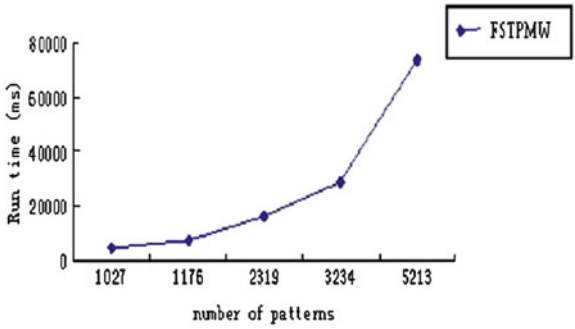


Fig. 56.1 Scalability with number of frequent sequential traversal patterns

56.1 Introduction

The World Wide Web is an immense source of data that can come either from the Web content, represented by the billions of pages publicly available, or from the Web usage, represented by the log information daily collected by all the servers around the world. Web Mining is that area of Data Mining which deals with the extraction of interesting knowledge from the World Wide Web [1].

More precisely, Web Content Mining is that part of Web Mining which focuses on the raw information available in Web pages; source data mainly consist of textual data in Webpages (e.g., words, but also tags); typical applications are content-based categorization and content-based ranking of Web pages [2]. Web Structure Mining is that part of Web Mining which focuses on the structure of Web sites; source data mainly consist of the structural information present in Web pages (e.g., links to other pages); typical applications are link-based categorization of Web pages, ranking of Web pages through a combination of content and structure [3], and reverse engineering of Web site models. Web Usage Mining is that part of Web Mining which deals with the extraction of knowledge from server log files; source data mainly consist of the (textual) logs that are collected when users access Web servers and might be represented in standard formats (e.g., Common Log Format, Extended Log Format, LogML) [4]; typical applications are those based on user modeling techniques, such as Web personalization, adaptive Web sites, and user modeling. Figure 56.1 shows the main application areas of WUM.

Srivastava et al. [5] systematically discuss the development of WUM and classify the content of WUM. Zhang and Liang [6] show the importance of data preprocessing in Web Usage Mining and present an algorithm called "USIA" which boasts high efficiency. Wang and Meinel [7] point out that user behaviors recovery and pattern definition play more important roles in web mining than other applications so they give a new insight on behavior recovery and complicated pattern definition. As current Web Usage Mining applications rely exclusively on the web server log files, Guo et al. [8] propose a system that integrates Web page clustering into log file association mining and use the cluster labels as Web page

content indicators in the hope of mining novel and interesting association rules from the combined data source.

Sequential pattern mining has been Apriori based sequential pattern mining was used based on the downward closure property. That is, if any length k sequential pattern is not frequent in a sequence database, superset sequential patterns can not be frequent. Using this characteristic,

56.2 Problem Definition and Related Work

56.2.1 Problem Definition

Let $P = \{P_l, P_2 \ldots P_n\}$ be a unique set of pages. A session S is an ordered list of itemsets, denoted as $(s_l, s_2,.., s_m)$, where s_j is an itemset which is also called an element of the session, and $s_j \subseteq P$. That is, $S = (s_l, s_2,..,s_m)$and s_j is $(x_1\ x_2 \ldots x_k)$,where X_t is an item. The brackets are omitted if an itemset has only one item. An item can occur at most one time in an element of a sequence but it can occur multiple times in different elements of a sequence. The size S of a sequence is the number of elements in the sequence. The length, l(s), is the total number of items in the sequence. A sequence with length 1 is called an 1-sequence. A sequence database, $D = \{S_l, S_2,.., S_n\}$, is a set of tuples (sid, s), where sid is a sequence identifier and S_k is an input sequence. A sequence $\alpha = (X_1, X_2,.., X_n)$ is called a subsequence of another sequence $\beta = (Y_1, Y_2,.., Y_m)$ $(n \leq m)$, and β is called a super sequence of α if there exist an integer $1 < i_1 < \ldots < i_n < m$ such that $X_1 \subseteq Y_{i1},\ldots, X_n \subseteq Y_{in}$. A tuple (sid, S) is said to contain a sequence S_a if S is a super sequence of S_a. The support of a sequence S_a in a sequence database D is the number of tuples in SDB that contains Sa.

56.2.2 Related Work

GSP mines sequential patterns based on an Apriori like approach by generating all candidate sequences. This is inefficient and ineffective. To overcome this problem, the database projection growth based approach, FreeSpan, was developed. Although FreeSpan outperforms the Apriori based GSP algorithm, FreeSpan may generate any substring combination in a sequence. The projection in FreeSpan must keep all sequences in the original sequence database without length reduction.

PrefixSpan, a more efficient pattern growth algorithm was proposed which improves the mining process. The main idea of PrefixSpan is to examine only the prefix subsequences and project only their corresponding suffix subsequences into

projected databases. In each projected database, sequential patterns are grown by exploring only local frequent patterns.

In SPADE, a vertical id-list data format was presented and the frequent sequence enumeration was performed by a simple join on id lists. SPADE can be considered as an extension of vertical format based frequent pattern mining. SPAM [9] utilizes depth first traversal of the search space combined with a vertical bitmap representation of each sequence.

Proposed Work In this section, we suggest an efficient sequential traversal pattern mining algorithm in which the main approach is to apply weight constraints into the frequent sequential traversal tree while maintaining the downward closure property. We discuss our algorithm in detail and show actual examples for sequential traversal pattern mining with weight constraint.

Definition 3.1 Weight Range A weight of a web page is a non-negative real number that shows the importance of each web page. The weight of each web page is assigned to reflect the importance of each web page in the session database.

Definition 3.2 Traversal sequence with Weight We can use the term, traversal sequence with weight to represent a set of sequential traversal patterns with weight.

Definition 3.3 Average Weight of Traversal We can use the term; average weight of subsequence is the sum of weight all pages in traversal divided by total number of pages in sequence.

Definition 3.4 Minimum and Maximum Weight of Subsequence Here we define the maximum and minimum weight of traversal is average weight. If the weight of sequence come under the maximum and minimum weight range than given sequence is frequent otherwise infrequent.

A. Sequential traversal pattern with weight constraint

In this paper, pages of traversals are assigned with weights to show their importance. For example, when users traverse web site, they may have different interest in each page, and therefore stay for different times. Web pages can be assigned with a weight standing for the user stay time, frequency of pages, content of pages and type of web site. This paper generalizes the mining problem to the case where pages of traversals are given such weights showing their importance. The weights are taken into account in the measurement of support, the ratio of traversals which contains a candidate pattern (Table 56.1).

In this section, we propose the concept of sequential traversal patterns with weight constraint, and show their importance. Example: In session S_1 the weight of P_2 is 0.2 and the support is 4. The weight range for P_2 is from 0.45 to 0.67. So, when we construct the frequent sequential traversal pattern tree P_2 is eliminated from session (Table 56.2).

B. Frequent Sequential Traversal Pattern Tree with weight constraint

In this section, a data structure called FSTP-tree is constructed. FSTP-Tree is a data structure, which must satisfy the following conditions. Firstly, it consists one

Table 56.1 A sequence database as a running example

Sid	Traversal Sequence			Weight
S1	P2 P4P1P5	P1	P3	0.2,0.3,0.12,0.34,0.6,0.3
S2	P1 P3 P4 P2	P2	P4	0.12,0.5,0.91,0.12,0.4,0.26
S3	P1 P3P6 P_7	P2	P1	0.6,0.2,0.32,0.56,0.45,0.7
S4	P2 P5 P1 P4	P3	P6	0.5,0.56,0.32,0.23,0.7,0.54

Table 56.2 The example of page with weight range

S.No.	Page	Support	Weight range
1	P1	4	0.12– 0.56
2	P2	4	0.45–0.67
3	P3	4	0.23–0.67
4	P4	3	0.12– 0.45
5	P5	2	0.34–0.67
6	P6	2	0.24–0.8
7	P7	0	0.12– 0.56

root "null", a set of item prefix subtrees as the children of the root, and a frequent - page head table. Secondly, every page in the page prefix subtree contains three fields: the name of the page, the support of the page, and a link to the next same page. Thirdly, every page in the frequent-page table contains three fields: the name of the page, the weight range and a link to the first node in the tree which denotes this page. The following algorithm to build the FSTP-tree:

Algorithm 1 (FSTP-tree Building: Building FSTP-tree of the SDB)
Input: A session database SDB, weights of pages and a minimum support
Output: corresponding FSTP-tree
Method:

1. Scan the whole SDB and find frequent pages from SDB based on support and weight range assign to the page. Here we add only those pages that come under the weight range and contribute to the support and those not come in given range consider as outlier and not contribute to support.
2. Create the root of the FSTP-tree, and label it NULL.
3. Scan the whole SDB for the second time. For each session in the SDB, we only preserve the pages which are frequent and have a weight in given weight range, and hold the traversal sequences of pages. The different branches of same prefix can be merged.

C. FSTPMW Algorithm

The divide-and-conquer strategy is used for finding frequent sequential traversal patterns. To handle the ordered problem, the FSTPMW uses a merging method. Each frequent ordered pattern whose first page is P_1 must be contained in one or more session. The merging process in fact is rebuilding a smaller FSTP-tree. This time, the relative sessions all contains P_1 as the first web page.

The complete algorithm given as:

Algorithm 2 (FSTPMW: Mining frequent sequential traversal pattern)
Input: FSTP-tree
Output: frequent sequential traversal pattern
Method: call FSTPMW (Weight range for each page, support, Minimum & Maximum Weight Range)

```
Procedure FSTPMW (FSTPtreeRootNode node,String prefix)
{
for each node x in the corresponding page head table do
if x.support less than minimum support then
calculate the average weight of prefix
if minimum weight<=average weight<=maximum weight{
output prefix;
}
return;
else if i.subs.count == 0 then prefix = prefix + i.content;
calculate the average weight of prefix
if minimum weight<=average weight<=maximum weight
{
output prefix;
}
return;
else
call CombineTree(i);
for each node j in i.subs do
call FSTPMW (j, prefix + i); end for
end if end for
}
```

56.3 Analysis and Performance Evaluation

In this section, we present our performance study over various datasets. We report our experimental results on the performance of FSTPMW in comparison with a recently developed algorithm; WSpan, which is the fastest algorithm for mining sequential patterns. The main purpose of this experiment is to demonstrate how effectively the sequential traversal patterns with weight constraint can be generated

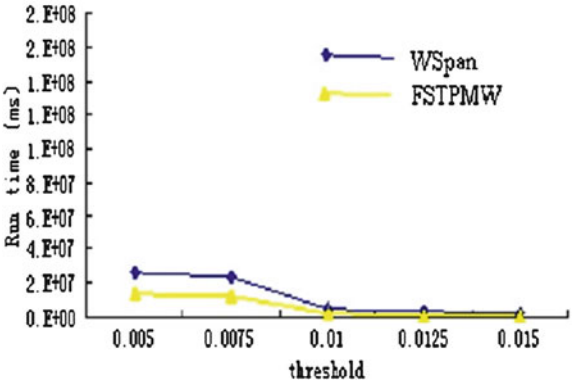


Fig. 56.2 Runtime

by incorporating a weight page, weight of sequence with a support. First, we show how the number of sequential traversal patterns can be adjusted through user assign weights, the efficiency in terms of runtime of the FSTPMW algorithm, and the quality of sequential traversal patterns. Second, we show that FSTPMW has good scalability against the number of sequence transactions in the datasets.

56.3.1 Environmental Results. Comparison of FSTPMW and WSpan

In this performance test, we focused on the efficiency of using a weight range. Our experiment shows that in most cases, FSTPMW outperforms WSpan. First, we evaluate the performance on the kosarak dataset (Fig. 56.2).

56.3.2 Further Extension

FSTPMW basically focuses on sequential pattern mining with weight constraint uses a weight range to adjust the number of sequential traversal patterns. Frequent sequential traversal pattern mining can be extended by considering levels of support and/or weight of sequential traversal patterns. There are many areas in which items have different importance and patterns with a similar level of support and/or weight are more meaningful.

56.4 Conclusion

Many studies exist on mining sequential frequent patterns. One of the main limitations of the traditional approach for mining sequential traversal patterns is that all items are treated uniformly, while each page of web site has different importance. Moreover, previous sequential traversal pattern mining generates a very large number of subsequence as the minimum support becomes lower. In this paper, we developed FSTPMW which focused on frequent sequential traversal pattern mining with weight constraint. A weight range is used to adjust the number of sequential patterns. The extensive performance analysis shows that FSTPMW is efficient and scalable in mining sequential traversal pattern.

References

1. Etzioni O (1996) The world-wide web: quagmire or gold mine? Commun ACM 39(11): 65–68
2. Kosala R, Blockeel H (2000) Web mining research: a survey. SIGKDD Explor 1:1–15
3. Brin S, Page L (1998) The anatomy of a large-scale hyper-textual web search engine. Comput Networks ISDN Syst 30(1– 7):107–117
4. Punin JR, Krishnamoorthy MS, Zaki MJ (2001) LOGML—log markup language for web usage mining. In: WEBKDD workshop 2001: mining log data across all customer touch points (with SIGKDD01), pp 88–112, San Francisco, Aug
5. Srivastava J, Cooley R, Mukund D (2000) Web usage mining: discovery and applications of usage patterns from web data. SIGKDD Explor 1(2):12–23
6. Zhang H, Liang W (2004) An intelligent algorithm of data pre-processing in Web usage mining. In: Proceedings of the world congress on intelligent control and automation (WCICA), v 4, p 3119–3123
7. Long W, Christoph M (2004) Behaviour recovery and complicated pattern definition in web usage mining. In: Web engineering: 4th international conference, ICWE 2004, Munich, July 26–30, pp 531–543
8. Jiayun G, Vlado K, Qigang G (2005) Integrating web content clustering into web log association rule mining. In: Advances in artificial intelligence: 18th conference of the canadian society for computational studies of intelligence, Canada
9. Ayres J, Gehrke J, Yiu T, Flannick J (2002) Sequential pattern mining using a bitmap representation. In: SIGKDD'02, pp 1–7
10. Agrawal R, Srikant R (1994) Fast algorithms for mining association rules. In: Proc.eedings of the 20th international conference on very large database, Chile, pp 487–499
11. Agrawal R, Srikant R (1995) Mining sequential patterns. In: Proceedings of the International Conference on Data Engineering (ICDE), Taipei March 1995

Chapter 57
A Blind Watermarking Algorithm for Audio Signals Based on Singular Value Decomposition

Ankit Murarka, Anshul Vashist and Malay Kishore Dutta

Abstract This paper proposes a blind digital watermarking scheme for audio signals based on Singular Value Decomposition (SVD) and Quantization Index Modulation (QIM). The process involves a watermark using a binary image file which has been readjusted using Arnold Transform before being embedded into the host signal. Synchronization code is embedded into the audio signal so that the watermark has the capability of self-synchronization against attacks. Experimentation is done to check the robustness of the proposed scheme in the presence of various attacks. Experimental results indicate that the proposed scheme provides good imperceptibility and robustness under various signal processing attacks.

Keywords Digital watermarking · Singular value decomposition · Perceptual transparency

57.1 Introduction

With the development of internet and transmission of digital data over the internet, it is easy to produce and distribute illegal copies of digital audio media. The fact that one can get any kind of information from the internet has given rise to numerous problems one of which is copyright infringement. Hence there is an increased need for mechanisms to protect the ownership of digital data [1]. Digital Watermarking is an effective technique which can deal with the problem of copyright infringement [2]. Watermarking is defined as adding some digital

A. Murarka (✉) · A. Vashist · M. K. Dutta
Department of Electronics and Communication Engineering,
Amity School of Engineering and Technology, Amity University, Noida 201301, India

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_57,

information to the host signal in a manner that it does not deteriorate the quality of the host signal [1]. To function as an effective tool to enforce ownership rights, the watermarking scheme must meet the requirements of good imperceptibility, strong robustness and high-level security [3–5]. A watermark has to remain imperceptible in order to maintain its secrecy [6, 7]. Robustness refers to the ability to detect the watermark after common signal processing attacks [3, 8, 7, 9]. The watermarking algorithm should be secure which means that the watermark can only be detected by the authorized person [4].

Many efficient audio watermarking algorithms are reported [10, 8, 11], which also meet the conflicting requirements of audio watermarking like perceptual transparency, robustness and watermark data rate. Watermarking using singular value decomposition (SVD) is given for image watermarking in [11]. In the proposed method the watermark is embedded with the aid of Singular Value Decomposition (SVD) and Quantization Index Modulation (QIM) which are elaborated in the next sections.

57.2 Singular Value Decomposition

Singular Value Decomposition [12] is a method for data reduction. SVD is a theorem of linear algebra which breaks a rectangular matrix, say $\{A_{ij}\}_{m\times n}$ into the product of three matrices—an orthogonal matrix *U*, a diagonal matrix *S* and transpose of the orthogonal matrix *V*. The theorem is expressed as:

$$A_{m\times n} = U_{m\times m} S_{m\times n} V^{T}_{n\times n} \tag{1}$$

The columns of U and V are mutually orthogonal unit vectors. S is a diagonal matrix containing square roots of Eigen values from U or V in descending order. The values are called SV's and are all non negative. The SVD watermarking scheme embeds the watermark bits by modifying the singular values (SV's).

57.3 Quantization Index Modulation

Quantization index modulation techniques (QIM) [13] are a class of watermarking techniques that can be used to achieve good trade-offs between the mutually conflicting parameters of audio watermarking like information embedding rate, watermark robustness and signal fidelity. In QIM, the embedding function, X (Q, W), is composed of an ensemble of functions of the host signal, H, accompanied by the watermark signal W such that:

$$X\ (Q,\ W) \approx Q \tag{2}$$

for all W. Therefore the distortion due to the watermark is minimal. Modulation of the index or the series of indices containing the watermark data is done and then host signal is quantized with the associated quantizer.

57.4 Proposed Scheme

A host audio signal, say A is taken for the watermark to be embedded in. The signal is divided into two equal parts. A 16 bit synchronization code is embedded in the first part. The second part of the audio signal is used for embedding the watermark data.

The watermark signal is taken as a binary image which is a square matrix (p × p). Before the image is embedded onto the audio signal, it is first rearranged using Arnold Transform [12]. Suppose that $(x, y)^T$ is the coordinate of a particular pixel in the watermark image, $\left(x', y'\right)^T$ is the coordinate of that pixel after transformation. Arnold Transform can be expressed as:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = (\text{mod P}) \tag{3}$$

where, $x, y \in \{0, 1 \ldots \ldots P-1\}$ The host audio signal A is divided into different non-overlapping partitions F_j, where j = 1, 2, 3,….., N × N. These frames are then reshaped into blocks B_j, j = 1,2,3,….., N × N (2-D square matrices), each of size r × r, so that SVD can be applied on them. Now SVD is performed on each block, represented as I.

$$\mathrm{I} = \mathrm{USV}^{\mathrm{T}} = \sum_{\mathrm{i}=1}^{\mathrm{r}} \mu_{\mathrm{i}} \mathrm{U_i V_i^T} \tag{4}$$

The Euclidean norms of singular values of the blocks are computed. Let $\mu^j = \left(\mu_1^j, \mu_2^j, \mu_3^j, \ldots, \mu_r^j\right)$ be the vector of SVs of block B_j. The norm of the vector is obtained as follows:

$$\mathrm{s_j} = \left| \sqrt{\sum_{\mathrm{i}=1}^{\mathrm{r}} \left(\mu_{\mathrm{i}}^{\mathrm{j}}\right)^2} \right| \tag{5}$$

A variable t_v is chosen to store the fractional parts of the Euclidean norm of all frames. Let $t_v = s_j mod 1$. Another variable t_w is taken to store the integer part of the norms.

A variable x is taken where, $\mathrm{x_j}$ = floor ($\mathrm{t_v}$ × 10).

Now according to the value of *"x_j mod 2"*, which may also be called as the *quantization coefficient*, and the watermark bit (w_j) the norms of each block are modified. If x_j *mod 2* is "0", then s_j is modified in the following manner:

$$s'_j = t_w + (x_j + 1)/10, \text{ if } w_j = 0$$
$$s'_j = t_w + x_j/10, \text{ if } w_j = 1 \quad (6)$$

If *x mod 2* is "1", then s_j is modified in the following manner:

$$s'_j = t_w + x_j/10, \text{ if } w_j = 0$$
$$s'_j = t_w + (x_j + 1)/10, \text{ if } w_j = 1 \quad (7)$$

The modified vectors of SV's of the blocks are calculated using the following equation:

$$\tilde{\mu}^j = \mu^j \times \left(\frac{s'_j}{s_j}\right) \quad (8)$$

The modified matrix of the block $\tilde{B}_j$ is obtained by applying inverse SVD to the modified SVs.

$$\tilde{B}_j = \sum_{i=1}^{r} \tilde{\mu}_i^j U_i(j) V_i^T(j) \quad (9)$$

The watermarked audio signal is obtained by combining all the modified blocks, $\tilde{B}_j$.

The extraction for this algorithm is blind. To extract the watermark, the watermarked audio signal is partitioned into frames which in turn are divided into blocks $\tilde{B}_j$, j = 1,2,3,……,p × p, of size r × r, where p × p is the number of bits in the watermark signal. Then SVD is applied to each block. The norms $\tilde{s}_j$ of the SVs of the blocks are computed. Let $t_v = \tilde{s}_j mod\, 1$. Then $x = floor\ (\ t_v \times\ 10)$ is calculated. If value of x is even, then the extracted watermark bit (w_j) is 1, else it is 0. Finally, Inverse Arnold Transform is applied on the extracted watermark image to obtain the unscrambled original image.

57.5 Experimental Results and Comparisons

The various audio samples used for testing were all sampled at a frequency of 44.1 kHz. The synchronization code used in the experiments was of 16 bits and size of the watermark signal was 32 × 32. The binary image used as watermark is given below (Fig. 57.1):

The audio samples were subjected to the *Subjective Listening Test* for evaluating the perceptual grades (PG). This test is carried out to check the audio quality

Fig. 57.1 Watermark signal

Table 57.1 Perceptual grades

PG	Perception of watermark	Quality of signal
1	Imperceptible	Excellent
2	Perceptible but not annoying	Good
3	Slightly Annoying	Fair
4	Annoying	Poor
5	Very Annoying	Bad

of the watermarked signal and see if it has undergone any significant changes. The grades are defined in Table 57.1.

Listeners of different age groups were provided with the audio signal and the watermarked signal and then asked to rate them in terms of perceptual grades (PG).

In order to test the effectiveness and robustness of the watermarking algorithm, several signal processing attacks were initiated [10]. The attacks performed on the watermarked signal are as follows:

i. A filter with a cut-off frequency of 11.025 kHz is used on the watermarked signal.
ii. The watermarked signal was resampled at 22.05 kHz and then restored back to 44.1 kHz.
iii. White Gaussian Noise is added to the signal so that the final signal has an SNR of more than 40 dB.
iv. Cropping: 500 samples are randomly cropped from the watermarked signal.
v. MP3 Compression: The MPEG–1 layer 3 compression with a bit rate of 64 kbps is applied to the watermarked signal.

The Signal to Noise Ratio (SNR) and Bit Error Rate (BER) were evaluated in order to check the robustness performance of the proposed scheme.

Table 57.2 gives a comparison of robustness and the perceptual grading performance of the proposed method.

Table 57.2 SNR, PG and BER of Audio Samples under Comparison

Attacks	Audio sample	SNR (dB)	Perceptual grades	Bit error rate (BER (%))		
				Scheme [17] (uniform DOE)	Scheme [5] (non-uniform DOE)	Proposed scheme
Low pass filtering	Classical 1	79.1269	1	0.12	0	0
	Country 1	62.4410	2	4.2	2.9	3.91
	Folk 1	65.2231	2	3..6	2.8	3.22
	Blues 1	77.7453	1	2.1	1.9	1.86
	Pop 1	67.0655	1	0.6	0.1	0.09
Re-sampling	Classical 1	79.1224	1	2.9	2.2	2.15
	Country 1	63.6981	2	2.8	1.9	2.15
	Folk 1	67.8933	1	1.1	0.9	0
	Blues 1	77.9915	1	0.8	0	0.98
	Pop 1	67.2795	1	0.1	0	0
Additive white gaussian noise (AWGN)	Classical 1	64.8506	1	0.6	0.1	0
	Country 1	57.2264	1	0.5	0	0
	Folk 1	59.0442	1	0.5	0	0
	Blues 1	68.2428	1	0.2	0	0
	Pop 1	53.4764	1	0.9	0	0
Cropping	Classical 1	77.4553	1	0.2	0.1	0
	Country 1	67.3345	1	0.6	0.3	0.20
	Folk 1	67.3902	1	0.2	0	0
	Blues 1	75.0911	1	0.2	0	0.09
	Pop 1	66.2892	1	0.1	0.04	0
MP3 compression	Classical 1	78.2243	1	0.4	0.1	0
	Country 1	61.1098	2	0.82	0.56	0.78
	Folk 1	66.4533	1	0.5	0.41	0.39
	Blues 1	77.8321	1	0.3	0.22	0.20
	Pop 1	67.0237	1	0.3	0.12	0.09

57.6 Conclusion

A blind watermarking scheme for audio signals is being presented in this paper. This technique makes use of SVD and QIM for embedding the watermark onto the host audio signal. The watermark embedding method is perceptually transparent which is found from the subjective listening tests and the SNR value which is

above 20 dB for all samples under test. The method was found to be resilient to various signal processing attacks. Apart from a very slight distortion there was no perceptible drop in the quality of the extracted watermark. The comparison of performance with other existing systems indicates that the proposed method is superior. From the experimental results it can be concluded that the method is robust and effective. Further this work can be extended to design better quantization formula and better quantization parameters to increase the robustness.

References

1. Bender W, Gruhl D, Morrimoto N, Lu A (1996) Techniques for data hiding IBM Syst J 35(3/4):313–336
2. Brassil JT, Low S, Maxemchuk NF,O'Gorman L (1995) Electronic marking and identification techniques to discourage document copying. IEEE J Selected Areas in Commun 13(8):1495–1504
3. Katzenbeisser S, Petitcolas F (2000) Information hiding techniques for steganography and digital watermarking. Artech House, London
4. Miller M; Cox I, Linnartz J-P, Kalker T (1999) A review of watermarking principles and practices. In: Parhi KK, Nishitani T (eds) Digital signal processing in multimedia systems, Marcel Dekker Inc., New York, pp 461–485
5. Voyatzis G, Pitas I (1996) Applications of toral automorphisms in image watermarking. In: Proceedings of international conference on image processing, vol 1, pp 273–240
6. Gordy JD, Bruton LT (2000) Performance evaluation of digital audio algorithms. In: Proceedings of the 43rd IEEE midwest symposium on circuits and systems, vol 1, pp 456–459
7. Shishkin AV (2011) Robust digital watermarks for audio signals. Radio Electron Commun Syst 54(3):138–146
8. Dutta MK, Pathak VK, Gupta P (2010) A robust watermarking algorithm for audio signals using SVD. In: International conference on contemporary computing, CCIS vol 94, Part 1, pp 84–93
9. Singh J et al. (2012) Audio watermarking based on quantization index modulation using combined perceptual masking. Multimedia tools and applications, Springer, Heidelberg (in press)
10. Dutta MK, Pathak VK, Gupta P (2010) An adaptive robust watermarking algorithm for audio signals using SVD. Transactions on computational science, Springer Verlag Publishers, Heidelberg, pp 131–153
11. Lee S-J, Jung S-H (2001) A survey of watermarking techniques applied to multimedia. In: Proceedings of IEEE international symposium on industrial electronics (ISIE'01), vol 1, pp 272–277
12. Changa C-C, Tsaib P, Lin C-C (2005) SVD based digital watermarking scheme. Pattern Recog Lett 26(10):1577–1586
13. Chen B, Wornell GW (2001) Quantization index modulation: a class of provably good methods for digital watermarking and information embedding. IEEE T Inform Theory 47(4):1423–1443

Chapter 58
Performance Evaluation of Web Browsers in Android

E. Harsha Prabha, Dhivya Piraviperumal, Dinesh Naik, Sowmya Kamath and Gaurav Prasad

Abstract In this day and age, smart phones are fast becoming ubiquitous. They have evolved from their traditional use of solely being a device for communication between people, to a multipurpose device. With the advent of Android smart phones, the number of people accessing the Internet through their mobile phones is on a steep rise. Hence, web browsers play a major role in providing a highly enjoyable browsing experience for its users. As such, the objective of this paper is to analyze the performance of five major mobile web browsers available in the Android platform. In this paper, we present the results of a study conducted based on several parameters that assess these mobile browsers' functionalities. Based on this evaluation, we also propose the best among these browsers to further enrich user experience of mobile web browsing along with utmost performance.

Keywords Android · Browsers · Performance · Mobile · Dolphin · Mozilla firefox · Skyfire · Opera mini

E. Harsha Prabha (✉) · D. Piraviperumal · D. Naik · S. Kamath · G. Prasad
Department of Information Technology, NITK Surathkal, India
e-mail: harshaprabha@gmail.com

D. Piraviperumal
e-mail: dhivya.piraviperumal@gmail.com

D. Naik
e-mail: dinnaik@gmail.com

S. Kamath
e-mail: sowmyakamath@gmail.com

G. Prasad
e-mail: chgauravprasad@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_58,

58.1 Introduction

With the advent of smart phones, mobile web usage is on the rise as more fully featured mobile browsers are available on the market. This is evident from the fact that smart phone sales showed a strong upward growth worldwide in the year 2011. As per the survey conducted by IDC, 31.8 % of the hand sets shipped in 2011 was smart phones [1].

According to StatCounter, 8.49 % of website hits/page views come from a handheld mobile device in January 2012 [2]. Currently, there are quite a few web browsers available for the Android market. We focussed on the most popular browsers in android platform [3]. We narrowed down our list of study to five browsers based on their market share, namely, Android browser, Dolphin HD, Mozilla Firefox, Opera Mini and Skyfire.

In this paper, we have proposed our study on performance analysis on most popular web browsers. The browser selection and parameter consideration is addressed in Sect. 58.2. Further, test phase and methodology is described in Sect. 58.3. Based on the implementation of the tests, the result of browser evaluation is presented in Sect. 58.4, followed by conclusion and references.

58.2 Browser Selection and Parameters Considered

In this paper, we concentrate on four full-featured mobile browsers and one mini browser. Consequently, we have chosen the five browsers on android platform based on two criteria: Market share of browsers and popularity of browsers. According to the data collected from netmarketshare.com [4], Android browser (ver. 4.0.3), Mozilla Firefox (ver. 10.0.2) and Opera Mini (ver. 6.5) were selected. Dolphin HD (ver. 7.4.0) and Skyfire (ver. 4.0.4) were selected based on data collected on their popularity from android play Google website [5]. The versions used are the latest versions dated 30th March 2012.

The parameters considered for the test phase were: Web Technology Support, Browser features, Plug in and Web feed Support, Add-on, Security, Accessibility Features, Benchmarks, Video Streaming, Gzip and Speed. Every parameter was given a base of five points each. Added points were given to parameters depending on the relevance and importance. For instance, security is an important parameter and hence eight points was assigned to it. Further, speed is the one which ultimately decides the performance of the browser and therefore allotted ten points.

58.3 Test Phase

The tests were conducted using an android emulator with a SDK version 4.03. The operating system used was Windows 7 Home Basic with an Intel i5 processor. For each browser selected, a detailed literature survey was conducted to have an in depth insight on their features and architecture. In addition, the android application package file (.apk file) corresponding to the chosen browsers were downloaded and installed in the android emulator. To assess the parameters selected, several tests and benchmark tools were used. Details of individual parameters is mentioned below.

58.3.1 Web Technology Support

The following attributes were considered under this category: XSLT, HTML-5, XHTML, XForms, Javascript, DOM, CSS3, WEB-GL and SVG. Detailed survey combined with benchmark tests like HTML5 and CSS3 were carried out based on which the points were given. For example, the browsers which showed CSS3 compatibility of more than 20 % and less than 50 % were awarded partial points while the browsers with above 50 % compatibility were awarded five points.

58.3.2 Benchmarks

To begin with, ACID3 test was conducted to check the browsers compliance with DOM and JavaScript [6]. Secondly, HTML5 test [7] and CSS3 test [8] was used to inspect the browsers' compatibility to the upcoming HTML5 and CSS3 standards and their related specifications. Thirdly, Sunspider [9], a popular JavaScript benchmark tool was used to evaluate Core JavaScript functionalities of the browsers. Finally, Web compatibility Test [10] for mobile browsers was conducted to expose web page rendering flaws in mobile web browsers.

58.3.3 Speed

CSS rendering, Cold Start and Warm Start were considered as parameters for speed. Cold start refers to the time taken to do a cold load of a browser. Once logged out and all background processes have been completed, browser was run as the first program. After completing a cold start, browser was closed, and time taken to start again was measured for the warm start test. This was done two more times and average was taken and corresponding points were awarded to each of the

browser. Furthermore, CSS rendering [11] test was performed and time taken for the DOM to rewrite the contents of a DIV was recorded. Average values from three trails were considered for the browser evaluation.

58.3.4 Security

One of the SQL injection defence mechanisms was implemented [12] through a web application and every browser was tested against the attack. This application was hosted inside a LAN network and the bandwidth used to render the page was maintained constant (a bandwidth of 418 Gbps was maintained) for all browsers. A timer was used to calculate the defending time for each browser. The browser with the maximum time for defending was given the minimum points. Other attributes under consideration were DOM session storage, DOM local storage, cookie set, SQL injection, Browser Security, Settings and Pop-up blocking.

58.3.5 Video Streaming

In the recent, the first feature that tops every smart phone users' checklist is video streaming. Hence, we considered support of video streaming as a sole parameter and awarded five points if the browser supported it.

The points to other parameters like browser features, Plug in and Web Feed Support, Add-on, Accessibility Features were given based on the survey conducted on all their respective attributes given below.

- *Browser features:* Password and Download Managing, Bookmarks, Search toolbar
- *Plug in and Web Feed Support:* RSS, Flash, PDF Support
- *Accessibility Features:* Tabbed Browsing, Zoom, Full text History, Sharing, Smart address bar, speed dial, sync, user agent switching, themes and gestures.

58.4 Results

The results obtained inferred that Dolphin HD browser scores above all the other selected browsers in all the above categories as shown in Fig. 58.1. On the other hand, Mozilla Firefox browser which has scored second in the overall rating shows poor performance in the basic features due to lack of plug-in support and video streaming support. Further, Opera Mini is the weakest link in the browser feature

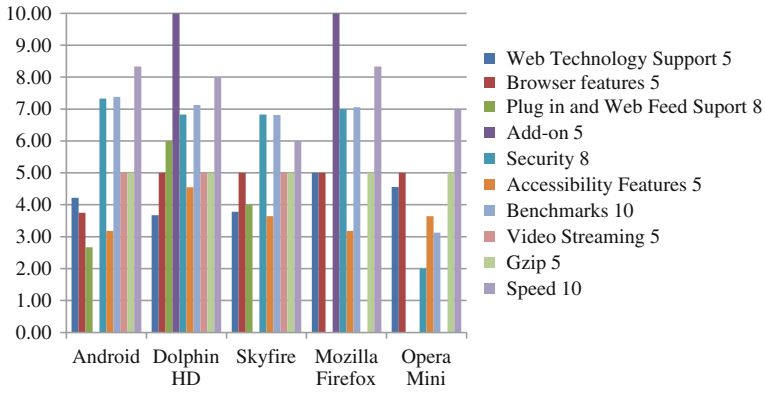


Fig. 58.1 Browser features, plugin and webfeed support, add-on, accessibility features, video streaming

which ranks below all the other browsers as it does not support plug-in, add-on and video streaming.

All the browsers had adequate basic security features required with Opera Mini as an exception. This is because, we were unable to conduct SQL injection test in Opera Mini as it doesn't allow browsing web pages from a local server under the default settings of the browser.

58.5 Conclusion

In this paper, we proposed the performance comparison of web browsers in Android platform. According to the results, Dolphin HD emerges as the best performance browser closely followed by Skyfire and Mozilla Firefox browser where the latter majorly lost the war in its video streaming incapability and plug-in support. It is notable that even though android browser stands in fourth place it's one of the stable and secure (tops in security) browser. Furthermore, Opera Mini in the fifth position lets us conclude that it is suitable for web browsing with data reduction, despite the fact that it does not support other major features like web technology and video streaming.

Further, due to these performance tests, we were also able to identify features that reduce the performance of the browsers (as indicated in Table 58.1). This finding paves the way for future developments in the functionalities of browsers to improve their performance. These can be implemented through separate add-ons for the browser which support it or its future version release for best possible performance.

Table 58.1 Parameter wise categorization of browser results

Parameters	Highest Scorer	Lowest Scorer
Web technology support	Mozilla firefox	Dolphin HD
Browser features	Dolphin HD, skyfire, mozilla firefox, opera Mini	Android
Plugin and webfeed support	Dolphin HD	Mozilla firefox, opera mini
Add-on	Dolphin HD, mozilla firefox	Android, opera mini, skyfire
Security	Android	Opera mini
Accessibility features	Dolphin HD	Android and mozilla
Benchmarks	Android	Opera mini
Video streaming	Android, dolphin HD, skyfire	Mozilla, opera mini

References

1. Global mobile statistics (2012) All quality mobile marketing research, mobile web stats, ad revenue, usage, trends. http://mobithinking.com/mobile-marketing-tools/latest-mobile-stats. Accessed on 8th April 2012
2. Stat Counter Global Stats (2012) http://gs.statcounter.com/#mobile_vs_desktop-ww-monthly-201101-201201. Accessed on 8th April 2012
3. Xianhua S, Zhenjun D, Rong C (2009) School of information science and technology. Research on mobile location service design based on android. Dalian Maritime University, Dalian
4. Mobile/Tablet Browser Market Share (2012) http://netmarketshare.com/browser-market-share.aspx?qprid=0&qpcustomd=1. Accessed on 28th March 2012
5. Android Apps on Google Apps. https://play.google.com/store?hl=en
6. TheACID3 test (2012) http://www.acid3.acidtests.org. Accessed on 18th April 2012
7. The HTML5 test (2012) http://www.html5test.comLast. Accessed on 18th April 2012
8. The CSS3 test (2012) http://www.css3test.com. Accessed on 18th April 2012
9. The sunspider javascript benchmark (2012) http://www.webkit.org/perf/sunspider/sunspider.htmll. Accessed on 18th April 2012
10. The web compatibility test for mobile browsers (2012) http://www.w3.org/2010/01/wctmb2/. Accessed on 18th April 2012
11. The CSS rendering (2012) http://nontroppo.org/timer/csstest.html. Accessed on 18th April 2012
12. William GJ Halfond, Jeremy V, Alessandro O A classification of SQL injection attacks and countermeasures

Chapter 59
FPGA Triggered Space Vector Modulated Voltage Source Inverter Using MATLAB/ System Generator®

L. A. Abishek Rajaraman, P. Ganesh, P. Geeth Prajwal Reddy and M. Senthil Kumaran

Abstract This paper involves the digital implementation of Space Vector Modulation (SVM) for a 3-phase Voltage Source Inverter (VSI). The System Generator (SG), which links both Xilinx and MATLAB, is used for constructing the Xilinx modules for the procedural implementation of the SVM. The SG converts the Matlab simulation of SVM done using Xilinx Blockset, into the corresponding Very High-Speed Integrated circuits Hardware Descriptive Language (VHDL) code. This VHDL code is compiled in Xilinx and the subsequent bit file generated, is loaded into the PROM of SPARTAN XILINX FPGA XCS3500e FG320. On execution of the bit file, the firing pulses are generated, which are applied to the VSI with Induction Motor Load.

Keywords Voltage source inverter · Field programmable gate array · System generator · Very high-speed integrated circuits hardware descriptive language

L. A. Abishek Rajaraman (✉) · P. Ganesh · P. Geeth Prajwal Reddy
SSN College of Engineering, Kalavakkam, Chennai, Tamil Nadu, India
e-mail: abishekrajaraman@gmail.com

P. Ganesh
e-mail: ganeshnet81@ymail.com

P. Geeth Prajwal Reddy
e-mail: geethprajwal@yahoo.com

M. Senthil Kumaran
Department of EEE, SSN College of Engineering, Kalavakkam, Chennai, Tamil Nadu, India
e-mail: senthilkumaranm@ssn.edu.in

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_59,

59.1 Introduction

Phase shifted Pulse Width Modulation (PWM) is one of the common triggering techniques used for three-phase VSI. However, in order to achieve lesser switching loss, reduced Total Harmonic Distortion (THD) and higher modulation index [1], significant research has been performed in PWM. This led to investigation into non-sinusoidal PWM techniques. Compared with sinusoidal PWM, non-sinusoidal PWM can increase the modulation index for line-to-line voltages. With the advent of non-sinusoidal PWM and Programmable Logic Devices, SVM technique evolved. SVM is an advanced control technique used in inverter and converter topologies. Nowadays, SVM is implemented in a cost effective way, using digital logic. FPGAs are used for this purpose because they are user reconfigurable and possess very high processing speed. Modern FPGAs are capable of operating at high clock frequencies and VHDL programs that require high execution rate can be easily implemented in the FPGA. SG is a utility in Matlab [2] that enables users to configure the system Simulink time according to the type of FPGA chosen. SG identifies the pin location of the clock in the FPGA and acts as an interface between Matlab and Xilinx by generating an equivalent VHDL code of the Matlab model made using the Xilinx Blockset of Simulink. The steps involved and methods used in the algorithm are explained in Sect. 59.3. Sections 59.4 and 59.5 presents simulation and experimental results.

59.2 Theory of Space Vector Modulation

59.2.1 Inverter Topology

Let us consider a Three-Phase VSI as shown in Fig. 59.1.

VSI is a power electronics topology that synthesizes, from a DC source, Three Phase Alternating Voltages, where amplitude, phase and frequency of the voltages are controllable. This VSI topology has eight realizable switching states [3]. At any instant, it is mandatory that only one switch in each leg of the VSI can be ON, to prevent short circuit of the source. Each switching state is represented by ON–OFF status (1-0) of the upper arm switches of the VSI, while the status of the lower arm switches would be their complement. When these Switching States are plotted onto the complex plane, they form a hexagonal pattern as shown in Fig. 59.2. The Three-Phase alternating voltage output of the VSI, using Clark transform [1], are represented uniquely by a rotating vector V_{REF} which is the vector sum of a pair of consecutive voltage switching state vectors, $V_0 \sim V_7$. The maximum amplitude of the reference vector, V_{REF}, such that over-modulation does not occur [4], is limited to 0.866 times the amplitude of the active switching state vector ($V_1 \sim V_6$).

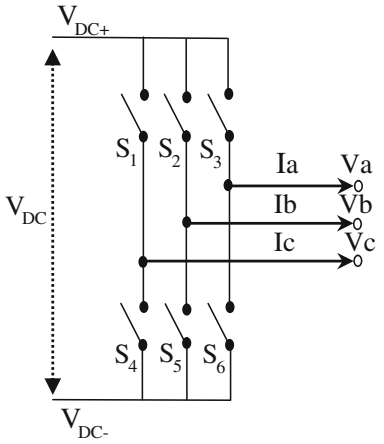


Fig. 59.1 Voltage source inverter topology

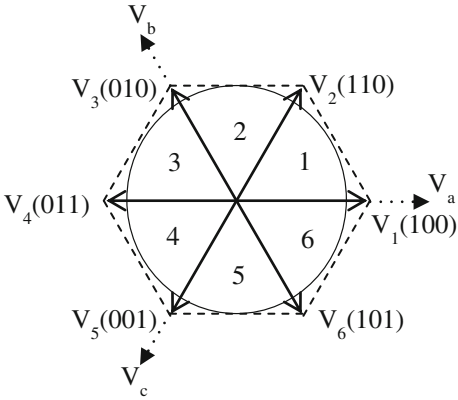


Fig. 59.2 VSI switching state locus

59.2.2 Reference Vector

Within a sector, V_{REF} is the vector sum of the adjacent vectors V_x and V_y with duty cycle δ_1 and δ_2 respectively as shown in Fig. 59.3.

The reference vector is expressed as

$$V_{REF} = \delta_1 . V_x + \delta_2 . V_y \tag{59.1}$$

While the active vectors ($V_1 \sim V_6$) are used to compute the direction of V_{REF}, the zero vectors (V_0, V_7) are used to alter the amplitude of V_{REF}. The zero vectors are applied with a duty cycle of δ_0. The duty cycles are calculated by

$$\delta_1 = m_v . \sin\left(\frac{\pi}{3} - \theta_r\right) \tag{59.2}$$

$$\delta_2 = m_v . \sin(\theta_r) \tag{59.3}$$

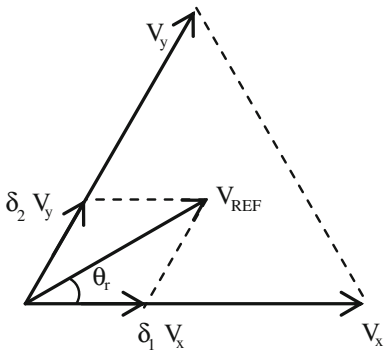


Fig. 59.3 Synthesis of reference vector V_{REF}

$$\delta_0 = 1 - (\delta_1 + \delta_2) \tag{59.4}$$

Where θ_r is the angle made by V_{REF} within the sector and m_v is the volt modulation index that defines the desired voltage transfer ratio as

$$m_v = \frac{\sqrt{3}.V_{REF}}{V_{DC}} \tag{59.5}$$

Thus based on the direction of V_{REF} in the complex plane, the sector and the corresponding pair of actives switching state vectors are determined. Once the duty cycles are calculated, a symmetric switching sequence is used to generate the gating signals. Thus active switching state vectors V_x, V_y and zero vectors are applied in the following sequence as shown in Fig. 59.4. The switching sequence is generated based on the PWM technique where a triangular wave, with frequency equal to the switching frequency of the VSI, is taken as the carrier wave and δ_1 and $\delta_1 + \delta_2$ are taken as the reference waves and pulses are generated by comparing the carrier and reference waves. The gating signals thus obtained from this algorithm are applied to the switches of the VSI to generate the sinusoidal Three Phase Voltages.

59.3 Digital Implementation of SVM Algorithm

59.3.1 Reference Angle Generation

As per the SVM algorithm, an equivalent rotating vector, V_{REF}, which varies from 0 to 2π during each cycle, represents the 3-Φ sinusoidal voltages. Since the m_v is constant, only the direction of the V_{REF} is required. This rotating vector is implemented using an up counter. The up counter has 'x' bits to count from 0 to 6.28 and 'y' binary point bits to provide adequate resolution. The clock period for the counter is set using the following empirical formula,

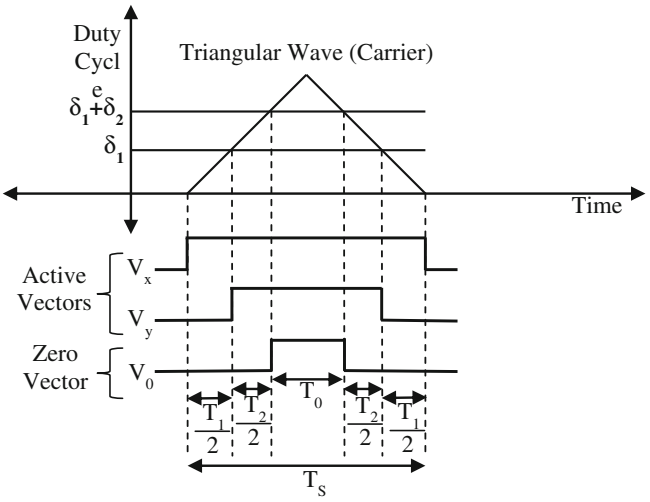


Fig. 59.4 Gating signals in symmetric switching

$$Required\ Time\ Period = \frac{Max\ Count\ value}{Resolution} \times Explicit\ Time\ Period \qquad (59.6)$$

Where count value, resolution and explicit time period are parameters of the blocks in the Xilinx Blockset [5].

For the required time period, the count value and resolution are chosen and an appropriate explicit period is determined.

59.3.2 Sector Number and Sector Angle

As shown in Fig. 59.2, there are six sectors present in the VSI switching state locus, each sector having a sector angle varying from 0 to $\frac{\pi}{3}$. Based on the direction of the V_{REF}, i.e. the current count value, the algorithm determines the sector number, S_n and the sector angle, θ_s.

59.3.3 Vector Pair Selection

In the VSI switching state locus, a pair of switching state vectors encloses each sector. With respect to the sector number of V_{REF}, the corresponding vector pair, V_x and V_y are deduced.

59.3.4 Gating Pattern Selection

Consider a switching state vector $V_1(100)$ whose sequence(100011) is treated as a binary number, 100011_2 and is represented by its equivalent decimal value, 35_{10}. Thus gating pattern of V_1 is given by 35_{10}. Similarly the gating pattern of all switching states, $V_0 \sim V_7$, are represented by their equivalent decimal values—$07_{10}, 35_{10}, 49_{10}, 28_{10}, 21_{10}, 28_{10}, 14_{10}, 56_{10}$ respectively. Based on the vector pair V_x and V_y, their corresponding gating patterns are selected and applied to the switches accordingly.

59.3.5 Duty Cycle Calculation

With reference to the Xilinx Reference Blockset of Simulink, CORDIC SINCOS block implements Sine and Cosine generator circuit to deduce the sine and cosine values of the input given to the block. As shown in Eq (59.2), (59.3) and (59.4), the values of the duty cycles δ_1, δ_2 and δ_0 are computed using CORDIC SINCOS block. A high frequency triangular wave is generated to act as the carrier wave in PWM and is compared with the values of δ_1 and $\delta_1 + \delta_2$ as shown in Fig. 59.4, to determine the sequence and time period for which switching state vectors, V_x, V_y and zero vector are applied.

59.3.6 Triggering Pulses

The PWM pulses are multiplied with their corresponding switching state vector's gating pattern and each bit of the resulting 6 bit-gating pattern, is applied as the gating signal to the corresponding switches of the VSI. Using the Slice block of the Xilinx Blockset of Simulink, the individual bits of the binary gating pattern are isolated.

59.4 Simulation

The MATLAB [6] model of the SVM algorithm implemented using Xilinx Blockset is show in Fig. 59.5. The simulation parameters are Input Voltage V_{in}: 230 V, Output line-to-line Voltage V_{out}: 230 V, Output Current I_{peak}: 9.2 A, Switching frequency: 10 kHz, Load Inductance: 27 mH, Load Resistance: 10Ω. Application of the triggering pulses, generated by SVM algorithm, to the switches of the VSI produced the following output waveforms as shown in Fig. 59.6 and Fig. 59.7. Using SG, the VHDL code is generated from the MATLAB model. This VDHL code

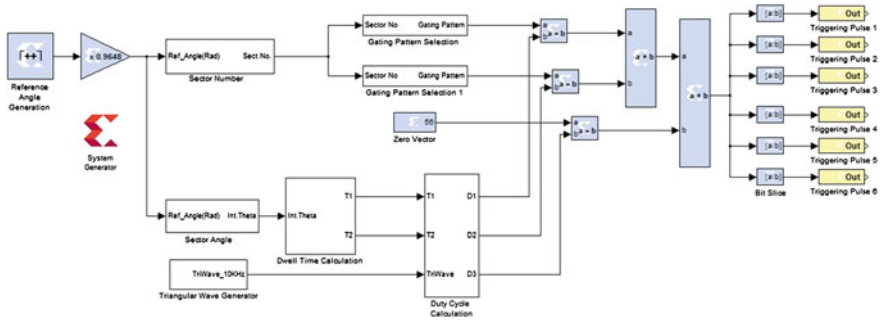


Fig. 59.5 MATLAB simulation block diagram

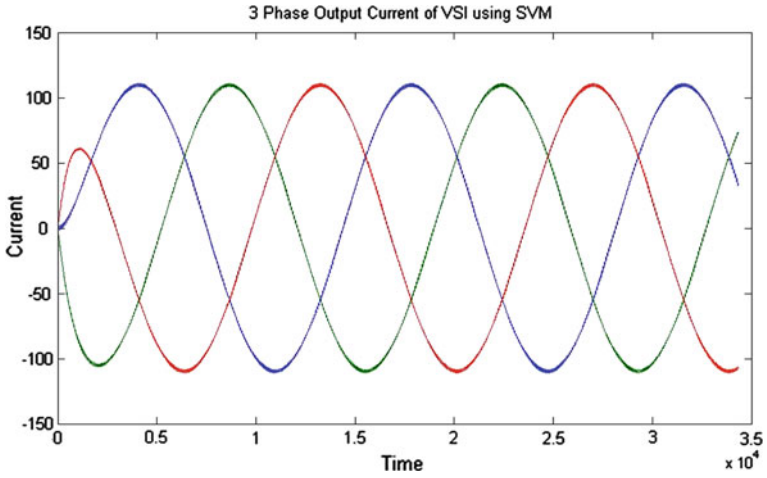


Fig. 59.6 3-Φ current output of VSI

is synthesized [7] and the design is implemented using ISE Xilinx software to generate the Bit file. The Bit file is loaded into the PROM of the FPGA.

The SVM algorithm is implemented using the FPGA hardware and the real time triggering pulses are generated. Fig. 59.8 shows the real time pulses in Agilent Digital Storage Oscilloscope (DSO). The triggering pulses generated by the FPGA are applied to a VSI with Induction Motor load and a 1-Φ current and voltage output waveform of the VSI are observed on the Agilent DSO. These waveforms are shown in Fig. 59.9.

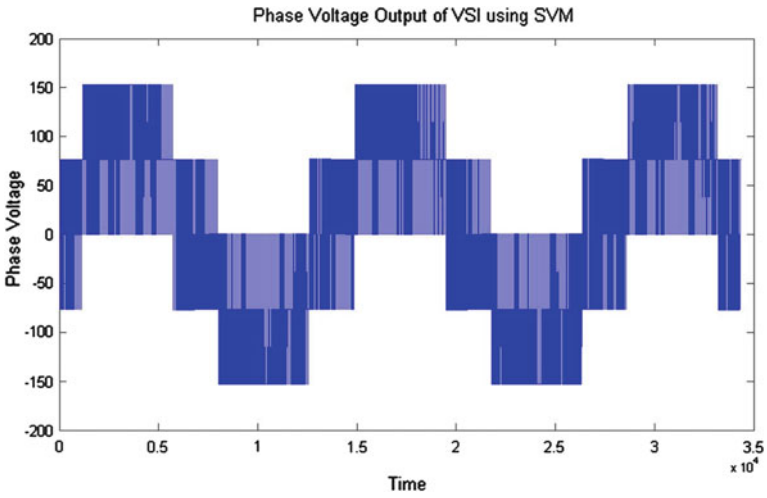


Fig. 59.7 Phase voltage output of VSI

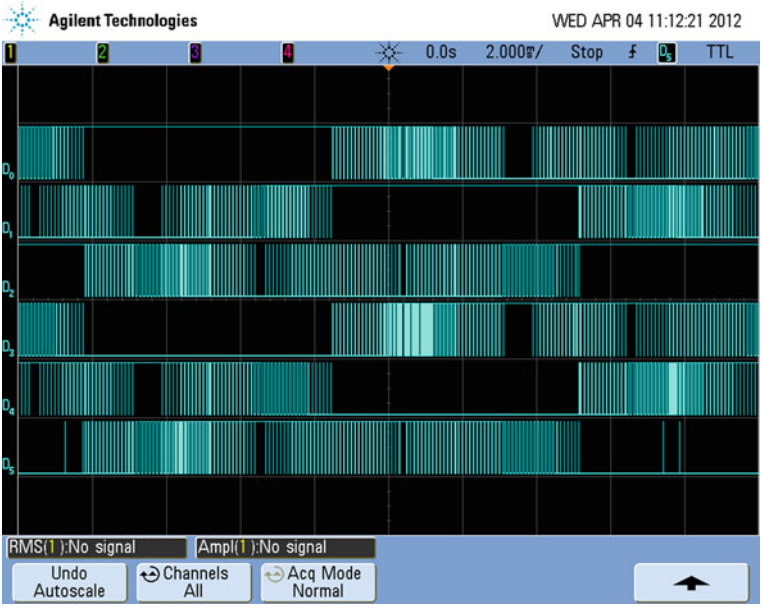


Fig. 59.8 Triggering pulses on agilent DSO

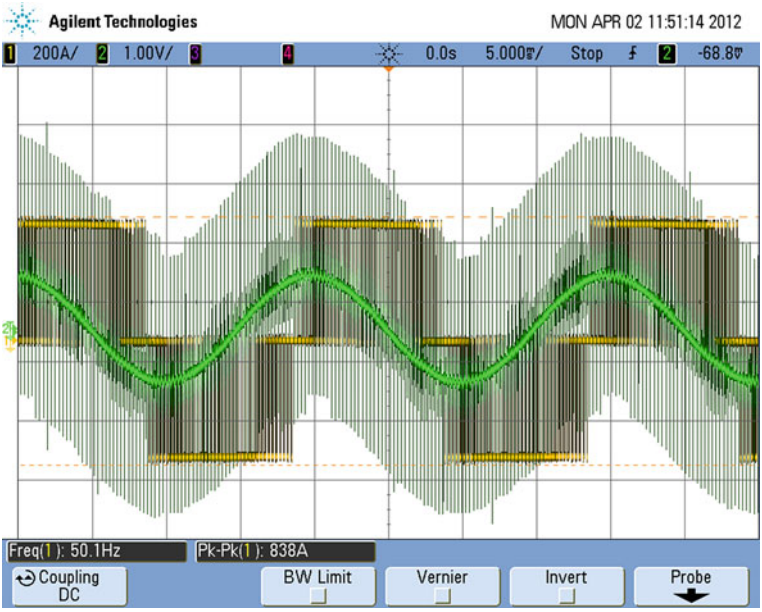


Fig. 9 1-Φ Current and voltage output of SVM triggered VSI on agilent DSO

Table 59.1 Comparison of results

	Matlab simulation result	Practical results
Line-to-Line voltages (peak)	460 V	400 V
Line current (rms)	1.1 A	0.8 A
Frequency	50 Hz	50.1 Hz

59.5 Simulation Results

On application of the 3-Φ output voltage of the inverter to the induction motor load, rated at 1HP, 1.8A, 50 Hz, 1430 RPM, the following results were tabulated(Table 59.1).

The Modulation Index for the simulation is set as 0.866.

59.6 Conclusion

The triggering pulses, generated from the digital implementation of SVM, applied to the switches of the VSI presents improved utilization and efficient conversion of the Direct Current (DC) source into Alternating Current (AC) output. SG helps in accurate building of an HDL algorithm with the help of appropriate blocksets and

generates the VHDL with minimum possible program lines, reducing need for brute-force coding methods. Since FPGA is a fast parallel processing unit, it can be programmed to implement motor speed control techniques such as Direct Torque Control (DTC), along with the SVM algorithm.

References

1. Dorin ON (2001) Space vector modulation-an introduction. In: Proceedings of IEEE/IECON, 2001, pp 1583–1592
2. Application guidelines—integrating xilinx system generator and simulink HDL coder
3. Jung J-W (2005) Space vector PWM inverter. In: DECE, The Ohio State University, Columbus
4. Bose BK (1986) Power electronics and AC drives, Prentice-Hall, New Jersey
5. System generator for DSP getting started guide
6. www.mathworks.com
7. Harrison CG, Jones PL, (1996) Xilinx FPGA design in a group environment using VHDL and synthesis tools. Digital system design using synthesis techniques (Digest No: 1996-029), IEE Colloquium On, pp 5/1–5/4 15 Feb 1996

Chapter 60
Face Recognition Using PCA and Bit-Plane Slicing

T. Srinivas, P. Sandeep Mohan, R. Shiva Shankar, Ch. Surender Reddy and P. V. Naganjaneyulu

Abstract The objective of the paper is face recognition using PCA and Bit plane slicing. It made a study on the dimensionality reduction on bit plane of images for face recognition. The proposed frame work would aid in robust design of face recognition system and addressed the challenging issues like pose and expression variation on ORL face database. It is in contrast to PCA on the image the design of PCA on bit plane reduces computation complexity and also reduces time. In the proposed frame work image is decomposed with the help of bit plane slicing, the feature have been extracted from the principle component analysis (PCA).

Keywords PCA · Bit-plane slicing · Feature extraction · Face recognition

T. Srinivas (✉) · P. Sandeep Mohan · R. Shiva Shankar
Sri Venkateswara Engineering College, Suryapet, India
e-mail: tipparthi@gmail.com

P. Sandeep Mohan
e-mail: psandeepp6@gmail.com

R. Shiva Shankar
e-mail: rssnits@gmail.com

Ch. Surender Reddy
R.R.S College of Engineering and Technology, Muthangi, India
e-mail: chsurender452@gmail.com

P. V. Naganjaneyulu
PNC and Vijai Institute of Engineering and Technology, Phirangipuram, India
e-mail: pvnaganjaneyulu@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_60,

60.1 Introduction

Basically face recognition can be used for verification and Identification. In the year 1988 Kirby and Sirovich applied PCA, a standard linear algebra technique for the face recognition problem. The technique was the landmark and considered as milestone because it requires only less than one hundred values in code a normalized face images accurately.

For the past few years, a systematic investigation has been going on to design a robust security/authentication mechanism. With the advent of miniaturized imaging systems the design process of security systems has been improved. The devices are application specific and present data (Biometric) to be incorporated into the design. Many researchers showed that the features extracted from face images aid in designing robust security/authentication systems. Successful face recognition system [1] is proposed utilizing Eigen face approach. This method is conventional, considers frontal and high contrast faces for implementing the system, but in real time faces may not be frontal and device intrinsic capture (illumination variation) properties pose difficulties in the process of detection. Thus in security and other computer vision applications, pose and variation in illuminations plays a critical role. The Eigen face approach is not satisfactorily addressing these problems.

In recent works [2–4], face recognition is carried out with PCA method and succeeded well, but it fails as input space increases and also suffers from the problem of discrimination between faces of similar persons like twins.

Face feature extraction suffers from

(a) Pose and expression variation,
(b) Resolution variation and
(c) Illumination problems

The methods designed using PCA [2–5] works well for either (a) or (b) but not on three issues altogether. Mainly in bio-metric home security applications, the above mentioned issues are obvious. Variant of principal component analysis is Kernel Principal Component analysis [5] (KPCA) and is nonlinear extension. In KPCA, input data is initially mapped into new feature space using non-linear mapping (kernel). PCA is performed on the kernel transformed data to extract feature vectors [2, 4]. The kernel mapping provides mechanism to address pose and expression variation (Figs. 60.1, 60.2).

The paper concentrated on face recognition by using bit-planes of a image. Section 60.2 explains about bit-plane Slicing.

To aid the process of recognition, nearest neighborhood classifier is used; this method finds an image to the class whose features are closest to it with respect to the Euclidean norm.

The performance of the proposed algorithm is verified on available databases on the internet, such as ORL face database [7] and YALE database [11]. ORL face database consists of 400 images of 40 individuals; each subject has 10 images in

Fig. 60.1 Sample Images of ORL database with different pose

Fig. 60.2 ORL face database

different poses. YALE face database consists of 5760 images of ten individuals in nine poses, and each pose in 64 illumination conditions. This paper has categorised into six sections. Section 60.2 is devoted to Bit-Plane Slicing, Sect. 60.3 Principle Component Analysis, Sect. 60.4 Nearest Neighborhood classifier, Sect. 60.5 Proposed Method, In Sect. 60.6 experimental results and discussions were compared the proposed method with PCA and showed accurate results on challenging databases and the Sect. 60.7 possible conclusion have been drawn out of technology.

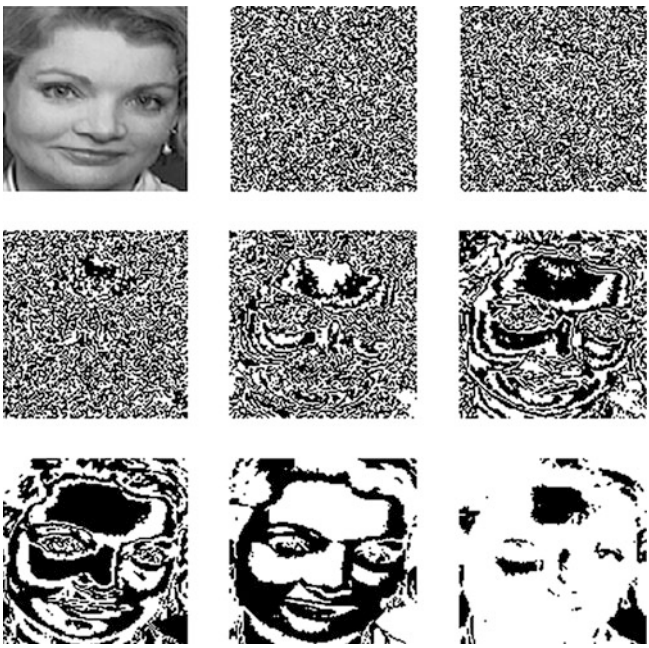


Fig. 60.3 8-bit planes of a image

60.2 Bit-Plane Slicing

Bit-Plane Slicing is a technique in which the image is sliced to different planes. It ranges from Bit level 0 which is the least significant bit (LSB) to Bit level 7 which is the most significant bit (MSB). The input of the method is an 8-bit per pixel image. This is a very important method in Image Processing.

The advantage of doing this method is to get the relative importance played by each bit of the image. It highlights the contribution made by specific bits. In this method, only in last 4 higher order bits planes significant data is visualized [7]. The lower level bit plane does not give much detail because they are made up of lower contrast. The bit level in bit plane 7 is equivalent to the bit level of the original image.The running time of the Bit-Plane algorithm for one image can range from 2 s to 1 min on a Pentium IV CPU using MATLAB code. Execution time will vary from one image to another, depending on the size of the image (Figs. 60.3, 60.4).

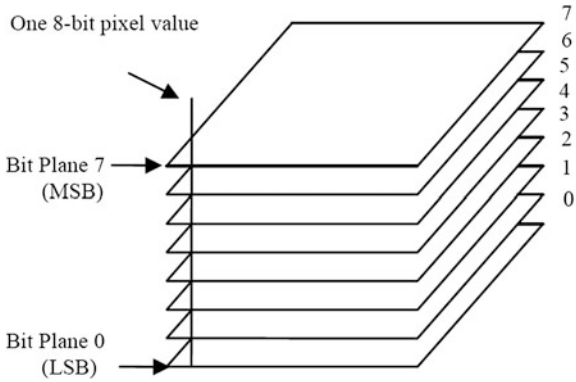


Fig. 60.4 Bit plane decomposition

60.3 Principle Component Analysis

A 2-D facial image can be represented as 1-D vector by concatenating each row (or column) into a long thin vector. Let's suppose we have M vectors of size N (= rows of image × columns of image) representing a set of sampled images. pj's represent the pixel values.

$$x_i = [p_1 ::: p_N]^T;\ i = 1, \ldots, M \tag{60.1}$$

The images are mean centered by subtracting the mean image from each image vector. Let m represent the mean image.

$$m = \frac{1}{M}\sum_{i=1}^{M} x_i\,(2) \tag{60.2}$$

And let w_i be defined as mean centered image

$$w_i = x_i - m \tag{60.3}$$

Our goal is to find a set of e_i's which have the largest possible projection onto each of the w_i's. We wish to find a set of M orthonormal vectors e_i for which the quantity

$$\lambda_i = \frac{1}{M}\sum_{n=1}^{M} (e_i^T w_n)^2$$

is maximized with the orthonormality constraint

$$e_l^T e_k = \delta_{lk}$$

It has been shown that the e_i's and , λ_i's are given by the eigenvectors and eigenvalues of the covariance matrix

$$C = WW^T \tag{60.4}$$

Where W is a matrix composed of the column vectors w_i placed side by side. The eigenvectors corresponding to nonzero eigenvalues of the covariance matrix produce an orthonormal basis for the subspace within which most image data can be represented with a small amount of error. The eigenvectors are sorted from high to low according to their corresponding eigenvalues. The eigenvector associated with the largest eigen value is one that reflects the greatest variance in the image. That is, the smallest eigen value is associated with the eigenvector that finds the least variance. They decrease in exponential fashion, meaning that the roughly 90 % of the total variance is contained in the first 5–10 % of the dimensions.

A facial image can be projected onto M' ($< M$) dimensions by computing

$$\Omega = [v1, v_{2\ldots}v_M]^T \tag{60.5}$$

60.4 Euclidean Classifier

Different distance metrics in 2-D are the Cityblock Distance between $(\mathrm{x1}, \mathrm{y1})$ and $(\mathrm{x2}, \mathrm{y2})$ is $|x1 - x2| + |y1 - y2|$. The CHESSBOARD DISTANCE is $\max(|x1 - x2|, |y1 - y2|)$. The QUASI-EUCLIDEAN DISTANCE [18] can be written as $\mathrm{abs}(\mathrm{x1} - \mathrm{x2}) + (\sqrt{2} - 1) * \mathrm{abs}(\mathrm{y1} - \mathrm{y2})$, in the case of $\mathrm{abs}(\mathrm{x1} - \mathrm{x2}) > \mathrm{abs}(\mathrm{y1} - \mathrm{y2})$ the QUASI-EUCLIDEAN DISTANCE written as, $(\sqrt{2} - 1) * \mathrm{abs}(\mathrm{x1} - \mathrm{x2}) + \mathrm{abs}(\mathrm{y1} - \mathrm{y2})$, the EUCLIDEAN DISTANCE can be written as $\sqrt{(x1 - x2)^2 + (y1 - y2)^2}$. In this work we have used nearest neighborhood classifier to recognize the image. This classifier comes under minimum distance classifiers. It is also called as Euclidean classifier. In this method the minimum the distance from test feature vectors to train feature vectors the correct the image is. If $\mathrm{X_i}$, $\mathrm{Y_j}$ represents test and train image features then

$$\|\mathrm{X_i} - \mathrm{Y_i}\| \equiv \sqrt{(\mathrm{X_i} - \mathrm{Y_i})^{\mathrm{T}}(\mathrm{X_i} - \mathrm{Y_i})} < \left\|\mathrm{X_i} - \mathrm{Y_j}\right\| \tag{60.6}$$

*Where‖. ‖ represents Euclidean norm

Because of its simplicity, it finds an image to the class whose features are closest to it with respect to the Euclidean norm.

60.5 Proposed Method

The proposed method starts with splitting the ORL database into two sections. One section is called as test database and other is train database. The test database consists of 200 images of all 40 subjects, in the similar way train database consists of 5 subjects each for 40 subjects. The algorithm is checked by picking a image

Table 60.1 Recognition rate versus number of eigen vectors

Features	Recognition Rate (%)
1	11.5
3	56.5
5	72.5
10	85.5
20	94
50	94
100	94

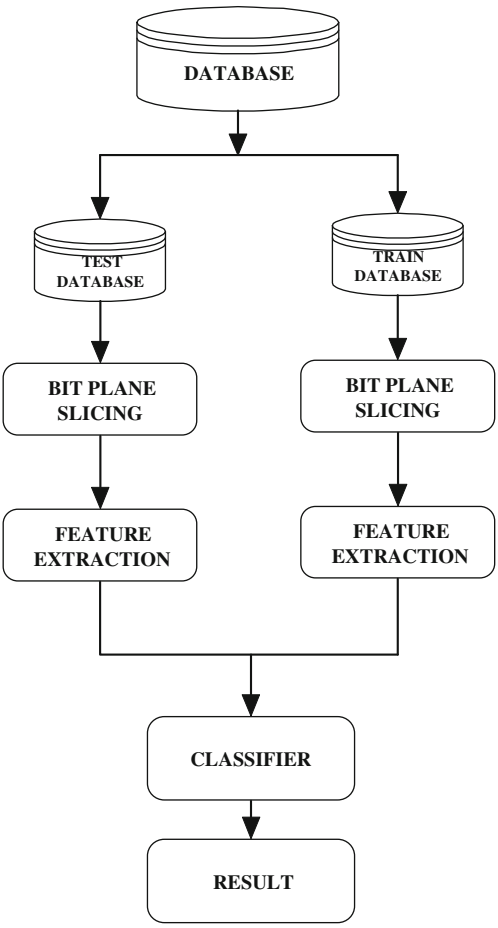


Fig. 60.5 Proposed algorithm

from test database and finding its nearest image from train database. After splitting the database next step is decomposing the image into 8 bit-planes. In previous works there is dimensionality reduction on image it self. In our novel method we applied the dimensionality reduction step after decomposing image into its bit-planes.

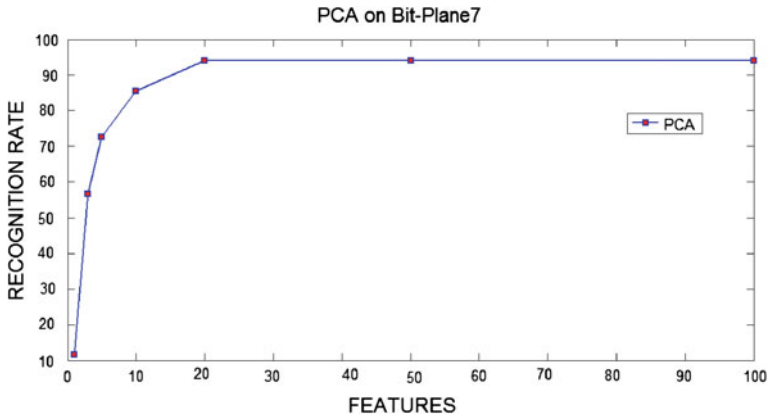


Fig. 60.6 Recognition rate versus features

60.6 Experimental Results

The paper results evaluated by using recognition rate parameter. The recognition parameter is calculated as ratio of no of successful attempts of algorithm to total number of attempts. All 200 test images are tested and recognition rate is calculated. The final recognition rates by varying the features are listed below. From the table it is observed that as the feature size increasing the recognition rate increases. It is also concluded that the optimum feature size for this algorithm is 20, because if the feature size increased beyond 20, the recognition rate is same and fixed.

60.7 Conclusion

The paper proposed a novel approach for face recognition by extracting features from a constant illumination and pose-variant image by using Bit-Plane Slicing and PCA method. All the images have been cropped from uneven size to 128 × 128 pixels. As a next step 7th Bit-Plane of image is mapped from input space of data to feature space is done by PCA method [4].

PCA is a powerful linear model for extracting non-linear features, the nearest-neighbor distance classifier which can enhance the recognition process.

Experimentation was done on ORL face database. Compared to existing approach [2], with marginal increase in computational cost and time, high recognition rate is reported in this paper. The Computed results were tabulated in Table 60.1 and graphically shown in Fig. 60.5. The experimental results in Table 60.1 shows that as the feature space increases as more than 10, the performance increases but at the same time cost of the algorithm increases. We are working out to apply the proposed method to deal with illumination variation, which is our future work remains (Fig. 60.6).

References

1. Patil AM, Kolhe SR, Patil PM (2009) Face recognition by PCA technique. In: Proceedings of the second international conference on emerging trends in engineering and technology, ICETET 2009, pp 192–195
2. Kokiopoulou E, Saad Y (2004) PCA and kernel PCA using polynomial filtering: a case study on face recognition
3. Meedeniya DA, Ratnaweera DAAC (2007) Enhanced face recognition through variation of principle component analysis (PCA). In: Second international conference on industrial and information systems, ICIIS 2007, pp 347–352
4. Kwang KI, Jung K, Kim HJ (2002) Face recognition using kernel principal component analysis. IEEE Signal Process Lett 9(2):40–42
5. Poon1 B, Ashraful Amin2 M, Yan H (2009) PCA based face recognition and testing criteria. In: Proceedings of the eighth international conference on machine learning and cybernetics, pp 2945–2949
6. AT&T Laboratories Cambridge, The Database of Faces, formerly ORL face database. Available at www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html
7. Sirovich L, Kirby M (1987) A low dimensional procedure for characterization of human faces. J Optical Soc Am A 4(3) 519–524
8. Punnam Chandar K, Mahesh Chandra M, Raman Kumar M, Swarna Latha B (2011) Preprocessing using SVD towards illumination invariant face recognition. In: Proceedings of IEEE RAICS 2011
9. Wu SQ, Wei LZ, Fang ZJ, Li RW, Ye XQ (2007) Infrared face recognition based on blood perfusion and sub-block DCT in wavelet domain. In: International conference on wavelet analysis and pattern recognition, vol 3, p 1252
10. Turk M, Pentland A (1991) Eigen faces for recognition. J Cognit Neurosci 3(1):71–86
11. Belhumeur PN, Hespanha JP, Kriegman DJ (1997) Eigenfaces versus fisherfaces: recognition using class specific linear projection. IEEE Trans Pattern Anal Mach Intell 19(7):711–720
12. Bartlett MS, Movellan JR, Sejnowski TJ (2002) Face recognition by independent component analysis. IEEE Trans Neural Netw 13(6):1450–1464
13. Moghaddam B, Nastar C, Pentland A (1996) A Bayesian similarity measure for direct measure for direct image matching. Proc Int Conf Pattern Recognit 2:350–358
14. Moghaddam B, Wahid W, Pentland A (1998) Beyond eigenfaces: Probabilistic matching for face recognition. In Proceedings of IEEE international conference on automatic face and gesture recognition, pp 30–35
15. Lanitis A, Taylor CJ, Cootes TF (1995) Automatic face identification system using flexible appearance models. Image Vis Comput 13(5):393–401
16. Shashua A, Riklin-Raviv T (2001) The quotient image: class-based rerendering and recognition with varying illuminations. IEEE Trans Pattern Anal Mach Intell 23(2):129–139
17. Zhang L, Samaras D (2003) Face recognition under variable lighting using harmonic image exemplars. In Proceedings of IEEE conference computer vision and pattern recognition, vol 1, pp 19–25
18. Chien J-T, Wu C–C (2002) Discriminant wavelet faces and nearest feature classifiers for face recognition. IEEE Trans Pattern Anal Mach Intell 24(12):1644–1649

Chapter 61
Operational Analysis, Performance Evaluation and Simulation of Solar Cell Powered Embedded EZ-Source Inverter Fed Induction Motor

K. C. R. Nisha and T. N. Basavaraj

Abstract This paper presents the operational analysis and performance evaluation of solar cell powered embedded EZ-source inverter fed induction motor. Embedded EZ-source inverter (EZSI) produces the same voltage gain as Z-source inverter (ZSI) but due to the DC sources embedded within the X-shaped impedance network it has the added advantage of inherent source filtering capability and also reduced capacitor sizing. This is attained without any extra passive filters. These advantages are significant for applications like photo electric and wind electric system. The operational analysis and simulation results exemplify that an EZSI is the most promising technique for renewable energy applications in order to reduce the overall system complexity and thereby improving the inverter efficiency.

Keywords EZ-source inverter · Harmonics · Shoot- through · Z-source inverter

61.1 Introduction

Nowadays, renewable energy applications are on greater demands, more particularly solar cell. A key component of PV generating system is the grid connected inverter. System performance depends on local climate, the orientation and inclination of PV array and inverter performance. The traditional photo electric

K. C. R. Nisha (✉)
Sathyabama University, Chennai, India
e-mail: nishashaji2007@gmail.com

T. N. Basavaraj
New Horizion College of Engineering, Bangalore, India

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_61,

systems contain VSI and CSI. They are either buck or boost, but not buck-boost converter [1].The common problem of this topology is that their main circuits cannot be interchangeable and also shoot through will occur when any two switches of the same phase leg is turned on which is a major killer to converter's reliability. Z-source inverter first proposed in paper [2] provides a feasible single stage power conversion concept. Z-source inverter is suitable in grid connection of alternative energy sources such as photo electric system as they usually produce low variable DC voltage.

As time progresses, developments related to Z-source inverters have elevated along many directions, covering its modulation [3], control [4] and other topological inventions. But unfortunately, a closer view at the existing network would reveal that it causes chopping current to be drawn from the source, if no explicit hardware filter is added. This chopping current not only raises the semiconductor current rating, but also complicates the maximum power point tracking (MPPT) objective set for most renewable energy sources [5]. In view, a new class of Z-source topologies, named as the embedded EZ-source inverters, was proposed in [6, 7], which however mainly focused on design of voltage and current type EZSI. A new single stage power conversion concept with implicit source filtering and reduced capacitor sizing for renewable energy applications is presented. This work also proposes EZSI for the control of 3 phase induction motor applied to solar electric systems.

61.2 Operating Principle of Embedded EZ-source Inverter

Two level voltage-type EZ-source inverter is shown in Fig. 61.1. It has its DC sources embedded within the X-shaped LC impedance network for filtering the currents drawn from the two DC sources of Vdc/2. Though the arrangement can occasionally interpret to a slightly higher cost, but the advantages exhibited by the EZ-source inverter outweigh the serious limitations [6]. These advantages are more clearly illustrated by analyzing the inverter operating principle. Based on switching states of the inverter, the EZSI can be classified into three modes.

Mode: **1** Inverter bridge is operating in one of six active states. In this mode the front-end diode D is forward biased and the Inverter bridge and external load is replaced by a current source. The capacitor is charged and energy flows to the load through the inductor. The inductor discharges in this mode.

Mode: **2** Inverter bridge is operating in any one of the two zero states as the inverter short circuits the load through either upper or lower three switching devices. The bridge can be viewed as an open circuit. The voltage of DC source appears across the inductor and capacitor but no current flows to load from DC source.

Mode: **3** The inverter is in one of the seven different ways of shoot through. In this mode the inverter bridge is short through and the diode D is reverse biased. In this mode no voltage appears across the load like in zero state operation, but the DC voltage of capacitor is boosted to required value based on shoot through duty ratio.

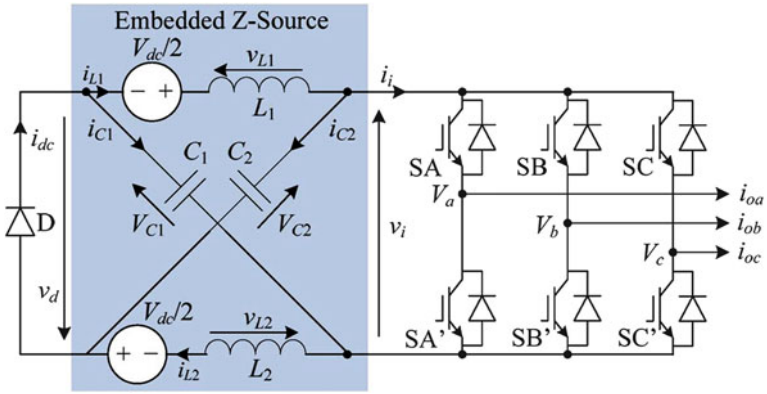


Fig. 61.1 Embedded EZ-source inverter circuit

61.2.1 Circuit Analysis

Assuming that the inductors L_1, L2 and capacitors C_1, C_2 have the same inductance (L) and capacitance(C) respectively, the E Z-source network becomes symmetrical.

$$vL1 = vL2 = vL; vC1 = vC2 = vC \tag{61.1}$$

Shoot-Through State ($S_x = S_{x1}$ = ON, x = A, B or C; D = OFF; time interval: T0)

$$vL = vC + \frac{vdc}{2}; vi = 0; vd = vD = -2vC \tag{61.2}$$

Nonshoot-Through State ($S_x \neq S_{x1}$, x = A, B or C; D = ON; time interval: T_1)

$$vL = \frac{vdc}{2} - vc; vd = vdc; vD = 0; vi = vc - vL = 2vC \tag{61.3}$$

Averaging the inductor voltage to zero, the capacitor voltage $v_{C,}$ peak DC link voltage v_{i1} and peak AC output voltage v_{x1} can be derived as:

$$\left\{ \begin{array}{l} vC = \dfrac{Vdc/2}{1 - 2T0/T} \\ vi1 = \dfrac{Vdc}{1 - 2T0/T} = BVdc \\ vx1 = \dfrac{MVdc}{2((1 - 2T0/T)} = B\left(\dfrac{MVdc}{2}\right) \end{array} \right\} \tag{61.4}$$

Where M is the modulation index ($M \leq 1$) and B is the boost factor ($B \geq 1$).

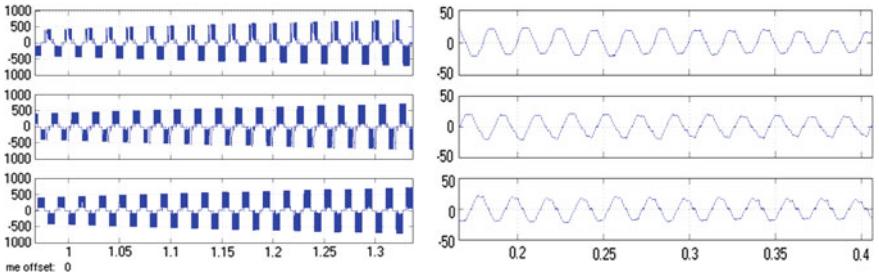


Fig. 61.2 Line voltage and current waveforms when M = 0.71 and $T_0/T = 0.41$

Analysis clearly shows that EZ-source inverter system produces same transfer gain as Z-source inverter system even though EZSI has two DC sources embedded within the impedance network for achieving source filtering. Also the capacitor voltage in (61.4) is greatly reduced when compared to Z-source inverter inferring that there is a significant reduction of capacitor voltage rating.

61.3 Simulation Results

To assess the operational and performance analysis of the solar cell powered EZSI fed induction motor, simulation model has been established using Matlab/Simulink package. The simulation is done up with the following parameters: $L_1 = L_2 = 2$ mH; $C_1 = C_2 = 2200$ μF; The purpose of the system is to control a 3-phase induction motor powered by solar cell whose input voltage is 110 V. Line voltages and currents are shown in Fig. 61.2.The spikes in the output voltage are due to the PWM switching pulses. The currents are smoothened by the inductance of the machine. Therefore the current harmonics are reduced. The rotor speed increases and the rotor settles at 1460 rpm and is shown in Fig. 61.3. Fourier analysis is done for both current and voltage.The spectrum is shown in Fig. 61.4.The THD value for current and voltage is 4.86. and 4.90 % respectively.

61.4 Conclusion

This paper has provided operational analysis, performance evaluation and simulation of EZ-source inverter fed induction motor powered by solar electric cell. The novel idea of this work is to control the 3-phase induction motor using embedded EZ-source concept for photo electric applications. The results of digital simulation are presented. Boosting is done with the help of Z-network and one leg of the inverter. Analyses show that EZ-source inverter produces the same voltage gain as their traditional Z-source inverter. The second advantage introduced by

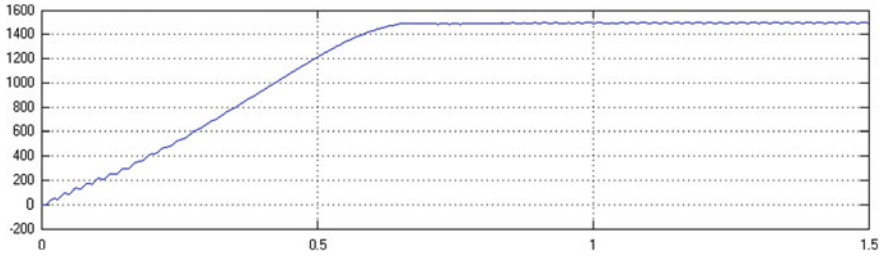


Fig. 61.3 Rotor speed in Rpm

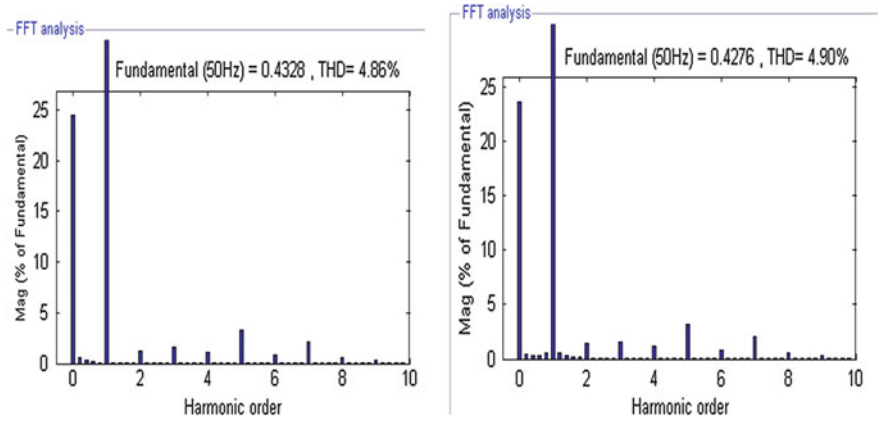


Fig. 61.4 THD for current and voltage waveforms

embedding the two sources is significant reduction in capacitor sizing. The additional advantages like lower voltage/current stresses and implicit source filtering are gained without requiring extra hardware. So EZ-source inverter system is a competitive secondary that can be used for solar energy harnessing applications where implicit source filtering is critical. The disadvantage of EZ-source inverter system is that, it requires two DC sources. Simulation results have also confirmed that EZ-source inverter system has improved harmonic performance than their counterparts.

References

1. Zope PH, Patil AJ, Somkuwar A (2010) Performance and simulation analysis of single-phase grid connected PV system based on Z-source inverter. In: IEEE conference PEDES 2010
2. Peng FZ (2003) Z-source inverter. IEEE Tran Ind Appl 39(2):504–510
3. Loh PC, Vilathgamuwa DM, Lai YS, Chua GT, Li Y (2005) Pulse-width modulated Z-source inverter. IEEE Trans Power Electron 20(6):1346–1355

4. Jung JW, Keyhani A (2007) Control of a fuel cell based Z-source converter. IEEE Trans Energy Convers 22(2):467–476
5. Gao F, Loh PC, Li D, Blaabjerg F (2011) Asymmetric and symmetric embedded EZ-source inverters. IET Power Electron 4(2) 181–193
6. Loh PC (2010) Embedded EZ-source inverters, IEEE Trans ind Appl 46(1):256–267
7. Gao F, Loh PC, Blaabjerg F, Gajanayake CJ 2008 Operational analysis and comparative evaluation of embedded Z-source inverters. In: Proceedings of IEEE power electronics specialists conference, pp 2757–2763

Chapter 62
Advanced DSP Based PLC Modem Over DC Lines for Real-Time Remote Monitoring of PV Plant Parameters

Atul Gupta, Venu Uppuluri Srinivasa, Devendra Paranjape and Nikhil Kashyap

Abstract Photovoltaic plants are becoming an inevitable option to meet the present energy requirements. This gives rise to the need of a smart monitoring system which can remotely provide real-time reliable information of each and every panel to the plant operators. Real-time monitoring at such a vast scale can be a great boon to the investors in improving the efficiency of photovoltaic (PV) plants as it can help to predict some basic faults in the individual panels, monitor their performance and suggesting corrective maintenance actions right away. In this paper, a remote monitoring system design using highly efficient low cost DSP's is proposed which implements Power Line Communication (PLC) concept over DC lines. It minimizes the use of any significant analog circuitry and discards the use of any separate dedicated cables for communication, thus making the system robust and significantly cheaper.

Keywords PLC · DSP · FSK · PV · THD · EMI distortion · Smart monitoring

A. Gupta (✉) · V. U. Srinivasa
Santerno India Design Center, Viman Nagar, Pune, Maharastra, India
e-mail: atulgupta2006@gmail.com

V. U. Srinivasa
e-mail: venuuppuluri@gmail.com

D. Paranjape · N. Kashyap
Department of EEE, BITS Pilani, K.K. Birla Goa Campus, Goa, India
e-mail: pddev4@gmail.com

N. Kashyap
e-mail: nikhilkashyap17edu@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_62,

62.1 Introduction

As the worldwide interest is shifting towards sustainable energy production, a lot of investment is being made into solar energy systems making photovoltaic plants a major area of investment in the energy sector today. In order to get high returns, the PV plants must perform at their peak efficiency. However, most of the PV plants suffer from critical performance issues pertaining mainly to their vast sizes. For instance, the maximum peak power point of a solar panel could decrease significantly due to dust particles settling on panels, impact of wind on panel's temperature, sudden shadowing, presence of clouds in some area of the plant, failure of a particular element of the plant, solar shading, inter row shading and many more. Therefore, a robust and efficient monitoring system which can provide us with values of different parameters of each solar panel, string inverter, grid etc. can vastly improve the performance of the plant. Apart from improving the amount of power generation the precise information of panel, inverters and grid parameters can also help the investor to estimate the cost involved in power generation accurately [1, 2].

In the proposed system, individual panels in a solar array can transmit diagnostic data over the existing DC power lines to an Array Diagnostic Unit Controller (ADUC) which can probe specific diagnostic information to a data monitoring console. This system makes use low cost DSPs and uses existing DC power lines as a communication channel which makes the design compact and cheap.

62.2 Circuit Design and Topology

The remote monitoring system discussed in this paper is applicable for PV plants employing String/Central inverters wherein a large number of PV panels arranged in series–parallel combination connect to a String/Central inverter (Fig. 62.1). Each of the panels and the String/central inverter in the PV plant has a Microcontroller Unit (MCU) mounted on it in a Slave and Master configuration. The Slave MCU updates itself with different parameters (like voltage, current, ambient & panel temperature, humidity etc.) values through sensor box present on each solar panel. The Master MCU communicates with the Slave MCU over existing DC power lines using Frequency Shift Keying (FSK) within frequency range of 100 kHz and data rate between 0 and 10 kbps following Narrowband PLC standards (IEC 61344, ANSI/EIA 709.1, 0.2, UPB as referred in Table 62.1) suitable for the control and command application. A C# (C Sharp) based GUI communicates serially with the master controller to show diagnostic data at the receiver's end using RS485 protocol which is very suitable for working under hard ambient and noisy conditions. The proposed system minimizes the use of hardware by exploiting TI's TMS320F28027 Piccolo MCU's peripherals which include the Serial Communication (SCIA) & Pulse Width Modulation (PWM) modules for

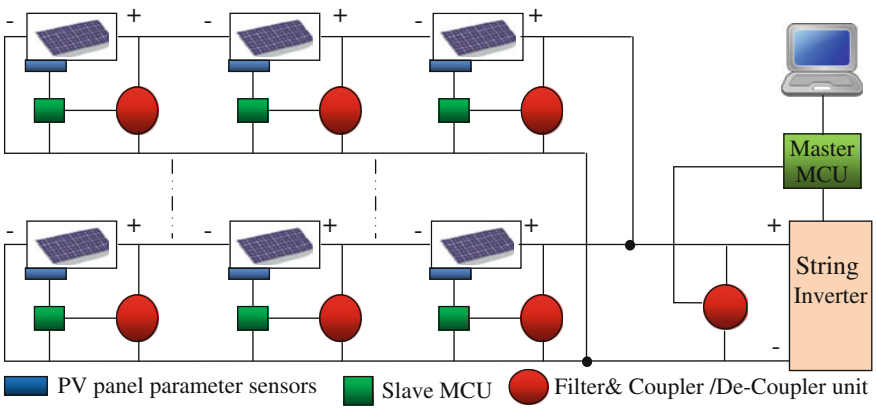


Fig. 62.1 Block diagram of proposed system

designing the Modulation circuit, Digital filter library & Comparator module for designing the Demodulation circuit.

The experimental results provided in this paper are for Space frequency, $F_0 = 18$ kHz & Mark frequency, $F_1 = 21$ kHz with baud rate of 400 bps for testing purposes. However, sine waves corresponding to space and mark frequencies in the range of 100 kHz will be used in the final design of the communication system.

62.3 Design Implementation

62.3.1 Modulation

This system implements FSK modulation (Figs. 62.2 and 62.3). The information data sent through GUI is received by the MCU via a UART port. Inside MCU, the received data is further converted into data packets and depending upon the value of data bits being '0' or '1', a PWM waveform is generated with 50 % duty cycle for the particular space & mark frequencies giving us a square wave FSK waveform as the output. Two such waveforms are generated with an optimized phase difference (as discussed in Table 62.2), superimposed and passed to a sallen-key low pass multi feedback filter to get a single frequency sinusoidal FSK waveform. A number of significant factors like Electromagnetic Interference (EMI), Total Harmonic Distortion (THD), etc. have been taken into consideration for generating sine waves from PWM.

$$H(f) = \frac{-\frac{R_2}{R_1}}{(j2\pi f)^2(R_2R_3C_1C_2) + j2\pi f\left(R_3C_1 + R_2C_1 + \left(\frac{R_2R_3C_1}{R_1}\right)\right) + 1}. \qquad (62.1)$$

Table 62.1 PLC technology classification based upon the data rate

	Low data rate	High data rate
Data rate	0–10 kbps	>1 Mbps
Modulation	BPSK,FSK, S-FSK, QAM	PSK + OFDM
Standards	IEC61344, ANSI/EIA 709.1,.2 , UPB	G.hn, IEEE 1901
Applications	Control and command applications	Broad-band via PLC,

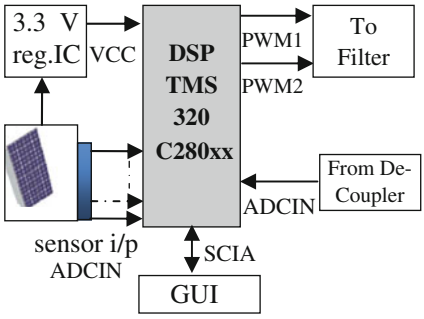


Fig. 62.2 DSP control board interface

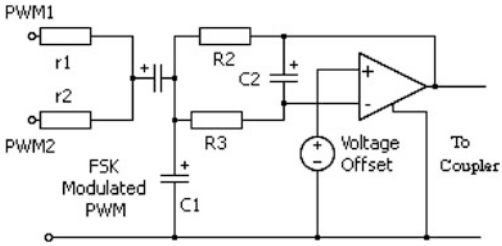


Fig. 62.3 Sallen-key low pass multi feedback filter circuit

$$Q = \frac{\sqrt{R_1 R_2 C_1 C_2}}{R_2 C_1 + R_3 C_1 (1 - K)}. \tag{62.2}$$

Equation (62.1) refers to the transfer function of the sallen-key multi feedback low pass filter and (62.2) refers to the Q-point wherein 'K' is the Gain factor of the filter. An optimization of the transfer function with respect to the Q-point and the cut-off frequency has been carried out subject to the minimum and maximum constraints of both these parameters for selecting mark & space frequencies symmetrically around the Q-point [3, 4].

$$f(t) = \frac{4}{T} \sum_{n=1}^{\infty} \sin\left(\frac{n\pi}{2}\right) \frac{\sin\left(\frac{n\pi x}{T}\right)}{\frac{n\pi x}{T}} \sin\left(\frac{2n\pi}{T} t\right). \tag{62.3}$$

Equation (62.3) represents the Fourier series of the added PWM signals wherein 'T' is the period and 'x' is the pulse width as shown in Fig. 62.4. The equation can

Table 62.2 THD's of FSK modulated signals at different values of 'x'

Phase diff.	x/T	THD	
		Without filtering effect	With filtering effect
144	0.1	0.4155	0.004275
108	0.2	0.3722	0.002562
72	0.3	0.362	0.0007017
60	**0.33**	0.3157	**0.0003064**
46.8	**0.37**	**0.2903**	0.0005937
36	0.4	0.3029	0.0009663
18	0.45	0.3718	0.001436

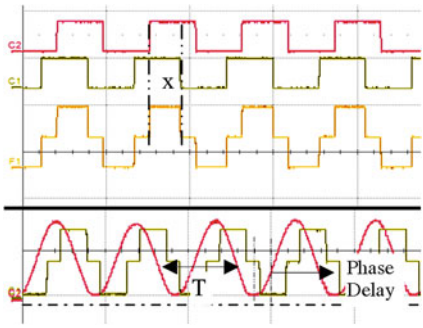


Fig. 62.4 (*Top to Bottom*) PWM signals, super-imposed PWMs, FSK sine wave signal

be split into the fundamental quantity and the harmonics to obtain the THD of the wave.

$$THD = \frac{\int_0^t [f(t, n=1)]^2 dt}{\sum_{n=2}^{\infty} \left(\int_0^t [f(t,n)]^2 dt \right)}. \tag{62.4}$$

Theoretically, it has been proved that the optimum value of 'x' is 0.37T. However, using (62.1), (62.3) and (62.4) it has been observed that the optimum value of 'x' is 0.33T when the transfer function of the filter is taken into account as confirmed by experimental data shown in Table 62.2.

62.3.2 Coupling/De-Coupling Circuit

The FSK sinusoid signal is fed to a coupling (High Frequency Transformer) circuit (as shown in Figs. 62.5 and 62.6) which superimposes the data on the DC signal via a capacitor at the transmitter's end and blocks the DC voltage [5, 6]. The coupling circuit makes use of a transformer to minimize the noise which occurs on account of common grounding. The circuit delivers maximum signal power using impedance matching circuit at a resonance frequency of 19.5 kHz.

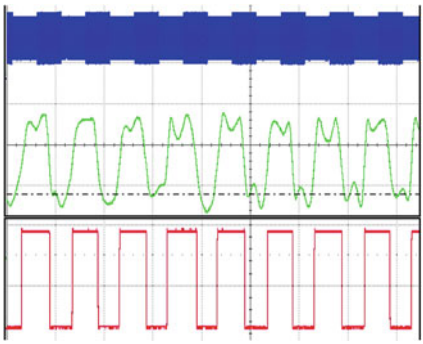


Fig. 62.5 C1: DC power line (15 V) coupled with FSK sinusoidal signal

62.3.3 Demodulation

The Demodulation technique used here exploits the multiplication property of two sine waves [7] as shown in Fig. 62.7 and can be implemented on DSP using RTOS [8].

The received modulated signal '*S(n)*' with a bit period '*Te*', is multiplied with its delayed form '*S(n−k)*'. If the time samples '*n*' and '*n−k*' belong to the same bit period, say corresponding to the frequency 'F_0' and delay '*k*' is smaller than the number of samples in a bit, then the product will be as follows :

$$v(n) = s(n) \times s(n-k) = A\sin(2\pi F_0 nT_e) \times A\sin(2\pi F_0(n-k)T_e). \tag{62.5}$$

$$v(n) = \left(\frac{A^2}{2}\right)[\cos(2\pi F_0 kT_e) - \cos(4\pi F_0\, nT_e - 2\pi F_0 kT_e)]. \tag{62.6}$$

In (62.5), the first term is a function of '*k*' only and hence, it is constant as long as '*n*' and '*n−k*' belongs to the same bit. The second term contains some harmonics which can be removed by filtering '*v(n)*' using a low pass filter. Thus, we get :

$\left(\frac{A^2}{2}\right)\cos(2\pi F_0 kT_e)$ for '0' & $\left(\frac{A^2}{2}\right)\cos(2\pi F_1 kT_e)$ for '1'.

To get optimum results (62.7) should be maximum.

$$d(k) = |\cos(2\pi F_0 kT_e) - \cos(2\pi F_1 kT_e)|. \tag{62.7}$$

In Fig. 62.8, the sampling frequency $F_S = 100$ kHz and bit period Te is 2.5 ms. The delay k = 1 which gives d(k) = 1.976 and the low pass filter is a simple IIR digital filter (order 4) with a cut-off frequency around 2 kHz.

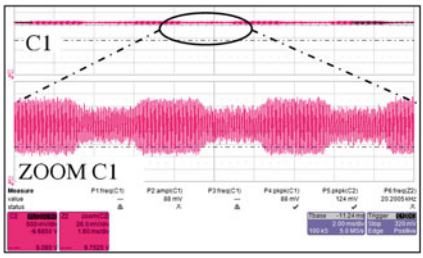


Fig. 62.6 Coupler/decouple circuitry

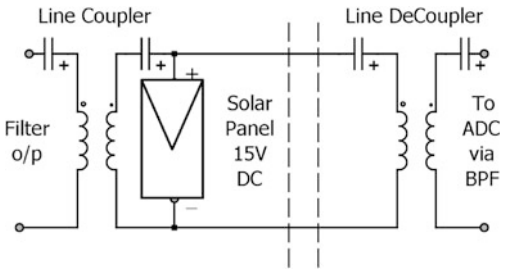


Fig. 62.7 Block diagram of demodulation circuit

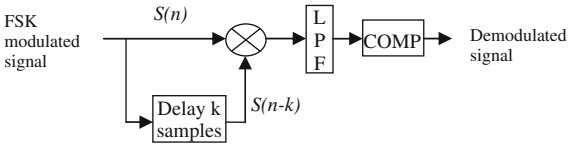


Fig. 62.8 (*Top to Bottom*) FSK data, demo-dulated signal filter o/p, comparator o/p

62.4 PLC Channel Noise Considerations

The noise associated with PLC can be classified into 5 principle types based upon the it's sources as shown in Fig. 62.9.

The model proposed in this paper makes use of the active filter to compensate for the loss in amplitude of the signal to remove noise. The use of transformer in the coupling unit (as shown in Fig. 62.6) minimizes the noise which occurs on account of common grounding. Also, the use of sine waves (approx. monotones) in FSK helps to filter out the noise at the receiver end considering the fact that the noise is spread over the entire frequency spectrum. It is also very beneficial because the noise sources at the receiving side significantly affects the data throughput as compared to the noise generated at the transmitting side thus limiting the baud rate.

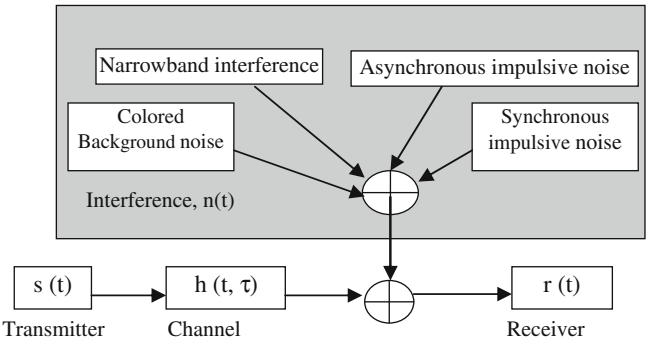


Fig. 62.9 Noise model for DC power line channel

62.5 Conclusion

This paper discusses the implementation of an advanced DSP based PLC modem for the remote monitoring system of a PV plant. The use of DSPs for generating FSK communication signals as well as for other objectives like reading sensor's values and performing control functions like MPPT facilitates for cost cutting and low power consumption and uses DSPs. The PLC technology used does not require any extra cabling and has a fairly simple circuitry unlike wireless networks which makes it very cheap and easy to implement. The solution addresses key PLC problems of contamination of communication signals and power consumption losses. The salient features of the system are its efficient conductive coupling, over voltage protection, and impedance matching for delivering maximum signal power.

However, one of the critical improvements to the system would be the use of OFDM which is the most preferred DSP technique in the PLC Modems pertaining to the most obvious advantages like extra robust coding mechanisms and the ability to deal with asymmetric interference generated on account of the transformers.

References

1. Sánchez-Pacheco FJ, Sotorrio-Ruiz PJ, Heredia-Larrubia JR, Pérez-Hidalgo F, Sidrach-de-Cardona M (2011) Low cost DC lines PLC based photovoltaic plants parameters smart monitoring communications and control module. In: IEEE, May 2011
2. Guo JC, Su S, Fu C (2011) Research on power quality monitoring system based on low-voltage PLC technology. In: APAP, 16–20 Oct 2011, IEEE
3. Park C-Y, Jung K-H, Choi W-H (2008) Coupling circuitary for impedance adaptation in power line communications using VCGIC. In: ISPLC, 3–4 Apr 2008
4. Kuo K-C, Guo J-W, Ou Y-H (2010) A fully digital modulator/demodulator for power line communication (PLC). In: APCCAS, 6–9 Dec 2010
5. Wang C-H, Chen C-Y, Sun T-P (2011) Circuit implementation of OOK modulation for low-speed power line communication using X10 standard. In: ICACT, 13–16 Feb 2011

6. Araneo R, Celozzi S, Lovat G (2009) Design of impedance matching couplers for power line communications. In: IEEE international symposium on electromagnetic compatibility(EMC), 17–21 Aug 2009
7. Baudoin G, Virolleau F, Venard O, Jardin P (1996) Teaching DSP through the practical case study of an FSK modem. In: ESIEE, Paris, Sept 1996
8. Gupta A, Srinivasa VU (2011) RTOS: a new approach in design and organization of high-speed power control applications. Springer

Chapter 63
Digital Security with Thermal Dorsal Hand Vein Patterns Using Morphological Techniques

V. K. Sree and P. S. Rao

Abstract Many biometrics such as face, finger prints and Iris have been developed extensively for human identification purpose and also to provide authentic input to many security systems in the past few decades. However verification using vein patterns of hand is less evolved and developed compared to other human traits. A new personal verification system using the thermal imaged vein pattern in the back of the hand that is the dorsal part is proposed in this paper. The hand vein pattern images are acquired using thermal tracer, enhanced using normalization and vein patterns are extracted using locally adaptive thresholding and skeletonization techniques. Similarity has been checked using Euclidean distance measure.

Keywords Dorsal hand vein patterns · Normalization · Adaptive thresholding · Skeletonization · Euclidean distance and identification

63.1 Introduction

Personal identification systems are gaining lot of demand due to increased threats and attacks from the terrorists. These can be prevented by tightening the security at important places. The traditional methods make use of smart cards or personal identification numbers etc. to identify a person. However these methods have

V. K. Sree (✉)
Department of ECE, VNRVJIET, Hyderabad, India
e-mail: kkrishnasree@yahoo.co.in

P. S. Rao
Department of ECE, Vignan Institute of Technology and Science, Hyderabad, India
e-mail: sparvatha@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_63,

limited security and are unreliable. Biometrics is the science of identifying a person using its physiological or behavioral features [1]. Compared to traditional methods biometric features are much harder for intruders to copy or forge and it has one more advantage that it is very rare for them to be lost. Hence for identification systems making use of biometric features offer a much more secure and reliable performance. Each of these biometric features has its strengths and weaknesses. Anatomically aside from surgical intervention the shape of the vascular patterns in the back of the hand is distinct from each other and it remains stable over a long period [2]. In addition as the blood vessels are hidden underneath the skin and are invisible to the human eye, vein patterns are much harder for intruders to copy as compared to other biometric features [3]. All these special properties of hand vein patterns make it a potentially good biometric to offer more secure and reliable features for personal verification [4]. Physical are related to the shape of the body like fingerprint, face recognition, DNA, palm print, hand geometry and iris of the eye. Behavioral are related to behavior of a person like voice, gait etc. The biometrics should have the certain characteristics like each person should have the said characteristic, it should distinguish individual from another, should be resistive to ageing, easy to acquire, should be accurate, robust, and should have acceptability. A biometric which possesses more number of characteristics is treated as a good bio-metric. A biometric system can operate in two modes, one is Verification and Identification. Verification is the one to one comparison of a captured biometric with a stored template to verify that the individual is who he claims to be. The identification only succeeds in identifying the individual if the comparison of the biometric sample to a template in the data base falls within a previously set threshold. Before the Verification or Identification an individual should enroll his information in the system to store it as template for subsequent uses. During the identification process, the Query image is matched with the set of templates available as data base with the help of some image processing algorithm. During this signature recognition, lot of comparisons need to be carried out with every template in the database.

63.2 Preliminaries

63.2.1 Materials

Veins are hidden underneath the skin, and are invisible to the naked eye and other visual inspection systems. However human superficial veins have higher temperature than the surrounding tissue. Based on this fact, the vein pattern in the back of the hand can be captured using a thermal camera. In this work NEC Thermal tracer is utilized to acquire thermal images of the back of the hand. The images collected from different people in a normal office environment between 20 and 25^{o}C.

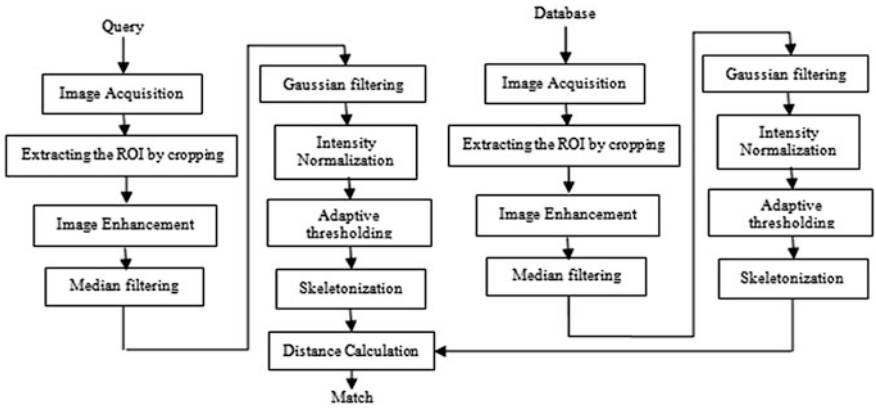


Fig. 63.1 Steps involved in identification of hand vein pattern architecture

63.2.2 Method

The Identification of hand vein pattern architecture is carried out from the given set of images according to the flow diagram given in Fig. 63.1. The thermal images are acquired for the backside of the hand. But the veins are more prominent on the back of the hand, and less in fingers.

Hence the region of fingers is removed and the remaining hand image will become the region of interest. The cropped images quality has been enhanced by applying median filtering and Gaussian filtering followed by normalization. Vein patterns are extracted by adaptive thresholding segmentation [5]. The extracted Vein patterns are skeletonized. With the help of line segment Euclidean distance, the Vein patterns are identified from the data base.

63.2.2.1 Preprocessing

A rectangular region of the hand is extracted as region of interest by removing the fingers from the image. The speckling noise present in the images during acquisition is removed by passing through a 5×5 median filter. The high frequency noise is removed by passing through a Gaussian low pass filter, with standard deviation $\sigma = 0.8$ and the kernel of Gaussian filter is given in Eq. (63.1)

$$H(u,v) = e^{\frac{-D^2(u,v)}{2\sigma^2}} \tag{63.1}$$

The possible imperfections in the image due to the sensor noise and other effects are reduced by normalizing the cropped hand Vein image. The normalization process is described by Eqs. (63.2) to (63.3).

$$I'(x,y) = \begin{cases} \mu_d + \sqrt{\frac{\sigma_d^2.(I(x,y)-\mu)^2}{\sigma^2}} \ldots\ldots\ldots I(x,y) > \mu \\ \mu_d - \sqrt{\frac{\sigma_d^2.(I(x,y)-\mu)^2}{\sigma^2}} \ldots\ldots otherwise \end{cases} \tag{63.2}$$

$$where = \mu = \frac{1}{N \times M} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} I(x,y), \sigma^2 = \frac{1}{N \times M} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} (I(x-y) - \mu)^2 \tag{63.3}$$

Where *I(x, y)* is the image sized $N \times M$, $\sigma_d =$ desired value for Variance and $\mu_d =$ desired value for Mean. The values of μ_d are selected as 0.5479 and the value of σ_d is selected as 0.0028 on experimentation, to achieve optimum results.

63.2.2.2 Segmentation

To segment the vein patterns from the background, locally adaptive thresholding is adopted. For every pixel in the image, its threshold value is set as the mean value of its 13×13 neighborhood. After segmenting the Vein patterns, their shapes are extracted by using skeletonization algorithm.

63.2.2.3 Skeletonization

The skeletonization of a binary object is a collection of lines and curves that encapsulate the size and shape of the object. Skeletonization is done using thinning algorithm in this work.

Thinning

Thinning is a morphological operation to reduce binary objects or shapes in an image to strokes that are a single pixel wide called skeletons. The thinning is performed by transforming the origin of the structuring element to each pixel in the image. Then it is compared with the corresponding image pixels. When the background and foreground pixels of the structuring element and images are matched, the origin of the structuring element is considered as background. Otherwise it is left unchanged.

The thinning of a set A by a structuring element B, denoted as $A \otimes B$.

$$A \otimes B = A - (A \otimes B) = A \cap (A \otimes B)^c \tag{63.4}$$

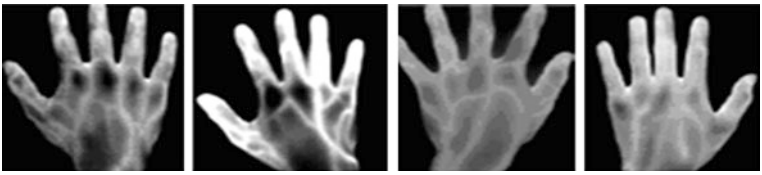


Fig. 63.2 Data base of input images

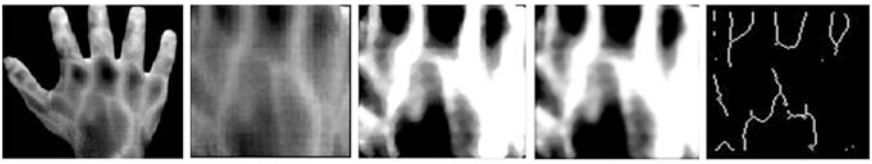


Fig. 63.3 **a** Hand pattern image, **b** cropped image, **c** image after using median filter, **d** image after using Gaussian filter **e** image after skeletonization

63.2.2.4 Similarity Detection

Vein pattern matching is done by measuring the line segment distance between a pair of Vein patterns. The Euclidean distance measure is considered for computing the similarity. Euclidean distance between $p(p_1,p_2,\ldots.p_n)$ and $Q(q_1,q_2,\ldots.q_n)$ in Euclidean n-dimensional is defined by the relation

$$D = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \ldots..(p_n - q_n)^2} = \sqrt{\sum_{1}^{n}(p_i - q_i)^2} \qquad (63.5)$$

An unknown pattern is assigned to the class to which it is closest in terms of distance measure.

63.2.3 Results and Discussions

The hand pattern images are formed as a data base after extracting the region of interest. The data base of input images are shown in Fig. 63.2. These cropped images become template images. The results obtained for various processing steps are shown in Fig. 63.3a, b, c, d, e. False acceptance rate and false rejection rate are the two parameters considered as evaluation parameters.

Table 63.1 Values of Euclidean distances measured

Sl. no	Input image	Distance between input image and other Images			
		Image 1	Image 2	Image 3	Image 4
1	Image 1	0	20.6155	3103050	8.2462
2	Image 2	37.3631	0	3703631	19
3	Image 3	21.1896	19.3132	0	7.0711
4	Image 4	66.7608	64.1249	64	0

Table 63.2 False acceptance rate and false rejection rate measured for the data base

Acceptance rate	100 %
False acceptance rate	0 %
False rate rejection	0 %

63.2.3.1 False Rejection Rate, False Acceptance Rate

False rejection rate refers to the total number of authorized persons not getting access to the system over the total number of people attempting to get the system.

False acceptance rate refers to the total number of unauthorized persons getting access to the system over the total number of people attempting to the system. Table 63.1 gives the values of Euclidean distances measured between each Query image with the images in the data base. Table 63.2 gives the false acceptance rate and false rejection rate measured for the data base (shown for 4 images).

63.3 Conclusion

The proposed method is able to identify all the images with 100 % acceptance rate. The Gaussian filter is able to remove the noise successfully because of the Adaptive selection of standard deviation. The skeletonization helped to extract the patterns much more accurately so that acceptance rate has been increased and this is 100 % because of small data base.

References

1. Jain AK, Ross A, Prabhakar S (2004) An introduction to biometric recognition. IEEE Trans circuits Syst Video Technol 14(1)
2. C-L Lin, K-C Fan (2004) Biometric verification using thermal images of palm-dorsa vein patterns. IEEE Trans circuits Syst Video Technol 14(2)
3. Tanaka T, Kubo N (2004) Biometric authentication by hand vein patterns. In: SICE annual conference in Sapporo, 4–6 Aug 2004

4. Badawi AM (2006) Hand vein biometric verification prototype: a testing performance and patterns similarity. In: Proceedings of the 2006 international conference on image processing, computer vision, and pattern recognition (IPCV'06), Las Vegas, 26–29 June 2006
5. Chen L, Zheng H, Li L, Xie P, Lui S (2007) Near-infrared dorsal hand vein image segmentation by local threshold using grayscale morphology. In: The 1st international conference on bioinformatics and biomedical engineering, pp 868–871

Chapter 64
Design of a Two Phase Inverter for a Specific Two Phase Induction Motor Through MATLAB Simulation

A. Y. Fadnis, R. M. Mohoril, D. R. Tutakne and Gaurav Gondhalekar

Abstract In this paper operation of two phase VSI supplying a two phase balanced cage induction motor is simulated using MATLAB simulation. The concept of switching function is used for simulation of balanced two phase motor for the first time. The simulation results give information regarding the voltages and currents of the various switches of the inverter have to withstand and hence help in choosing components with proper ratings for a specific motor for a two phase motor drive.

Keywords Two phase inverter · Sinusoidal pulse width modulation (SPWM) switching function

64.1 Introduction

The operation of three phase voltage source inverters for supplying three phase motors is widely studied.[1, 2].The use of switching functions for the simulation studies of inverters is explained and implemented by Lee and Ehasani [2]. It is

A. Y. Fadnis (✉) · R. M. Mohoril · G. Gondhalekar
Department of Electrical Engineering, Y.C.C.E, Wanadongri, Nagpur, India
e-mail: ayfadnis@gmail.com

R. M. Mohoril
e-mail: mm_ycce_ep@yahoo.com

G. Gondhalekar
e-mail: gauravgondhalekar@hotmail.com

D. R. Tutakne
Department of Electrical Engineering, S.R.K.N.E.C, Nagpur, India
e-mail: dhananjaydrt@rediffmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_64,

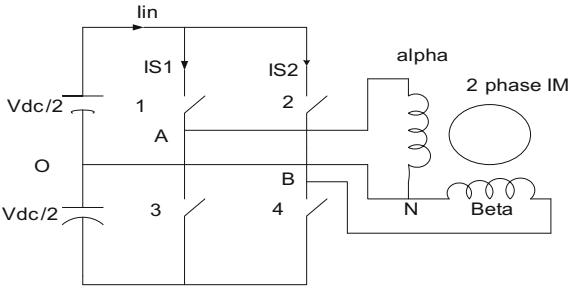


Fig. 64.1 The system used for simulation

however to be appreciated that like three phase cage type induction motors, two phase balanced cage type induction motors are also used in association of the appropriate two phase VSIs [3, 4].

64.1.1 Simulated System

The simulated system is shown in Fig. 64.1. The circuit used is a H bridge inverter. It consists of four switches and a centre—tapped d.c. link supplied by a diode rectifier The circuit also shows the two phase symmetrical motor with two windings α and β. The two windings are arranged spatially at 90° electrical on the stator. Two terminals of the two windings are connected at point N which is connected to O. The switches are operated in a sequence decided by sinusoidal pulse width modulation (SPWM). The sequential operation of switches give rise to two voltages 90° electrically apart. The carrier frequency for SPWM is 1 kHz, modulation index is 8, input voltage = 300 V and the load is a series combination of R = 3.87Ω and L = 20 mH.

64.1.2 Simulation of the Output Voltage of the Inverter

In the simulation implemented in this paper, switches are not simulated as physical switches but as transfer functions which process the input voltages into output voltages just as physical switches do. A switching function is like transfer function such that

$$\text{Vout} = [\text{SF}] * \text{Vin} \tag{64.1}$$

Switching functions SF_1 and SF_2 correspond to SPWM strategy. The switching functions SF_1 and SF_2 will differ from one PWM technique to another. Here SF_1 and SF_2 are defined and implemented for the specific SPWM technique. Figure 64.2a and b show the block connections for SF_1, SF2 voltages and the simulated voltages V_α and V_β

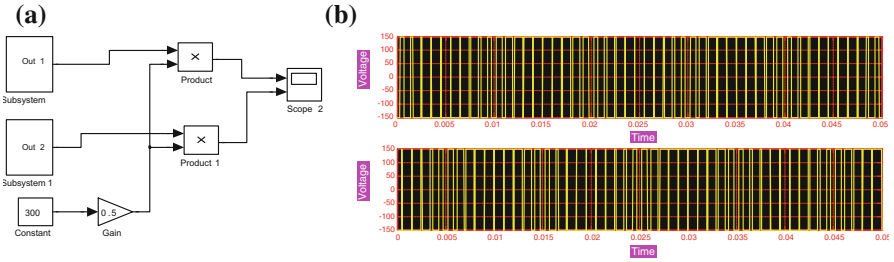
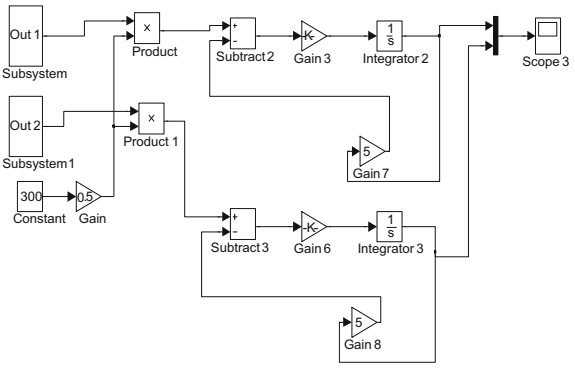


Fig. 64.2 **a** Connected blocks **b** Simulated V_α and V_β

Fig. 64.3 Blocks for I_α and I_β



64.1.3 Simulation of Motor Current Waveforms i_α and i_β

For calculation of motor currents i_α and i_β the motor is represented by a series combination of R and L, corresponding to the parameters of the motor under stalled condition (since the stalled condition corresponds to the maximum current that will have to be fed to the motor). The currents in these phases α and β are

$$V_\alpha = Ri_\alpha + Li_\alpha/dt \tag{64.2}$$

$$V_\beta = Ri_\beta + Li_\beta/dt \tag{64.3}$$

The block simulating the currents I_α and I_β are shown in Fig. 64.3 and the currents in Fig. 64.4

64.1.4 Simulation of Switch Currents

Individual switch of the inverter is made of parallel combination of controlled switch and diode. For finding out current in controlled switch the simulation of switching functions SF_3 and SF_4 is required. Figure 64.5 shows blocks for SF_3 and SF_4 and the concerned simulated waveform.

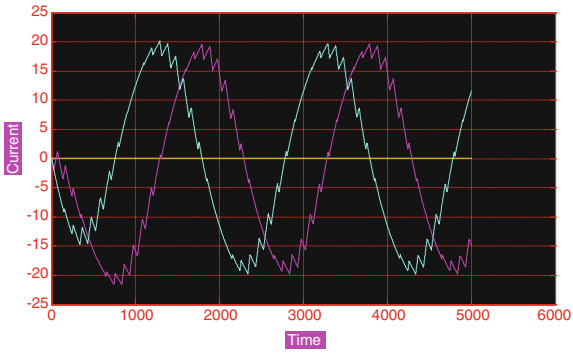


Fig. 64.4 Currents in motor windings

Calculation of IS_1 and IS_2 shown in Fig. 64.1 is done with the help of following equation

$$IS_1 = [SF_3]\ I_\alpha \tag{64.4}$$

$$IS_2 = [SF_4]\ I_\beta \tag{64.5}$$

The connected blocks and the simulation of IS_1 and IS_2 are shown in Fig. 64.6. IS_2 is similar to IS_1 with 90° phase shift.

The positive half wave represents the current through controlled switch whereas the negative half wave represents the current through the diodes. The separation of the switch current and the diode current is shown in the following Fig. 64.7

64.1.5 Calculation of Average, Rms, and Peak Value of Switch Currents

The peak value is easily noted from the waveform as 22 amps. The rms, and average values calculated from MATLAB blocks are 8.9 A rms and 4.7 A average respectively. The use of switching functions has made it possible to realize that the rms value of the switch current is only half the rms of the total load current. Use of switching functions has also brought out the fact that the switch current has a unidirectional average component.

64.1.6 Discussion and Analysis of the Simulation

The analysis shows that it is possible to evaluate the maximum currents that the inverter will have to feed to the motor. Hence it is useful to determine the rating of the inverter switch (both the controlled switch and diode parts of the component switch). The analysis also shows that the rms value needed to be fed through the controlled switch can be separately calculated because of the use of switching function.

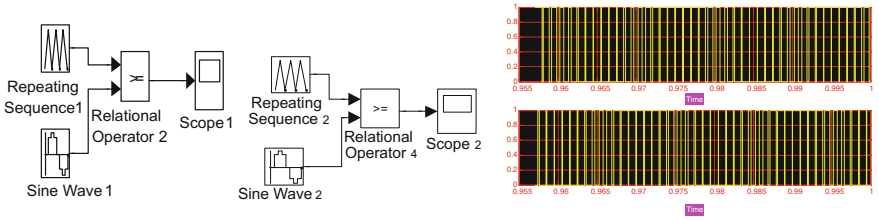


Fig. 64.5 Blocks for SF_3 and SF_4 and the simulation of SF_3 and SF_4

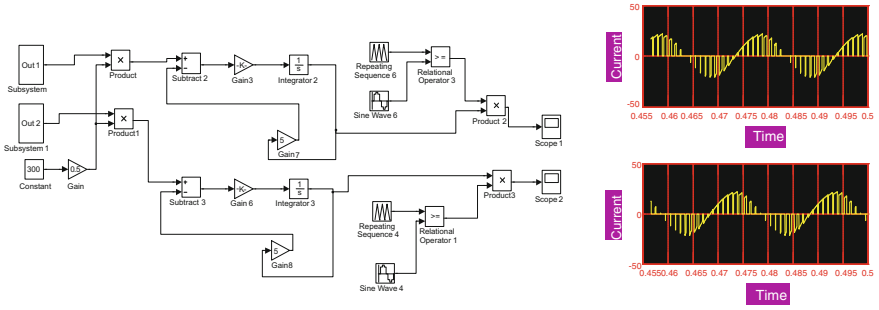


Fig. 64.6 Blocks connected for simulation of IS_1 and IS_2 and the simulated waveforms

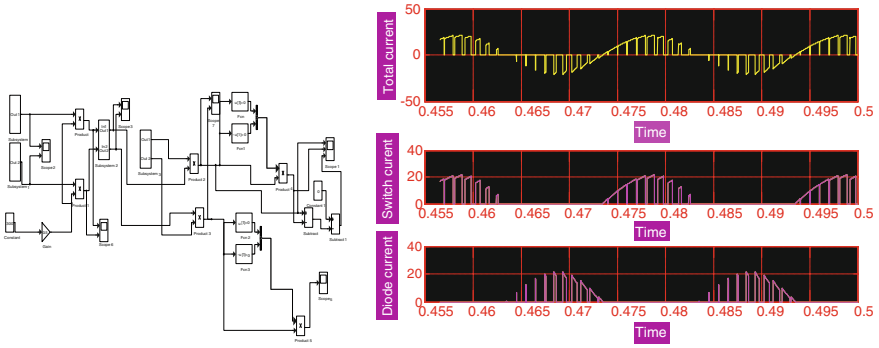


Fig. 64.7 - Blocks for separation of diode and controlled—switch and the diode and switch current

64.2 Conclusion

It is seen that the knowledge of the load parameters is essential for calculating the ratings of inverter switches. It is also seen that the use of switching function techniques gives a more realistic estimate of rms ratings of the diode and the controlled switch.

References

1. Ziogas P, Wiechmann EP, Stefanovic VR (1985) A computer—aided analysis and design approach for static voltage source inverter. IEEE Trans Ind Appl 1 A, 21(5):1234–1241
2. Lee BK, Ehasani M (1999) A simplified functional model for three phase voltage source inverter using switching function concept IEEE, pp 462–467
3. Hyun jang D (2007) PWM methods for two phase inverter. IEEE Ind Appl Mag Mar/Apr 2007/ww.IEEE ORG.IAS
4. Popescue M Analytical prediction of the electromagnetic torque in single phase and two phase motors. Doctoral thesis (abstract) Helsinki University of Technology, pp 10–13

Chapter 65
Artificial Neural Network Based Power System Stability Analysis

S. Kumari Lalitha and Y. Chittemma

Abstract In this paper, an Artificial Neural Network (ANN) approach for the analysis of a power system stability has been proposed and proved to be effective. Here the main consideration is the power system voltage stability i.e. static voltage stability. With instance of 9-Bus [3] power system, also worked on IEEE-57 Bus [4] system and it is verified that the method is effective for power system voltage stability assessment.[3, 4, 8] The implementation of these structures is shown through Mat lab and by the use of ANN approach [5, 6] and the above two methods are compared for the test system. The network would be a useful tool to assess power system voltage stability quickly.

Keywords ANN · Power system voltage stability · VCPI · Newton–Raphson method · Load flow · BP neural network

65.1 Introduction

Power System Voltage Stability [8] is the ability of a power system to maintain acceptable voltages at all buses in the system under normal conditions and after being subjected to disturbance. Power system Voltage stability assessment is to determine whether the power system is voltage stable or not [7]. Power system is

S. K. Lalitha (✉) · Y. Chittemma
Department of Electrical and Electronics Engineering, G.M.R. Institute Of Technology, Rajam, Andhra Pradesh, India
e-mail: lalliyuvaraj@gmail.com

Y. Chittemma
e-mail: chitti.kgr@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_65,

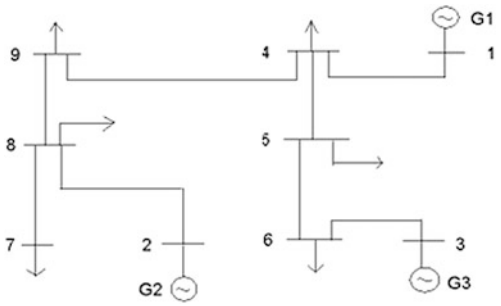


Fig. 65.1 9-Bus system

inherently complex, non-linear, uncertain and so on. As a result, it is difficult to use conventional techniques and mathematical models to describe power system voltage stability assessment. For this, the voltage collapse proximity indicator (VCPI) [2] is used. This paper presents a Back Propagation Neural Network approach [5, 6] for the assessment of power system voltage stability with the VCPI [2] as assessment index. The static model is based on load flow calculation. With instances of 9-Bus, IEEE-57 Bus power system, it is verified that the method is effective to voltage stability assessment [3, 4] on power system.

65.2 The Voltage Collapse Proximity Indicator

Voltage collapse proximity indicators [2] are considered as measures to estimate whether the voltage of a system collapses or not. It varies in the range between 0 and 1 with $L < 1$ for stable state and $L = 1$ for voltage collapse state. The VCPI for a node j can be calculated by

$$L_j = \left|1 + \frac{V_{0j}}{V_{2j}}\right| = \left|\frac{S_j}{V_j^2 Y_{jj}}\right| = \frac{S_j}{V_j^2 Y_{jj}}$$

Where S_j is the transformed power, Y_{jj} is the transformed admittance $= \left(1/Z_{jj}\right)$, and V_j is the consumer node voltage.

It is difficult to calculate the index L for each load bus directly by the mathematical analysis. As a result, this paper presents a BP neural network to estimate the index L, sequentially to assess the power system voltage stability rapidly and timely. Since the power system voltage stability is affected mostly by load characteristics of the system, that is, with the power of load bus increasing, the system becomes more close to voltage instability.

Table 65.1 Indicators of Every Load Bus

Active power						Reactive power						VCPI					
P4	P5	P6	P7	P8	P9	Q4	Q5	Q6	Q7	Q8	Q9	L4	L5	L6	L7	L8	L9
1.25	1.2	0.9	1.9	1.22	0.92	1.25	0.8	0.0	0.0	0.0	0.0	0.202	0.2696	0.462	0.1014	0.2645	0.3256
1.25	1.2	0.9	1.9	1.22	0.92	1.25	0.2	0.0	0.0	0.0	0.0	0.174	0.2606	0.388	0.0984	0.2558	0.3081
1.25	1.2	0.9	1.9	1.22	0.92	1.25	1.2	0.0	0.0	0.0	0.0	0.246	0.2851	0.546	0.1042	0.2726	0.3428
1.25	1.2	0.9	1.9	1.22	0.92	1.25	0.8	0.0	0.0	0.0	0.0	0.202	0.2698	0.462	0.1014	0.2645	0.3256
1.25	1.2	0.9	1.9	1.22	0.92	1.25	1.5	0.0	0.0	0.0	0.0	0.302	0.3098	0.642	0.1072	0.2818	0.3632
1.25	1.2	0.9	1.9	1.22	0.92	1.25	1.5	0.0	0.0	0.0	0.0	0.302	0.3098	0.642	0.1072	0.2818	0.3632
1.25	1.2	0.9	1.9	1.22	0.92	1.25	1	3.05	0.0	0.0	0.0	0.276	1.0015	0.685	0.4492	0.3121	0.3878
1.25	1.2	0.9	1.9	1.22	0.92	1.25	1	3.05	0.0	0.0	0.0	0.297	1.265	0.758	0.5642	0.3259	0.4069
1.25	1.2	0.9	1.9	1.22	0.92	1.25	1	3.05	0.0	0.0	0.0	0.303	1.3263	0.778	0.5911	0.3294	0.4119

Table 65.2 Indicators of every load bus for IEEE 57-bus system

L2	L5	L6	L9	L10	L13	L14	L15	L16	L17	L18	L19	L20	L23
0.05129	0.04952	0.26108	1.2559	0.719	0.4974	0.02028	0.098	0.3962	0.703	0.0377	0.2499	0.0108	0.00115
0.05168	0.04942	0.26069	1.2655	0.541	0.3286	0.02019	0.098	0.3953	0.701	0.0378	0.2494	0.0107	0.00115
0.05129	0.0495	0.26104	1.2477	0.711	0.489	0.02022	0.098	0.3962	0.703	0.037	0.2496	0.0108	0.00115
0.05129	0.04950	0.26103	1.2468	0.711	0.4882	0.02022	0.098	0.3962	0.703	0.0377	0.2495	0.0108	0.00115
0.05176	0.04939	0.26059	1.2588	0.495	0.2809	0.02011	0.098	0.395	0.701	0.0378	0.2490	0.0107	0.00114
0.05118	0.04955	0.26118	1.2443	0.764	0.5394	0.02025	0.098	0.3965	0.703	0.0377	0.2498	0.0108	0.00115

Next 12 load buses (43,44,47,49,50,51, 52,53,54,55,56,57)

L25	L27	L28	L29	L30	L31	L32	L33	L35	L38	L41	L42
0.003828	0.00348	0.02152	0.011535	0.06159	0.02096	0.07832	0.00168	0.00293	0.04655	0.095242	0.047476
0.003807	0.00346	0.021487	0.011483	0.06127	0.02085	0.077903	0.00167	0.00291	0.04635	0.094596	0.047161
0.003814	0.00347	0.021479	0.011493	0.06132	0.02086	0.077946	0.00167	0.00292	0.04636	0.094505	0.047112
0.003812	0.00347	0.021474	0.011488	0.06129	0.02085	0.077905	0.00167	0.00299	0.04634	0.094424	0.047072
0.003789	0.00345	0.021435	0.011429	0.06093	0.02074	0.077443	0.00166	0.00287	0.04613	0.093701	0.04672
0.003821	0.00344	0.021492	0.011507	0.06142	0.02093	0.078087	0.00167	0.00296	0.04643	0.094685	0.0472

L43	L44	L47	L49	L50	L51	L52	L53	L54	L55	L56	L57
0.026453	0.00540	0.00645	0.08843	0.052973	0.051032	0.066845	0.00259	0.086613	0.14494	0.10494	0.007574
0.026209	0.00537	0.00643	0.08798	0.052596	0.050674	0.066523	0.00263	0.085644	0.03756	0.1043	0.007552
0.02624	0.00538	0.00642	0.08806	0.052714	0.050773	0.06655	0.00257	0.085843	0.14111	0.10416	0.007516
0.026216	0.00538	0.00642	0.08802	0.052685	0.050745	0.066518	0.00258	0.085759	0.14099	0.10407	0.007509
0.025937	0.00534	0.00640	0.08751	0.052251	0.050333	0.066149	0.00261	0.08464	0.00754	0.10335	0.007486
0.026306	0.00539	0.00643	0.08820	0.052823	0.050877	0.066636	0.00255	0.086081	0.17426	0.10434	0.007522

65.2.1 Training the Neural Network

The proposed method has been tested in 9-Bus system [3] shown in Fig. 65.1.

The patterns used for training network are given in Table 65.1 (only lists 9 set of data).

And $L5 = 1$ and $L5 > 1$ loading levels also shown in Table 65.1.

65.2.2 Generation of Training Data

The method of generating the training data [1] is as follows:

a. Only change the active power of certain bus, the others remain the same;
b. Only change the reactive power of certain bus, the others remain the same;
c. Only change the active and reactive power of certain bus, the others remain the same;
d. Change the active and reactive power of all buses.

65.3 Case Studies

The method also tested on IEEE 14, 30 and IEEE 57-bus system [4] which are shown in Table 65.2.

65.4 Discussion and Conclusion

A BP neural network approach [5, 6] for the voltage stability assessment of a power system has been proposed and proved to be effective. The test results of the three-layer BP neural networks indicate that if the sample data are accurate, reliable, and the model parameters are appropriate, the network would be a useful tool to assess voltage stability quickly on line.

References

1. Han X, Zheng Z Voltage stability assessment based on BP neural network. Nannan TIAN College of Electrical and Power Engineering Taiyuan University of Technology Taiyuan China
2. Kessel P, Glavitch H (1986) Estimating the voltage stability of power systems. IEEE Trans Power Deliv 1(3):346354
3. Subramani C, Sekhar Dash S, Jagadeesh kumar M (2009) Voltage stability based collapse prediction and weak cluster identification. Int J Electr Power Eng 3(2):124–128
4. Kamalasadan S, Srivastava AK, Thukaram D (2006) Novel algorithm for online voltage stability assessment based on feed forward neural network. 2006 IEEE
5. Chen D, Mohler RR (2003) Neural-network-based load modeling and its use in voltage stability analysis. IEEE Trans Control Sys Technol 11(4):460–470
6. Chen X, Guang PX (2003) Artificial Neural network technology and its application. China Electric Power Press, Beijing
7. Shuangxi Z, Lingzhi Z, Xijiu G, Xiaohai W (2003) The voltage stability and its controlling of power system. China Electric Power Press, Beijing
8. Salama MM, Ebtsam MS et al (2001) Estimating the voltage collapse proximity indicator using artificial neural network. Energy Convers Manage 42:6979
9. Abdul Rehman M, Musirin I, Othman MM (2008) Evolutionary programming based technique for secure operating point identification in static voltage stability assessment. J Artif Intell 1(1):12–20
10. Anderson PM, Fouad AA (2003) Power system control and stability, 2nd edn. IEEE Press
11. Tamura Y, Mori H, Iwamoto S (1983) Relationship between voltage instability and multiple load flow solutions in electric power systems. IEEE Trans Power Apparat Sys PAS-102:5
12. Jarjis J, Galiana FD (1981) Quantitative analysis of steady state stability in power networks. IEEE Trans Power Apparat Sys PAS-100:1
13. Martin T, Howard B, Demuth MH, Beale H (1996) Neural networks design. PWS pub, pp 170–178
14. Simon H (2004) Neural networks a comprehensive foundation. Pearson Education, India pp 30–35

15. Dinavahi VR, Srivastava SC (2001) ANN based voltage stability margin prediction. IEEE PES Summer Meet 2001 2:1275–1280
16. Charabarti S, Jeyasurya B (2004) On-Line voltage stability monitoring using artificial neural network. Large engineering system conference on power engineering, 2004, LESCOPE 2004, pp 71–75

Part III
Poster Papers

Chapter 66
Efficient Bandwidth Utilization in Client–Server Models

Alex Antony Arokiaraj

Abstract The amount of data sent in the network conspicuously affects the network performance and it also adds significant latency to the applications. Various data compression techniques have been in use for decades providing both storage efficiency as well as transmission efficiency. For applications that communicate over a network with limited bandwidth, efficient bandwidth utilization not only depends on the amount of data sent in the network, but also on the number of calls made between the applications, especially in client–server models. Therefore something apart from the techniques of data compression has to be ordained to achieve the latter. Mitigating the number of calls made between the client and the server should not affect the data consistency between them, thereby making the applications unreliable. I propose a model to be incorporated in the client–server frameworks to achieve efficient bandwidth utilization, by constricting the number of calls made between the two applications, and also the amount of data sent. Since there is a considerable amount of information sent on each call, reduction in the number of calls results in a substantial reduction in data transmissions. The information sent on each call, not only refer to the TCP/IP setup, but also the server's original response to a request from another client. I confer the "eBUCS" protocol, which will be intricately tied up with the client–server frameworks, and also the scenarios under which the protocol will limit itself in order to avoid its adverse effects on the latency.

Keywords eBUCS · Client–Server frameworks · Bandwidth utilization · Traffic analysis and optimization · Latency

A. A. Arokiaraj (✉)
Ericsson Inc, 11, Udhayam Nagar, Kumbakonam 612001, Tamilnadu, India
e-mail: alexantony1988@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_66,

66.1 Introduction

Most network applications can be divided into two pieces: a client and a server. They utilize the network bandwidth in terms of original data that is transferred, and the calls made between them. These calls are made through numerous protocols, such as NetBIOS®, RPC (Remote Procedure Call) ®, DCOM®, Pipes®, IPC (Interprocess Communication) ® [1]. Every call is initiated by a request from the client or server.

Clients most probably do not share any of their resources, but requests a server's content or service function. Therefore a '*request*' can be interpreted as something that intends to know some information or data, which the client does not know. In most of the architectures, the server serves data to its clients from its database. Hence, in any application, that exploits the client–server model, a request can be classified into two major categories. 1. A request for information. 2. A request for a communication link to be established ex: TCP/IP® setup.

66.1.1 Predicaments Eliciting the Need of a Protocol

Request handling is done by the application developer himself, which opens up the possibility of 'requests and information overflow' causing the latency of the applications and the network performance itself to plummet. Some examples include (a) Creating a request that provides the exact information that the client already has (b) Generating frequent requests for operations on the same set of data, from the server; Rather the developer would have chosen to copy the data from the server to the client and perform recurring requests within the client itself. There is no common straight-forward solution to all these delicate situations, because a straight-forward solution would count on the efficiency in terms of other factors. For example, copying the information from the server to client and then performing operations on the client-side data would be beneficial, only if there are more probable operations in the client-side that operate on the same set of data. If the probability of the operations working on the same set of data is less, then copying the entire information would be a mere waste of bandwidth. A subtle inspection on these drawbacks shows that, more than a common solution, an intelligent protocol, that would possess higher acclimatization capabilities is required. The following sections of the paper will describe the architecture and the call flow set up of the proposed protocol. It will be followed by the simulation results to show that, the proposed methods supports the arguments put forth in this paper.

66.2 Virtual Client Database

In many commercial applications (like Access Management System), the client will utilize only a portion of the server's database that is allotted to it. For example, a client named 'X' will use only data content that relates to 'X' in the server's database. Any other content in the server's database is irrelevant to client 'X'. Practically, there are applications where clients utilize other user's information and all of them share a common data, but for the better understanding of the architecture, we will consider the generic Access Management System where the client 'X' uses only the content related to the 'X' in the servers database and all other users have permission to add information to the content of 'X'. Although, we are primarily dealing with the probability of operations that operate on the same dataset, considering this generic access management system as described above forms an excellent base for the architecture, and an acute introspection shows that, it can ideally be used as a model to build the intelligence of the protocol, and eventually be amended to support all situations.

In the model we have considered, there is limited information for a particular user. Hence if there are multiple operations that operate on the same dataset, the developer can choose to copy the dataset specific to the particular user in the client, and the succeeding operations will be performed in the client-side without reaching the server. Such a database that is created on the client's side is called as a '*Virtual Client Database*' (VCD). Therefore 'eBUCS' protocol will create a VCD of the client-specific information from the server's database, in the client if the value T_n, exceeds a certain value. Discussion on the eBUCS threshold value and its equation is described in the simulation results (Sect. 62.3).

66.2.1 Operations on the VCD

The eBUCS protocol will mediate all the requests that are originating from the client. The originating requests from the client can be classified into two categories 1. *Modifying requests* 2. *Non-modifying requests*. Modifying requests are those that tend to change the information contained in the server's database. Non-modifying requests are those that will not affect the information contained in the database. For example, operation 'sort' is a non-modifying request while 'delete' is a modifying request. All the modifying requests will be routed to the server while the non-modifying requests will be routed to the VCD running in the client itself. Therefore the frameworks have to be intricately tied up with the eBUCS protocol. eBUCS protocol will have all the possible operations defined under its own library and will provide an interface to define user-defined operations as well.

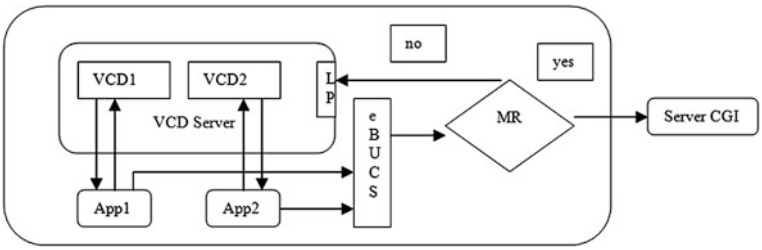


Fig. 66.1 Model showing the operation of the VCD Server

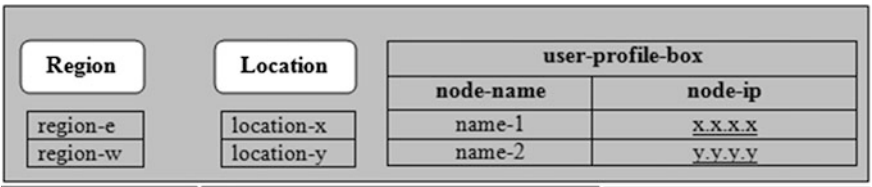
66.2.2 Virtual Client Database Server

In order to provide support for multiple applications running on the same client machine, or multiple groups of operations working on a subset of different datasets, a VCD server with multiple VCDs is created as shown in Fig. 66.1. It will in turn serve requests from multiple applications running on the same client, and all the requests will be handled by the child processes created by the VCD server. eBUCS operates on top of the application Layer in the OSI model. It affects the network as well as the transport layers. Any client application initiating a request forwards the request to eBUCS. eBUCS with its own set of library functions and the user-defined functions classifies it into a *modifying request* or a *non-modifying request*. For non-modifying requests, eBUCS replaces the destination IP address in the network layer with the local host address and the port number with the listening port of the VCD Server (Fig. 66.1).

Once the non-modifying request is routed to the VCD server, there is a persistent connection between the application and the specific VCD. The connection is closed, once the operation returns the result to the application. For modifying requests, eBUCS acts as a transparent medium and passes the packet on to the presentation layer. There are some exceptions to this. If the eBUCS threshold value T_n, is not reached, then eBUCS handles the non-modifying request as though it is a modifying request.

66.3 Simulation and Results on Bandwidth Utilization

Let us consider the generic *Access Management System* with three related tables for a user and his associated profile information. The server database for the client-application contains three related tables. The Fig. 66.2 represents the interface that is presented to the user. The interface contains two operations. (a) *Filter based on region* (b) *Filter based on location*. The first operation fetches all the related node information for the user from the server database, and fills the location and the user-profile-box. The second operation fetches the same information, but fills only



user-information	
u-id	u-name
user-a	username-a
user-b	username-b

node-information				
node-id	node-ip	region	location	nodename
1	x.x.x.x	e	x	name-1
2	y.y.y.y	w	y	name-2

access-information	
u-id	node-id
user-a	1

Fig. 66.2 Sample dataset format on which the simulation was performed. The tables user-information, node-information, access-information are present in the server and the user-profile box is presented to the client

the user-profile box. The two important factors to note here is that, both the operations will initiate a non-modifying request from the client and they operate on the subset of a same dataset. The dataset is the entire information specific to the user in all the three related tables.

A simulation was performed to analyze the bandwidth utilization, by following the normal strategy of initiating a request to the server, in response to a client call and then, by creating the VCD and routing the non-modifying requests to it. The simulation involved the following considerations. (a) Approximate length of a TCP/IP datagram is 20 bytes long. (b) The number of bytes required for connection setup and connection close is approximated to 20*6 = 120 bytes. (c) The number of bytes associated with each database query is approximated to 1 byte. (d) From points b and c, the total approximate number of bytes transferred during a non-modifying request from the client is calculated as (20*6) + 1 = 21 bytes. (e) The total amount of information associated with user-a on all the three related tables, is 4 K = (4*1024) = 4096 bytes. (f) All the non-modifying requests results in a dataset that is a subset of the 4,096 bytes, as mentioned above.

The graph in Fig. 66.3 shows the cumulative amount of data transferred during each event. An event is any operation that might initiate a request at any time ‘t ‘. The normal curve is plotted for the amount of bytes transferred on each event along with the server’s response, circumventing the VCD. The eBUCS curve is plotted for the amount of bytes transferred on each event, but now the non-modifying requests, passing through the VCD. There are some interesting observations on this graph. (a) The points of intersection of the two curves, are referred to as ‘*eBUCS threshold points*’ designated by T_n. (b) It is obvious that, the cumulative amount of data transferred is higher for the normal curve after the points of intersection T_n. (c) However, there is one exception at event 8, where the cumulative amount of data transferred is higher for the eBUCS curve after T_n. This is the point when an update took place in the server and the eBUCS engine had to transfer all the updated information to the client. (d) A subtle observation shows that, at event 8, the normal curve gains a massive increase in the slope which provides us a common solution.

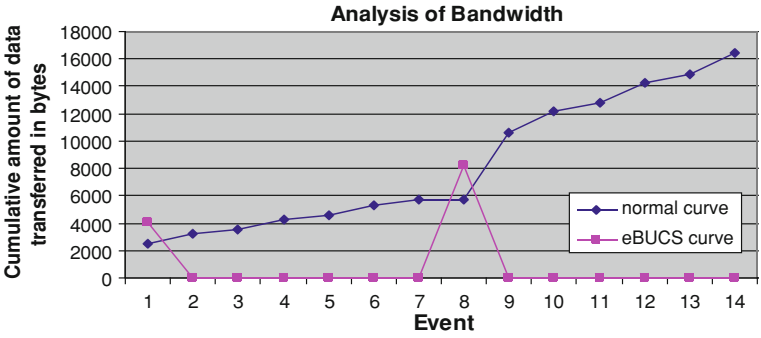


Fig. 66.3 Simulation results showing the cumulative bandwidth utilization over a period of time

Beyond all the eBUCS threshold points, the tendency of the application to utilize the bandwidth is far higher for normal curves (Fig. 66.3).

66.4 Intelligent eBUCS and eBUCS Equation

The Fig. 66.4 is a representation of the protocol flows and gives an insight of how eBUCS provides an intelligent approach to varying situations by using the value of T_n. It is time to generalize the observations made from the basic model we considered and create intelligence in the protocol to work efficiently in all scenarios. The graph clearly shows that the concept of creating a virtual client database is not efficient in all cases. Let us consider an application, where both the curves don't intersect at all. Therefore $T_n = \infty$. Creating a VCD will create adverse effect on the B/W utilization in this scenario. In fact, this is the worst adverse effect, eBUCS will pose on the B/W utilization. A Solution would be to pass the non-modifying requests to the VCD only when an event occurs after T_n is reached. However, it is very unfair for an application developer to do all this pre-analysis and calculate T_n, before the application is developed. It is also to be noted that the value of T_n is likely to change post deployment, as it greatly depends on the probability of non-modifying requests that occur over a period of time. Hence eBUCS should consider calculating this value dynamically so that it possesses enough intelligence to be embedded in all the client–server frameworks. The following section focuses on calculating the value of T_n. Let 'N' be the total amount of information, the subsets of which forms the base dataset for non-modifying requests in the application. For example, if A is the dataset which an event 'e' operates on, then $A \subset N$. Any event (a non-modifying request) that operates on this subset results in a response, whose amount of bytes will be N-γ, where γ is the factor by which the total amount of information is reduced by, to produce the response for the operation. Γ ranges from 0 to N. Let 'm' be the amount of bytes transferred during an initial request setup/termination and 'o' be the amount of bytes transferred for the query.

Let us assume that the VCD is not is effect, and 'n' events occur over a period of time, such that

$$a(n_1) = N - \gamma_1 + m + o_1. \tag{66.1}$$

$$a(n_2) = N - \gamma_2 + m + o_2. \tag{66.2}$$

$$a(n_n) = N - \gamma_n + m + o_3. \tag{66.3}$$

and so on, where a(n) is the total amount of information transferred during an event n. Considering the fact that 'm' is constant and 'o' is very small when compared to a(n), when an operation iterates over a period of time, the two curves intersect at this point

$$T_n = N/N - \gamma. \tag{66.4}$$

If there are two operations that iterate over a period of time and are equally likely to occur, then the two curves intersect at this point,

$$T_n = 1/2[1/(1 - \gamma_1/N) + 1/(1 - \gamma_2/N)]. \tag{66.5}$$

For 'n' operations that iterate over a period of time and are equally probable,

$$T_n = 1/n[1/(1 - \gamma_1/N) + 1/(1 - \gamma_2/N) + \ldots\ldots\ldots + 1/(1 - \gamma_n/N)]. \tag{66.6}$$

These equations, however do not consider the updates that are likely to happen in the server database. They hold good only when the probability of updates that might occur over a period of time is lesser than the probability of operations that might occur in the client-side over a period of time. This is also seen clearly from the graph. At event 8, the cumulative amount of data transferred for eBUCS curve is higher than the normal curve, however the normal curve suffered a sudden increase in slope, because the probability of updates in the server database was lesser in the simulation done. Hence it is vital to rephrase the equation. Let 'p' be the probability that a non-modifying request occurs. Therefore the probability that it might occur over a period of time t is given by $p_t = 1-(1-p_t)^{\circledR}$. If 's' is the probability that an update might occur in the server database, then for $p_t > s_t$, 'n' operations that are equally probable, will have

$$T_n = 1/n[1/(1 - \gamma_1/N) + 1/(1 - \gamma_2/N) + \ldots\ldots + 1/(1 - \gamma_n/N)] \tag{66.7}$$

This intelligence makes eBUCS to be incorporated in all client–server frameworks and is handled in the eBUCS Probability Engine (EPE).

66.4.1 eBUCS Reference Model

See Fig. 66.4

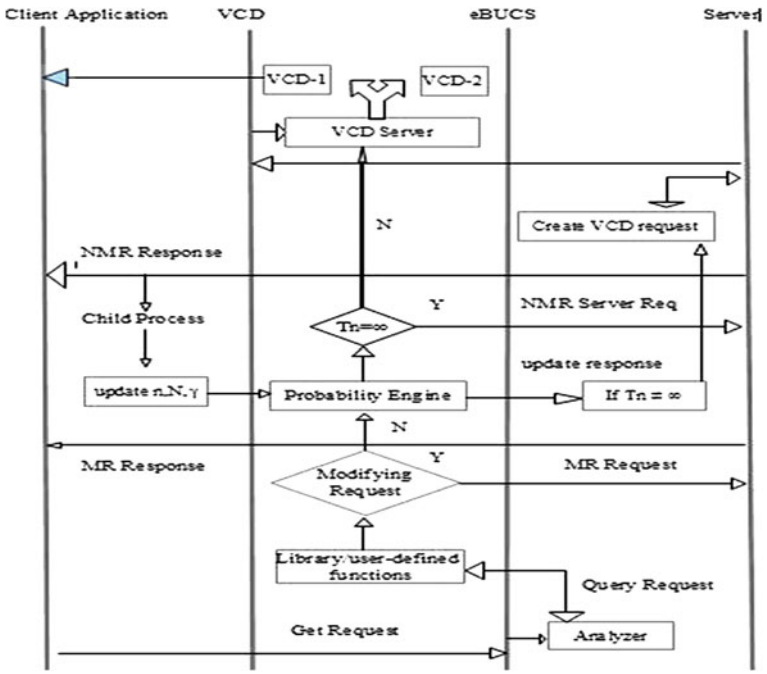


Fig. 66.4 eBUCS reference model showing the call flows

Reference

1. Rajinder Y (2007) Client/server programming with TCP/IP sockets 1

Chapter 67
Vlsi Approach for Four Quadrant Analog Multiplier for 2.4 Ghz to 2.5 Ghz

Sanjay Tembhurne and L. P. Thakare

Abstract Wireless communication world is revolutionized because of the design of the multiplier in the 2.4–2.5 GHz band because of the unlicensed use of the frequency band to the entire user. Multiplier is a key block of analog communication system. Multiplier mixes the signals of different frequencies or signals of different types, to strength signal for long distance communication, which emphasises the designing of more efficient and low power mixers or multipliers for RF applications. MOS RF (radio frequency) multiplier with reduces on chip area operate at ISM Band frequency with high linearity. A 2.4–2.50 GHz band (ISM BAND) multiplier designed and simulated on tanner tool 13. The simulations results presented here are for 1–10 GHz. The circuit is implemented using 180 nm level 3 models and simulated in TSPICE simulator. The transistor operating in linear region reduces the drain current and also the power consumption with large input range. Power consumption is reduces to 68.57 μW.

Keywords Analog multiplier · ISM band · Quadrant · RF (radio frequency)

67.1 Introduction

The industrial, scientific and medical (ISM) radio bands were originally reserved internationally for the use of RF electromagnetic fields for industrial, scientific and medical purposes along with communications. In the design uses the cascaded

S. Tembhurne (✉)
VLSI, G. H. Raisoni College of Engineering, Nagpur, India
e-mail: sanjaybtembhurne@hotmail.com

L. P. Thakare
Department of Electronics, G. H. Raisoni College of Engineering, Nagpur, India
e-mail: laxmanthakre@yahoo.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_67,

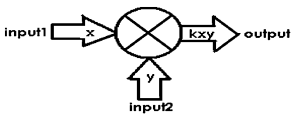


Fig. 67.1 Multiplier block

topology which makes circuit larger as number of MOS increase so the layout become complicated, at the same time body effect problem is having the effect on result. This body effect reflects on linearity of the circuit. Due to the reduction in body effect the linearity of the circuit increase [4]. In the generalized mixer the local oscillator radio frequency section has to design into different but here in the proposed circuit the only the multiplier circuit is used for multiplication of the signal.

Analog multipliers constitute a field of active research due to their usefulness in analog signal processing [2]. Frequency translation, phase detection, correlation, convolution, adaptive filtering, etc., are usually achieved using these circuits. So the need of energy efficient with high linearity multiplier increases rapidly. Different structures for performing the multiplication of two signals are proposed [1–6, 7, 8]. MOS transistor behavior in saturation region can be expressed with the square-law equation in its ideal form. But, in practice, there are some second order effects that make its behavior far from the square-law equation. These effects are: body effect, channel length modulation, mobility reduction and etc. [2] this problem of second order effect is minimize in the proposed multiplier because MOS devices are operating in linear region.

67.2 Multiplier Principle

$$\text{Eq. } x.y = kxy \tag{67.1}$$

Multiplier is the circuit which can multiply different parameter of the input signals either frequency or amplitude or phase or combination of these in the proposed multiplier circuit the output is the multiplication of the two signal x and y. resultant is the kxy where k is the gain of the circuit x and y are the two input. The amplitude of the signal will get multiplied and also the frequency of the signal will also get multiplied.

67.3 Methodology

Design methodology is based on the four quadrants designing here to two x and y are applied with out of phase of both at different input terminals. Design is focusing on the parallel topology which is helpful to reduce the number of MOS

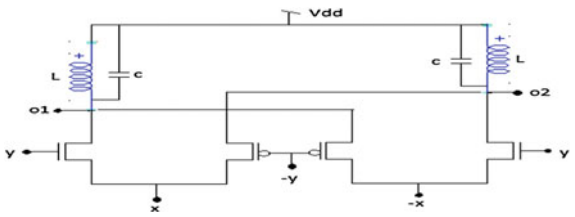


Fig. 67.2 Proposed multiplier

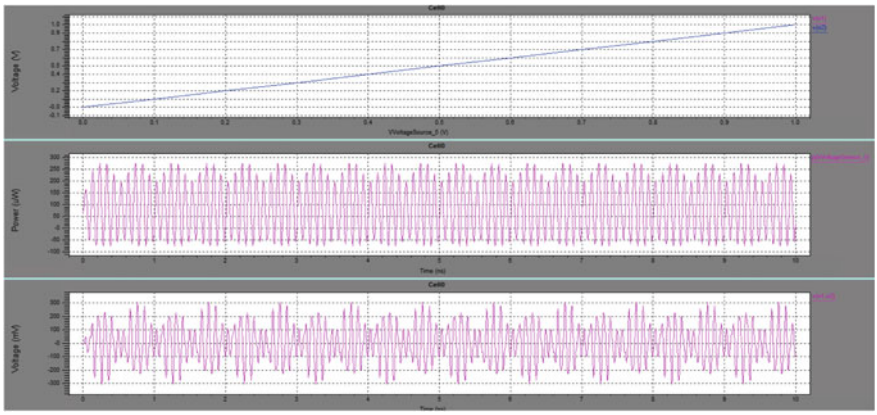


Fig. 67.3 DC analysis, transient analysis

device in the design this directly reduce the chip area. In this paper parallel topology with one NMOS and one PMOS connected in parallel topology. The circuit is in differential output mode. Output is the result of the difference between two output o1 and o2. The advantage of differential mode is the common mode noise will get cancelled so the output noise will reduces. This type of multiplier exploits symmetry to remove the unwanted RF & LO output signals from output by cancellation. From the above Fig. 2 of the proposed multiplier here is the four drain current of four MOS device which is flowing through the load and get added at the output node o1 and o2. As in proposed multiplier NMOS and PMOS are connected in parallel but the output of the PMOS2 is connected to the drain of NMOS1 of symmetric half of the multiplier and drain of PMOS1 is connected to the drain of NMOS2. So the outputs are cross coupled.

$$I_{DN} = \beta\left[(V_{GS} - V_{THN})V_{DS} - - V_{DS}^2/2\right] \tag{67.2}$$

$$I_{DP} = \beta\left[(V_{SG} - V_{THP})V_{SD} - - V_{SD}^2/2\right] \tag{67.3}$$

From above Fig. 67.2 and using equation (67.1) and (67.2) total current can be calculated i.e.

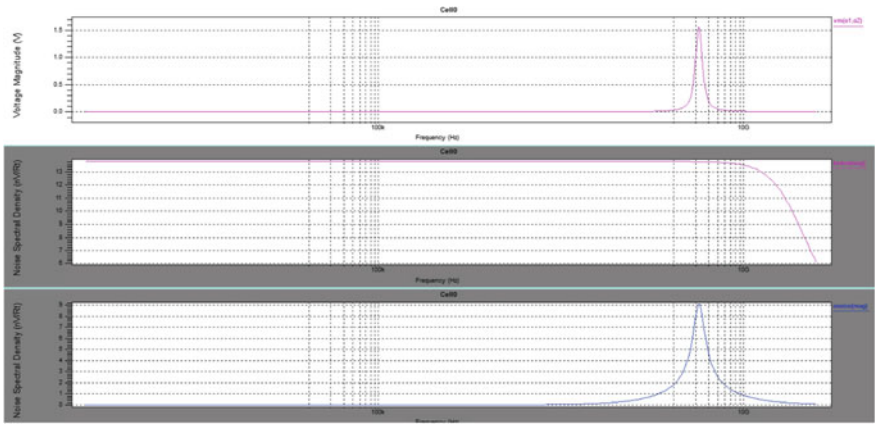


Fig. 67.4 AC analysis, noise analysis

Table 67.1 Multiplier specification

Sr.No	Parameter	Value
1	Supply voltage	0.8 V
2	Input frequency	X = 2.4 GHz, Y = 10 GHz
3	Input amplitude	X = 1 V, Y = 1.5 V
4	Output amplitude	1.5 V
5	Power dissipation	94.57 μW

$$I_{D1} = I_{DN1} + I_{DP2} \tag{67.4}$$

$$I_{D2} = I_{DN2} + I_{DP1} \tag{67.5}$$

Form Eqs. (67.4) and (67.5)

$$I_{Total} = I_{D1} + I_{D2} \tag{67.6}$$

Power of the proposed multiplier is given by

$$P_{Total} = I_{Total} * V_{DD} \tag{67.7}$$

67.4 Simulation Results

In Fig 67.3 first result show the linearity result of the multiplier, linearity prove that the circuit is having very less effect of harmonics distortion so the circuit gets on as soon as the input is applied to the MOS device so circuit noise reduce it is the result of DC analysis. Second shows result of transient analysis and the power dissipation of the circuit having the average value of 94.57 μW. Third result shows

the amplitude modulated output of the circuit which is nothing but the differential output of the circuit giving the multiplied output of 1.5 V from the input 1 and 1.5 V.

In Fig 67.4 first result shows AC analysis having the bandwidth between 2.4 and 2.5 GHz same as Bluetooth operating frequency. So the multiplier is useful for designing of the Bluetooth device. Circuit can do the modulation and demodulation both operations with small parameter change in circuit. Second and third result shows the input and output noise respectively (Table 67.1).

67.5 Conclusions

Multiplier in this paper is design for ISM band so that multiplier can be used in the different ISM band using circuit design directly. The power consumption of the circuit is also low so that battery life of the new device like Bluetooth and cordless phone can be directly used. Also the range of the device is also increase because circuit is design using the higher frequency range of 10 GHz.

References

1. Sawigun C, Demosthenous A (2007) Compact low-voltage CMOS four-quadrant analog multiplier", 2007 IEEE, pp 751–754
2. Ebrahimi A, Naimi HM (2010) 1.2 V single supply and low power CMOS four-quadrant analog multiplier. In: 2010 XIth international workshop on symbolic and numerical methods, modeling and applications to circuit design (SM2ACD), pp 978–983
3. Hidayat R, Dejhan K, Moungnoul1 P, Miyanaga Y (2007) A GHz analog multiplier for UWB communications. In: Proceedings of Asia-Pacific conference on communications, pp 55–58
4. Akshatha BC, Vijay Kumar A (2009) Low voltage, low power, high linearity, high speed CMOS voltage mode analog multiplier. In: Second international conference on emerging trends in engineering and technology, ICETET-09, pp 149–154
5. Lau KT, Lee ST, Ong VKS (1998) Our-quadrant analogue CMOS multiplier cell for VLSI signal and information processing. IEEE Proc Circ Device Sys 145(2):132–134
6. Chen C, Li Z, (2006) A low-power CMOS analog multiplier. IEEE Trans Circ Sys Exp Briefs 53(2):100–104

Chapter 68
Intelligent Enterprise Application Servers: A Vision for Self-Managing Performance

G. Ravi Kumar, C. Muthusamy and A. Vinaya Babu

Abstract The recent trends in computing are posing performance challenges to Enterprise Application Server environments. Such emerging challenges triggered the development of self-managing and self-correcting computing and application server systems. There are attempts in building self-correcting Application Servers such as Web, EJB Servers. But majority of such mechanisms are solutions to specific components of the Application Servers. In this paper a vision of an Intelligent Enterprise Application Servers is discussed and intelligent control system based framework is proposed.

Keywords Intelligent control · Control system · Enterprise application servers

68.1 Introduction

IT Resources usage is rapidly increasing and becoming complex [1] due to the recent trends in computing environments. The users of compute resources are at individual and enterprise levels using for utility bill payments, online banking,

G. R. Kumar (✉)
HP Bangalore, JNTUH, Hyderabad, India
e-mail: ravikgullapalli@yahoo.co.in

C. Muthusamy
Yahoo, Bangalore, India
e-mail: chelgeetha@yahoo.com

A. Vinaya Babu
JNTUH Coll of Engineering, Hyderabad, India
e-mail: dravinaybabu@yahoo.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_68,

retailing and auctioning [2]. Majority of the software applications and services are deployed into the Enterprise Application Servers. There is a constant challenge of managing the performance SLAs [3] in such complex environment. Self-managing and self-correcting architectures emerged from such requirements. In this paper we focus on building Intelligent Application Servers capable of self-managing the performance at runtime.

68.2 Problem and Related Work

Application Servers are runtime platforms to host and run distributed software systems. They require dealing with performance challenges considering recent trends in computing such as high volume of users, rich internet applications, heterogeneity of computing and software systems [2]. The challenge of providing higher performance is further complex, which triggered self-managing application server systems. Towards building them an Automated Tuning System was proposed to tune Application Server QoS dealing with micro level adaptation in [4]. Feedback control system based solutions explored in JEE Server components such as Web Servers [5], EJB Servers [6]. PI, PID Controllers, Fuzzy and Neural Controllers, hybrid controllers [7, 8] are applied to self-manage the performance specific to Server components such as EJB, Web, JMS Providers, but no end-to-end solution to enrich Application Servers with self-managing abilities is present. Feedback control as a service is in investigation by Tharindu [9] for cloud environments. Online model tuning is suggested by Viraj et al. [10], Loris et al. proposed selfware platforms for autonomic management of J2EE Servers in [11], which are the motivations for the proposed solution. In this paper a vision of an Intelligent Enterprise Application Servers capable of self-managing performance is discussed and an intelligent control system based framework is proposed for Enterprise Application servers.

68.3 Vision

Intelligent Enterprise Application Servers such as JEE Servers [12] require the techniques to incorporate the self-managing capabilities during the server modeling and design time. The components are typically implementation of object oriented models. It is essential to appropriately model using ARMA [13] models or other similar techniques, in order to self-manage their performance in runtime. This requires a generic modeling framework that enhances UML [14] tools to add self-managing features during system modeling and design. The server component implementations need the necessary aspects that are executed at runtime. The self-managing features will be potentially executed at the runtime of the Application Servers. The runtime self-managing capabilities have to initialize the component

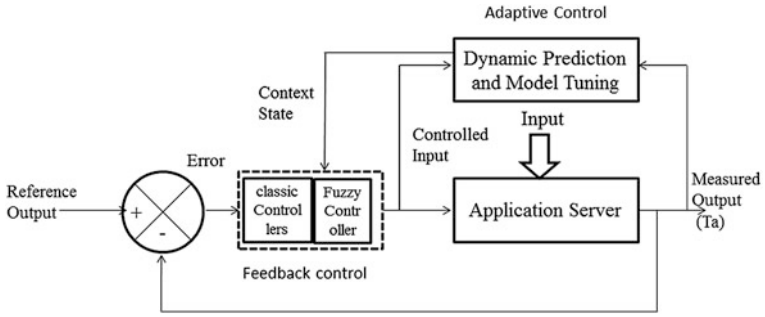


Fig. 68.1 Intelligent control system

models, compute the model parameters, dynamically adjust the models, able to predict the server behavior for the future periods of time accurately, provide a set of prediction algorithms as the same algorithm cannot be applied to different data sets. A set of classic, hybrid feedback controllers have to be applied to tune the server performance. Advanced adaptive controllers are also required to improved performance management. In order to avoid human intervention, the self-managing system requires the ability to select right controller based on the system current and predicted dynamics. A complete realization of the discussed vision would aid in building Intelligent Application Servers.

68.4 Intelligent Control Solution

An Intelligent control solution is proposed for self-managing the performance of Enterprise Application Servers in an automated manner. Control Systems concepts have inherent feedback and adaptive capabilities, this solution aims at incorporating them as first class elements within the Enterprise Application Server components. The solution has two distinct parts. The first part provides design constructs for various server components modeling and design through UML modeling [15]. The second part of the solution provides a dynamic modeling and predictive framework to provide self-managing abilities during the runtime of the server. The current paper focusses on the second part of the solution as shown in the Fig. 68.1 below which is the proposed Intelligent control solution. The Adaptive control has a set of prediction algorithms used to predict the state of the server component called as ContexState, for a given number of periods. The predicted state helps in choosing a right action or a controller. ARMA based models are considered to the JEE server component modeling as represented by the (1) below:

$$y(t+1) = ay(t) + bu(t). \tag{68.1}$$

Where

```
<ContextState>
  <LongPattern>
   <predictedValue>SuddenVariations</predictedValue>
   <periodsAhead>10</periodsAhead>
   <ControllerType>Fuzzy</ControllerType>
   <CPU>75</CPU>
   <MEM>68</MEM>
   <error>15</error>
  </LongPattern>
 <ImmediatePattern>
   <predictedValue>SuddenVariations</predictedValue>
   <periodsAhead>2</periodsAhead>
   <ControllerType>Fuzzy</ControllerType>
   <CPU>75</CPU>
   <MEM>68</MEM>
   <error>15</error>
 </ImmediatePattern >
  <OutParam>MessageThroughput</OutParam>
  <OutParamVal>50</OutParamVal>
  <InParam>Smax</InParam>
  <InParamVal>75</InParamVal>
  <InParamModel>a</InParamModel>
  <InParamValue>0.5</InParamValue>
 </ContextState>
```

Fig. 68.2 ContextState

$y(t)$ = current output, $u(t)$ = current input
a = model parameter, b = model parameter
$y(t + 1)$ = output in the next step

The adaptive control triggers dynamic model adjustment to tune model parameters when the prediction error increases beyond a predefined error limit.

68.4.1 ContextState

The Fig. 68.2 below shows the sample XML description of the ContextState for a JMS Provider. It consists of the LongPattern which consists of the data predicted for a given number of periods ahead, along with other parameters predicted for the given component. The same set of values is part of ImmediatePattern which is one period ahead. The action to be taken or controller to be chosen is decided by a Two-Level Prediction Algorithm that compares these Long and Immediate Pattern values. The context state also contains the server model, its input and output values, model parameters.

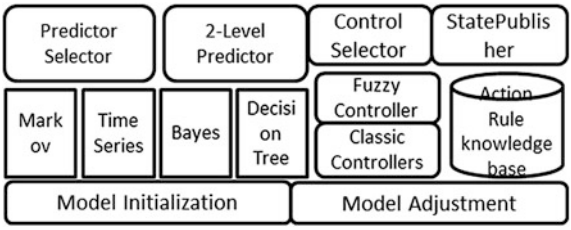


Fig. 68.3 Intelligent control architecture

68.4.2 Architecture

The Fig. 68.3 below shows the architecture of the proposed solution. The classic and fuzzy controllers are present in the feedback path of the control system. The remaining functional blocks are present in the Adaptive control part of the control system.

68.4.3 Adaptive Control

The adaptive control loop of the Fig. 68.1 consists of the functional blocks capable of performing dynamic prediction and tuning the models as explained below:

Predictor Set: The adaptive control functional block consists of a predictor set implementations such as Markov model, Bayes's rule, Decision Tree, Time-Series algorithms. The Predictor Selector chooses suitable predictor at runtime based on the data set

Two-Level (2-level) Prediction: There is a two-level prediction algorithm that runs to predict the context states. The algorithm at any point of time predicts the ContextState for 'm' periods ahead and stored as LongPattern data. After 'm-(m−1)' periods, the prediction is again run to generate ImmediatePattern data. If the both these patterns are same, then the ImmediatePattern ContextState is passed to the feedback control. If there are not same, then an average of the context state properties of both patterns is computed and is sent as ContextState to the feedback control.

```
program TwoLevelPrediction
  LongPatternPeriod = 10; ContextState ctxState;
  ImmediatePatternPeriod = LongPatternPeriod - 1;
  ContextState ctxStateLongPatt;
  ContextState ctxStateImmPatt;
while (true)
begin
  ctxStateLongPatt = predict(LongPatternPeriod);
  ctxStateImmPatt = predict(ImmediatePatternPeriod);
```

Table 68.1 Context state and control selector rules

Context state	Controller action
Sudden variations	Fuzzy
Gradual increase	PID
Constant	PI
Quick adaptation	PID
Model adjust	Param estimate

```
        if (ctxStateLongPatt equals ctxStateImmPatt)
        then
         ctxState = ctxStateImmPatt
        else ctxState = avg(ctxStateImmPatt,
              ctxStateLongPatt)
        publish(ctxState)
        end if
end while
```

Model Adjustment: The Model Tuner is responsible for tuning the model parameters based on the context state predicted. When the prediction error exceeds a pre-defined threshold then the re-estimate of model parameters is triggered, but initial values of are estimated using the training data.

StatePublisher: StatePublisher is responsible in passing the predicted context state to the feedback element of the control system to actually perform the system tuning.

Action Rule Knowledgebase: This is IF–THEN rules database used by Fuzzy control to take actions based on the context state received from the StatePublisher. It also contains the type of controller required or action to be taken based on the context state.

Control Selector: The control selector chooses the controller to be employed to tune the server component based on the context state predicted using pre-defined control selector rules as shown in the Table 68.1 stored in Action Rule knowledgebase.

68.4.4 Feedback Control

Classic Controllers: P, PI and PID Controllers are the classic feedback controllers available as a set. The control selector chooses the controller to be actuated based on the context state predicted.

Fuzzy Controller: The Fuzzy controller is a set of IF–THEN rules that are captured initially using training data set. During the runtime, a new set of rules are identified and captured into the Action Rule knowledge base.

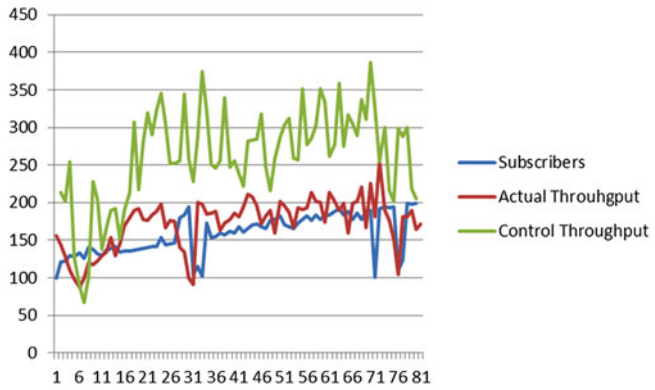


Fig. 68.4 Message throughput

68.5 Implementation

A preliminary implementation of predictor and control selector, rules and intelligent controller is available on a simulated data set and the Fig. 68.4 below shows the Message throughput. Based on the context state, the controllers selected by the control selector are PI and Fuzzy controllers. The average controlled message throughput is much higher than average Actual throughput against the number of subscribers over a period of time.

68.6 Discussion and Future Work

The proposed solution in order to build Intelligent Enterprise Application Servers defines the server component state as a ContextState. It gives detailed information about the server current and the predicted states, enabling appropriate tuning. A set of predictors are used rather than using a single prediction algorithm for ContextState, providing flexibility to choose the suitable algorithm based on data set. A few building blocks of the proposed solution implemented for JDBC driver [16] and JMS Providers [17]. The majority of the proposed architecture is currently under implementation, there are efforts to progress and evaluate further. There is a necessity of performance analysis using SASO properties [18] for the controllers and measure overhead introduced by the proposed solution.

References

1. Giovanna F, Ezhilchelvan P, Mitrani I (2006) Performance modeling and evaluation of E-business systems. IEEE ANSS

2. Cornel K et al (2008) A survey of context adaptation in autonomic computing. IEEE ICAS
3. Chen Y, Iyer S, Liu X, Milojicic D, Sahai A (2007) SLA decomposition: translating service level objectives to system level thresholds. ICAC
4. Giovanna F, Shrivastava S, Ezhilchelvan P (2004) An approach to adaptive performance tuning of application servers. In: IEEE international workshop on QoS in application servers
5. Tarek A, Lu Y, Zhana R, Henriksson D (2004) Practical application of control theory to web services. In: American control conference
6. Yan Z, Qu W, Liu A (2006) Adaptive self-configuration architecture for J2EE-based middleware. HICSS 9
7. Palden L, Zhou X (2010) Autonomic provisioning with self-adaptive neural fuzzy control for end-to-end delay guarantee. IEEE international symposium on modeling, analysis and simulation of computer telecommunication systems
8. Ravi Kumar G, Muthusamy C, Vinaya Babu A A study of intelligent controllers application in distributed system. INDJCSE 2(4):589
9. Tharindu P, Colman A (2010) Feedback controllers in the cloud. Swinburne University, Cloud Workshop
10. Viraj B et al Enabling self-managing applications using model-based online control strategies. ICAC
11. Loris B et al (2008) Autonomic management of J2EE servers
12. JEE: http://www.oracle.com/technetwork/java/javaee/tech/index.html
13. ARMA: http://en.wikipedia.org/wiki/Autoregressive_moving_average_model
14. UML: http://en.wikipedia.org/wiki/Unified_Modeling_Language
15. Ravi Kumar G, Muthusamy C, Vinaya Babu A (2012) Design and modeling autonomic aware software in UML—a control system solution. ICCIT (Unpublished)
16. Ravi Kumar G, Muthusamy C, Vinaya Babu A, Marndi RN A feedback control solution in improving database driver caching. IJEST 3(7):5722–5727
17. Ravi Kumar G, Muthusamy C, Vinaya Babu A (2012) Self-regulating message throughput in enterprise messaging servers—a feedback control solution. IJACSA 3(1):148–155
18. Hellerstein JL, Diao Y, Parekh S, Tilbury (2004) Feedback control of computing systems. Wiley, New York

Chapter 69
A Review of Disc Scrubbing and Intra Disk Redundancy for Reducing Data Loss in Disk FileSystems

Genti Daci and Aisa Bezhani

Abstract Because of high demand that applications and new technologies have today for data storage capacity, more disk drives are needed, resulting in increased probability to inaccessible sectors, referred as Latent Sector Errors. Aiming to reduce data loss by LSE, two main techniques are extensively studied lately: Disk Scrubbing, which performs reading operations during idle periods on systems to search for errors and Intra Disk Redundancy which is based on redundancy codes. Based on previous evaluation results, we discuss and introduce the benefits on using both schemes simultaneously: combining different IDR coding schemes with different scrubbing strategies in particular regions that store crucial data. Finally, we apply a dynamic scheduling algorithm for a minimum impact on system performance.

Keywords Latent sector errors · Disk scrubbing · Intra disk redundancy

69.1 Introduction

LSEs leave the HDD functioning but corrupts the data. There are several reasons on why we focus on this subject: The usage of cheaper disk drives, which at the same time are less reliable and the need for more storage and faster performance.

G. Daci (✉) · A. Bezhani
Faculty of Information Technology, Polytechnic University of Tirana,
Tirana, Republic of Albania
e-mail: gdaci@ieee.org

A. Bezhani
e-mail: aisa_bezhani@live.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_69,

Also, during Raid reconstruction, if the reading operation finds a damaged sector, the data will be lost. IDR adds an extra redundancy level into disk drives, while Disc Scrubbing scans periodically to find media errors and so to prevent data loss, having a positive impact on MTTDL. Disc scrubbing and IDR, were first evaluated based on the assumption that LSE is similar to the Poisson distribution [1]. IPC was thought to give the same reliability on data as an OS without irrecoverable sectors [2]. After an expanded analyze of data in [3], a better statistical approach was adopted, the Pareto distribution. As a result, a reconsideration of the older IDR techniques like: SPC, IPC [4], MDS is done and new techniques are created: hybrid SPC and MDC, and CDP. The same thing applies to Disk scrubbing: Localized, accelerated, staggered [5], and accelerated staggered scrubbing. As a result of new strategies, we reconsider the simultaneous usage of methods, to detect errors and then to correct data, preventing it from loss.

69.2 LSE

Error bursts: By error burst we mean a series of contiguous errors in a logical block space. 90–98 % of error bursts was created by a single error. Another factor is the distance of error bursts. Depending on the disc type, 20–50 % of all errors happened on the first 10 % of the disk, on others, this percentage was focused at the last sectors. On 60 % of the cases the problem was identified by disk scrubbing.

Spatial and local correlation: If we want to determine that an IDR scheme is appropriate for a certain system, despite of the distance that two errors occur, will the system have the necessary time to recover from the first error before the second one occurs? Over 90 % of burst errors were discovered in 2 weeks, and over 95 % were detected in a 1 month distance.

69.3 Intra Disk Redundancy

Logic structure: IDR has a simple logic: every stripe is divided into segments. There are data bits and parity bits inside each segment. Parity bits are created using different parity codes. The schemes proposed change the security level according to disc space and overhead penalties. IPC uses an internal scheme which is based on simple XOR operations. This scheme guarantees security in Raid systems, without a noticeable growth in overhead. The way IPC works is: N sequential data sectors are put in a certain way to create a matrix. When data is updated, the parity sectors are updated too. According to Poisson distribution it reached the same reliability state comparable to a system without inaccessible sectors. To reach this state of security, [2] increased the disk size by 6 % in order to store the same amount of data. MDS is compound of k data sectors along with m parity sectors.

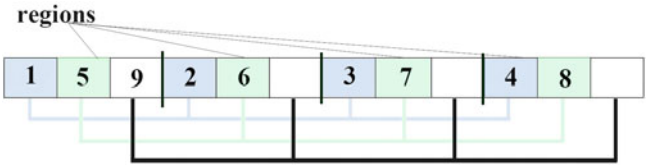


Fig. 69.1 Staggered scrubbing

This scheme can tolerate the loss of m sectors per segment. It can recover the data for m errors, where m is the number of sector that the parity bits occupy. Obviously there is higher reliability, but also higher overhead and parity bits take a considerable part of the hard disc space not usable for actual data [1]. CDP can detect more errors and it has less overhead if compared to MDS. This code is based on RDP which is used to recover from double failures on disk and it works with IDR too.

69.4 Disk Scrubbing

Schemes: Sequential scrubbing is a process which scans and reads disc sectors one after another. The moment it scans the last sectors, the process repeats. Random scrubbing is similar to sequential scrubbing; with the only difference that accessing is done in a random order. The problem in this case is that several sectors may be scanned many times, while others may not be checked at all.

Staggered scrubbing, divides the disc in m regions. Each division is compound of r segments. It reads the first segment of every single region according to Logical Block Address (LBA) as shown in Fig. 69.1. After that it starts reading the second segments of the regions, and so on. Accelerated scrubbing: When an error is detected, the rest of the disk sectors are scanned with an accelerated rate, while Accelerated staggered scrubbing combines both techniques, as the name implies. It detects the erroneous sectors, and then scans the entire disk with accelerated rate.

69.5 Our Approach

The idea is to use simultaneously Disc Scrubbing to detect media errors on disc and IDR to rebuild the corrupted data. To avoid the overhead we use both, disc scrubbing and IDR as background activities, so it can impact as less as possible the normal operation of the system [6]. It is proven that the cooperation of the two gives a high reliability, more than their linear when they operate alone in a system. We consider adapting Busy Bee [7]. Its scheduling algorithm adapts dynamically,

according on the current workload. Also, it has a crucial feature: this policy doesn't leave the background operation starve.

```
do {
If (high probability of short_foreground_job_coming)
  { if (scanner_head < 0.1*n)//n-total_number of sectors
Apply_sequential_scanning
  else   if   (scanner_head > = 0.1*n)   Apply_staggered_
scrubbing
   else if (LSE detected)
{ Apply_accelerated_staggered_scrubbing
    Call     intradisk_redundacy_operation_on_affected_
sector}
else wait for foreground job }}
While (idle_time)
```

The above pseudo-code reveals the concept on which this theory is based. The scrubbing operation happens during idle time, and because it is infinite, we continue until a foreground job is present. If the foreground job is statistically thought to be short, the system continues to scrub. In the case it detects a LSE, weather the foreground job is short or long, the system doesn't stop scanning. In this case we prefer to secure our data, not minding the overhead in this extreme case of risking data loss. Otherwise, the scrubbing stops. In order to catch a LSE as soon as possible, we apply three scanning schemes that we apply. A sequential scrub is done on the first 10 % of the total number of sectors, knowing that it has the greatest probability of developing LSE. After that we chose staggered scrubbing, knowing that it is the one to detect the errors first, while not affecting performance as much as the staggered strategies. Last, if we encounter an error, we want to prevent losing data in any cost. Accelerated staggered scrubbing is applied, not minding its overhead. In this moment, IDR is called to perform a reading of the damaged sector, and also a recalculation, so it can locate the error, and fix it. Our options are: SPC which is simple, but it is not appropriate since bust errors have spatial locality. This means that if two errors occur in the same logical space, our data can't be reconstructed, and the scrubbing would be pointless. Next approach would be MDS, but using RS codes imply high overhead, so we use CDP, which tolerates a considerable number of failure patterns and implies a smaller processing time. While doing so, we scan at an accelerated rate, to detect any other possible erroneous areas. In real life, there is no way for us to predict if the coming job is going to be long or short. In this case, if we are expecting a short job the scrubbing process will start. If a long job comes, the scrubbing will stop, giving precedence to the foreground job, if and only if an error is not detected at that moment.

69.6 Conclusions

Based on previous analyses our aim was to find the best solution while using disc scrubbing and IDR simultaneously. We proposed a combination of disk scrubbing techniques by which we can detect LSE in the fastest way possible. By doing this we can allow the usage of IDR parity codes that will not imply much overhead. Opting for a minimal impact on performance we use both strategies as background operations, applying a dynamic algorithm.

References

1. Bairavasundaram LN, Goodson GR, Pasupathy S, Schindler J (2007) An analysis of latent sector errors in disk drives. In: Proceedings of the 2007 ACM SIGMETRICS international conference on Measurement and modeling of computer systems. California
2. Iliadis I, Haas R, Hu XH, Eleftheriou E (2011) Disk Scrubbing versus intradisk redundancy for RAID storage. In: ACM transactions on storage (TOS), Vol 7 No 2, pp 1–42
3. Schroeder B, Damouras S, Gill P (2010) Understanding latent sector errors and how to protect against them. In: ACM Transactions on storage (TOS), Vol 6, no 3, pp 1–23
4. Dholakia A, Eleftheriou E, Hu X-Y, Iliadis I, Menon J, Rao KK (2008) A new intra-disk redundancy scheme for high-reliability RAID storage systems in the presence of unrecoverable errors. In: ACM TOS, Vol 4, 1, pp 1–42
5. Oprea A, Juels A (2010) A clean-slate look at disk scrubbing. In: Proceedings of the 8th USENIX conference on File and storage technologies, pp 5–5, California
6. Mi N, Riska A, Smirni E, Riedel E (2008) Enhancing data availability in disk drives through background activities. In Proceedings of DSN
7. Yan F, Riska A, Smirni E (2012) Busy bee: how to use traffic information for better scheduling of background tasks. In: ICPE '12 proceedings of the third joint WOSP/SIPEW international conference on performance engineering, pp 145–156
8. Grawinkel M, Schafer T, Brinkmann A, Hagemeyer J, Porrmann M (2011) Evaluation of applied intra-disk redundancy schemes to improve single disk. In: IEEE 19th annual international symposium on modeling analysis and simulation of computer and telecommunication systems, pp 297–306

Chapter 70
Homogeneous and Heterogeneous Energy Schemes for Hierarchical Cluster Based Routing Protocols in WSN: A Survey

M. Jagadeeswara Reddy, P. Suman Prakash and P. Chenna Reddy

Abstract Recent technological advances in WSN's have led to so many new hierarchical cluster based protocols especially designed for sensor networks where energy metric is essentially concerned to increase the network life time. Therefore the energy of the sensor node is a crucial parameter in the protocol design to increase the lifetime of a sensor node in the network. In this concern much research already done in recent years, analysis different aspects like as, energy efficiency, power consumption, routing, Quality-of-service, cluster formation and so on. Each routing protocol is examined and explore under hierarchal type. In this paper we mainly focus on surveying hierarchal cluster based routing protocols for two different environments such as homogeneous and heterogeneous.

70.1 Introduction

Recent advances in wireless communications technology and MEMS and low power and integrated digital electronics have enabled the development and deployment of low energy, tiny, micro-sensors in a Wireless sensor Networks (WSN) [1–5]. WSN's are emerging technologies for current century [6, 7]. These huge no of sensors are distributed anonymously to detect and monitor physical or environmental conditions such as vibration, pressure, temperature, sound etc. [1] and provides opportunities for a wide variety applications like as battle field monitoring, civilian and military applications [8]. Initially wireless sensor networks are motivated by military applications now used in many application areas

M. Jagadeeswara Reddy (✉) · P. Suman Prakash · P. Chenna Reddy
CSE, Jawaharlal Nehru Technological University College of Engineering, Pulivendula, India

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_70,

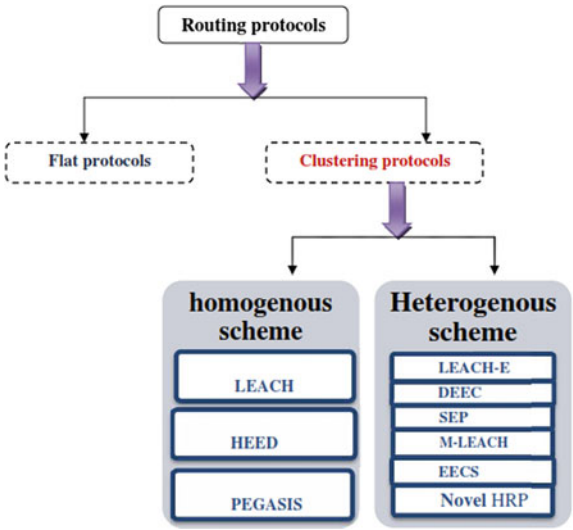


Fig. 70.1 Classification of clustering protocols

wild animal monitoring, home automation, health care applications, asset tracking. Clustering topology [9] is requires for organizing nodes, efficient data gathering, data aggregation [10], and reduced redundant transmissions and so on.

70.2 Related Work

WSN's are emerged research area from both the real-time users and research community. Previous researcher are focused on several routing protocols in the order of LEACH [9], TEEN [10], APTEEN [11], PEGASIS [12], SEP [13], HEED [14], EECS [15], DEEC [16], H-HEED [17], ACE [18], TPC [19], E-LEACH [20], TL-LEACH [21], M-LEACH [22], LEACH-C [23] and V-LEACH [24] and so on.

70.3 Cluster-Based Routing Protocols in WSN

The Clustering Protocols [3] can be also categorized into two subclasses: the clustering algorithm **homogeneous schemes** (same energy levels), **heterogeneous clustering schemes** (different energy's and environments), where all the node of the sensor network are equipped with different amount of energy. This classification is shown in Fig. 70.1.

Clustering Approach	scheme	Cluster formation & Cluster head selection Based on
LEACH-E	Heterogeneous	Same as LEACH but uses different threshold values **&**residual energy and initial energy of node
DEEC		Same as LEACH but uses different threshold values **&** ratio b/w residual energy of each nodes and the average energy of the network
SEP		with different energies **&** weighted election probabilities of each node to become cluster head according to the remaining energy in each node
M-LEACH		Randomized clustering **&** residual energy of existing cluster heads, distance & hold back value
EECS		Using intra-cluster communication cost and cost of communication between CH and BS **&** elects the cluster-heads with more residual energy through local radio communication
Novel HRP		Same as LEACH but while advertisement phase different **&** based on node residual energy, distance), number of transmission bits and transmission power
LEACH	Homogeneous	By using localized control and coordination **&** randomly and based on threshold value
HEED		clusters can be formed in several iterations **&**based on node residual energy &intra-cluster communication
PEGASIS		Greedy method **&** randomly

Fig. 70.2 Survey table on homogeneous & heterogeneous energy schemes

70.3.1 Homogeneous Schemes

Under Homogeneous environment most of the clustering protocols (LEACH with some versions, HEED, PEGASIS, ACE, CBRP, EECH and so on) are energy efficient, apart from LEACH is a basic to efficient technique to form a clusters is as follows

70.3.1.1 Low Energy Adaptive Cluster Based Hierarchy (LEACH)

This is the first hierarchical clustering protocol for WSN, which provides efficient load balancing, cluster head formation. In this protocol, every node select one random number, which is less than the threshold value T(n), they become cluster heads for that round.

$$T(n) = \left\{ \begin{array}{cc} \frac{p}{1-p*\left(r \bmod \frac{1}{p}\right)} & if n \in G \\ 0 & otherwise \end{array} \right\} \tag{70.1}$$

70.3.2 Heterogeneous Schemes

In Heterogeneous environment nodes are having different energy levels, transmission and environments. So clustering is a challenging task, some of them are energy efficient protocols with direct communication (E-LEACH, DEEC, SEP, EECS, Novel HRP and so on), M-LEACH with Multi-hop communication. Mobility is fixed for all except in DEEC; it allows micro mobility. Fig. 70.2 illustrates the comparison between homogeneous and heterogeneous environments.

70.4 Conclusion

In WSN's Energy efficiency is achieved by using efficient clustering mechanism in order to reduce energy consumption, improve link stability and increase network life time. Efficient clustering is easy in homogeneous but little bit tough in heterogeneous considerations. This paper illustrates the comparison of protocols in both environments. Apart from all some protocols improves network lifetime and minimize energy consumption.

References

1. Akyildiz IF (2002) Wireless sensor networks: a survey. Comput Netw 38:393–422
2. Sohrabi K (2000) Protocols for self-organization of a wireless sensor network. IEEE 7:16–27
3. Min R (2001) Low power wireless sensor networks. In: Proceedings of international conference on VLSI design, Bangalore
4. Rabaey JM (2000) PicoRadio supports ad hoc ultra low power wireless networking. IEEE Comput 33(7):42–48
5. Katz RH, Kahn JM, Pister KSJ (1999) Mobile networking for smart dust. In: (MobiCom_99), Seattle
6. Chong CY, Kumar SP (2003) Sensor networks: evolution, opportunities, and challenges. IEEE 91:1247–1256
7. Neil G (1999) 21 ideas for the 21st century: business week, pp 78–167
8. Lindsey S, Raghavendra C, Sivalingam KM (2002) Data gathering algorithms in sensor networks using energy metrics. IEEE Trans Parallel Distrib Syst 13(9):924–935
9. Heinzelman W, Chandrakasan A (2000) Energy-efficient communication protocol for wireless micro sensor networks. Hawaii, 4 Jan
10. Agarwal DP, Manjeshwar A (2001) TEEN: a protocol for enhanced efficiency in wireless sensor networks. In: IEEE, San Francisco
11. Agarwal DP, Manjeshwar A (2002) APTEEN: a hybrid protocol for efficient routing and comprehensive information retrieval in wirelesssensor networks. In: IPDPS, Fort Lauderdale
12. Lindsey S, Raghavendra C (2002) PEGASIS: power efficient gathering in sensor information systems. In: IEEE Aerospace Conference, Big Sky
13. Matta I, Bestavros A (2004) SEP: Astable election protocol for clustered heterogeneous wireless sensor networks

14. Younis O, Sonia (2004) F HEED: A hybrid, energy-efficient, distributed clustering approach for ad-hoc sensor networks, IEEE 3(4):369–379
15. Mao Y, Chengfa L (2005) EECS: an energy efficient clustering scheme in wireless sensor networks. In: IEEE
16. Qing L, Mingwan W (2006) Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks. In: IEEE
17. Ajay Sharma K, Hammet K (2010) Heterogenous-hybrid energy efficient distributed routing protocol
18. Haowen C, Adrian P ACE: an emergent algorithm for highly uniform cluster formation, Carnegie Mellon University. Pittsburgh
19. Choi W, Shah P (2004) Framework for energy-saving data gathering using two-phase clustering. In: WSN, MobiQuitous'04
20. Xiangning F, Song Y (2007) Improvement on LEACH protocol of wireless sensor network. In: Proceedings of the improvement on LEACH protocol of wireless sensor network (VLEACH)
21. Loscrì V, Morabito G, Marano S () A two-levels hierarchy for low-energy adaptive clustering hierarchy
22. Hang Z, Jiang Zhe (2006) Study and design on cluster routing protocols of wireless sensor networks
23. Heinzelman WB (2002) An application-specific protocol architecture for wireless microsensor networks. In: IEEE
24. Karaki AL (2004) Data aggregation in wireless sensor networks-exact and approximate algorithms. IEEE, pp 241
25. Zheng J, Jamalipour A (2009) Wireless sensor networks: a networking perspective. Wiley-IEEE Press, Chichester
26. Dechene DJ, El Jardali A, Luccini M, Sauer A (2010) A survey of clustering algorithms for wireless sensor networks, University Of Western Ontario London

Chapter 71
Itinerary Management System

Janhavi Baikerikar, Saket Bhat, Vaibhav Baliga, Alfred Almeida, Abhay Tripathi and Lionel D'souza

Abstract Usually, individual tourists tend to plan their itinerary well ahead they arrive, either through tourist guide or online information sources. With the enhancement of information technology, it is expected that mobile computed itinerary planning would be a complete supplementary to the hard copy travel guides and magazines in the future. However, current online itinerary planner overlooks the transportation link and optimum travel plan. It is this place where this paper plays its role. This paper aims to develop a guiding solution based on the B+ Trees Algorithm that helps in preparing an itinerary for tourists visiting any city on an individual basis. Our aim is to minimize the travelling time and maximize the time of sightseeing, shortest travelling time between tourist spots.

Keywords B+ Trees · Itinerary · Mobile computing

J. Baikerikar (✉)
Information Technology, Don Bosco Institute of Technology, Mumbai, Maharashtra, India
e-mail: janhavibaikerikar@gmail.com

S. Bhat · V. Baliga · A. Almeida · A. Tripathi · L. D'souza
Information Technology, Don Bosco Institute of Technology, Mumbai, Maharashtra State, India
e-mail: ssbhat06@yahoo.co.in

V. Baliga
e-mail: vaibhav.baliga@yahoo.in

A. Almeida
e-mail: alfredalmeida07@yahoo.com

A. Tripathi
e-mail: abhay.trips@gmail.com

L. D'souza
e-mail: lioneld91@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_71,

71.1 Introduction

This paper basically provides itinerary for tourists with android based mobile phone that supports online maps. Combined with the information present in the database, this application enhances the way users interact with the physical world, adding additional information about tourist spots, ac, travel guidance, in order to evoke previous memories or complement present stories. This paper uses B+ Trees Algorithm which is used for selecting optimal path between two nodes. i.e. (cost efficient and shortest distance) [7]. Presently, itinerary management systems are web based applications. Here, our basic aim is to use online maps and provide a pocket based application. The paper is organized as follows—Sect. 71.2 describes itinerary followed by itinerary management system in Sect. 71.3. Section 71.4 describes B+ tree algorithm followed by system workflow and results in Sects. 71.5 and 71.6. The conclusion future work is described in Sect. 71.7.

71.2 Itinerary

Itinerary may include geographic location, transport, shopping, dining or any other information of tourist spots limited to a particular area [5].

71.2.1 Information Gathering

Many different options like books are available to the tourists to know about the places to visit and attractions nearby but they may not suffice their need of subject to change. But with changing times, the people also these days may not buy the books, instead they may prefer visiting sites or just get information on the go through a mobile application.

71.2.2 Tourist Spot Info

We have taken into account Mumbai city in Maharashtra state of the Indian Province which is one of the busiest cities in the world. We have tried putting all the tourist spots which should interest the tourist together, which was a herculean task.

71.2.3 Transport

Mumbai Transport network is very large and hence extensive study of all viable routes and different modes of transport and their costs was made through various sources available online.

71.3 Itinerary Management System

Our aim is to design an itinerary planner for PDA's which work on android operating system. This application which is to be developed using different open source technologies is aimed at providing a better solution than the present web based itinerary planner for the ever broadening world of mobile applications. The basic idea is to give the user a schedule planned with the shortest possible routes and the cost effective one. The idea revolves around extracting information from the databases and representing the info to user on mobile interface using online maps thus making the application use possible on the go without wi-fi or internet connection.

71.4 B+ Trees Algorithm

If there are multiple routes then the shortest route will be processed using the following method. Also cost efficient and the time constraint will be met. In a B+ tree, all records are stored at the leaf level of the tree; only keys are stored at the interior nodes. The primary value of a B+ tree is in storing data for efficient retrieval in a block Oriented storage context—in particular, file systems. This is primarily because unlike Binary Search Trees, B+ trees have very high fanout (typically on the order of 100 or more), which reduces the number of I/O operations required to find an element in the tree. The order, or branching factor *b* of a B+ tree measures the capacity of nodes (i.e. the number of children nodes) for internal nodes in the tree. The actual number of children for a node, referred to here as *m*, is constrained for internal nodes so that [b/2 < = m < = b]. The root is an exception: it is allowed to have as few as two children. The algorithm to perform a search for a record r follows pointers to the correct child of each node until a leaf is reached. Then, the leaf is scanned until the correct record is found (or until failure). (Figs. 71.1, 71.2).

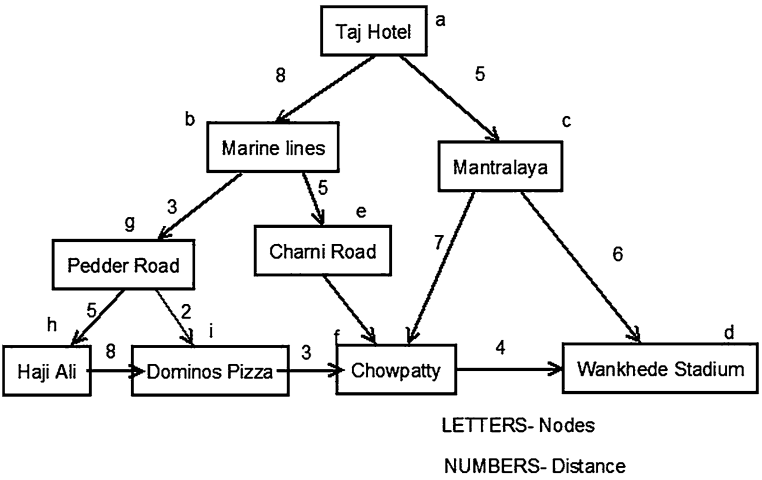
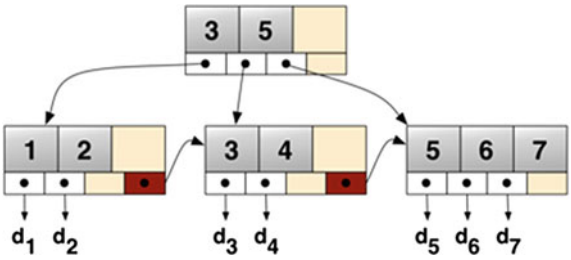


Fig. 71.1

Fig. 71.2



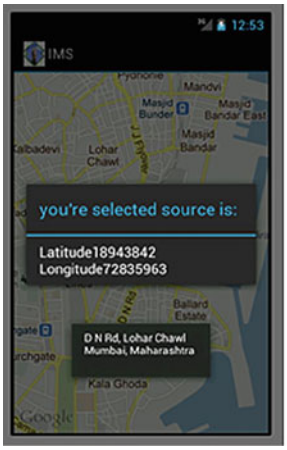
71.5 System Workflow

The Application mainly being designed for portable devices like mobiles and tabs, with android as operating system, we are making the use of analogous database engine SQlite which is integrated with the eclipse idk and has classes defined for android coding. Sqlite is used as it uses very little memory for execution as low as 250 kb. All the queries can be written in the Sqlite database browser which is to be downloaded.

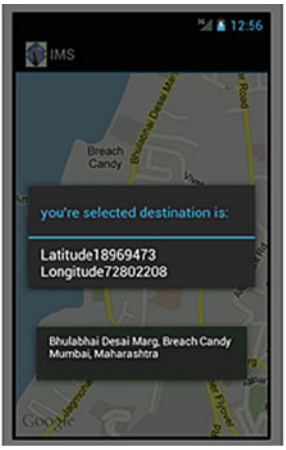
It takes four steps to build the customized itinerary:

- Start point should be chosen first since it is the first and last node of the itinerary.
- Secondly, the arrival and departure time would let the system know how much time the user would be make himself available each day.
- User should then select the tourist spots and the fares.
- The system would then process the information and generate the itinerary using the algorithm specified in the section above.

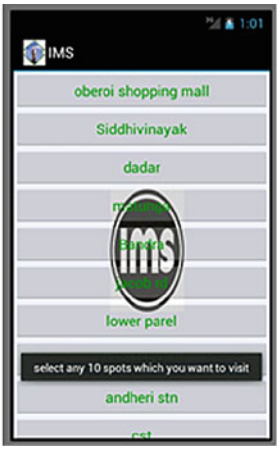
71.6 Results



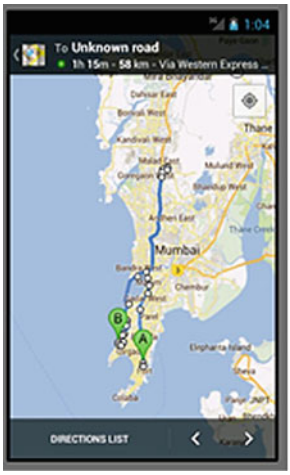
The Eclipse compiler compiles the code and displays the home screen. Here the user selects his Starting location and the latitude and longitude of it is displayed.



Here the user selects the destination he intends to go to and its corresponding latitude and longitude is displayed.

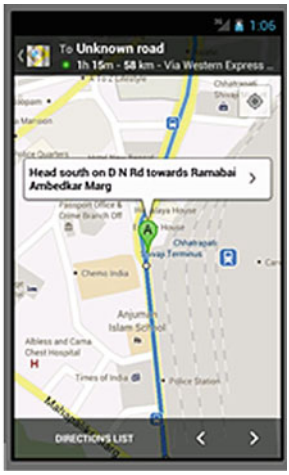


Here the database retrieves the spots available and the user selects the spots he would like to visit.



Here the map overlay algorithm draws a line from the source to the destination, thereby showing the route to the user and also displays the distance between the source and destination.

In our application the user can further scroll down and request for directions. Therefore the user receives detailed step by step directions to reach his destination.



This snapshot displays a user reaching his destination.

71.7 Conclusion and Future Work

The application has a wide scope with the flourishing of mobile phones in the decade and rise of demand for android applications. The application can be used by tourist as well as people living in that city to enhance their knowledge about different spots. We have restricted the area to the city of Mumbai with further research and more efficient algorithms this system can be transformed into an application on the go which can be used online. It can be also used to share tourist experiences and reviews about the spots for any city in any country.

References

1. Android SQLite Database and Content Provider—Tutorial http://www.vogella.de/articles/AndroidSQLite/article.html
2. Tam H, Pen Chung L Evaluation of online Itinerary planner http://www.isprs.org/proceedings/XXXVIII/part2/Presentations/S8/Tam.pdf
3. Wikipedia http://en.wikipedia.org/wiki/B %2B_tree

Chapter 72
Trust and Reliability Based Scheduling Algorithm for Cloud IaaS

Punit Gupta, Mayank Kumar Goyal, Prakash Kumar and Alok Aggarwal

Abstract Trust Models are used to enhance secure and reliable scheduling in Distributed, Grid and Cloud environment. Trust models that are being proposed or implemented in Distributed and Grid environment, does not fully fit in cloud computing environment. Since the parameters that have being taken into consideration in these trust models, does not fit in the cloud Infrastructure As A Service, a suitable trust model is proposed based on the existing model that is suitable for trust value management for the cloud IaaS parameters. Based on the above achieved trust values, a scheduling algorithm is also proposed that may further enhance the QOS of services been provided to the users.

Keywords Cloud · QoS · Trust management · Cloud IaaS · Scheduling · VMM · Distributed applications · Distributed programming

P. Gupta (✉) · M. K. Goyal · P. Kumar · A. Aggarwal
Department of Computer Science Engineering,
Jaypee Institute of Information Technology, Noida, India
e-mail: Punitg07@gmail.com

M. K. Goyal
e-mail: mayankrkgit@gmail.com

P. Kumar
e-mail: prakash.kumar@jiit.ac.in

A. Aggarwal
e-mail: alok.aggarwal@jiit.ac.in

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_72,

72.1 Introduction

Distributed environment has being in use since long time, but with the evolution of cloud computing, it can be used by everyone. With distributed environment, there comes the problem of security and reliability. Different trust management models are being proposed to solve these problems, to categorize the datacenters on the basis of their trust value, being calculated on the basis of the few parameters taken into account. The task of these models is to categorize the datacenters, not only on the basis of the one parameter, but on the basis of multiple parameters collaboratively.

There are many models being proposed for cloud computing, especially in software as a service (SaaS) which are referred in this paper to identify the difference between the trust management models. The main difference between the cloud SaaS and IaaS is that in SaaS there are many domains and data types which has different properties. But on the other hand, in IaaS there is only one data type that is the virtual machine (VM) image. Hence we consider the processing capabilities of different virtual machine monitors (VMMs) to calculate the trust value of the datacenter or the node.

In current scenario of cloud IaaS, we use scheduling algorithm to schedule the request and allocate a VM to the user. But there are different types of user who have fewer resources and has paid less, user with large resources and has paid more, and a free user or a public user. So taking into consideration basic scheduling algorithm where a user request is allocated a VM based on the load on a datacenter and the cost of datacenter. But not taking into consideration the properties of the datacenter. Due to this a datacenter with High Quality of service (QoS) is been allocated to a public user and the request from the other user who has paid more will be allocated a datacenter with low QoS. This problem is been tried to solved using cost based scheduling algorithms [1] but it has not taken into consideration the properties of a VMM in a datacenter.

So we propose a trust management model to overcome this problem, by taking into consideration VMM characteristics which vary from datacenter to datacenter. Then these trust value are been used by the scheduling algorithm proposed to improve the scheduling of the resources.

72.2 Proposed Work

As from above scheduling algorithms they only take into consideration either the scheduling algorithm based on power or simple resource allocation algorithm, to overcome that and map them to real paradigm the cost based algorithm was proposed and rank based algorithm were proposed [2, 3]. At the same time algorithm based on deadline and budget were proposed. But the problem with these algorithm is that the take into only cost or only QoS.

The algorithm proposed here is to increase the QoS being provided to the use. The problem as explained in Sect. 72.2 can be solved by using proposed model. Trust model are used to calculate trust values for the datacenters based on the parameters which are as follows:

a) ***Initialization time****: Time taken to allocate the resources requested and deploys them.*
b) ***Machine instruction per second (MIPS)****: Number of instruction computed per second.*
c) ***Fault rate****: This number of faults in a period of time.*

Trust values are being calculated for each datacenters. Then these trust values are being updated after a fixed period of time which can vary because if trust is call updated frequently then there will be no considerable changes in the trust value.

Whole cycle include following steps:

72.2.1 Initialization

In this the trust value of the datacenter and the client are being initialized.

1) First the datacenter trust is initialized based on the trust management model which is used to initialize the trust value and updating the trust value periodically. In this if the datacenter is newly introduced, it is initialized with the default trust value.
2) From user a request along with the user-budget is been taken. Then the cost estimation for the request is done for all the corresponding datacenters. On the basis of the user-budget and the user-request cost for each datacenter classification of datacenters is done. This set of datacenters is then ordered based on their trust value i.e. their trustworthiness and reliability. Here client trust is simply the budget of the client.

72.2.2 Trust Evolution

This step includes updating the datacenter trust value periodically after a fixed interval of time. (Fig. 72.1).

Let there be datacenters Di and its attribute is given by Ai1 (initiation time), Ai2 (MIPS), Ai3 (Fault rate) and there corresponding trust value be Ti.

$$D_i,\ \ i = 1 \cdots k, \quad T_i,\ i = 1 \cdots k,$$

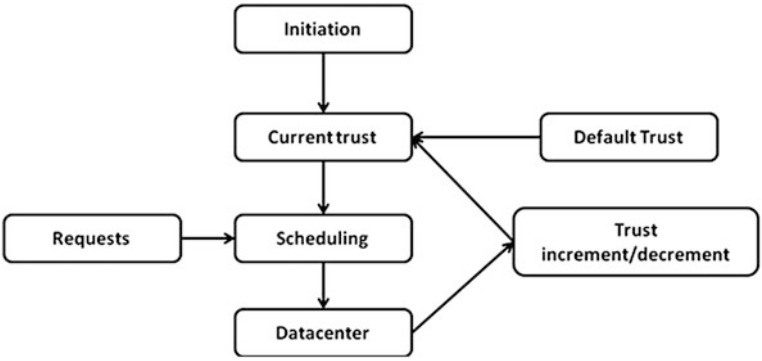


Fig. 72.1 Trust resource scheduling

Table 72.1 Result with trust model and scheduling algorithm

Data center ID	VM ID	Time	Start time	Finish time	d-trust	c-trust
Datacenter_0	0	400	0	400	9	8
Datacenter_1	2	400	0	400	10	12
Datacenter_2	1	400	0	400	12	11
Datacenter_3	3	400	0	400	8	10

The value to be updated in trust is calculated on the basis of following equation given by:

$\text{Diff} = (A_{i1}\_\text{OLD}–A_{i1}\_\text{NEW}) + (A_{i2}\_\text{OLD}–A_{i2}\_\text{NEW}) + (A_{i3}\_\text{OLD}–A_{i3}\_\text{NEW})$.

$T_i = T_i \pm \text{Diff}$.

72.3 Experimental Results

In this for simulation CloudSim API is used. Firstly, the CloudSim API does not support trust based concept, so the datacenter trust value was introduced as the attribute of the datacenter. Based on this attribute we have computed the result to show the real problem. For this we have considered two classes of budget range, with 3 datacenters and 5 user requests.

Table 72.1 is the result of the computation with trust model and scheduling algorithm being implemented in it. As the result after the request being submitted to the scheduling algorithm and based on trust value of client and datacenter they are being scheduled accordingly. A request with higher trust value is being allotted to the datacenter with higher trust value. These computations are tested for seventy and hundred user requests.

72.4 Conclusion

In this paper different type or trust models and scheduling algorithm are been discussed with their drawbacks. To overcome the drawbacks a trust based cost efficient algorithm is proposed which perform better them scheduling algorithm implemented in CloudSim. For future work this trust model may be compared with other models and see the improvement in the QoS.

References

1. Yang Z, Yin C, Yang L (2011) A cost-based resource scheduling paradigm in cloud computing. In: PDCAT 2011–2012, pp 417–422, January 2011
2. Yang Z, Yin C, Yan L (2011) A cost-based resource scheduling paradigm in cloudcomputing. In: PDCAT 2011, pp 417–422, Oct 2011
3. Selvarani S, Sadhasivam GS (2010) Improved cost-based algorithm for task scheduling incloud computing. In: Computational intelligence and computing research (ICCIC), pp 1–5, Dec 2010
4. Buyya R, Ranjan R, Calheiros RN (2009) Modeling and simulation of scalable Cloud computing environments and the CloudSim toolkit: Challenges and opportunities. In: High performance computing and simulation, pp 1–11, June 2009
5. Castelfranchi C (2004) Trust mediation in knowledge management and sharing. In: Proceedings of second international conference on trust management, pp 304–318

Chapter 73
Hybrid Covert Channel an Obliterate for Information Hiding

Rajeshwari Goudar and Pournima More

Abstract In a prisoners' problem there are two individuals attempt to communicate covertly without alerting a "warden" who controls the communications channel. This problem becomes more or less difficult because of various assumptions or requirements. One assumption which makes the problem considerably more manageable is that the participants are allowed to share some secret information such as an encryption key prior to imprisonment. Another assumption, which makes the problem much more difficult, is that the warden be allowed to modify as well as read the messages sent between the prisoners. This paper describes Hybrid Covert Channel techniques, in which no secret information needs to be shared before imprisonment. In this case if the warden is not allowed to modify the contents of the channel, a modification of an existing protocol will be shown to admit pure steganography. Then, a technique is described that allows pure steganography between two prisoners in the presence of an active content-modifying warden. This technique is possible through the use of two distinct channels rather than one: the subliminal channel for steganographic communication which is augmented by a supraliminal channel, one in which information is not hidden from the warden but cannot be modified.

Keywords Subliminal channel · Covert channel · Noisy channel · Hybrid covert channel

R. Goudar (✉) · P. More
Computer Department, MAE Alandi, University of Pune, Maharashtra, 411015, India
e-mail: rmgoudar66@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_73,

73.1 Introduction

In prisoner's problem, two people usually named Sender and Receiver, are in prison and intend to an escape plan. The problem is that all communication between them is through Warden. Sender and Receiver must send hidden information which is inconspicuous and cannot be noticed by Warden. Inconspicuous data is used to hide the real message which is usually referred as cover data.

73.2 Literature Review

Hybrid covert channel is composition of two or more variants of Covert Channel, either active at same instance or at different instances of time. It is impossible to completely assess the number of covert channels involved in a hybrid composition and it may be more complex if hybrid covert channel behaves as multi-trapdoor and protocol hopped, Here, Hybrid Covert Channel is visualizing as a combination of noisy channel in TCP and subliminal channel in SSL, both being transport layer protocols.

73.3 Proposed System: Hybrid Covert Channel

Covert channels can be considered as one of the main sub disciplines of data hiding. In data hiding, the two communicating parties are allowed to communicate with each other based on the security policy of the system while exploiting the features as associated with covert channel. A covert channel is primarily used for information transmission, but that is not designed nor intended for communications [1]. An overt channel can be utilized to act as a covert channel by having embedding and detection processes incorporated at the source and the receiver, respectively. Covert channel are piggy-backed on legitimate overt channel.

Hybrid Covert Channel is a covert channel where there is co-existence of two or more different variants of covert channels as shown in Fig. 73.1 There are various possibilities of hybrid channeling, for example noisy covert channel in transport layer with subliminal channel and supraliminal channel in network layer or application layer. Because of these variations Hybrid Covert Channel is very difficult to detect as it behaves like a single channel [2, 3].

In this paper we propose a new Hybrid Covert Channel an obliterate which is a combination of noisy covert channel in the transport layer and subliminal channel and supraliminal channel in the Secure Socket Layer/Transport Layer Security (SSL/TLS).SSL is a basic security service that any network can provide and transport layer is essential layer for communication over packet switched network. Such composition of hybrid channeling can create an uncertainty in the legitimate network environment. The components of Hybrid Covert Channel are described as follows:

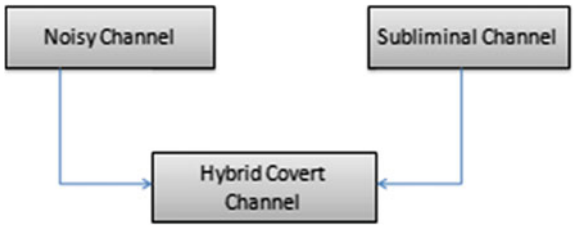


Fig. 73.1 Hybrid covert channel (hybrid NC/SC)

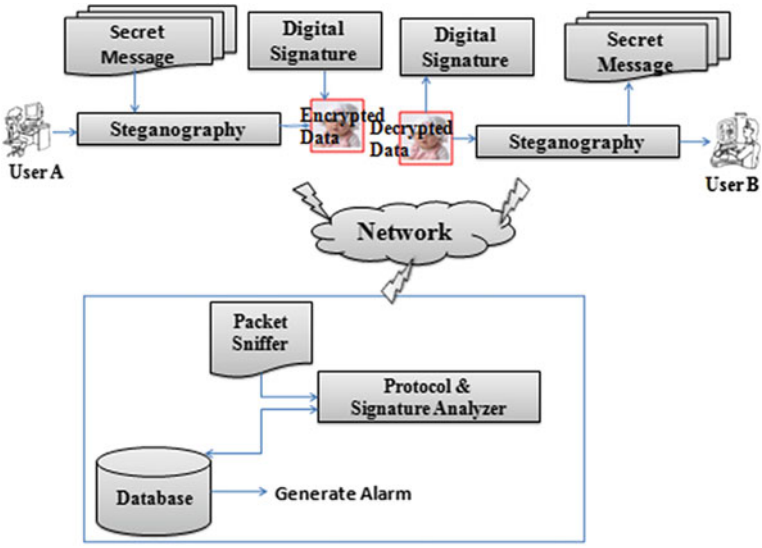


Fig. 73.2 Architecture model

73.3.1 Noisy and Noiseless Channels

Covert Channel can be Noisy or Noiseless. A channel is noiseless if symbols transmitted by the sender are the same as those received by the receiver. Channel is noisy if the symbols transmitted by the sender are different from those received by the receiver. Linguistic information can be involved in the decoding processes applied by noisy channel.

73.3.2 Supraliminal Channel

Aim of Supraliminal channel is to hide a message to the recipient by embedding it in image. In this case, image can be meaningless, but message's content is

important. But, it is possible for a warden to identify the channel without considerably altering the image if information is embedded in an existing image.

73.3.3 Subliminal Channel

Subliminal channels are covert channels that can be used to communicate secretly over an insecure channel with the help of digital signature. In this situation, subliminal transmitter and subliminal receiver must be mutually trustworthy parties. This channel uses Digital Signature Algorithm for generation of secrete key. In case of using Digital Signature Algorithm, there exist parameters which have to be set randomly. The signature remains verifiable and indistinguishable from a normal signature as the algorithm's signature creation procedure is unchanged [4].

73.3.4 Hybrid Channel: A Combination of Supraliminal Channel and Subliminal Channel

In case of Hybrid Covert Channel we are using both supraliminal and subliminal channels in combination. Thus it provides both steganography as well as digital signature policies and it becomes more secured way of communication. In order to communicate Sender applies stenography on data which he wants to send. For this, he implements LSB modification for lossless image technique on data. Digital Signature Algorithm on Stego image is applied and this processed data is then sent to the receiver. During transmission of data, it is captured by Warden who analyzes this data to state whether data is secure or not. For this he checks the content of TCP packet header fields & IP packet header fields. If the data is secure, then it is forwarded to receiver. Otherwise it will be discarded. At receiver end, receiver implements Digital Signature Algorithm and Stego Algorithm to get original data.

73.4 System Architecture

Architecture module consists of three Sub modules network packet sniffer, Protocol & Signature analyzer and Database to detect channel is as shown in Fig. 73.2 Packet Sniffer Module is used to monitor and capture TCP and IP packet's header that are being sent during communication on the network, saving them for later analysis. After capturing TCP and IP packet's header fields are stored in MySQL database. Protocol and Signature Analyzer module checks whether the TCP and IP Packet headers of ongoing communication those are

stored in the MySQL database, are according to the rules stored in the MySQL database. If so, then communicating parties are termed as secured else an alarm is generated indicating an insecure communication. Database module is used to store rules that are being used by Protocol and Signature Analyzer module as well as stores TCP and IP Packet header fields that are being sent during communication on the network.

73.5 Conclusion

In Hybrid Covert Channel we are using both supraliminal and subliminal channels in combination, it provides both steganography as well as digital signature policies and becomes more secured way of communication. Thus Hybrid Covert Channels find interesting applications in network security and in facilitating various network processes.

References

1. Steven JM, Lewis S (2005) Embedding covert channels into TCP/IP, Draft for information hiding workshop 2005
2. Rowland CH Covert channels in the TCP/IP protocol suite. First Monday Peer-Rev J Internet
3. Meredith LP, Sassaman L (2007) Subliminal channels in the private information retrieval protocols, 28th symposium on information theory in the Benelux 2007
4. Simmons GJ (1984) The prisoner's problem and the subliminal channel, advances in cryptology. In: Proceedings of crypto

Chapter 74
Steganography and Its Technique: Technical Overview

Gulshan Shrivastava, Aakanksha Pandey and Kavita Sharma

Abstract This paper is basically overview on the steganography. In this paper, we mainly present current, past and future work done in the field of steganography. Here we give a brief description on different techniques that are being used in steganography. These techniques do also have different types which are briefly explained here. This paper also contains the detection process used in different techniques of steganography. Furthermore, we also described the application of steganography in different field.

Keywords Steganography · Steganographic systems · Steganalysis · Steganography technique

74.1 Introduction

It is very important to send a message with security from source and destination. In each and every field, information plays very critical role. From a human being to an organisation, everyone wants to make his/her data or information confidential. So, for that one has to provide security to his/her information.

G. Shrivastava (✉) · A. Pandey · K. Sharma
Department of Information Technology, Dronacharya College of Engineering, Gr. Noida, U.P., India
e-mail: gulshanstv@gmail.com

A. Pandey
e-mail: aakankshapandey.1991@gmail.com

K. Sharma
e-mail: kavitasharma_06@yahoo.co.in

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_74,

Steganography helps to provide to essential security to the information. Steganography is the art of covered or hidden writing [1]. So, its main purpose is to hide the confidential information from unauthorised user or from third party. Basically, in steganography, we hide the information through an image, audio or with any multimedia file. In such a manner, that if any person views that image or that multimedia file cannot be able to predict whether there is any hidden information behind the data or not. Therefore, he cannot be able to decrypt that data. These multimedia files through which we hide the information can be an image, any audio clip, text or any video file. Steganography is an important part for the purpose of providing security to data. It makes the information hidden in such a way that apart from sender and receiver no other person can even realize that if there is any hidden information or message. In this paper, we have tried to give review on steganography. This paper concludes the past, present and future, different techniques, detection and evaluation on steganography [2].

74.2 Steganographic

Steganography comes from Greek word Steganos that is covered and graptos that means writing. The origin of steganography is biological and physiological. The word steganography comes into use in 1,500 s after the appearance of 'Trithemius' book on the subject 'Steganography'. A short overview in this field can be divided into three parts:-

74.2.1 Past

The first recorded uses of steganography can be traced back to 440 BC when Herodotus mentions two examples of steganography in his *Histories*. Demaratus sent a warning about a forthcoming attack to Greece by writing it directly on the wooden backing of a wax tablet before applying its beeswax surface. Wax tablets were in common use then as reusable writing surfaces, sometimes used for shorthand.

74.2.2 Present

Now days, many steganographic systems are using multimedia subjects as the cover media because most of the people are using multimedia approach over the network communication like computers, laptops, mobile phones, notebooks etc. [3].In today's approach, steganography can be classified into five different fields. So in modern generation in different cases we use different tools and techniques for the privacy and confidentiality of the information with the help of steganography as well as with cryptography.

74.2.3 Future

In this modern world, we all are familiar with a word 'Hacking'. Hacking is a process under which some unauthorised user tries to access our private and confidential information. If we talk about steganography, so in terms of steganography, this problem of hacking is known as Steganalysis [4]. Therefore, in steganalysis, a steganalyzer uses different tools and techniques to crack the cover object or page so that he/she can be able to get and use the information hidden behind the cover page or object. So, in future, keeping security in mind, different tools and techniques will get developed [5].One solution to solve the above problem can be Dual Steganography. Dual steganography is the process of performing steganography along with cryptography.

74.3 Steganography Techniques

For performing steganography, there are techniques which can hide the confidential or secret information inside a multimedia files. There are five techniques which we will discuss here.

74.3.1 Text

Text steganography is the most difficult kind of steganography due to the lack of redundancy in text as compared to image or audio or other. However, it requires less memory and provides for simpler communication. One method that could be used for text steganography is data compression. Data compression encodes information in one representation, into another representation. The new representation of data is smaller in size. One of the possible schemes to achieve data compression is Huffman coding.

74.3.2 Image

Image steganography is mainly used to hide the information through an image file. We can insert or encode each bit of information in the image directly. It can also be done by inserting each bit of information at the noisy area of the image where a normal human being pays less attention- those areas where there is a great colour variation [4]. We can also disjoint the information over the image. Different image steganography approaches are as follows:- **i.** Least Significant Bit Insertion **ii.** Masking and Filtering **iii.** Redundant Pattern Encoding **iv.** Encrypt & Scatter **v.** Algorithm and Transformation.

74.3.3 Audio

Audio steganography is quite difficult to other digital steganography like digital image. In this steganography important information is hiding inside a digital sound. Sounds file in which steganography can be performing are MP3, AU & WAV. This can be performed by alternating or changing the sequence binary sequence of an audio file. For hiding the information in an audio file no. of methods is used Form using simple algorithm which manipulate the information is signals to the advance signal processing techniques. Techniques used in audio steganography are:-. **i.** LSB Coding **ii.** Parity Coding **iii.** Phase Coding **iv.** Spread Spectrum **v.** Echo Hiding.

74.3.4 Video

Basically, video files are the combination of both image and sound. So, the technique used in image and audio steganography can also be applies to video steganography. There is a large amount of data that can be hidden inside the video. This is the advantage of video steganography. Video files are in a moving flow so a small deformation might not be observed by a normal human being. All the attention of the person is on the images and audio that is flowing on the screen.

74.3.5 Protocol

While using a network transmission, every person either in an organisation or at home wants to protect his/her information for the sake of confidentiality of their information. For that purpose, protocol steganography is used. The information is encoded and network control protocol like TCP/IP protocol is used for the transmission purpose over the network.

74.4 Detection

As for performing steganography, different techniques are available for hiding the information. So as for detecting these steganographic systems also different steps are followed for different techniques.

74.4.1 Text

Text steganography encrypt the information in such a way that other then receiver and sender, no one can find that there is some hidden information. Different methods can be implementing like substitution or transposition. Two keys are used

to encrypt the information i.e. public key and private key. If someone from any source gets these keys, then the information can no longer be confidential or private.

74.4.2 Image

It is quite difficult to find out the stego-image with naked eyes but it usually leave some fingerprints or statistical hint which shows that there is some modification in the image. There are also some tools which helps to find out whether the image is stego-image or not. These tools are also known as statistical tools. So for, statistical analysis can be performed to check whether the image is stego or not [6].

74.4.3 Audio and Video

We can use statistical analysis for audio files as well as LSB modifications techniques on sounds files. We have some other tools also which helps us to convert inaudible sounds for a human ear to an audible sound. And if that sound be audible to human then he or she can easily retrieve it and use it in some other or unauthorized work. It can also be implemented on video files if the information came directly on the screen or in audio, then it can easily be visible or audible. And if this happens, then its advantage becomes the disadvantage. The large amount of precious information gets revealed in front of other persons and they can easily be able to use it.

74.5 Conclusion

Steganography is an art through which we cover all the hidden writing. Steganography can help to hide the information which is important for an organisation as well as for a normal user. As the confidentiality and privacy is always necessary for an organisation. Steganography is not a new technique, it comes in light since 440 B.C. when Histieaus shaved his most trusted slave and tattooed it with a message which disappeared when the hair re-grown. Through this he would be able to hide the information. Now days, steganography is a developing field. It uses multimedia objects like image, audio, video to hide or cover the private information. In future, for cracking the cover image and retrieving the information, steganography can be used. The process in which the steganalyzer cracks the cover image is steganalysis. This will be done with the help of different tools and techniques.It can also tend to some limitations. These limitations are done by government and they claimed that criminals can use these techniques for the

communication purpose. But if this will be done by keeping in mind that steganography helps everyone to keep the information secure, protected and private, steganography in terms plays an important role.

References

1. Johnson, N. F., Jajodia, S. (1998) Exploring steganography: Seeing the unseen. Computer 31(2):26–34
2. Mohanty, S. P.: Digital Watermarking: A Tutorial Review (1999), http://informatika.stei.itb.ac.id/∼rinaldi.munir/ Kriptografi/WMSurvey1999Mohanty.pdf
3. A.C. Brainos II: A Study Of Steganography And The Art Of Hiding Information, http://www.infosecwriters.com/ text_resources/pdf/steganographyDTEC6823.pdf
4. Shrivastava, G., Sharma, K., Dwivedi A. : Forensic Computing Models: Technical Overview, CCSEA, SEA, CLOUD, DKMP, CS & IT 05, pp. 207–216, 2012
5. J.R. Krenn: Steganography and Steganalysis, January 2004
6. DebashisGanguly. Swarnendu Mukherjee, MohitMundhra: Digital Watermarking: A New Approach. In Proc. 41st Annual conference, CSI'06, paper no. : 11 (2006)
7. Maitra, I.K., Nag, S.,Datta, B.,Bandyopadhyay,S.K.:DigitalSteganalysis: Review on Recent Approaches,Journal of Global Research in Computer Science, Vol. 2, No. 1,pp.1-5, January 2011

Chapter 75
Data Stream Mining: A Review

S. Pramod and O. P. Vyas

Abstract: In the data stream model the data arrives at high speed so that the algorithms used for mining the data streams must process them in a very strict constraints of space and time. This raises new issues that need to be considered when developing association rule mining algorithms for data streams. So it is important to study the existing stream mining algorithms to open up the challenges and the research scope for the new researchers. In this paper we are briefly discussing the different issues and challenges in the data stream mining.

Keywords Data stream mining · Association rule mining · Data mining · Online stream mining

75.1 Introduction

Frequent Itemset mining [1] is an important step for the majority of important data mining tasks. The majority of existing works concentrated on frequent itemsets from the static databases. The importance of data stream mining and its applications [2] lead us to study the frequent itemset generation from stream data. There are many challenges in frequent itemset mining on stream data. It is impossible to keep the entire stream in the main memory or even in the secondary storage since

S. Pramod (✉)
Associate Professor, Information Technology, Christian College of Engineering, Bhilai, C.G, India
e-mail: pramodsnair@yahoo.com

O. P. Vyas
IIIT-Allahabad, Allahabad, U.P, India
e-mail: 22dropvyas@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_75,

the data stream comes continuously and the amount of data is unpredictable. Another challenge is to do the mining in single scan since the multi scan is impossible on stream data. Adjust with the fast arrival of data is another challenge. Due to these constraints many research studies conducted in the direction of approximate mining with the reasonable guarantees on the result. There are many algorithms available for mining Frequent itemsets, frequent maximal itemsets [3], or frequent closed itemsets [4] over data streams.

The tentative nature of frequent itemset mining normally results in a large number of frequent itemset generations. The increase in the number of frequent itemset generated will degrade the mining efficiency. The frequent closed itemset mining(FCI) is the solution for the above said problem. The FCI is a non redundant representation of the sets of frequent itemsets [5]. The commendable reduction in the size of the result set leads to improved performance in the speed and memory usage. We noticed that the FCIs approach could not be applied over land mark window since the number of FCIs approaches that of frequent itemsets when the window becomes very large.

There is a frequent itemset mining technique called Frequent Maximal Itemset [6]. Compare with the Frequent Closed Itemset mining technique it will generate comparatively less number of itemsets, due to this reason it is significantly more efficient in terms of both CPU and memory. But the disadvantage of FMI mining is that it loses the frequency information of the subset of FMIs so the error bound will also increased. There are many concise representations of frequent itemsets are proposed [7–9], these are significantly saving memory space, CPU and shows better accuracy. This technique could be applied in stream mining with the efficient incremental technique and batch processing.

75.2 Issues in Data Stream Mining

One of the main issues is to find the data *extraction model* to extract the data for the frequent itemset generation for association rules. The derived model should support and handle the data streams that come continuously and unboundedly. There are three stream data processing models available that are Landmark window model, Damped window model and Sliding Window model.

Load Handling in Data Streams: The data arrival rates of online data are very high and often busty. When the load of mining exceeds the capacity of the system, load shedding is required to keep up the arrival rates of the streams. When the incoming rate of a stream is higher than the number of transactions the system can handle at a particular time, then we have to develop a technique to handle this situation. The different techniques published in this direction are [10, 11] sampling and window reduction.

Multiple Source Handling: The stream data comes from different sources at fast rate. The computational overhead of handling such a situation is very high. To reduce the computational over head, the different researches put forwarded some techniques like parallel and incremental techniques [11–14] and deal with approximate and precise solutions.

Itemset Storage in Memory: The offline association rule mining algorithms are using the static data. The nature of online data is kept on changing due to the flooding of data. The influence of the huge amount of data inflow in a lesser time span will change the association rules as well as the technique used for the storage. The characteristic influence of the stream data is in two important areas that are storage and the technique to give an acceptable association result with least information as possible. The different algorithms developed by different authors are not sufficient to cater the purpose. One of the algorithms [15] is working by storing the most frequent items and the count in the main memory. But the problem with this technique is that the infrequent item may become frequent in future. Another algorithm [16] is useful to generate maximum of 3 data items in the frequent itemsets and it cannot have more than 1,800 data items. This is actually the limitation of this algorithm.

Data Structure: Due to the restrictions in the memory space as well as the continuous inflow of data streams we required an efficient storage mechanism to deal with it. The failure of such a mechanism will drastically reduce the overall performance of the entire mining system because even if we store it in disk the i/o operation will increase the processing time. The designed data structure needs to be incrementally maintained since it is no way possible to rescan the already passed through streams and due to the requirement of the continuous in flow of data streams. Lattice or prefix [17, 18] data structure is useful in approximate savings of the frequent itemsets. The FP-Tree [19] is another useful approach for the data structure for stream mining. Jin and Agarwal [20] developed a memory resident summary data structure that implements a prefix tree using the hash table.

75.3 Types of Association Rule Mining Algorithms

Precise/Exact Mining Algorithm: This type of mining keeps track of all the itemsets and their actual frequency for monitoring of infrequent itemsets that may become frequent in future. The number of items to be stored in the memory is very large and it is computationally high. This mining technique is applicable for small sliding windows and streams where relatively low data arrival rate. This type of algorithms generates all the itemsets that having the support greater than or equal to the support value. There are some algorithms that fall in this category [15, 16, 20–22]. The algorithm used in [15] takes more than one scan to generate the precise results and the algorithm [16] generate the exact results but it is not useful to generate large itemsets. Closed Frequent Itemset mining is a method for the precise mining. In this approach only special frequent itemset that are closed frequent itemsets will be maintained in the memory [21, 22].

Approximation Algorithm: This type of algorithms generates approximate results with or without an error guarantee. This is a widely accepted technique for frequent itemset mining in most of the stream processing applications [17]. In most of the time we have to provide a result in such a way that throwing the light to the unknown out of the ordinary patterns, in such situations, approximate methods

with reasonable results is useful. The algorithms fall in this type of approaches can again divided into different based on the possibility of having the non frequent itemset in the frequent sets and vice versa, that are false positive [21] oriented or false negative [23] oriented. In the case of false positive includes some infrequent itemsets in the result sets and in the case of false negative misses some frequent itemsets in the result sets. In the existing false positive approaches the set of generated sub frequent itemsets must be very large to obtain highly accurate answer. The problem of extra memory usage due to the increase of sub FIs is solved by Yu [20]. In his approach at first the lowered minimum support is used to find the set of potential FIs and then slowly increases the lowered minimum support to control the frequent itemsets in the memory.

There are algorithms to handle the concept drifting. Concept Drifting is a process of losing the importance of the old data due to the massive inflow of the new data [24] in the online stream mining. One algorithm proposed by chi [21] to handle the concept drifting by boundary movements in the closed enumeration tree (CET). Few of the algorithms in this category [17–19, 23, 25, 26] are followed the approximation method. There is another algorithm [20] available for stream mining from the potential frequent 2-itemsets and it used the apriori property to reduce the number of candidate itemsets. This algorithm is approximate and false positive.

Temporal Online Mining [19, 27]*:* In many applications user may be interested to find the association rules of a certain span of time. In such cases the data structure to be adjusted dynamically to bring out the necessary itemset during the requested time span. Storing the stream data with time and retrieving it as and when it required is a challenging task.

Stream Data in Multidimensional Nature: If we need to deal with the multidimensional nature data like the sensor data in stream mining where multi dimensional nature like temperature, humidity etc. involved, the challenges are how to efficiently store, update, and retrieve the multidimensional information to retrieve the association rules. One paper [28] proposed a method to integrate multidimensional analysis and sequential data mining, and another [14] proposed an algorithm to find sequential patterns from d-dimensional sequence data, where $d > 2$.

Online Interactive Processing: In many cases of online mining there will not be any stoppage point for the mining, in such cases the modification of mining parameter should be made on the fly. So making the online mining interactive will be another challenge. In one paper [29] presented some techniques to do the frequent itemset updation as well as to do the changes in the parameters without re executing the entire process. In another paper [30] proposes a selective updates rather than the entire model. Ghoting and Parthasarathy [31] proposed a scheme in which gives controlled interactive response times when processing distributed data streams.

75.4 Windowing Approach to Data Stream Mining

Land Mark Window: The Landmark Model mines all frequent itemsets over the entire stream data from a specific time point called Land mark to the present.

There are many papers published [15, 16, 18, 20, 23, 32, 33] in this direction. The disadvantage of this model is that the applications that interested only in the most recent data streams could not use this model.

Sliding Window Model: The sliding window model processes only [21, 24, 25, 27] the items in the window and maintains only the frequent itemsets. The size of the sliding window can be decided according to the applications and the system resources. The recently generated transactions in the window influence the mining result of the sliding window, otherwise all the items in the window to be maintained. The size of the sliding window may vary depends up on the applications it use.

Damped Window: This model is also known as Time fading model. This model finds the frequent itemsets by assigning a weight to each transactions and the weight will decreases as time goes by. This type of algorithms [4, 19, 32] are useful for the applications in which new transactions have the effect in mining result. The performance of the algorithm can be examined by looking at the accuracy, processing speed, memory usage and the handling of the concept drift. In one paper [22] an algorithm proposed to monitor the requirement of the updation with the sampling technique. In this approach the sampling technique will compare the old association rules with the new association rules and when the difference increases a particular threshold the algorithm request for the updation. In another algorithm [34] the author calculates the metric distance to decide when to do the updation. The disadvantage is that it is useful only with the streams with little concepts drifting.

75.5 Conclusion

In this paper we have discussed some of the issues to be considered to develop an effective, performance oriented algorithm for data stream mining. And we reviewed how the existing important algorithms could handle these different issues. We have discussed some of the application oriented issues for the better understanding of the existing algorithms and the way the different algorithms handle the data stream mining so that the researchers can analyze and study further for the research work.

References

1. Agrawal R, Imielinski T, Swami AN (1993) Mining association rules between sets of items in large databases. In: Proceedings of the SIGMOD
2. Garofalakis M, Gehrke J, Rastogi R (2002) Querying and mining data streams: you only get one look. In: Tutorial of SIGMOD
3. Li H-F, Lee S-Y, Shan M-K (2005) Online mining (recently) maximal frequent itemsets over data streams. In: Proceedings of the 15th international workshop on research issues in data engineering. Stream data mining and applications (RIDE-SDMA'05)

4. Pasquier N, Bastide Y, Taouil R, Lakhal L (1999) Discovering frequent closed itemsets for association rules. In: Proceedings of ICDT
5. Chi Y, Wang H, Yu PS, Muntz RR (2006) Catch the moment: maintaining closed frequent itemsets over a data stream sliding window. KAIS 10(3):265–294
6. Lee D, Lee W (2005) Finding maximal frequent itemsets over online data streams adaptively. In: Proceedings of ICDM
7. Calders T, Goethals B (2002) Mining all nonderivable frequent itemsets. In: Proceedings of PKDD
8. Xin D, Han J, Yan X, Cheng H (2005) Mining compressed frequent-pattern sets. In: Proceedings of VLDB
9. Bonchi F, Lucchese C (2005) On condensed representations of constrained frequent patterns. KAIS 9(2):180–201
10. Zaki M, Parthasarathy S, Li W, Ogihara M (1997) Evaluation of sampling for data mining of association rules. In: Proceedings of RIDE
11. Srivastava U, Widom J (2004) Memory-limited execution of windowed stream joins. In: Proceedings of VLDB
12. Otey ME, Wang C, Parthasarathy S, Veloso A, Meira W Jr (2003) Mining frequent itemsets in distributed and dynamic databases. In: IEEE international conference on data mining
13. Otey ME, Parthasarathy S, Wang, Veloso A, Meira W Jr (2004) Parallel and distributed methods for incremental frequent itemset mining. IEEE Trans Sys Man Cybernet
14. Assaf S, Wolff R, Trock D (2003) Distributed algorithm for mining association rules. IEEE international conference on data mining
15. Richard M.K, Shenker S (2003) A simple algorithm for finding frequent elements in streams and bags. ACM Trans Database Sys
16. Li Y, Sanver M (2004) Mining short association rules with one database scan. In: International conference on information and knowledge engineering
17. Manku GS and Motwani R (2002) Approximate frequency counts over data streams. In: Proceedings of the VLDB
18. Li H, Lee S, Shan M (2004) An efficient algorithm for mining frequent itemsets over the entire history of data streams. In: Proceedings of the first international workshop on knowledge discovery in data streams
19. Chris G, Han J, Pei J, Yan X, Yu PS (2003) Mining frequent patterns in data streams at multiple time granularities. Data mining: next Generation challenges and future directions, AAAI/MIT
20. Jin R, Agrawal G (2005) An algorithm for in-core frequent itemset mining on streaming data. In: Proceedings of ICDM
21. Chi Y, Wang H, Yu PS, Muntz RR (2004) Moment: maintaining closed frequent itemsets over a stream sliding window. In: Proceedings of ICDM
22. David WC, Han J, Ng VT, Wong CY (1996) Maintenance of discovered association rules in large databases: an incremental updating technique. In: IEEE international conference on data mining
23. Yu J, Chong Z, Lu H, Zhou A (2004) False positive or false negative: mining frequent itemsets from high speed transactional data streams. In: Proceedings of VLDB
24. Haixun W, Fan W, Yu PS, Han J (2003) Mining concept-drifting data streams using ensemble classifiers. In: ACM SIGKDD International conference on knowledge discovery and data mining
25. Chang JH, Lee WS (2004) A sliding window method for finding recently frequent itemsets over online data streams. J Inform Sci Eng 20(4)
26. Li L (2009) Mining frequent itemsets over data streams using efficient window sliding techniques. Sci Direct Expert Sys Appl 36:1466–1477
27. Chih-Hsiang L, Chiu D-Y, Wu Y-H, Chen ALP (2005) Mining frequent itemsets from data streams with a time-sensitive sliding window. In: SIAM International conference on data mining

28. Pinto H, Han J, Pei J, Wang K, Chen Q, Dayal U (2001) Multi- dimensional sequential pattern mining. In: International conference on information and knowledge management
29. Parthasarathy S, Zaki MJ, Ogihara M, Dwarkadas S (1999) Incremental and interactive sequence mining. In: International conference on information and knowledge management
30. Adriano V, Otey M, Meira P Jr (2003) Parallel and distributed frequent itemset mining on dynamic datasets. In: International conference on high performance computing
31. Ghoting A, Parthasarathy S (2004) Facilitating interactive distributed data stream processing and mining. In: IEEE international symposium on parallel and distributed processing systems
32. Chang JH, Lee WS (2003) Finding recent frequent itemsets adaptively over online data streams. In: Proceedings of the KDD
33. Graham C, Muthukrishnan S (2005) What's hot and What's not: tracking most frequent items dynamically. ACM Trans Database Sys
34. Qingguo Z, Xu K, Ma S (2003) When to update the sequential patterns of stream data. In: Pacific-Asia conference on knowledge discovery and data mining

Chapter 76
Comparison of Various Harmonic Mitigation Techniques in Induction Furnaces

Arvind Dhingra and Ashwani Kumar Sharma

Abstract Harmonics are a necessary evil associated with non linear loads such as Induction furnaces. As the amount of non linear loads in power system is increasing due to pro-filtration of non linear devices such as inverters, inductive loads etc.; it is leading to a deterioration of power quality. Non linear loads tend to inject harmonics in the power system. Induction furnaces are a major non linear load as far as our power system is concerned. This paper is an attempt to compare the various techniques of harmonic mitigation employed for mitigating harmonics in induction furnace.

Keywords Harmonics · Voltage and Current Harmonics · Harmonic Mitigation

76.1 Introduction

Harmonics are multiples of fundamental frequency which are generated due to presence of non linear loads. Harmonics could be current or voltage harmonics. There are prevalent standards which give permissible limits for harmonics in a system.

A. Dhingra (✉) · Ashwani. K. Sharma
NIT, Kurukshetra, India
e-mail: arvinddhingra@hotmail.com

Ashwani. K. Sharma
e-mail: assignmentee@gmail.com

A. Dhingra
AP, GNDEC, Ludhiana, India

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_76,

Current harmonics [1] are caused by non linear loads such as thyristor drives, induction furnaces etc. The effect of these loads is the distortion of the fundamental sinusoidal current waveform alternating at 50 Hz. Current harmonics affect the system by loading the distribution system as the waveforms of the other frequencies use up capacity without contributing any power to the load. Besides, harmonic currents load the power sources such as transformers and alternators.

Voltage harmonics are caused by current harmonics which distort the voltage waveform. Their impact depends on the distance of the load causing the harmonics from the power source. The current and voltage harmonic distortion causes several problems in electrical power systems, such as incorrect operation of devices, premature ageing of equipment, additional losses in transmission and distribution networks, overvoltages and overcurrents [3].

Induction furnaces are used in variety of metal processing applications and can range in size from a few hundred kW to 3000 kW [2]. These furnaces typically use a six pulse phase controlled rectifier feeding a frequency converter to deliver 300–1000 Hz alternating current to the furnace coil. This variable frequency A.C. allows for more efficient operation of furnace and better control of heating process. While offering these benefits the frequency converters also generate harmonic currents that propagate through the supply transformer on to the utility distribution system. These harmonic currents cause distortion problems at both the customer facility and distribution system. Electrical load can also be varying widely as high power is needed when the material is being heated. Adding power factor capacitors to the furnace aggravates the problem by creating parallel resonance between the capacitors and source inductance.

Various approaches to harmonic mitigation are prevalent and have been in use since long. We shall discuss each one of them along with their merits and demerits.

76.2 Line Reactors

Line Reactors [4] are the simplest and low cost means of attenuating harmonics, connected in series with individual non-linear loads. By inserting series inductive reactance into the circuit, they attenuate harmonics as well as absorb voltage transients that may otherwise cause a voltage source to trip on over-voltage. Line reactors offer the advantage of low cost and they can achieve a significant reduction in harmonics when the appropriate percent impedance is utilized. For reasonable harmonic attenuation, a 5 % impedance line reactor should be installed ahead of the motor drive or other 6-pulse non-linear load. Their disadvantages are that they cause a voltage drop, increase system losses and normal impedance values do not achieve current distortion levels much below 35 % THD-I. Additionally, the harmonic mitigation capabilities of the reactor reduce as load current is reduced because the reactor's effective percent impedance is reduced.

76.3 Isolation Transformers

The isolation transformers [4] can be used effectively to reduce harmonic distortion. The typical configuration of isolation transformer, for power quality purposes, is delta primary and wye secondary. The inductive reactance of isolation transformer is low enough at the fundamental frequency to easily pass fundamental current, but increases proportionately for harmonic frequencies and can achieve performance similar to that of a line reactor. Due to the capacitive coupling between each winding and the shield, a low impedance path is created to attenuate noise, transients and zero sequence currents. The shield helps to mitigate the common mode disturbances to their originating side (primary or secondary) of the transformer.

76.4 Passive Filters

Conventional passive filters consist of inductance, capacitance, and resistance elements configured and tuned to control harmonics of a particular frequency (notch filters) [5, 7]. Common types of passive filters are single tuned, 1st order high pass, 2^{nd} order high pass and 3^{rd} order high pass. Another popular type of passive filter is the high-pass filter (HPF) [6]. A HPF will allow a large percentage of all harmonics above its corner frequency to pass through. The passive filters have low cost but are bulky. They are used to suppress lower order harmonics. Filtering is not perfect over the variable frequency range so the harmonic problem is not completely solved and is variable [8, 9]. One of the important effects of passive filters is sharp parallel resonance point at frequency below the notch frequency. Some of the deficiencies of passive filters are overcome by use of Particle swarm optimization technique [20]. A particle swarm optimization method with nonlinear time-varying evolution (PSO-NTVE) is employed in the planning of large-scale passive harmonic filters for a multi-bus system under abundant harmonic current sources. The objective is to minimize the cost of the filter, the filters loss, the total harmonic distortion of currents and voltages at each bus simultaneously. This approach helps in better control of harmonics.

76.5 Active Filters

The basic principle of APF is to utilize power electronics technologies to produce specific currents components that cancel the harmonic currents components caused by the nonlinear load. Basic filter design equations neglect the finite bandwidth of amplifiers. Amplifiers consume power and inject noise into a system. Power handling capability is limited by the amplifier stages [11]. Voltage stresses and

power losses are a problem in active power filters (APF) [10]. The power losses reduce the harmonic filtering efficiency, which increases the filtering costs. Active filters can also be used in conjunction with predictive control strategy [12]. Predictive control has the advantage of selecting the best possible switching state for each future value minimizing the cost function.

76.6 UPQC

UPQC is Unified Power Quality Conditioner [13]. UPQC usually consists of two voltage-source converters sharing the same capacitive DC link [14–17]. One of the converters is an active rectifier (AR) or shunt active filter while other is a series active filter (SF). Also, at the point of the load connection, passive filter banks are connected. The UPQC has the capability of improving power quality at the point of installation of power distribution systems or industrial power systems [18, 24]. UPQC is able to compensate current's harmonics, to compensate reactive power, voltage distortions and control load flow but cannot compensate voltage interruption because of not having any sources.

76.7 APLC

Active power line conditioner (APLC) is an advanced concept in the field power quality control [21, 22]. It is based on integration of series and shunt power converters sharing a common DC link. APLC includes two voltage source inverters (VSIs) that are connected to a DC energy storage capacitor. One of these two VSIs is connected in series with AC line while the other one is connected in shunt with load bus. Active power line conditioner can be applied in power systems for unbalance and harmonic compensation (UPQC) [23, 24].

76.8 SHE–PWM

In this approach a set of transcendental equations are solved to determine the switching angle of SHE-PWM waveform [19]. Differential Evolution algorithm is used. The objective function of DE is designed to minimize the selected harmonics to near zero. Furthermore the fundamental component of the output voltage can be controlled independently. The difficulty in SHE-PWM approach is to calculate its switching angles. This is because the equations are non-linear and transcendental.

76.9 DVR and D-STATCOM

Dynamic voltage restorer (DVR) [26] is a custom power device used for voltage compensation of sensitive loads against voltage disturbances in power distribution lines. In order to ensure constant voltage without any harmonic distortions, the load voltage should follow a reference sinusoidal waveform of desired amplitude.

The main challenge that faces the series mitigation devices like DVRs is that it is difficult to detect and track the flicker disturbances that encompass characteristic and non characteristic harmonics, inter harmonics, and sub harmonics with unknown frequencies [29]. The distribution static compensator or D-STATCOM is an excellent cost-effective solution to reduce the impact of non-linear load operation in power systems [27]. The principal advantages of this FACTS device is that it has the capability to adapt by itself to changes that may occur in the electric power system, such as voltage flickers and harmonics, generated by loads like the electric arc furnace. The D-STATCOM also has the advantage over other FACTS devices that it minimizes the possible occurrence of resonances, reduces harmonics, and balances the voltage at the PCC. The problem with D STATCOM is that induction furnace operation during melting makes variation of instantaneous power fast.

76.10 Conclusion

Other strategies for mitigation of harmonics include use of Fuzzy logic, neural networks, use of specialized equipments like harmonic mitigation transformers, phase shifting transformers etc. The fuzzy logic or neural network techniques can be used for design of active filters for better control and mitigation of harmonics. All the strategies considered above have their own advantages and disadvantages. Depending upon the need and requirement of a particular installation, the best suited strategy may be chosen for mitigation of harmonics. The right choice is always dependent on a variety of factors, such as the activity sector, the applicable standards, the power level.

References

1. (2010) The difference between current and voltage harmonics, Electro-technik. A magazine on electrical engineering
2. McGranaghan M (2006) Controlling induction furnace harmonics, signature.Spring 6(2):2
3. Sousa J, Correia de Barros MT, Covasand M, Simões A (2009) Harmonics and , flicker analysis in arc furnace power systems. IEEE Trans Power Quality
4. Gonzalo S, John H (2005) A review of harmonic mitigating techniques. IEEE Trans

5. Dugan R, McGranaghan M, Santoso S, Wayne Beaty H (2003) Electrical power systems quality. 2nd edn
6. Ludbrook A (1988) Harmonic filters for notch reduction. IEEE Trans Indust Appl 24:947–954
7. Das JC (2004) Passive filters: potentialities and limitations. IEEE Trans Ind Appl 40(1):232–241
8. Mindykowski J, Tarasiuk T, Piotr R, Problems Of passive filters application in system with varying frequency. In: 9th international conference on electrical power quality and utilization, Barcelona
9. Chang Gary W, Hung-Lu W, Gen-Sheng C, Shou-Yung C (2009) Passive harmonic filter planning in a power system with considering probabilistic constraints. IEEE Trans Power Delivery 24(1):208-219
10. Parkatti P, Tuusa H (2008) Comparison of active power filters with series-connected capacitor. In: 2nd IEEE international conference on power and energy (PECon 08). Johor Baharu, Malaysia
11. Muhammad HR (2010) Microelectronic circuits: analysis and design, engage learning ISBN 0495667722 pp 804
12. Paredes AE, Simpson JI, Pontt J, Silva C (2010) Predictive current control of a multilevel active filter for industrial installations. IEEE Trans Power Sys 8(1):39–44
13. Hideaki F, Hirofumi A (1988) The unified power quality conditioner: the integration of series and shunt active filters. IEEE Trans Power Elect 13(2):494–501
14. Chen G, Chen Y, Sanchez LF, Smedly KM (2000)Unified power quality conditioner for distribution system without reference calculation. In: In proceeding of IEEE international Conference on Power Quality pp 123–127
15. Zheng Peng Fang, Hirofumi Akagi, Akira Nabae (1993) Compensation characteristics of the combined system of shunt passive and series active filters. IEEE Trans Ind Appl 29(1):144–152
16. Graovac D, Kati V, Rufer A, Kneevi J (2001) Unified power quality conditioner based on current source converter topology. EPE pp 1–9
17. Khadkikar V, Agarwal P, Chandra A, Barry AO, Nguyen TD (2004) A simple new control technique for unified power quality conditioner (UPQC). In: Proceeding of 11th international conference on harmonics and quality of power pp 289–293
18. Rezaeipour R, Kazemi A (2008) Review of Novel control strategies for UPQC. Inter J Power Elect Eng 2(3):185–198
19. Salam Z, Bahari N (2010) Selective harmonics elimination PWM (SHEPWM) using differential evolution approach. IEEE Trans 1–5
20. Chia-Nan Ko, Ying-Pin Chang, Chia-Ju Wu (2009) A PSO method with nonlinear time-varying evolution for optimal design of harmonic filters. IEEE Trans Power Sys 24(1):437–445
21. Aredes M, Heumann K, Watanabe EH (1998) An universal active power line conditioner. IEEE Trans Power Delivery 13(2):545–551
22. Banaei MR, Hosseini SH (2006) Mitigation of current harmonic using Adaptive neural network with active power line conditioner. IPEMC
23. Ghosh A, Ledwich G (2001) A unified power quality conditioner (UPQC) for simultaneous voltage and current compensation. Elect Power Energy Sys 59:55–63
24. Fujita H, Akagi H (1998) The unified power quality conditioner: the Integration of series-and shunt-active filters. IEEE Trans Power Elect 13(2):315–322
25. Bajpai RS, Gupta Rajesh (2011) Series compensation to mitigate harmonics and voltage sags/swells in distributed generation based on symmetrical components estimation. IEEE Trans 496:1639–1645
26. Sharmeela C, Uma G, Mohan M (2005)Multi-level distribution statcom for voltage sag and swell reduction. In: Proceeding IEEE power engineering society general meeting 1:278–282

27. Cerrada AG, González PG, CollantesR, Gómez T, Anzola J (2000) Comparison of thyristor-controlled reactors and voltage-source inverters for compensation of flicker caused by arc furnaces. IEEE Trans Power Delivery 15:1225–1231
28. ElnadyA, Noureldin A (2011) Mitigation of arc voltage flicker using an innovative scheme of adaptive notch filters. IEEE Transa Power Delivery 26(3):1326–1337
29. Siahi M, Najafi M, Hoseynpoor M, Ebrahimi R (2011) Design and Simulation of UPQC to Improve Power Quality and Transfer Power of photovoltaic array to grid. Aust J Basic Appl Sci 5(3):662–673

Chapter 77
Real Time Remote Monitoring and Measurement of Loss due to Dry Flue Gas for an Industrial Boiler

C. L. Chayalakshmi, D. S. Jangamshetti and Savita Sonoli

Abstract The loss due to dry flue gas is quite considerable in boilers. This paper presents a novel method of wireless and real time monitoring of this loss in an industrial boiler. The proposed system consists of measuring the percentage of CO_2 present in the flue gas, temperature of flue gas and the ambient temperature. All these signals from the sensor are suitably conditioned and then transmitted to the central station using ZigBee communication. ZigBee is a popular wireless protocol used in process industries. At the central station the loss due to dry flue gas is calculated and displayed. At the central station, an ARM7TDMI-S is used for data acquisition and for calculating the boiler loss. In this work, not only the parameters are measured, but also wirelessly communicated to the central station for finding the loss due to dry flue gas.

Keywords Dry flue gas loss · ARM7TDMI-S · Boiler efficiency · ZigBee

C. L. Chayalakshmi (✉) · D. S. Jangamshetti
Basaveshwar Engineering College, Bagalkot, Karnataka, India
e-mail: chayalakshmi.cl@gmail.com

D. S. Jangamshetti
e-mail: dsjshetti@rediffmail.com

S. Sonoli
Sir M Visveswaraya Institute of Technology, Bangalore, Karnataka, India
e-mail: savitachitriki@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_77,

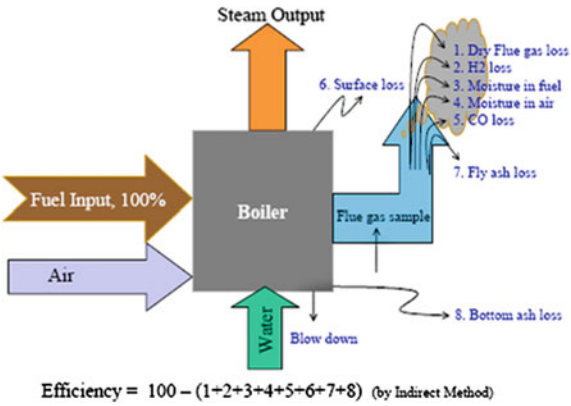


Fig. 77.1 Block diagram of a boiler efficiency calculation [3]

77.1 Introduction

Productivity and quality are the major emphasis of any industry, whereas energy cost is considered as a second priority [1]. However, due to the alarming increase in energy cost, every effort should be made to minimize it. Boiler system is an integral component of a process industry like sugar, cement, textile etc,. Performance of the boiler, like efficiency and evaporation ratio reduces with time. This is due to poor combustion, heat transfer performance of boiler. Efficiency testing helps to find out how far the boiler efficiency drifts away from the best efficiency. Any observed abnormal deviations could therefore be investigated to pin point the problem area for necessary corrective action. Hence it is necessary to find out the current level of efficiency for performance evaluation, which is a pre requisite for energy conservation action in industry.

There are two ways of determining the efficiency of a boiler: direct method and indirect method. The direct method of evaluating the efficiency of a boiler is based on fuel consumption and steam generation for a particular time period as per standards.

In direct method, efficiency is calculated using Eq. (77.1)

$$\eta = \frac{m_s i_s - m_w i_w}{B H_u} \tag{77.1}$$

where m_s is the steam flow rate (kg/s), m_w is the water flow rate (kg/s), i_s is the steam enthalpy (kJ/kg), i_w is the feed water enthalpy (kJ/kg), B is the fuel flow rate (kg/s) and H_u is fuel lower heating value (kJ/kg) [2]. Even though the efficiency can be easily found, direct method will not give clues to the operator about the lower efficiency levels of a system due to various losses. Considering the disadvantages of direct method, one can prefer indirect method of evaluating the efficiency.

The generalized block diagram required for calculating the boiler efficiency using indirect method is as shown in Fig. 77.1 and Table 7.1 gives various losses that contribute for the boiler loss.

Table 77.1 Types of boiler losses

Boiler loss
Loss due to dry flue gas (sensible heat)
Loss due to hydrogen in fuel (H_2)
Loss due to moisture in fuel (H_2O)
Loss due to moisture in air (H_2O)
Loss due to carbon monoxide (CO)
Loss due to surface radiation, convection and other unaccounted
Unburnt losses in fly ash (Carbon)
Unburnt losses in bottom ash (Carbon)

In indirect method, efficiency is calculated using Eq. (77.2).

$$\eta = 100 - Z \tag{77.2}$$

where Z is the sum of various losses.

Loss due to dry flue gas is the major contributor to the overall loss in boilers. If we measure this loss effectively, then efforts can be made to reduce the same. At present, the efficiency is calculated from data mining technique [4], which provides necessary information to the manager to perform the work effectively and efficiently. Data mining is the extraction of meaningful information by the computer from stream of data. This stream of data is generated by the computerized sensors. The extracted data helps to reduce the cost and to increase the revenue. Data mining finds the correlations or patterns among dozens of fields in large relational database. Data mining is one of the off line solutions to get the efficiency of a boiler.

With the rapid advancement in real time embedded systems and development of network and communication technology, the inconvenience of wiring is resolved with wireless sensor network (WSN). Wireless sensor network has many advantages over wired network. ZigBee is one of the wireless communication technologies which has several features suitable for industrial environment such as low cost, low power dissipation, high-capacity networks, safe and reliable data transmission [5–7].

77.2 Proposed Work

The block diagram of the proposed work is as shown in Fig. 77.2. The system proposed uses temperature sensor and CO_2 sensor, signal conditioning circuits, ARM processor, ZigBee transceiver pair.

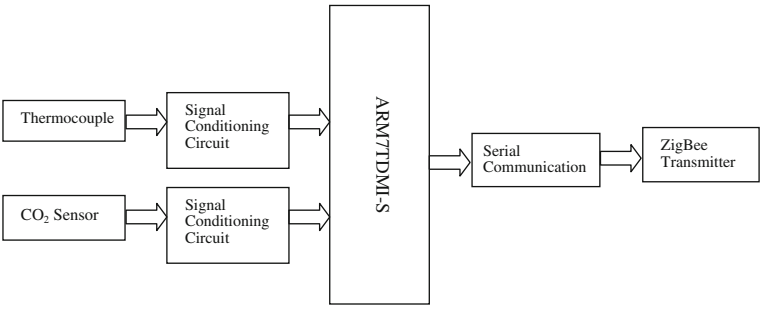


Fig. 77.2 Block diagram of transmitter section

77.2.1 Sensing Module

The requirement here is to measure the flue gas temperature. This temperature represents the major portion of the energy not converted to usable output. The higher the temperature, the less energy transferred to output and the lower the boiler efficiency. To measure this temperature, thermocouple is more suitable. According to the basic theory, the temperature of the flue gas is the sum of the saturation temperature of steam and 40 °C. Hence, J type thermocouple is most suited, as its maximum temperature range is up to 750 °C and it has highest degree of linearity amongst all thermocouples. The sensitivity of 40 μV/°C is quite acceptable in analog signal processing. The sensor output voltage is amplified to suit to the requirement of an inbuilt ADC of ARM processor. If LPC2129, which is one of ARM7TDMI processor is used, then the resolution of the in-built ADC is 4.88 mV, as the in-built ADC is of 10 bits. If one degree Celsius is to be measured, then 40 μV/°C should be converted to 4.88 mV with the help of appropriate signal conditioning circuit.

77.2.2 ARM Processor Module at the Transmitter

A 4-channel, 10-bit inbuilt ADC is used for conversion of analog signal obtained from the sensors to digital bits. ARM processor, LPC2129 is used for transmitting the digital bits corresponding to temperature and percentage of CO_2 from the output of an ADC. The bits are sent serially to the ZigBee module for transmission.

77.2.3 ZigBee/IEEE 802.15.4 Module

Among the various choices of wireless transmission methods like Radio frequency, optical communication (Laser) and Infrared, Radio Frequency (RF) based communication is suitable for most of the wireless sensor network applications.

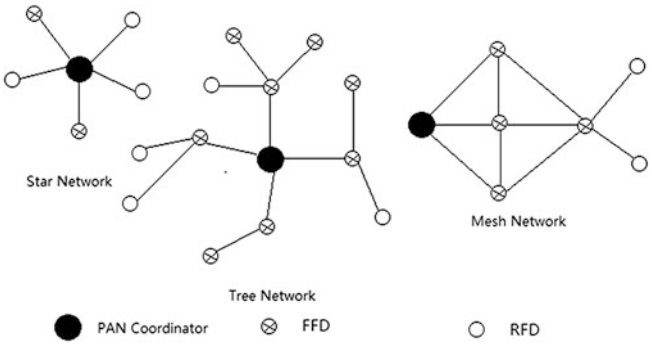


Fig. 77.3 The architecture of ZigBee network

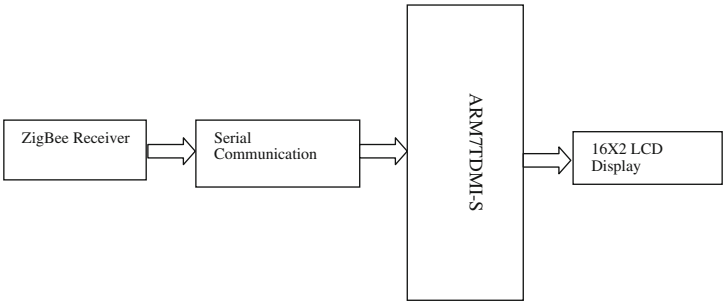


Fig. 77.4 Block diagram of receiver section

ZigBee uses the communication frequency between 433 MHz and 2.4 GHz. One pair of ZigBee module is used for transmission and reception. ZigBee supports multiple network structures, like star, tree, and mesh network, as shown in Fig. 77.3. ZigBee modules are composed of the coordinator, the router, and the end device [4].

77.2.4 ARM Processor

The block diagram of the receiver section is as shown in Fig. 77.4. An ARM processor based on ARM7TDMIS is used at the central station for data acquisition and processing to calculate loss of a boiler due to the dry flue gas [8–10]. The ARM processor incorporates LCD controller, Ethernet Medium Access Controller (MAC), UARTs, Controller Area Network (CAN) channels, an SPI interface and other resources [7]. The ARM7 family of processors is a range of low-power, 32-bit RISC cores optimized for cost and power-sensitive applications. ARM7TDMI supports 16-bit thumb instruction set, enabling high code density to

be achieved with 32-bit performance levels. A synthesizable version of the ARM7TDMI core is ideal for modern design flows and its key factors are portable and flexible [11]. It has on-chip RAM, Flash memory with a facility to expand the memory, an ARM processor is suitable for the data acquisition from sensors placed at remote places. High performance, low power consumption, small die size, high code density, real time debug facilities, coprocessor interface are the common features of all ARM7TDMI processor.

77.3 Software Design

ARM processor which is at the central station receives the signals from ZigBee receiver through serial communication UART. It collects the data from CO_2 sensor and the thermocouple. Then the software has to be designed on ARM platform for performing the steps to calculate the loss of a boiler because of dry flue gas. For coal fired boiler, moisture content, carbon content, hydrogen content, oxygen content, Gross Calorific Value of the fuel is collected. Then using the following steps, the loss due to dry flue gas is calculated [3].

Step 1: Find theoretical air requirement

$$\begin{aligned}\text{Theoretical air required } = &[(11.6 \times \text{C}) + \{34.8 \times (\text{H}_2 - \text{O}_2/8)\} \\ &+(4.35 \times \text{S})]/100 \quad \text{kg/kgof coal}\end{aligned} \tag{77.3}$$

Step 2: Find theoretical $\%CO_2$ ($(\%CO_2)_t$)

$$(\%\text{CO}_2)_t = \text{ Moles of C}/[\text{Moles of N}_2 + \text{ Moles of C}] \tag{77.4}$$

After measuring the actual CO_2 in flue gas ($(\%CO_2)_a$),

Step 3: Find % Excess Air supplied (EA)

$$\begin{aligned}\%\text{Excess Air supplied (EA)} = \{7900 \times [(\text{CO}_2\%)_t - (\text{CO}_2\%)_a]\}/[(\text{CO}_2)_a\% \times [100 \\ -(\text{CO}_2\%)_t]]\end{aligned} \tag{77.5}$$

Step 4: Find Actual mass of air supplied

$$\text{Actual mass of air supplied } = \{1 + \text{ EA}/100\} \times \text{theoretical air} \tag{77.6}$$

Step 5: Find Actual mass of dry flue gas

$$\begin{aligned}\text{Actual mass of dry flue gas } = &\text{ Mass of CO}_2 + \text{ Mass of N}_2\text{content present in fuel} \\ &+ \text{Mass of N}_2\text{inthe combustion airsupplied} \\ &+ \text{Mass of oxygen in flue gas}\end{aligned} \tag{77.7}$$

After measuring the flue gas temperature and ambient temperature,

Step 6: % Heat loss in dry flue gas is calculated using the following equation:

$$\%\ \text{Heat loss in dry flue gas} = \left[\text{m x} C_p \text{x}\ (T_f - T_a) / \text{Gross Calorific Value of fuel}\right] \times 100 \tag{77.8}$$

Where, m = Mass of dry flue gas calculated from step 5, C_p = Specific heat of flue gas in kCal/kg, T_f = Flue gas temperature in °C, T_a = Ambient temperature in °C.

77.4 Conclusion

This paper presents a novel method of real time monitoring of a loss in the boiler so that the boiler efficiency can be calculated and displayed on-line. The method proposed is of low cost when compared with the existing method of finding real time boiler efficiency. The difficulty of connecting a wire from the field to the central station is eliminated. ARM processor is suggested in place of PLCs which are of low cost.

References

1. Tyagi SK, Sharma BK, Sanjeev S, (2011) Boiler efficiency evaluation and monitoring through data mining techniques in textile industry. Int J Comput Sci Eng ISSN: 0975-3397, vol 3
2. Ramalingam KK (2002) Power plant engineering. SciTech Publications, Chennai
3. Energy Performance Assessment of Boiler, http://www.em-ea.org/Guide%20Books/book-4/4.1%20Boiler.pdf.
4. Durmus K, Muharrem E (2010) Energy conservation opportunities in an industrial boiler. J Energy Eng 136(1):18–25 ISSN 0733-9402/2010
5. Wei W, He G, Wan J (2011) Research on ZigBee wireless communication Technology, 978-1-4244-8165-1/11/, IEEE The National Natural Science Foundation of China (No.51077079)
6. SrideviVeerasingam, SaurabhKarodi, SapnaShukla, MeharChaitanyaYeleti: Design of Wireless sensor Network node on ZigBee for Temperature Monitoring, IEEE International conference on Advances in Computing and Telecommunication Technologies, 978-0-7695-3915-7/09, (2009).
7. Pang, N (2011) ZigBee Mesh network for greenhouse monitoring. In: IEEE international conference on mechatronic science, electric engineering and computer 978-1-61284-722-1/11
8. Lu P, Liang G, Yu Q (2011) A data acquisition system for methane drainage based on ARM and FPGA. In: IEEE international conference on computer distributed control and intelligent environmental monitoring, 978-0-7695-4350-5-5/11
9. Zhu Q, Zhu D, Su X (2010) Distributed remote temperature monitoring and acquisition system based on Can bus. In: IEEE prognostics and system health management conference, 978-4244-4758-9/10

10. Daogang P, Hao Z, Kai Z, Hui L, Fei X (2009) Research and development of the remote I/O data acquisition system based on embedded ARM platform. In: IEEE International conference on electronic computer technology, 978-0-7695-3559-3
11. ARM, The architecture for the digital world, http://www.arm.com

Chapter 78
Characterization of Electrical and Thermal Properties of Enamel Filled with Carbon Nanotubes

D. Edison Selvaraj, C. Pugazhendhi Sugumaran and A. SivaPrakash

Abstract The last decade has witnessed significant developments in the area of nanoparticles and nanoscale fillers on electrical, thermal and mechanical properties of polymeric materials. The dielectric and thermal properties of standard (Polyamide-imide) and nanoscale filled samples were detailed and analyzed. Carbon nanotubes have been tested as filler. Carbon nanotubes were synthesized by the process called chemical vapour deposition (CVD). The basic properties such as dielectric loss tangent (tanδ), dielectric constant(ε), dielectric strength, partial discharge inception voltage, surface resistivity, quality factor, phase angle, dielectric conductivity, dielectric power loss and thermal withstand strength of the enamel filled with carbon nanotubes were analyzed and compared with the properties of the standard enamel. The experimental results show that there was a significant improvement in the properties of the enamel by the addition of carbon nanotubes.

Keywords Carbon nanotubes · Chemical vapour deposition · Dielectric strength · Partial discharge · Dielectric spectroscopy

D. Edison Selvaraj (✉) · A. SivaPrakash
Department of EEE, Mepco Schlenk Engineering College, Sivakasi, India
e-mail: edisonsivakasi@gmail.com

A. SivaPrakash
e-mail: sivahp@mepcoeng.ac.in

C. Pugazhendhi Sugumaran
Division of High Voltage Engineering, College of Engineering, Guindy, Chennai, India
e-mail: cpsugumar@gmail.com

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_78,

78.1 Introduction

In the last few years, a great deal of attention has been given to the application of nanodielectrics in the field of electrical insulating materials. It has been reported that the use of nano particles in the matrix of polymeric materials can greatly improve the thermal, mechanical and electrical properties of polymeric nanocomposites [1]. Insulating materials play a significant role in the design and performance of high voltage systems. They can be used for insulation purposes, cooling purposes and mechanical support [2]. Despite the basic understanding of electrical breakdown of materials, electrical surface flashover phenomena, physical mechanisms responsible for the initiation of such unwanted electrical activities within an insulation system composed of such advanced materials must be investigated before they can be commercially available [3]. The findings of such studies were essential for the development of nano-electric and other advanced materials and the techniques to predict the reliability of the advanced electrical systems which utilize these materials. The nanostructured polymeric materials are object of great interest by the researchers. The reasons of this interest were well-known: several mechanical, thermal and electrical properties can be improved by adding few percent of inorganic "nanofiller". But as regards barrier properties these materials gave the best results [4]. This paper focused on the characterization of dielectric and thermal properties of standard enamel and carbon nanotubes filled enamel. There was a significant improvement in the properties of the enamel by the addition of carbon nanotubes.

78.2 Proposed Work

78.2.1 Sample Preparation

The nanocomposites were prepared by radical initiator curing method. 80 % of enamel and 20 % of epoxy resin were taken. Diamino Diphenyl Methane (DDM) was used as curing agent. For 1 g of resin, 0.27 g of DDM was taken. The DDM was melted at 60–80 °C for 10 min. The enamel, resin and melted DDM were mixed in a beaker. The mixture was poured into the die coated by a Teflon sheet. The die was heated at 120 °C for 3 h. Then, the die was taken away from the oven and it was cooled for 1 h. The carbon nanotubes were mixed with the enamel by ultrasonic vibrator at different proportions (1, 3 and 5 wt%). Four different samples were produced by this method.

Fig. 78.1 Quartz boat

Fig. 78.2 Experimental setup of CVD system

78.2.2 Synthesis of Carbon Nano Tubes

Chemical vapour deposition (CVD) process was done in an experimental set up consisting of a horizontal reaction furnace, quartz tube, PID controller, flow meters, control valves, gas sources and thermocouple as shown in Figs. 78.1 and 78.2.

Approximately 200 mg of catalyst powder was taken in the quartz boat and was placed in the central region of the furnace. The furnace was flushed with argon gas and was heated at a rate of 5 °C/min till it attains 800 °C. After the attainment of the desired temperature, H_2 gas was introduced into the furnace at a flow rate of 100 cc/min for 60 min so as to generate active metallic(Fe, Mo) or bimetallic (Fe-Mo) nanoparticles on Alumina support. Subsequently, C_2H_2 gas was introduced into the furnace for 1 h. After that it was allowed to cool at a rate of 5 °C/min. It was then washed with distilled water. Then the particle size of the powder was analyzed by using the SEM characterization techniques.

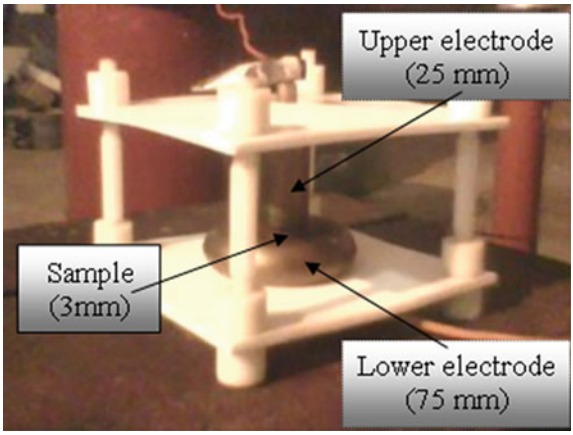


Fig. 78.3 Electrode setup for BD and PD

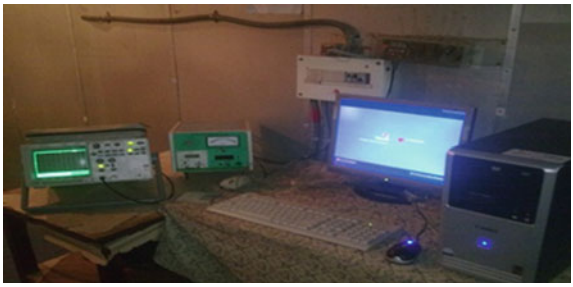


Fig. 78.4 Set up for partial discharge measurement

78.2.3 Dielectric Spectroscopy Analysis

Dielectric spectroscopy was used to measure the inductance, capacitance and impedance of the component. It was used to measure the dielectric properties of a medium as a function of frequency.

78.2.4 Partial Discharge and Dielectric Strength Measurements

The partial discharge experiment was carried out inside the shielded room to avoid the external noises. The different samples were placed between the electrodes and the whole electrode setup was kept inside the oil to avoid gliding discharge between contacts. The voltage was applied gradually. The initial discharges occurring in the samples were captured by a high quality oscilloscope. The inception and extinction voltages were noted. A standardized testing arrangement with electrode setup for the determination of the breakdown (BD) voltage and partial discharge (PD) inception and extinction voltage of solid samples as per standard (IEC 60243—1) was shown in Figs. 78.3and 78.4.

The dielectric strength test was conducted with alternating voltage, which should be increased from zero to the breakdown value. The voltage was applied to the samples by means of a high voltage transformer. The value of the voltage at which breakdown occurs in the sample was noted. The sample thickness was 3 mm and the diameter of upper electrode was 25 mm and the diameter of lower electrode was 75 mm [5]. The entire arrangement was immersed in an insulating liquid with higher dielectric constant.

78.2.5 Thermo Gravimetric Analysis

Thermo gravimetric analysis (TGA)was a simple analytical technique to measure the weight loss or weight gain of a material as a function of temperature. As materials were heated, they can lose weight by drying or by liberating some gases. Some materials can gain weight by reacting with the atmosphere in the testing environment. The TGA results have been obtained from diamond TG/DTA 6,000 instrument system. The sample of 0.1 mg–10 g was taken and the heat was applied at a rate of 0.1–50 °C/min. The temperature was maintained in the range of 50–900 °C to maintain consistent heating rate and gas flow. Sampling purity, reaction rate, identification, activation energy and heat of reactions were measured using this instrument.

78.3 Results and Discussions

78.3.1 Analysis of Nano-Scale Structure

Figure 78.5 shows the SEM analyzed image results. These results show that particles were in the form of nano metric range. The sizes of the particles were in the range from 50 to 120 nm size. The XRD results for the carbon nanotubes were shown in Fig. 78.6. The particle size of the carbon nanotubes was found using XRD. From the graph, it was clear that the value of 2θ lies between 30 and 40°. Hence the results show that the particle size of the carbon particles were in the nanometer range.

78.3.2 Dielectric Spectroscopy Analysis

The dielectric characteristics of the samples 0 (pure), 1, 3 and 5 wt% were analyzed by dielectric spectroscopy instruments from 50 to 5 MHz range. The loss factor and quality versus frequencies at 90° C with different samples were shown

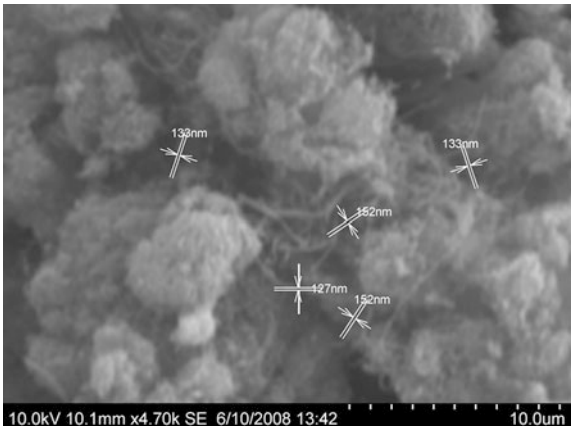


Fig. 78.5 SEM analysis of Carbon nanotubes

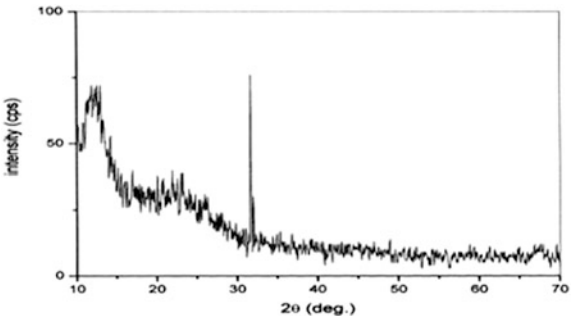


Fig. 78.6 XRD pattern for carbon nanotubes

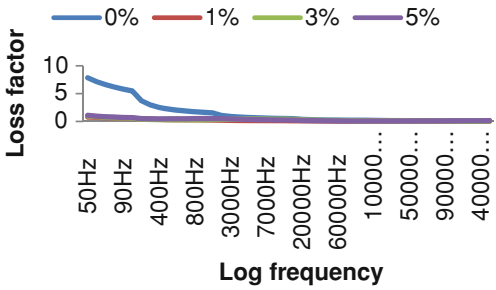


Fig. 78.7 Frequency dependence of loss factor for the enamel and the different (1, 3 and 5 wt%) proportions of carbon nanotubes filled enamel at 90° C

graphically in the Figs. 78.7 and 78.8. A good insulating material should have higher permittivity value.

The permittivity versus frequency for different samples was represented in Fig. 78.9 and 78.10. At 50 Hz the permittivity was higher. At higher frequencies there was not much variation in permittivity. At lower frequencies the dissipation factor was higher because of higher settling time in polarization. At higher frequencies, settling time was low in polarization so the dissipation factor was low. At 10 kHz, there was minute variation in permittivity.

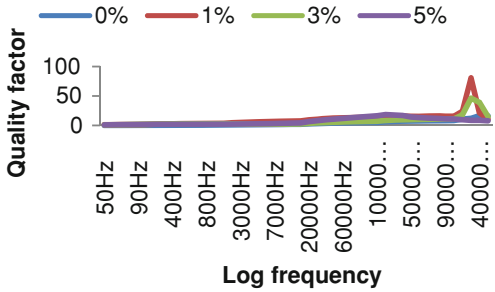


Fig. 78.8 Frequency dependence of quality factor for the enamel and the different (1, 3 and 5 wt%) proportions of carbon nanotubes filled enamel at 90 °C

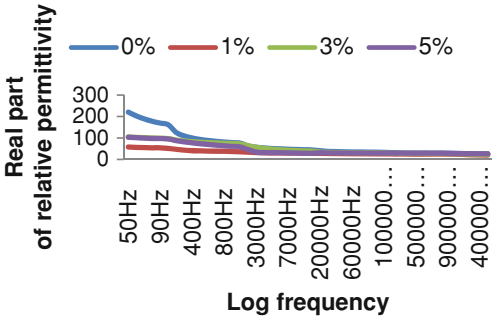


Fig. 78.9 Frequency dependence of the real part of the Relative permittivity for the enamel and the different (1, 3 and 5 wt%) proportions of carbon nanotubes filled enamel at 90 °C

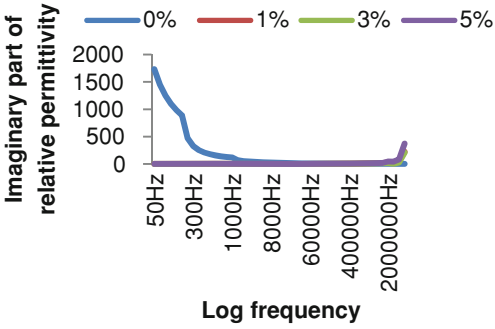


Fig. 78.10 Frequency dependence of the imaginary part of the relative permittivity for the enamel and the different (1, 3 and 5 wt%) proportions of carbon nanotubes 500

78.3.3 Partial Discharge and Dielectric Strength Measurement

Partial discharges were in general a consequence of local electrical stress concentrations in the insulation or on the surface of the insulation. The partial discharge measurement was carried out in uniform field electrode configurations. The breakdown voltage shows an increasing dependence on the nature and smoothness of the electrode material. The breakdown strength reduces considerably due to the presence of impurities. The different values of Partial discharge and dielectric strength for uniform field configurations were shown in Table 78.1. The 1 wt% nanocomposite sample has higher inception and extinction voltages when compared to other samples.

Table 78.1 Partial discharge and dielectric strength values for different samples

% wt of carbon Nano tubes	Inception voltage (kV)	pC	Extinction voltage (kV)	pC	Breakdown strength (kV/mm)
0	4.74	55	4.10	1.3	2.56
1	3.20	65	2.29	1.4	1.91
3	4.21	62	3.73	1.4	2.34
5	4.31	66	3.73	1.4	2.35

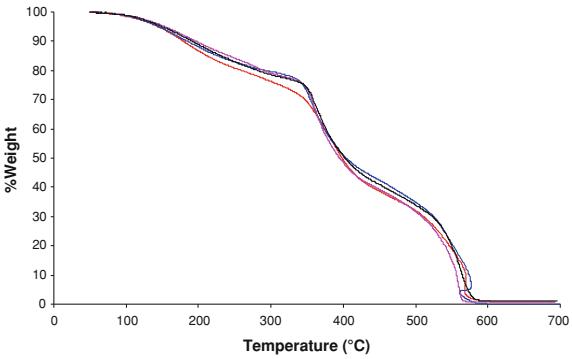


Fig. 78.11 TGA results

Table 78.2 TGA result for various samples

Sample	On set temp (°C)	Peak temp(°C)	End temp (°C)
Pure (wt%)	504.90	551.98	593.09
1	517.76	569.59	595.16
3	519.76	589.79	600.26
5	531.26	595.63	609.30

78.3.4 Thermo Gravimetric Analysis (TGA)

The TGA result of the 0 % (pure enamel), 1, 3 and 5 % were shown in the Fig. 78.11. The graph was plotted by taking the TGA signal (actual weight loss or gain converted to percent weight loss) on the Y-axis and the sample temperature in °C on the X-axis. The melting point temperatures for various samples were given in the Table 78.2. From the result the 5 wt% sample has the higher melting point when compared to other samples.

78.4 Conclusions

SEM analysis showed that the prepared carbon particles were appearing in the form of nano metric size. The various dielectric properties were analyzed by dielectric spectroscopy instrument at 90 °C for the frequency range of 50–5 MHz. The thermal property of the various nanocomposite samples were analyzed by thermo gravimetric analysis (TGA) for the temperature range from 50 to 700 °C at the heating rate of 0.1–50 °C/min. The 5 wt% sample was having better thermal performance when compared to other samples. These results show that the additions of few weight percentages of carbon nanotubes to the enamel improved the dielectric and thermal behaviour of the enamel.

References

1. Pugazhendhi Sugumaran C, Mohan MR, Udayakumar K (2010) Investigation of dielectric and thermal properties of nano-FILLER (ZrO_2) mixed enamel. IEEE Trans Dielect Electr Insul 17(6)
2. Nguyen M et al (2009) Investigations on dielectric properties of enameled wires with nanofilled varnish for rotating machines fed by inverters. IEEE Electr Insul Conf
3. Takahiro I et al (2008) Nano- and micro-filler combination enabling practical use of nanocomposite insulating materials. In: Proceedings of international symposium on electrical insulating materials
4. Hulya K, Serkan M, Koppisetty K (2005) Nano-dielectric materials in electrical insulation application. IEEE
5. Guoqin Z et al (2005) Study of nano TiO_2 filler in the corona -resistant magnetic wire insulation performance of inverter-fed motor. In: Proceedings of international symposium on electrical insulating materials
6. Guastavino F et al (2007) Characterization of nanofilled epoxy varnish subjected to surface partial discharges. In: IEEE annual report conference on electrical insulation and dielectric phenomena
7. Inuzukal K, Inanol H, Hayakawal N (2006) Partial discharge characteristics of nanocomposite enameled wire for inverter-fed motor. In: Annual report conference on electrical insulation and dielectric phenomena
8. Kozako M, Norikazu F, Yoshimichi O (2004) Surface degradation of polyamide nanocomposites caused by partial discharges using IEC(b) electrodes. IEEE Trans Dielectr Electr Insul
9. Naoki H, Okubo H (2008) Lifetime characteristics of nanocomposite enameled wire under surge voltage application. IEEE Electr Insul Mag
10. Takahiro I et al (2006) Effects of nano- and micro-filler mixture on electrical insulation properties of epoxy based composites. IEEE Trans Dielectr Electr Insul 13(1)
11. Takahiro I et al (2008) Improving epoxy-based insulating materials with nano-fillers toward practical application. IEEE

Chapter 79
Tuned Fuzzy Logic Control of Switched Reluctance Motor Drives

Nessy Thankachan and S. V. Reeba

Abstract The switched reluctance motor (SRM) has gained much attention in the past few years over other types of electric motors in the drive applications due to its simple structure, ruggedness and inexpensive manufacturing potential. However, these merits are overshadowed by its inherent high torque ripple, acoustic noise and difficulty to control. When the exact analytical model of the controlled system is uncertain or difficult to be characterized, intelligent control arts such as fuzzy logic control (FLC) may allow better performance compared to conventional controllers. In this paper a PI-like fuzzy logic speed controller with output scaling factor tuned, by an updating factor, based on fuzzy logic reasoning, is applied to an SRM drive system. A reduced rule base is used to simplify the program complexity of the controller without losing the system performance and stability. The nonlinear modeling of SRM is done based on look up tables with data obtained by finite element analysis.

Keywords Fuzzy logic controller · Scaling factor · Switched reluctance motor

79.1 Introduction

Switched reluctance motor (SRM) can be a potential alternative to other conventional ac machines commonly used in various industries due to its unique characteristics in the aspects of mechanical simplicity in construction, high

N. Thankachan (✉) · S. V. Reeba
Department of EEE, College of Engineering, Trivandrum, India
e-mail: nessythankachan@gmail.com

S. V. Reeba
e-mail: reeba_pradeep@yahoo.co.in

V. V. Das (ed.), *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, Lecture Notes in Electrical Engineering 150,
DOI: 10.1007/978-1-4614-3363-7_79,

efficiency, fault tolerance ability, high reliability and robustness in operation. However, due to the doubly salient structure and magnetic saturation, SRM acquaints with torque ripples, vibration and noise, which limits its application. It is however, difficult to control, due to its non-linear nature [1, 2].

Therefore, precise control of SRM is not easy using conventional methods like PI or PID controls as its flux linkage, inductance, and torque possess mutual coupling with rotor position and phase current. Hence, analytical or computer-based experimental determinations are often required to characterize the magnetization curves of the SRM. When the analytical model of the controlled system is vague or difficult to model, intelligent control techniques such as Fuzzy Logic Controller (FLC) gives better control performance [3].

The success of fuzzy logic controllers mainly lies in their ability to cope with knowledge represented in a linguistic form instead of representation in the conventional mathematical framework. The advantages of fuzzy logic includes robustness, a much wider range of operating conditions, operation with noise and disturbances of different natures, cheap and customizable[4]. In essence, FLC is a nonlinear controller which is suitable for the control of nonlinear, structure variation, parameter variation and transfer function unknown plant.

In this paper a Tuned Fuzzy logic speed Controller (TFC), with tuning factor based on fuzzy logic reasoning is designed and simulated for application to SRM drives. The TFC uses the speed error and change in speed error as inputs and generates an equivalent control term, which improves system performance in steady state. Section 79.2 describes the non linear modeling of SRM, Sect. 79.3 gives an overview of PI fuzzy controller and TFC. Section 79.4 illustrates the design of a TFC with reduced number of control rules. Section 79.5 depicts the simulation and results of the work. Conclusions are given in the last section.

79.2 Switched Reluctance Motor

79.2.1 Non Linear MODEL of SRM

The SRM drive system simulation is much more complex than ac & dc motor due to the nonlinear B–H characteristics of the magnetic material, the dependence of phase flux linkages on both the rotor position and current and a single source of excitation [1, 2].

SRM models has three parts: the electrical model, torque characteristics and mechanical model Fig. 79.1 [5–9].

The data for the look-up table which approximates the inductance and torque as a function of current and rotor position are obtained by conducting Finite Element Analysis. The specifications of the designed 8/6 SRM used are shown in Appendix

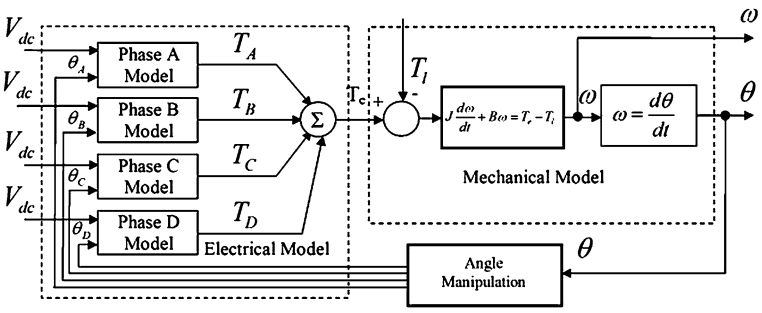
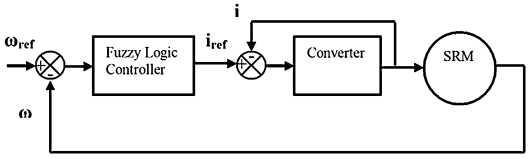


Fig. 79.1 Four phase SRM

Fig. 79.2 SRM with FLC



79.2.2 Control of SRM

In SRM, the inductance depends on both the position of the rotor and the excitation current. For motoring operation the windings are excited at the onset of increasing inductance. The average torque of the machine can be controlled by adjusting the magnitude of current in the winding using a current controller to reduce the torque ripple and ensure safe operation of the machine.

Figure 79.2 shows the block diagram of the drive with FLC. The speed error is given to the FLC to adjust the value of reference current to the motor to maintain the speed of the drive at its reference value.

79.3 PIFLC and TFC

A PI type FLC generates incremental control output from error and change in error and is a velocity type control unlike a PD type is a position type control [6]. The main difference between a PIFC and a TFC is that the TFC includes another control rule base for the gain tuning factor α (Fig. 79.3). FLC tuning implies the handling of a great quantity of variables like the shape, number and ranges of the membership functions (MFs), the percentage of overlap among them and the design of the rule base [8].

The Scaling factor (SF) Ge, $G\Delta e$, and $G\Delta u$, perform the specific normalization of input and output variables and determine controller stability and performance [8]. For conventional FLCs, the controller output (Δu_N) is mapped onto the

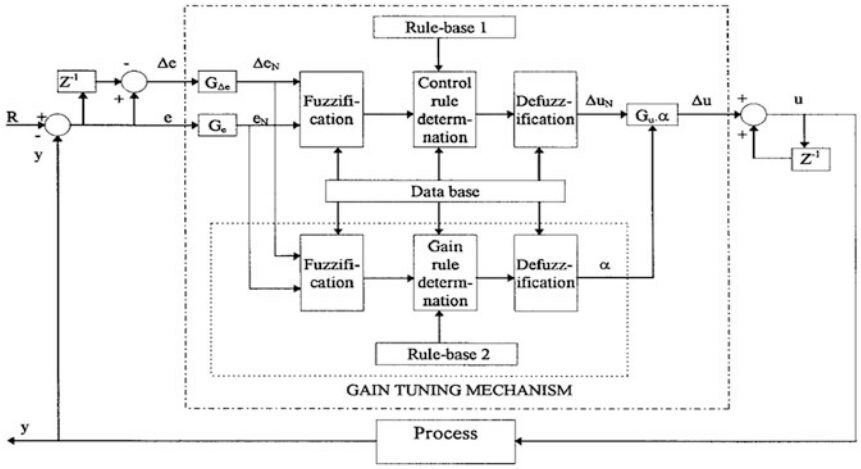


Fig. 79.3 TFC

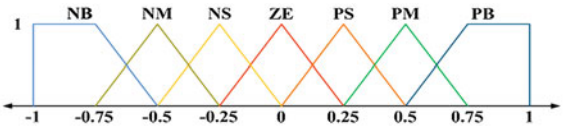


Fig. 79.4 Membership functions of e, Δ*e*, and Δ*u*. NB—negative big, NM—negative medium, NS—negative small, ZE—zero, PS—positive small, PM—positive medium, PB—positive big

respective actual output (Δu) domain by the output SF ($G\Delta u$) while in TFC, the actual output is obtained by using the effective SF ($\alpha G\Delta u$). A triangular membership function is chosen for inputs and output (Figs. 79.4, 79.5). The rule bases for controller output Δu and α is designed with a 2-D phase plane. (Tables 79.1 and 79.2). Here mamdani fuzzy inference and the center of area defuzzification method is used Table 79.2.

79.4 Tuned FLC with Control Rule Reduction

In all industrial applications, the control algorithm must be implemented on a micro controller with limited memory space and fast computational capability. The rule base of TFC uses 98 rules causing a challenge on the performances request of the used DSP on the cost factor. Therefore a TFC with control rule reduction is used for the derivation of α. (Fig. 79.6) keeping the same rule base (Table 79.1) for ΔU. Here input variables (error and change of error) uses 3 fuzzy sets each (Negative, Zero, Positive), defined on a normalized domain and the output variable

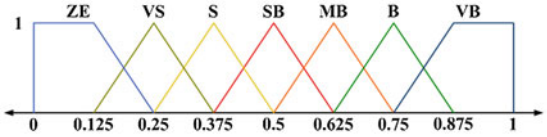


Fig. 79.5 Membership function of α 515. *Note* ZE—zero, VS—verysmall, S—small, SB—small big, MB—medium big, B—big, VB—very big

Table 79.1 Rule base for ΔU

e ΔU Δe	NB	NM	NS	ZE	PS	PM	PB
NB	NB	NB	NB	NM	NS	NS	ZE
NM	NB	NM	NM	NM	NS	ZE	PS
NS	NB	NM	NS	NS	ZE	PS	PM
ZE	NB	NM	NS	ZE	PS	PM	PB
PS	NM	NS	ZE	PS	PS	PM	PB
PM	NS	ZE	PS	PM	PM	PM	PB
PB	ZE	PS	PS	PM	PB	PB	PB

Table 79.2 Rule base for α

e α Δe	NB	NM	NS	ZE	PS	PM	PB
NB	VB	VB	VB	B	SB	S	ZE
NM	VB	VB	B	B	MB	S	VS
NS	VB	MB	B	VB	VS	S	VS
ZE	S	SB	MB	ZE	MB	SB	S
PS	VS	S	VS	VB	B	MB	VB
PM	VS	S	MB	B	B	VB	VB
PB	ZE	S	SB	B	VB	VB	VB

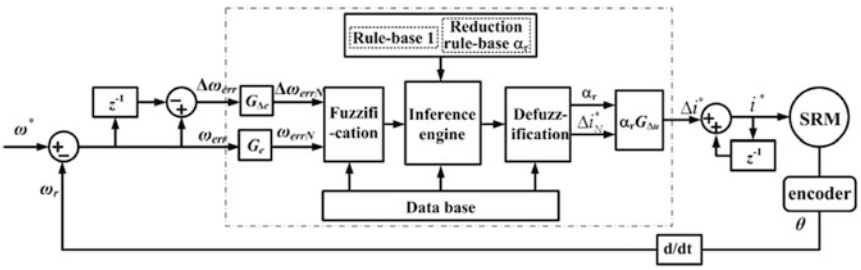


Fig. 79.6 A tuned FLC with a reduced control rule base

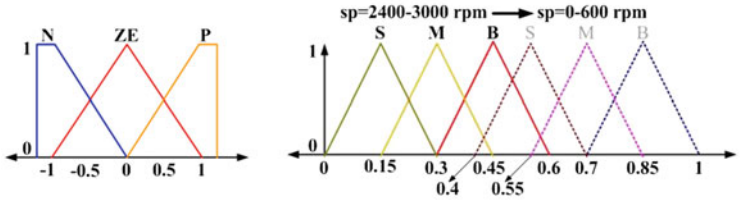


Fig. 79.7 **a** ERROR & change. **b** Gain factor

Table 79.3 Rule base for deriving $\alpha_{reduced}$

	N	ZE	P
N	B	M	S
ZE	M	S	M
P	S	M	B

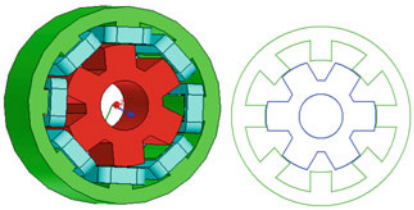


Fig. 79.8 Design of 8/6 SRM

α with 3 fuzzy sets (Small, Medium, Big) with different domains for different operating conditions (Fig. 79.7). Thus in total there will be only 58 rules as opposed to the 98 rules of previous case Table 79.3.

79.5 Simulation and Results

Figure (79.8) shows the geometry of 8/6 SRM designed in Maxwell software (a 4-phase SRM with 6rotor poles and 8 stator poles)

After designing an 8/6 SRM (data given in Appendix), simulation in full-load condition gives the following results:

Figure 79.9 shows the non linear magnetization characteristics of the low power SRM in 10° steps from the aligned position to the unaligned position.

Figure 79.10 shows air-gap inductance reaches to the highest value of 40 mH and varies non linearly with rotor position. Unlike an ideal inductance profile which is trapezoidal, this curves near the top due to saturation.

The phase current reaches to the highest value of 30 A (Fig. 79.11).

The designed motor has an efficiency of 73 % (Fig. 79.12).

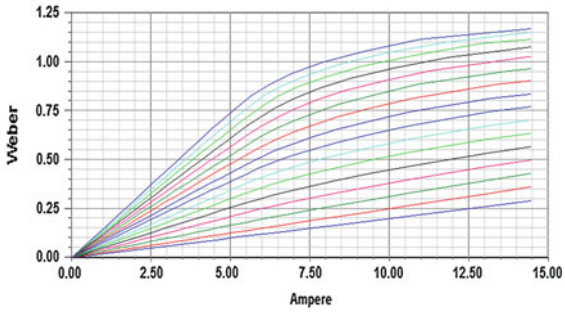


Fig. 79.9 Flux linkage versus current

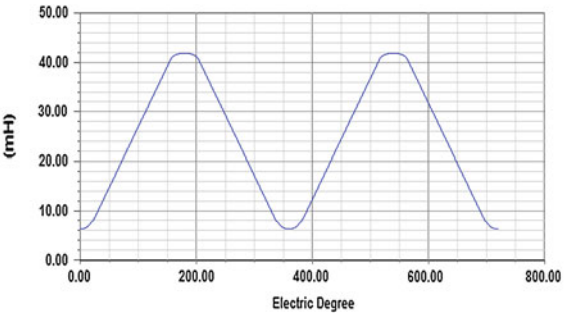


Fig. 79.10 Air gap inductance

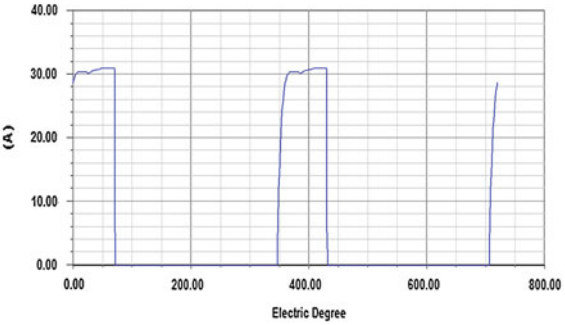


Fig. 79.11 Max phase current

The machine designed in Maxwell.software is simulated for different currents, with position varied and the data for inductance profile as well as the torque is obtained (Fig. 79.14). The dynamic model of SRM is constructed in the SIMULINK environment (Fig. 79.13). The power converter is built using logic operators. Hysteresis controller limits the current in each phase. The 4 phases are designed, and synchronized to each other by providing the fixed values of on and off values along with the reference current, as inputs and the individual instantaneous torques are summed to obtain the total torque developed by the motor.

Under no load condition the various characteristics are obtained for the model.

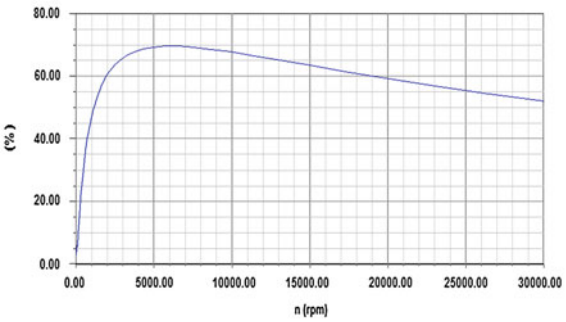


Fig. 79.12 Efficiency versus speed

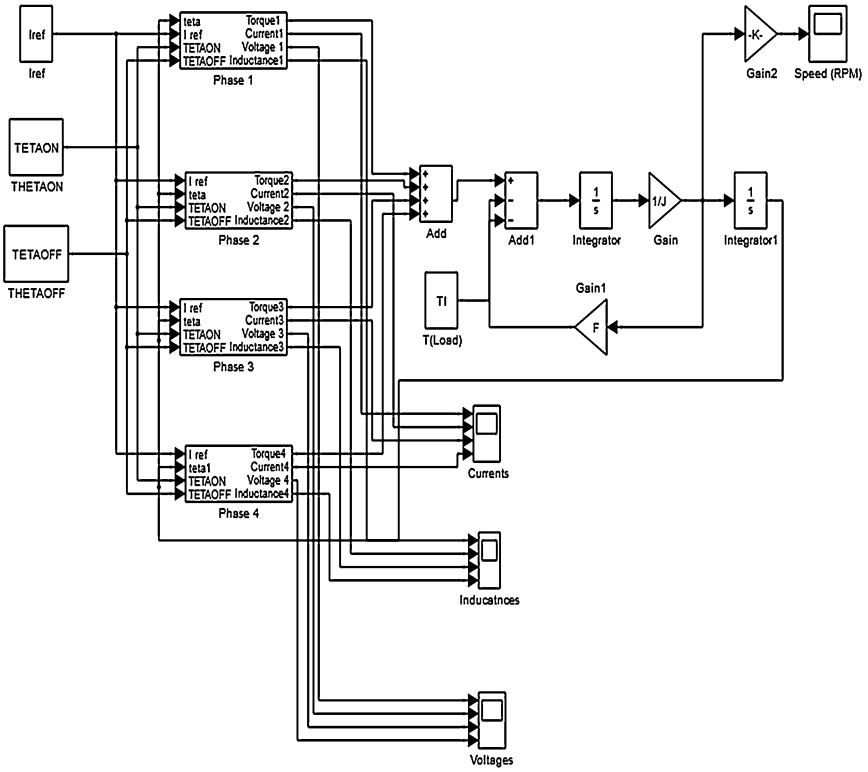


Fig. 79.13 8/6 Non linear SRM

Figure 79.15 gives the sequence of voltage application for different phases and the corresponding variation of current in each phase. On application of voltage, the current rises and then it falls back at the removal of voltage.

Figure 79.16 shows the motor speed response with large oscillations at the final steady state.

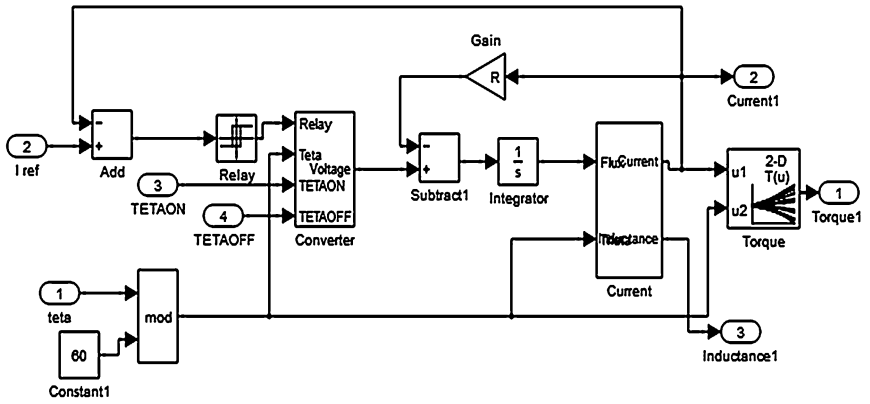
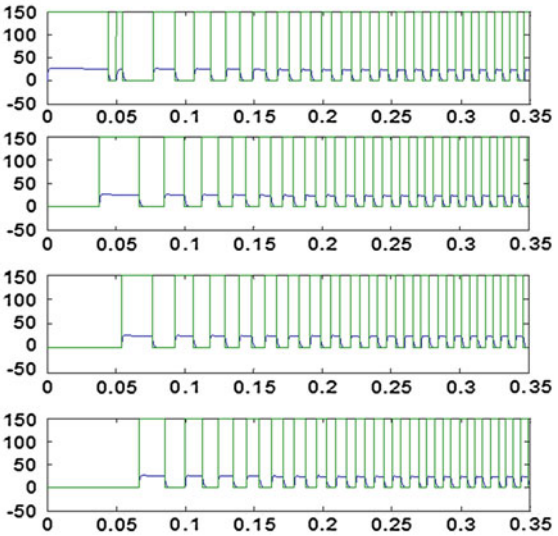


Fig. 79.14 Single phase winding of SRM

Fig. 79.15 Voltage and current



The inductance profile for different phases (Fig. 79.17) is nonlinear and varies depending on the position of the rotor and the excitation current in the winding.

One of the drawbacks of SRM is the high torque ripple for the machine. The total torque output of the machine is given in Fig. 79.18. It is the sum of instantaneous torques produced by different phases. The variation of the torque is in accordance with the inductance profile of different phases Fig. 79.19.

The motor speed and torque responses for a reference command of 900 rpm under a load of 0.75 Nm for various controllers are shown below.

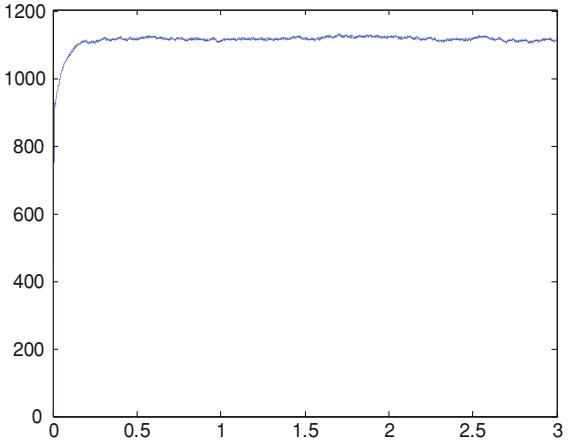


Fig. 79.16 Speed versus time

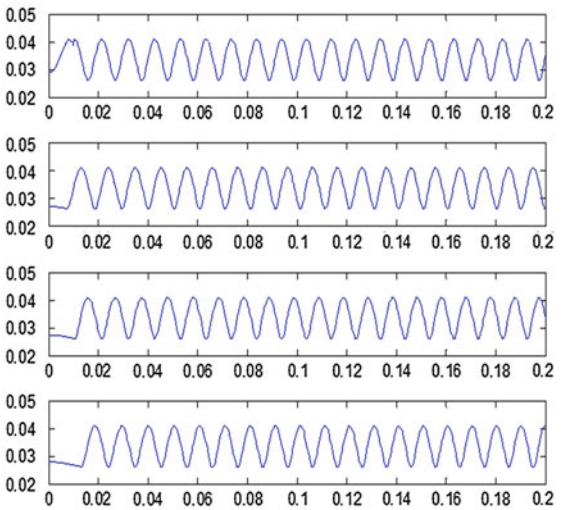


Fig. 79.17 Inductance profile

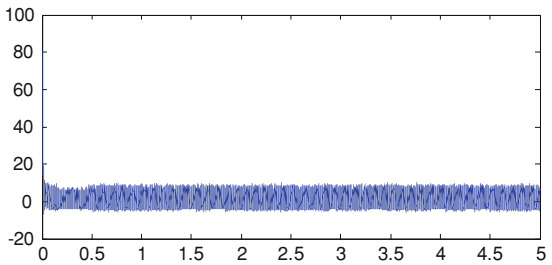


Fig. 79.18 Total torque versus time

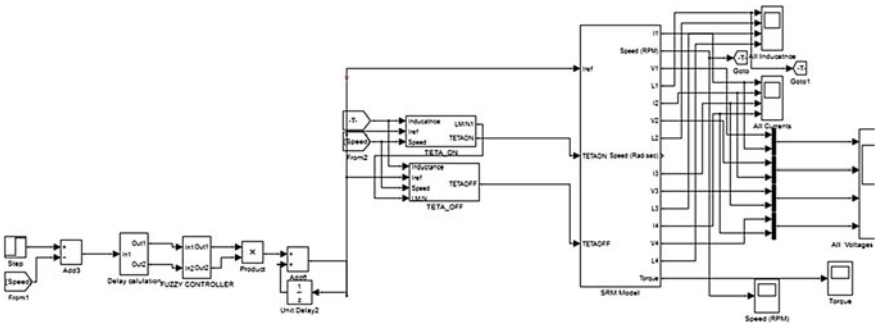


Fig. 79.19 SRM with TFC

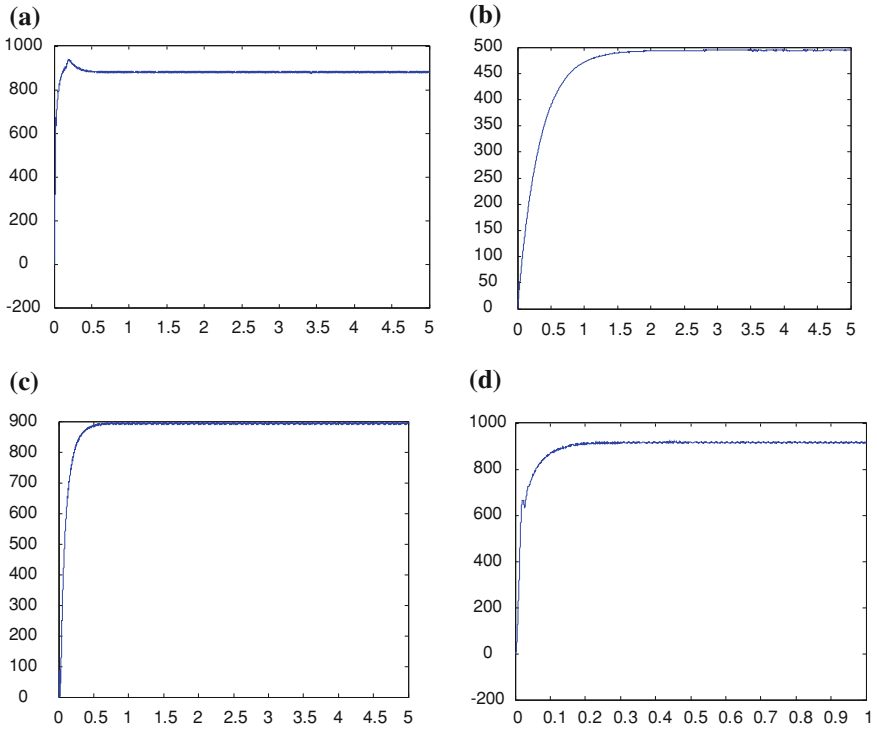


Fig. 79.20 **a** With PI controller. **b** With conventional fuzzy. **c** With TFC. **d** With TFC reduced

The SRM drive with PI controller (Fig. 79.20a). shows speed response with overshoot, the settling time is long, showing oscillations. This demonstrates that the conventional PI controller is not a good choice for the speed control of SRM.

The FLC with fixed gain cannot exactly follow the given command (Fig. 79.20b).

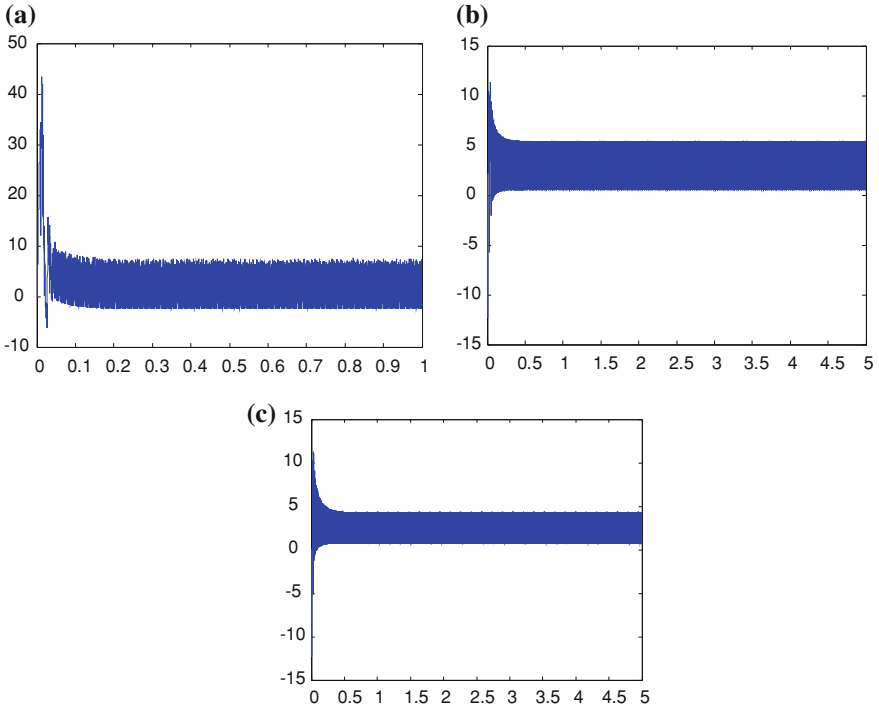


Fig. 79.21 **a** With PI controller. **b** With conventional FC. **c** With TFC

The rotor speed response curve illustrates that the performance of the TFC is satisfactory. The motor responds quickly with very little overshoot. But there is slight oscillation around the reference speed. This is caused by using the look up table method (Fig. 79.20c). The TFC shows a strong characteristic of robustness.

The Tuned Fuzzy Controller with reduced rule base can give the same system performance and stability with reduced program complexity for the controller (Fig. 79.20d).

The torque ripple is very high under a loaded condition of 0.75 Nm with PI control (Fig. 79.21a).

The torque response is much improved with conventional fuzzy when compared to PI controllers (Fig. 79.21b)

The torque ripple has reduced considerably for the drive with TFC (both cases) compared to its conventional counterparts.

79.6 Conclusions

The importance of TFC is highlighted here by comparing the performance of various control approaches, including PI and conventional FLC for the speed control of 4 phase SRM drives. From the simulation results, the performance of 8/6 SRM for a reference speed under loaded condition is analyzed. It is found that the TFC produces no overshoot and has a faster settling time when compared to its conventional counterparts. The system presents little error in steady state. Also the torque ripple is reduced considerably. The TFC with reduced control rules can simplify the program complexity of the controller by reducing the number of fuzzy sets of the MFs without losing the system performance and stability. Thus TFC presents a flexible, expert knowledge based, robust and model free control.

Appendix

Voltage	150 V
Max current	30 A
Stator diameter	143 mm
Rotor diameter	69 mm
Air gap	0.4 mm
Stack length	143 mm
Stator tooth arc	0.328 rad
Rotor tooth arc	0.427 rad
Stator yoke thickness	12.1 mm
Rotor yoke thickness	9 mm
Stator tooth height	24.5 mm
Shaft diameter	26 mm
Coil turns	180

References

1. Miller TJE (2001) Electronic control of switched reluctance motor, Newnes
2. Krishnan R (2001) Switched reluctance motor drives, modelling, simulation, analysis, design and applications, CRC Press, New York
3. Rodrigues MG, Suemitsu WI, Branco P, Dente JA, Rolim LGB (1997) "Fuzzy logic control of a switched reluctance motor." In: Proceedings of the IEEE international symposium on industrial electronics, vol. 2. pp 527–531
4. Wadnerkar VS, Bhaskar MM, Das TR, RajKumar AD (2010) "A new fuzzy logic based modeling and simulation of a switched reluctance motor". J Electr Eng Technol 5(2):276281
5. Chancharoensook P, Rahman MF (2002) "Dynamic modeling of a four phase 8/6 switched reluctance motor using current and torque look-up tables," IEEE transactions industrial applications, pp 491–496

6. Chowdhuri S, Biswas SK, Mukherjee A (2006) "Performance studies of fuzzy logic based PI-like controller designed for speed control of switched reluctance motor." In: Proceedings of IEEE international conference industrial electronics application, pp 1–5
7. Soares FS, Costa Branco PJ (2001) "Simulation of a 6/4 switched reluctance motor based on matlab/simulink environment," IEEE transactions on aerospace and electronic systems, vol. 37, no. 3, July 2001
8. Wang S-C, Liu Y-H (2011) A modified pi-like fuzzy logic controller for switched reluctance motor drives. IEEE Trans Ind Electr 58(5):1812–1825
9. Srinivas P, Prasad PVN (2010) "Torque ripple minimization of 8/6 switched reluctance motor with fuzzy logic controller for constant dwell angles", international conf on power electronics, drives and energy systems (PEDES) & 2010 Power India, pp 1–6