Volume 9 • Issue 3 • May / June 2015

Editor-in-Chief
Dr. Chen-Chi Shing

INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND SECURITY (IJCSS)

ISSN : 1985-1553

Publication Frequency: 6 Issues / Year

CSC PUBLISHERS
http://www.cscjournals.org

INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND SECURITY (IJCSS)

VOLUME 9, ISSUE 3, 2015

EDITED BY
DR. NABEEL TAHIR

ISSN (Online): 1985-1553

International Journal of Computer Science and Security is published both in traditional paper form and in Internet. This journal is published at the website http://www.cscjournals.org, maintained by Computer Science Journals (CSC Journals), Malaysia.

IJCSS Journal is a part of CSC Publishers
Computer Science Journals
http://www.cscjournals.org

INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND SECURITY (IJCSS)

Book: Volume 9, Issue 3, May / June 2015
Publishing Date: 30-06-2015
ISSN (Online): 1985 -1553

IJCSS Journal is a part of CSC Publishers
http://www.cscjournals.org

Typesetting: Camera-ready by author, data conversation by CSC Publishing Services – CSC Journals, Malaysia

CSC Publishers, 2015

EDITORIAL PREFACE

This is *Third* Issue of Volume *Nine* of the International Journal of Computer Science and Security (IJCSS). IJCSS is an International refereed journal for publication of current research in computer science and computer security technologies. IJCSS publishes research papers dealing primarily with the technological aspects of computer science in general and computer security in particular. Publications of IJCSS are beneficial for researchers, academics, scholars, advanced students, practitioners, and those seeking an update on current experience, state of the art research theories and future prospects in relation to computer science in general but specific to computer security studies. Some important topics cover by IJCSS are databases, electronic commerce, multimedia, bioinformatics, signal processing, image processing, access control, computer security, cryptography, communications and data security, etc.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Started with Volume 9, 2015, IJCSS appears with more focused issues. Besides normal publications, IJCSS intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

This journal publishes new dissertations and state of the art research to target its readership that not only includes researchers, industrialists and scientist but also advanced students and practitioners. The aim of IJCSS is to publish research which is not only technically proficient, but contains innovation or information for our international readers. In order to position IJCSS as one of the top International journal in computer science and security, a group of highly valuable and senior International scholars are serving its Editorial Board who ensures that each issue must publish qualitative research articles from International research communities relevant to Computer science and security fields.

IJCSS editors understand that how much it is important for authors and researchers to have their work published with a minimum delay after submission of their papers. They also strongly believe that the direct communication between the editors and authors are important for the welfare, quality and wellbeing of the Journal and its readers. Therefore, all activities from paper submission to paper publication are controlled through electronic systems that include electronic submission, editorial panel and review system that ensures rapid decision with least delays in the publication processes.

To build its international reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJCSS. We would like to remind you that the success of our journal depends directly on the number of quality articles submitted for review. Accordingly, we would like to request your participation by submitting quality manuscripts for review and encouraging your colleagues to submit quality manuscripts for review. One of the great benefits we can provide to our prospective authors is the mentoring nature of our review process. IJCSS provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

Editorial Board Members
International Journal of Computer Science and Security (IJCSS)

EDITORIAL BOARD

Assistant Professor Vishal Bharti
Maharishi Dayanand University
India

Dr. Parvinder Singh
University of Sc. & Tech
India

Assistant Professor Vishal Bharti
Maharishi Dayanand University,
India

TABLE OF CONTENTS

Volume 9, Issue 3, May / June 2015

Pages

Computer-Aided Disaster Recovery Planning Tools (CADRP)

Omar H. Alhazmi *ohhazmi@taibahu.edu.sa*
Department of Computer Science
Taibah University
Medina, Saudi Arabia

Abstract

Information Technology Disaster Recovery Plans (DRPs) are becoming an essential component for any organization with IT infrastructure. However, DRPs varies in performance and cost; therefore, based on requirements and resources, an organization can design their DRP. Typically, DRPs depends on data and/or system replication, data needs to be backed up frequently, and a plan to restore the system to running state within the allowed time. Hence, DRP designer must know the needed business requirements in terms of recovery time objective (RTO) and recovery point objective (RPO). Then, the appropriate technical requirements will be set. At the same time, the cost factor can play a role in choosing the appropriate DRP. The industry has a widely accepted seven-tier system of how DRP can be designed. In this work, we design and implement a software tool that can simulate the IT DPR systems and therefore help designers to design, optimize, and test their design before it is physically implemented. This tool will run a simulated system with DRP specific design and the designer can exercise with the system to show it's RTO, RPO, and cost that can significantly improve DRP design.

Keywords: Disaster Recovery, Business Continuity, RTO, RPO, Simulation.

1. INTRODUCTION

The disaster recovery plan (DRP): "is a documented process or set of procedures to recover and protect a business IT infrastructure in the event of a disaster", [1]. Therefore, it contains manual procedures usually performed by IT professionals and automated procedures performed by the IT system. Of course, minimizing manual procedures and maximizing automated procedures will reduce recovery time an important factor referred to as Recovery Time Objective (RTO). Moreover, data and system backups are typical part of any disaster recovery plan; more frequent backups will improve another factor which is Recovery Point Objective (RPO) which can also mean lost data.

The lower RPO and RTO, the better the disaster recovery plan; however, the cost also goes up. Therefore, some organizations go for DRP for critical systems and another DRP for non-business essential systems.

In this work, we introduce a Computer Aided Disaster Recovery Planning tools (CADRP) that will:

1) Design and test: Help disaster recovery specialists test and design different plans and to be able to compare them using safe simulation environment.

2) Choose, compare and optimize: Help CIOs and disaster recovery specialists choose among different alternatives which vary in RPO, RTO and Cost; moreover, and also DRP engineer with all these choices and different technologies available in the market, especially cloud services and choose the right solution wither DRaaS or Platform as a service (PaaS) or system as a service (SaaS).

3) Research and develop: to help researchers from industry and academia to conduct research affordably on CADRP platform in order to conduct quality research by using it as a virtual lab.

CADRP is a software tool that will take the environment parameters and allow the DRP specialist or administrator with some intermediate expertise in disaster recovery planning to build a virtual system and a virtual disaster recovery system with some specific given specifications and scenario. Then, CADRP will analyze system's components and calculate some statistics about the system. Moreover, CADRP will run a simulation of a system along with its disaster recovery system and at a certain point the system will assume that a disaster has struck and the original system will stop while the disaster recovery system will run and a "dynamic" analysis will be performed to calculate some metrics about the system including the critical factors of RPO and RTO. Finally, CADRP will produce a detailed report about the DRP plan.

The CADRP system will also consider systems with various tier levels 1-7 which can also make it a valuable research tool for researchers interested to work in this area, it also support cloud DR solutions.

In section 2, we will preview related work, disaster recovery tiers scheme, and disaster recovery cost analysis; next, in section 3, we preview CADRP in details; later in section 4, we will test the system; finally section 5 will discuss the conclusion and point to some future research.

2. DISASTER RECOVERY PLAN

In this section we will preview two aspects of DRP, the performance in DRP tier and the DRP cost. The choice of disaster recovery tier will have direct impact on cost. Some equations will help to test if DRP cost should be justified financially or not.

2.1 Related Work

Before the 1990s several disaster recovery solutions existed, however, they varied in their sophistication, cost and performance; therefore, by 1990 a need to categorize these solutions became necessary. Hence, a tier system of DRPs where established by IBM [5] and later over the past decades others also suggested different way of classification like Novell 4-tier system [6], Hitachi's system [7] also Webornatr [8] and Xiotech [9]. These schemes have similarities and differences; however, IBM's is the oldest and has got some acceptance in the industry. Therefore, when we designed the CADRP tools we considered IBM's to be the main reference. In addition to that, new emerging technologies being introduced to disaster recovery challenging classical DRPs tier schemes [4], such as Disaster Recovery as a Service (DRaaS) provided by major cloud service providers will also be considered by CADRP.

In searching for a research about disaster recovery simulation tools, we have found SYMIAN by Bartolini et. Al. which is a discrete event simulator, it basically simulates an incident happening to a particular application of a system to consider corrective measures [10], we have not came across any other tool that shall serve the purpose we are aiming at.

2.2 Disaster Recovery Plans Tiers

Table 1 below shows Share/IBM scheme, it is simple and yet flexible; this can explain its popularity over other schemes. Table 1, briefly explains about each disaster recovery tiers; we can ignore tier 0 which means that there is no DRP at all. Tier 1, is simple with minimum cost and can be ideal solution cost-wise for data of small and non-critical nature, it what can also be done at the personal level when backing up mobile phones or photo albums. Tier 2, has a little more readiness for recovery, a stand by system that needs to be built back with all necessary software installations, configuration and data restoration; therefore, RTO would be ranging typically from hours to days, while RPO will heavily depending on the frequency of backup (which is manually done at this tier).

Starting from tier 3, will be having a disaster recovery with more predictable RTO and RPO, automated backups at this level allows more frequent backups this improved RPO, also RTO improves with less manual work done. So, as we go to tier 4, 5 and 6. RTO and RPO improve as more frequent backups and more automation of the DRP are done at these tiers, at the same the cost gets higher. Moreover, at tier 7, is concerned with having the system mirrored with fully operational disaster recovery site. In case of a disaster or disruption, the disaster recovery site replaces the original site automatically; this significantly reduces the impact of human factor which usually causes significant delays in the recovery process from the prospective of RTO and RPO.

Disaster recovery plans traditionally fall in one of the seven tiers on IBM's7-tiers system (see Table 1), [4]. In these 7-tiers system RPO and RTO get lower (i.e. improves) as we go up in tier number.

Tier	Technology	Description
0	No off-site data	No saved information, no recovery plan at all
1	Data backup with no hot site	Data are packed up and taken to a remote location for storage, also called PTAM; *the "Pick-up Truck Access Method."*
2	Data backup with a hot site	Same as tier 1; however, the remote site has ready infrastructure capable of restoring operation to the latest backup within hours/days
3	Electronic vaulting	Same as tier 2; however, backups are done via electronic vaulting, and high speed communication (no PTAM)
4	Point-in-time copies	Same as tier3; however, data are backed up more frequently; thus, better estimation of data loss and recovery time.
5	Transaction integrity	This application level tier ensures that original site and backup site are consistent; thus, minimizing loss to zero or near zero level.
6	Zero or little data loss	This tier requires site mirroring, two sites working in sync
7	Highly automated, business integrated solution	Same as tier 6; plus the recovery process is automated; therefore, the system will recover itself with no or minimum intervention.

TABLE 1: Share / IBM Disaster Recovery Tiers.

2.3 Disaster Recovery Plans Cost

For these DRP plans there are different kinds of cost which are:

- The initial cost (Ci) which is the cost to establish the DRP
- The Ongoing cost (Co) the operational overhead including human resources, hardware and software
- The cost of a disaster (Cd) the cost incurred by the incident
- The annual cost (CT) which is the annual cost of DRP

Here we overview some of the equations used in calculating the cost and study the feasibility of the DRP. The total annual system cost *(C_T)* is the sum of the: Initial cost *(C_i)*, ongoing cost *(C_o)* and Cost of disaster *(C_d)*, therefore, [4]:

$$C_T = C_i + C_o + C_d \qquad (1)$$

On the other hand, the total cost caused by disasters *(C_d)*, is affected by the cost of an incident *(C_λ)* and the probability of a disaster happening *(P_d)*, [1]:

$$C_d = C_\lambda * P_d \qquad (2)$$

DRP	C_i	C_o	C_d	$\mathbf{C_T}$	C_λ	P_d	$\mathbf{C_d}$	$\mathbf{C_d \geq C_T}$
1	5k	200	1k	**6.2k**	100k	.01	**1k**	**No**
2	19k	800	200	**21.2k**	1m	.02	**20k**	**Yes**
3	100k	5k	0	**106k**	2m	.03	**60k**	**Yes**

TABLE 2: Examples of Different DRPs for Different Systems.

In order for a disaster recovery planner to decide if a certain DRP is cost effective the equation below should be true, [2]:

$$C_d \geq C_T \qquad (3)$$

To clarify this, let's take those examples shown in Table 2 above. In case 1, the system will cost 6,200$ while the disaster will cost 1,000$, therefore the chosen DRP is too expensive for the system. In case 2, the system will cost about the same as disaster. On the other hand, case 3 show that the chosen DRP will save the organization 102k-60k = 42k.

3. COMPUTER AIDED DISASTER RECOVERY PLANNING

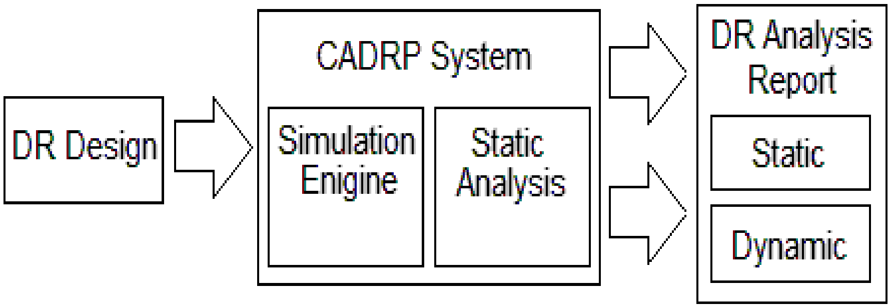


FIGURE 1: CADRP System's Outline.

Computer aided Disaster Recovery Planning (CADRP) system should accommodate disaster recovery system design ranging from the lowest tier (0) to the highest (7) on IBM tier system. CADRP should provide visual drag and drop interface. Then the system should be analyzed statically and dynamically (simulation). Figure 1 above shows an outline of the CADRP system.

3.1 CADRP Design Overview

In this part the DRP designer should design the original system and the DRP system and set the environment factors (See Figure 2): the recovery system may be absent (tier 0), or it can be a memory card, a hard disk, a tape like in lower tiers 1 to 3 or a server like in higher tiers 4 to 7; furthermore, a cloud server or storage can also be selected. Moreover, some data must be entered in order for CADRP system to analyze the DRP and generate correct reports (see Figure 3), these data is mainly about the environment to determine the weight of some factors, and this will help for the feasibility analysis and calculating RTO.

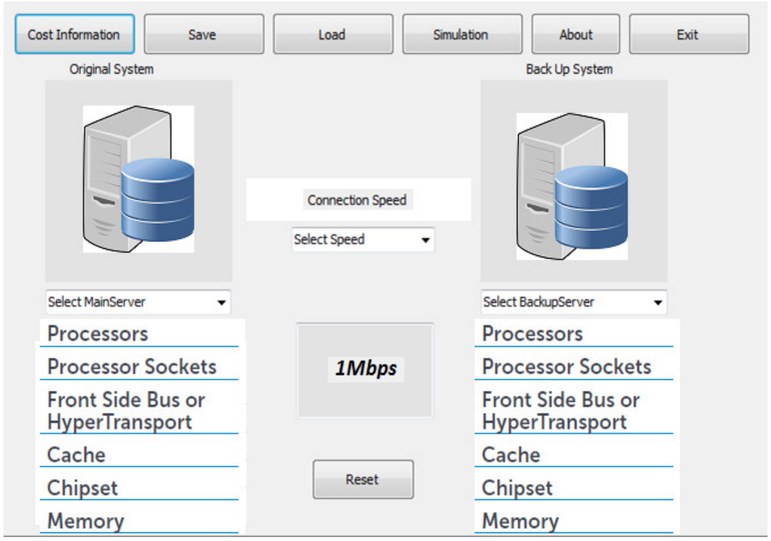
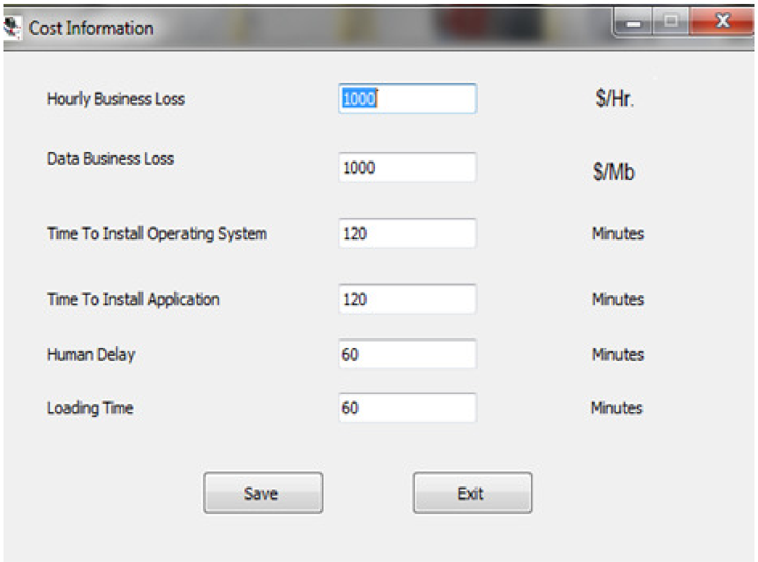


FIGURE 2: CADRP Design Screen.


Cost Information

Hourly Business Loss	1000	$/Hr.
Data Business Loss	1000	$/Mb
Time To Install Operating System	120	Minutes
Time To Install Application	120	Minutes
Human Delay	60	Minutes
Loading Time	60	Minutes

FIGURE 3: Some Data Collected by CADRP.

3.2 Simulation Engine

This module will be responsible for simulating two systems, the original system and the disaster recovery system. After the system is designed, and also the appropriate parameters entered; this module will run a hypothetical application that will run in cycles, at each cycle the original system will keep processing the current transaction depending on its CPU speed, if the transaction is fully processed then the system will process the next transaction and the old transaction will be sent to the disaster recovery back up system (it can be tape, disk, server or virtual server. So, depending on the connection speed from the original site to the backup site sometimes there is some delay; in addition to that there is a speed also for tape, disk or server to process or store the coming transactions. There is one important assumption that must be made, is to have a *sync DR* or *async DR*, each one has advantages and disadvantages, as in sync systems both systems must be in the same transaction, so the slower of the two systems will slow the other, while the async let the DR system work on its own pace, without causing the original system to wait (see Figure

4). On the other hand the sync DR preserves the integrity of the transaction as it will not move to the next one until it is processed and stored on both systems.

```
int Simulate()
{
long i=0; //transaction processed at the original system
long j=0; // transaction being transferred
long k=0; // transaction processed at the DR system
long cycle =0;
do{
  cycle++;
  Process (Orig_system, transaction[i]);
  // process portion of the transaction in this cycle
  If processed (Orig_system,transaction[i]) {i++,j=i};
  // if the transaction is completed, transfer it
  Transfer (Speed,transaction[j]);
  //transfer portion of the transaction during the current cycle
  If transferred (transaction [j]){ j++,k=j};
  // if transaction is transferred process it at DR
  Process (DR_system, transaction [k]);
  If processed (DR_system,transaction [k]){ j++,k=j};
  // if transaction is transferred process it at DR
  If (disaster_triggered) disaster =1;
  } while (disaster == 0)
RPO= k-i; // number of lost transactions
Return (RPO)
}
```

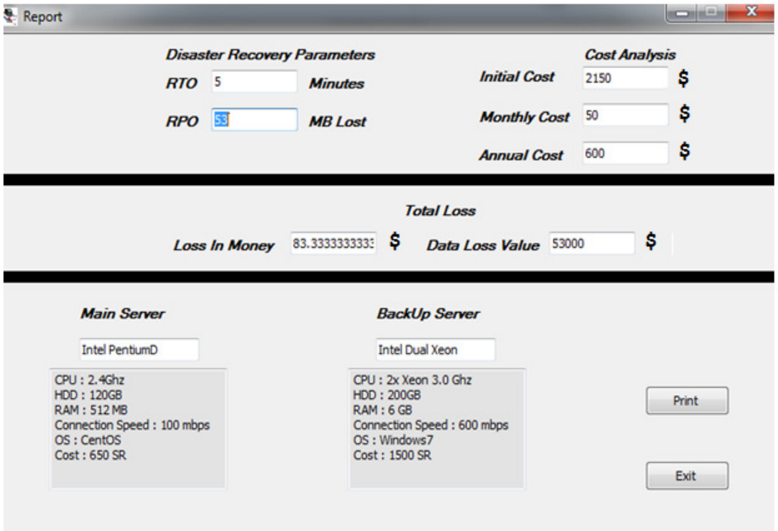
FIGURE 4: Basic Algorithms for Disaster Recovery Simulation Engine.

3.3 Static Analysis Module

In this part CADRP will calculate the parameters which they do not go through the simulation, including cost, ongoing cost, storage size, RTO. The static provided system can also help developers by estimating the optimal cost of a certain DRP solution.

3.4 Data Analysis Report

The generated report would take the format shown in Figure 5, so the disaster recovery planner would have results coming simulated system


Report

Disaster Recovery Parameters
RTO 5 Minutes
RPO 53 MB Lost

Cost Analysis
Initial Cost 2150 $
Monthly Cost 50 $
Annual Cost 600 $

Total Loss
Loss In Money 83.333333333: $
Data Loss Value 53000 $

Main Server
Intel PentiumD
CPU : 2.4Ghz
HDD : 120GB
RAM : 512 MB
Connection Speed : 100 mbps
OS : CentOS
Cost : 650 SR

BackUp Server
Intel Dual Xeon
CPU : 2x Xeon 3.0 Ghz
HDD : 200GB
RAM : 6 GB
Connection Speed : 600 mbps
OS : Windows7
Cost : 1500 SR

FIGURE 5: The Final Generated Report.

4. TESTING THE SYSTEM

We will evaluate RPO and RTO eight cases. By testing each case on the tool (see Figure 5) and then running the simulation, we will see how the RPO changes when giving different solution and parameters. The cases tested here are:

- Tape backup : cases 1 and 2, (tier 1, Table 1)
- Hard disk back: cases 3 and 4 (tier 3, Table 1)
- Mirrored servers; cases 5 and 6 (tier 6, Table 1)

- Cloud servers; cases 7 and 8 (tier 6, Table 1)

Here we assumed that each transaction has a fixed size of 1 Megabyte, we find that when the original system completed 10,000 transaction the backup system has just finished storing the 9970th transaction, this when the hypothetical disaster happens and therefore we have lost the last 30 transactions (about 30 Megabyte of data); so depending on criticality of the data this can be translated in loss of money.

Case	Disaster Recovery System	Network bandwidth (mbps)	Number of transactions before disaster	Lost transactions	Lost data(mb)	RTO(minutes)
1	Tape	.25	10000	30	30	240
2	Tape	.5	10000	21	21	240
3	Disk	.5	10000	24	24	120
4	Disk	1	10000	8	8	120
5	Server	10	10000	2	2	0-5
6	Server	100	10000	1	1	0-5
7	Virtual Server	0.5	10000	20	20	0-5
8	Virtual Server	100	10000	1	1	0-5

TABLE 3: Simulation Results.

In cases 5 to 8 we can notice the improvement in performance by significant reduction in RPO; however, connection speed is a main bottleneck when we tried low connection speed we have lost 20 transactions. In this simulation we can see which factor has more impact on RPO, we can see that CPU is important; however, the network bandwidth might be the main factor.

In Table 3 above, we can also see the RTO. When calculating an RTO we assumed that there is a 60 minutes operating system restore and a 60 minutes configuration and 120 minutes of system and data restoration. Therefore, the total for worst case RTO is 240 minutes; here, we ignored any impact of absence of system admins during the disaster, most of the time there is logistic delays of traffic and other factor that can add hours or days to manual system restoration. On the other hand, for automated mirrored systems recovery solution the backup server the RTO can be within few minutes.

5. CONCLUSION AND FUTURE WORK

In this work, a computer aided disaster recovery planning tool was presented to be used in practice and in research. Basically, this tool will help to design a recovery plan and also to compare different disaster recovery plans in order to find an ‘optimal DRP’, which will shall be effective cost-wise and performance wise and can help show the trade-off between RPO, RTO and Cost.

The tool will run a simulation to produce a report showing RPO, RTO, Systems cost and disaster cost, it will also analyze all entered data such as cost per lost megabyte and also cost of losing business time and use it to produce a helpful report.

One limitation that needs to be worked on is that RTO is calculated using some preset parameters and the analysis is straight forward of applying some formulas.

Suggested Future work would be to add more resolution and functionality to CADRP; for example, to take policy and procedures into account and to include more external environment factors. Also the tool can be extended to support more than one disaster recovery sites. In future, we plan to incorporate this tool in system courses to teach students about disaster recovery plans. Moreover, the system can also incorporate COBIT and ITIL business continuity maturity levels into CADRP’s next update [11].

6. REFERENCES

[1] Abram, Bill (14 June 2012). "5 Tips to Build an Effective Disaster Recovery Plan". Small Business Computing. http://www.smallbusinesscomputing.com/News/ITManagement/5-tips-to-build-an-effective-disaster-recovery-plan.html.

[2] Lars Albrecht, Bernd Baier, Designing a bullet-proof Disaster Recovery Architecture for business-critical Applications, White paper, http://www.libelle.com/fileadmin/Public/Whitepaper/WhitePaper_Bullet-Proof_DRArchitecture.pdf.

[3] O. H. Alhazmi and Y.K. Malaiya, “Evaluating Disaster Recovery Plans Using the Cloud”, Proc. Reliability and Maintainability Symposium (RAMS 2013), Orlando, January 2013, pp. 37-42.

[4] O.H. Alhazmi and Y.K. Malaiya, “Are the Classical Disaster Recovery Tiers Still Applicable Today?”, Proc. 25thIEEE Int. Symposium on Software Reliability Engineering Workshop, Nov. 2014.

[5] Robert Kern, Victor Peltz, “Disaster Recovery Levels”, IBM Systems Magazine, November 2003.

[6] Novell, “Consolidated Disaster Recovery”, http://www.novell.com/docrep/2009/03/Consolidated_Disaster_Recovery_White_Paper_en.pdf, March 2009.

[7] Roselinda R. Schulman, Disaster Recovery Issues and Solutions, A White Paper, Hitachi Data Systems, September 2004.

[8] Montri Wiboonratr and Kitti Kosavisutte, “Optimal strategic decision for disaster recovery,” Int. Journal of Management Science and Engineering Management, V ol. 4 (2009) No. 4, pp. 260-269.

[9] XiotechCorporation ,Tiered Data Protection and Recovery, , May 2006.

[10] Claudio Bartolini, Cesare Stefanelli, Mauro Tortonesi, SYMIAN: A Simulation Tool for the Optimization of the IT Incident Management Process, Lecture Notes in Computer Science Volume 5273, 2008, pp 83-94.

[11] Melita Kozina,COBIT - ITIL mapping for Business Process Continuity Management, Proceedings of the 20th Central European Conference on Information and Intelligent Systems,pp113-119,2009.

Digital Image Watermarking Techniques: A Review

Pushpa Mala .S. *pushpasiddaraju@gmail.com*
Research Scholar, Jain University, Bengaluru, India
& Sambhram Institute of Technology, Bengaluru, India

D. Jayadevappa *devappa.22@gmail.com*
Dept. of Electronics & Instrumentation Engineering,
JSS Academy of Technical Education, Bengaluru, India

K.Ezhilarasan *murali981983@gmail.com*
Research Scholar, Jain University, Bengaluru, India
& Sambhram Institute of Technology, Bengaluru, India

Abstract

Advancements in science and technology have introduced the need to protect data, authenticate data, integrate data, assert ownership, content labelling and security. Digital Watermarking schemes protect all forms of digital data. Digital Image Watermarking can be applied to gray scale, halftone, color, medical and 3D images. The process of watermarking can be broadly classified into three phases namely embedding, attacking, and decoding for typical scenarios. Some of the watermarking schemes adopted in the past include vector quantization, spread spectrum, SVD, DCT, DFT, etc. It was observed that the spread spectrum was more robust and it had also been applied for patenting. In spite of this, the method could not withstand high amplitude noise. Hence, later DCT, DFT and Wavelets were used. These schemes were not robust to collusion attacks. In this review, we have identified the embedding and detection schemes of the existing watermarks over the past decade and analyzed the robustness of each of these methods. The different parameters considered to analyze the performance of the existing watermarking schemes are also discussed. Research under watermarking is a great field of interest involving multimedia security, forensics, data authentication and digital rights protection. This paper will be useful for researchers to implement a robust watermarking scheme.

Keywords: Digital Image Watermarking, Watermark Embedding, Watermark Detection, frequency domain, robust, reversible

1. INTRODUCTION

In this modern era, a tremondous growth in science and technology is noticed. This has led to a large number of e-commerce sites and applications. Intellectual property protection, data authentication, ownership and security are of great concern to the owners of a document. Every document includes digital information in some form or the other. Some of the information included may be pictures, others video etc. There is a need to protect these information from hackers. It is known that the hacker is always one step ahead of the creator. Cryptography and Stegenography are such scemes used where the former process the message and the later coceals the existence. These methods are not widely used since they are either less robust or partially robust to digital data modifications. Digital Watermarking was developed to achieve better robustness. A Watermark is a design impressed during creation and is used for copyright protection, data authentication, identifying the source, creator, owner, or authorised consumer of the document or image. It is also used to identify a document or an image that is modified or illegally distributed.

The general characteristics of Digital Image watermarking schemes are robustness, imperceptibility, capacity and security. Most of these characteristics are contradictory and make a trade off to achieve robustness. The life cycle of a watermarking process includes embedding, attacking and extraction. In the course of literature studies, it is noted that several schemes to embed, detect and recover the watermark exist. These schemes adopted additive and multiplicative approaches in time domain. Different algorithms were developed adopting SS(Spread Spectrum)[6], DCT(Discrete Cosine Transform)[2], DFT(Discrete Fourier Transform), DWT(Discrete Wavelet Transform)[10], SVD(Singular Value Decomposition)[13], Ridgelets[19] and Contourlets[56] in frequency domain. Each scheme was evaluated for its performance by using image quality metrics such as Scaling, Cropping, AWGN (Additive White Gaussian Noise), PSNR (Peak Signal to Noise Ratio), SSIM (Structural Similarity Index Measure), Compression, Wrapping and Histogram Equalization. Although several such schemes performed well during evaluation there is still a need to develop more and more robust schemes due the rapid growth of the web technologies. In due course, it is also noted that some of the watermarking schemes can also be developed into protocols.

A watermark could be designed either for the source or for the destination. Watermarking at the source reduces piracy. Although most of them are designed for the source, some watermarking schemes are based for the destination instead of the source. While designing such watermarking schemes the features that were considered were transparency, robustness and capacity. Transparency here refers to the fact that the watermark be made visible in the image. It was also necessary to consider that the watermark stay unaltered when the imaged was illicitly tampered. As the number of watermarks increased, capacity could be defined as the ability to detect the watermarks with a low probability of error. The earliest techniques used were placing a watermark in the least significant bits or in the high frequency components. Such watermarks could be destroyed with simple quantization or low pass filtering. This process would sometimes affect or degrade the image quality.

2. FEATURES OF DIGITAL WATERMARKING SCHEMES

The following are some important features that a Digital Watermark exhibit.

1) **Robustness**- The watermark embedded must be robust refers to the capability of the watermark to survive a large number of signal processing operations, intentional and unintentional attacks.

2) **Imperceptibility**- Imperceptibility refers to the watermark being invisible to the HVS [Human Visual system]. This feature plays a vital role in content authentication.

3) **Security**- The watermark must be secure such that the hacker cannot remove the watermark. This can be accomplished by developing sophisticated algorithms and hence the watermark remains accessible only to authorized person.

4) **Verifiability**- The Watermark must be capable in determining ownership information.

5) **Computational Cost**- Computational simplicity is a feature that must be adopted so that the computational cost reduces.

6) **Watermark Detection**- Identifies the successful detection of the watermark at the detector end.

7) **Capacity**- Capacity describes how many information bits can be embedded. It can also be described as the possibility of embedding multiple watermarks.

8) **Tradeoff parameters**- There always exists a tradeoff between robustness, imperceptibility and capacity (Figure.1). Any watermarking scheme should be capable to overcome these tradeoffs.

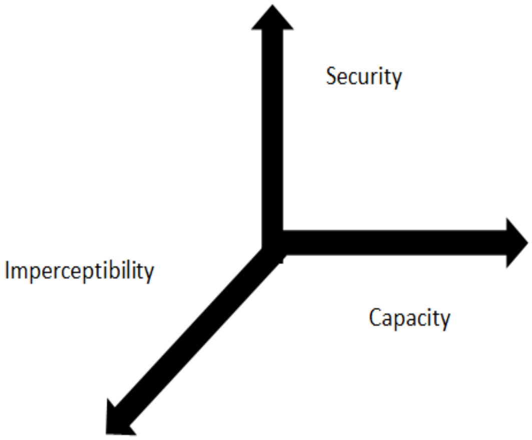


FIGURE 1: Tradeoffs in Digital Image Watermarking.

3. DIGITAL IMAGE WATERMARKING APPROACHES

In this section an attempt is made to classify watermarking schemes according to the characteristics of the embedded watermark.

3.1. Visible and Invisible Watermarking

Prior to 1998, watermarking schemes were widely classified as visible and invisible. A visible (perceptible) watermark is one which is visible to the HVS (Human Visual System). Visible watermarking scheme is that wherein a secondary image (watermark) is embedded into the primary image (host), such that the watermark is visible to the human eye. Huang and Tang[40] proposed a digital visible watermarking scheme wherein the intensity of the watermark was varied in the different regions of the image depending on the underlying content of the image and visual sensitivity to spatial frequencies. The proposed scheme modified the DWT coefficients of the host image. These schemes require large bit rates and strong strength than invisible watermarking schemes.

An invisible(imperceptible) watermark is one which is not visible to the HVS. Tzeng et al. [15] proposed a watermarking scheme which used optimization techniques to embed invisible watermarks in an image. It was assumed that any attack creates an invisible modification in the image. The proposed scheme was both blind and nonblind. The watermarked image was subjected to several blind attacks using programs like Photoshop Version 6, Photo impact Version 5 and malicious attacks like spreading noise, copy attacks and distortion. Invertibility and quasi-invertibility are such properties of invisible watermarking schemes for resolving the rightful ownership of the image. An invertible watermarking scheme is one that is susceptible to an attack that creates multiple claims of ownership for the same watermarked content. Noninvertible watermarking schemes, and subsequently the examples of such schemes are believed to be nonquasi-invertible as proposed by Craver *et.al* [1]. Wong and Memom [55] proposed a public watermarking scheme that could be made either visible or invisible based on the users choice for uncompressed images in spatial domain. This scheme could be further extended to compressed images in frequency domain too.

3. 2. Private and Public Watermarking

A private watermark can be detected by only authorized users i.e. normally the owner has private control over the watermark. Wang *et al.*[16] proposed an scheme in which they claimed that the owner had a private control over the watermark. Images were represented in vector form than in pixel format. Randomly generated orthonormal filter banks were used as private key. The watermark was embedded into the mid-frequency range, was invisible and robust to compression. The watermark was an image of real numbers obtained from wavelet coefficients. An algorithm was developed to scramble the watermark so that it differed from image to image. Piva *et al.* [37]

also proposed a public watermarking scheme, which did not need access to the original data for extracting the watermark.

3.3. Robust and Fragile Watermarking

Robust watermarks are generally used for copyright protection and ownership verification whereas fragile or semi-fragile watermarks are used in content authentication and integrity verification. Izquierdo and Guerra [21] proposed a block wise fragile watermarking scheme. Block wise watermarking was performed by partitioning the $N \times N$ image into small blocks A^{k}, where k = 1, 2…*l,* each of size $P \times Q$. The watermarked block was defined as À. The authentication procedure detected changes in À. The receiver of the image knew the secret keys to determine whether the block was authentic or fake. Zhao *et al.*[30] proposed a semi-fragile multipurpose watermarking scheme for cultural heritage imagery using DCT-DWT dual domain algorithm. The embedding process took place in Haar DWT domain and was generated in DCT domain. There was a trade-off in being robust to content preservation and the scheme was fragile towards malicious attacks. As will be discussed in the later sections, most of the watermarking schemes are robust.

3.4. Blind and Non-blind Watermarking

Blind watermarking schemes detect the watermark without accessing the original image. They require only the secret keys and are also public in nature. Nonblind watermarking schemes access the host image during the detection phase and are private in nature. Wong *et al.* [22] proposed multiple blind watermarking schemes for images namely SWE (Single Watermark Embedding), MWE (Multiple Watermark Embedding) and IWE (Iterative Watermark Embedding). SWE used two correlated secret keys to embed the watermark bit sequences into the image, MWE embedded multiple watermarks and minimized distortion energy, IWE embedded watermarks into JPEG compressed images. Another DCT based watermarking scheme was proposed by Chu [23] wherein the watermark was inserted in the sub images obtained by sub sampling. Lin *et al.* [45] also proposed a blind watermarking scheme for copyright protection. This method was based on the difference of the maximum wavelet coefficient and the second maximum wavelet coefficient to embed the watermark. Extraction of the watermark was done by designing a threshold value. It was noted by Meerwald *et al.* [50] that the above scheme had neglected security under intentional attack exploiting knowledge of implementation. Although modifications to the quantized coefficients location was adopted, the later too was prone to targeted attacks.

3.5 Spatial and Frequency Domain Watermarking

In Spatial domain, the watermark can be inserted into the host image by altering the pixel values of the host image while in frequency domain, the watermark is inserted into the coefficients of the transformed host image. Modasseri and Berger [39] have proposed an algorithm that was directly applied to the bit streams in compressed domain. The watermark was embedded into the bit stream as forced bit errors. Coltuc and Chassery [43] proposed another spatial domain watermarking scheme adopting reversible contrast mapping-an integer transform that applies to a pair of pixels. The watermark was embedded into the LSB's. This scheme seemed to be the lowest in computational capacity and was appropriate for real time applications. The approaches discussed later adopting DCT, DFT, DWT, ridgelets and contourlets are some frequency domain approaches.

4. FREQUENCY DOMAIN WATERMARKING APPROACHES

Watermarking algorithms have a long history dating prior to the 1990's. In this paper, the scope of review is limited to frequency domain techniques from the 1990's. A watermarking scheme suitable for one domain may not work well for the other. For example, fragile/semifragile watermarks are usually used for image authentication to verify whether the received image was modified during transmission or not, while robustness refers to content preservation.

Most of the work on robust digital watermarking was based on SS techniques. SS refers to a technique of transmitting a narrow bandwidth signal over a larger bandwidth. The signal energy present over a wide range of frequencies was undetectable. During digital watermarking, the watermark was spread over a range of frequencies such that the energy in a single frequency range was so small that the watermark was not detected quiet easily. Such watermarks could be destroyed if high amplitude noise was added to all the frequency ranges. Applying frequency transformations to the data, significant regions of the spectrum could be highlighted. Any unintentional effects must not alter the significant regions of the spectrum else the image gets degraded. In order to place an n-length watermark into an $N \times N$ image, the $N \times N$ DCT of the image was computed and watermark was placed into the n highest magnitude coefficients (which are data dependent) of the DCT transformed image. Cox *et al.*[6] proposed a DCT-based spread spectrum watermarking technique. The watermark was developed from Gaussian distribution with zero mean and unit variance. The watermark was spread over all the frequencies and was not detectable since the energy in a single frequency was very small. This method was very popular and many early researchers adopted it. Altun *et al.* [47] considered optimal formulations of spread spectrum watermarking and proposed an algorithm for optimal embedding of the watermark that combined projections onto convex sets with a bisection parameter to determine the optimum watermarked image. They demonstrated optimal watermarking schemes to maximize its robustness to additive noise, compression, distortion minimizing the visibility of the watermark.

Podilchuk and Wenjun Zeng [2] have proposed two schemes based on a block- DCT framework where the typical block size for the DCT is 8×8. The two schemes are 1) a block-DCT scheme which has the advantage of direct watermark encoding of JPEG bit streams and 2) a wavelet-based scheme. The watermark encoder for the IA-DCT (Image Adaptive Discrete Cosine Transform) scheme is described by

$$X^*_{u,v,b} = \begin{cases} X_{u,v,b} + t^C_{u,v,b} w_{u,v,b} & \text{if } X_{u,v,b} > t^C_{u,v,b} \\ X_{u,v,b} & \text{otherwise} \end{cases} \quad (1)$$

where $X_{u,v,b}$ refers to the DCT coefficients, $X^*_{u,v,b}$ refers to the watermarked DCT coefficients, $w_{u,v,b}$ is the sequence of watermark values, and $t^C_{u,v,b}$ is the computed JND(Just Noticeable Difference) calculated from the visual model described by Watson [3].
The watermark insertion for IA-W (Image Adaptive Wavelet) is described by

$$X^*_{u,v,l,f} = \begin{cases} X_{u,v,l,f} + t^F_{l,f} w_{u,v,l,f} & \text{if } X_{u,v,l,f} > t^F_{l,f} \\ X_{u,v,l,f} & \text{otherwise} \end{cases} \quad (2)$$

where $X_{u,v,l,f}$ refers to the wavelet coefficient at position (u,v) in resolution level l and frequency orientation f, $X^*_{u,v,l,f}$ refers to the watermarked wavelet coefficient, $w_{u,v,l,f}$ is the watermark sequence, and $t^F_{l,f}$ corresponds to the computed frequency weight at level l and frequency orientation f for biorthogonal filters. The IA-DCT scheme was not robust to misalignments and JPEG compression while the IA-W scheme was comparatively robust. The watermark was embedded into DC components [54] in order to make the invisible watermark more robust by incorporating the feature of texture masking and luminance masking of HVS.

Kumsawat *et al.* [38] proposed spread spectrum watermarking scheme using discrete multiwavelet transform for copyright protection. Multiwavelets could simultaneously possess properties like orthogonality, symmetry and compact support. The watermark was embedded into the DMT (Discrete Multiwavelet Transform) coefficients by performing three level multiwavelet decomposition. Genetic algorithms were applied to achieve optimum performance. As discussed earlier, to reduce the probability of the watermark being detected the watermark signal must be wide band and noise like. Chen and Leung [34] proposed a chaotic system to generate the noise-like signals. It was a nonlinear system and was robust to synchronization errors. An ergodic demodulator was developed to detect the chaotic watermark.

Furon and Duhamel [20] proposed an asymmetric watermarking scheme as an alternate to the SS approach. Another asymmetric approach was proposed by Kim *et al.*[31] which accommodated many embedded watermarks and had only one detection watermark. The proposed method adopted Phase Shift Transform and the watermark was embedded with reference keys. The detection was performed using a reference watermark.

Langelaar and Lagendigk [35] proposed the DEW (Differential energy watermarking) scheme. This scheme selectively discarded high frequency DCT coefficients in compressed domain. Hence the image was considered as a set of small blocks having size 8×8. Each block was interpreted as a collection of 64 prequantized DCT coefficients. The set was then divided into different groups (*lc*- region) each containing blocks. A particular *lc*-region was divided into two sub regions A, B each containing n/2 blocks. Das *et al.*[36] presented two modifications for the DEW scheme. The first was where the energy difference was created by changes in low frequency DCT coefficients. The second was where a random permutation of blocks were used in such a way that in any *lc*-region, the energy of the *lc*-region A and B differ by a small quantity. The later scheme was more robust to the former scheme.

Watermarking modulation is a technique wherein the values of the transformed coefficients are replaced by watermark coefficients. Higher correlation values between the original image and the watermarked image indicates a genuine watermark. It was noted by Lu *et al.*[7] that this could be achieved if the transformed coefficients were along the same direction during the embedding and attacking process. They also noted that both [6] and [2] had adopted a random modulation technique. They had not considered the relationship between the signs of the modulation pair and hence they could not sustain most of the watermark attacking techniques. They proposed a complementary modulation strategy. Here, two watermarks which play complementary roles, in resisting various kinds of attacks were embedded. The cocktail watermark encoding algorithm proposed by them adopted complementary modulation rules and considered the sign of its wavelet coefficient and its watermarked value.

Lu and Liao [10] further adopted cocktail watermarking in another work where they embedded robust and fragile watermarks simultaneously that could be blindly extracted without access to host image. This was the first scheme which combined fragile and robust watermarking schemes. The host image's wavelet coefficients were quantized as masking threshold units. This method could not detect color changes if the color was modified. This could be overcome by randomly selecting positions among Y, Cb and Cr for watermarking. It was also noted that the watermark embedded using negative modulation was more robust to compression than that using positive compression. Liu and Chou [49] designed a robust and transparent scheme for color images. They transformed the host image into CIELAB color space, estimated the JND profile of the color components Y, C_b, C_r. These were used to embed the watermark by modulating the quantization indices of the coefficients in the significant portion of the color image. This scheme could be made more robust to malicious attacks and compression if the perceptual redundancy of color images was used to accurately locate embedding coefficients.

Lin and Chen [8] proposed a DCT watermarking scheme where the watermark of 64x64 was embedded into the LSB of the DCT coefficients of the host image. It was resilient to some image processing operations like cropping, uniform noise and JPEG compression to some degree. Unlike pseudo-random permutation of the watermark adopted by [8], Lu and Liao [9] used a watermark with visual recognizable patterns. The original image was decomposed into wavelet coefficients. Multi energy watermarking scheme based on qualified significant wavelet tree (QSWT) was used to achieve robustness against JPEG compression, image cropping, sharpening and median filtering. Suhail and Obaidat [29] too adopted DCT and proposed a scheme which was resistant to geometric manipulations and withstood cropping attacks. The proposed scheme segmented the image based on Voronoi diagrams and the traditional pseudo random sequence was embedded in the DCT domain of each segment. Briassouli *et al.* [33] adopted symmetric alpha stable family of distributions to model the heavy tailed DCT coefficients. The watermark detector was designed based on Cauchy distribution. Cauchy detectors are robust in heavy tailed environments.

Solachidis and Pitas [11] proposed a watermarking scheme based on circularly symmetric watermarks applied in DFT domain. The authors considered a grey scale image I (n1, n2) of size $N \times N$.

The DFT of the image was given by

$$I\left(k_1,k_2\right)=\sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} i(n_1,n_2)e^{-j2\pi n_1 k_1/N_1-j2\pi n_2 k_2/N_2} \qquad (3)$$

The magnitude and phase of the image were given by $M(k_1,k_2)=\left|I(k_1,k_2)\right|$ and $P(k_1,k_2)$ respectively. The watermark used, $W(k_1,k_2)$ was embedded in the DFT domain. Circular shifts in spatial domain did not affect the magnitude of the FT. It is also noted that scaling in spatial domain causes inverse scaling in frequency domain. Also rotation in spatial domain causes the same rotation in frequency domain. Since compression affected the high frequencies of FT, the watermark was added in the middle frequency range. This method was robust to filtering, noise addition, scaling, rotation, cropping and compression. It was also noted that there were practically zero errors since the detector output was always bigger than the chosen threshold for a watermarked image and smaller for a non watermarked image.

Tsui and Zhang [46] proposed two vector watermarking schemes, the Quaternion Fourier Transform (QFT) and Spatiochromatic DFT (SCDFT), done in the frequency domain of the chromatic components. The SCDFT utilizes chromatic information ignoring luminance which would result in destroying the watermark and was not robust. Using QFT the watermark was spread uniformly along the chromatic and luminance components of the image and was robust.

Initially, research in digital image watermarking had been focused on grey scale images. As technology advanced, extension to color images was done considering image luminance or processing each color channel independently. A DCT domain watermarking is proposed by [12] where the watermark is hidden by modifying the subset of DCT coefficients of each color channel. The first ‘k’ coefficients are skipped to obtain an improved invisible watermark. The watermark was casted three times to obtain high robustness. The RGB bands were extracted from a given image, the DCT of each band was computed and the coefficients recorded in zigzag pattern. Three vectors V_r, V_b, V_g were obtained and modified to embed the watermark. The modified vectors were reinserted in the zigzag scan of the corresponding channel and inverse DCT was performed to obtain the individual watermarked bands. This watermarking scheme provided a tradeoff between accuracy and effectiveness.

Singular Value Decomposition (SVD) watermarking scheme was proposed by Liu and Tan [13]. Here the SVD ‘$\mathbf{A}$’ of an $N \times N$ image was computed to obtain two orthogonal matrices ‘$\mathbf{U}$’, ‘$\mathbf{V}$’ and a diagonal matrix ‘$\mathbf{S}$’. The watermark matrix ‘$\mathbf{W}$’ was added into the matrix ‘$\mathbf{S}$’ and SVD was performed on the new matrix to obtain ‘$\mathbf{U_W}$’, ‘$\mathbf{V_W}$’ and ‘$\mathbf{S_W}$. The watermarked image ‘$A_\mathbf{W}$’ was then obtained by multiplying the matrices ‘$\mathbf{U}$’, ‘$\mathbf{S_W}$’, ‘$\mathbf{V_W}^\mathbf{T}$’ through the following steps

$$\mathbf{A} = \mathbf{U\,S\,V^T} \qquad (4)$$

$$\mathbf{S}+\boldsymbol{\alpha}\mathbf{W} = \mathbf{U_W S_W V_W}^\mathbf{T} \qquad (5)$$

$$A_\mathbf{W} = \mathbf{U\,S_W V^T} \qquad (6)$$

The watermark was extracted by reversing the above steps. For a given scale factor α, the error between the original and the watermarked image was controlled by the spectral norm of the watermark. This method was compared with [6] and it was found that the method was more robust than the later. This method was filed for patenting too due to its high robustness than [6]. Zhang and Li [41] stated that the extracted watermark using the above method was not the embedded watermark. It was the reference watermark. During watermark detection, the SVD matrices depend on the reference watermark which biases false detection.

Wavelet transforms are used in digital watermarking schemes. It is important to know how the wavelet transform can be used in watermarking schemes. The basic idea of the DWT for a two-dimensional image is described as follows. The image is first decomposed into four parts of high, middle, and low frequencies (i.e., LL1, HL1, LH1, HH1 subbands) by critically subsampling horizontal and vertical channels using subband filters, where LL1 represent approximate wavelet coefficients. The subbands labelled HL1, LH1, and HH1 represent the detail wavelet coefficients. To obtain the next coarser scaled wavelet coefficients, the sub- band LL1 is further decomposed and critically subsampled. This process is repeated several times and further, from these DWT coefficients, the original image can be reconstructed. This reconstruction process is called the Inverse DWT (IDWT). The coefficient computation is complicated and time consuming for a common wavelet filter and this can be overcome by using reversible wavelet transform that maps integers to integers [14] based on lifting framework. Lu and Liao [10] adopted wavelets in their cocktail watermarking scheme as discussed earlier. Wei *et al.*[5] proposed a method wherein the watermark was inserted in wavelet coefficients and its amplitudes were controlled by the wavelet coefficients so that the watermark noise did not exceed the JND of each wavelet coefficient. Celik *et al.*[44] proposed four different approaches for securely embedding spread spectrum watermarks at the client side using one time pads, stream switching, joint decryption and LUT(Look up Table) based ciphers.

Watermarking can be incorporated into image capture pipeline or geometric properties of an image. Geometric properties based watermarking schemes were still in the early areas of research during the 1990's. Tang and Hang [25] proposed a robust watermarking scheme that was robust to geometric distortion and signal processing attacks adopting Mexican Hat Wavelet Scale interaction for feature extraction. Other methods for feature extraction are the Harris detector and the Achard-Rouque adopted by Bas *et al.*[26].

The Marr wavelet [27], [28] is rotation invariant. The mother wavelet function is defined by

$$\begin{aligned}&\psi\overrightarrow{(x)} = (2-\left\|\vec{x}\right\|^2 e^{-\left\|\vec{x}\right\|^2/2},\\&\text{where}\left\|\vec{x}\right\| = \left(x^2+y^2\right)^{1/2}\end{aligned} \tag{7}$$

The 2D FT is given by

$$\psi^*\overrightarrow{(k)} = \left\|\vec{k}\right\|^2\ e^{-k\left\|\vec{k}\right\|^2/2} \tag{8}$$

where $\left(\vec{k}\right)$ represents the 2D spatial frequency.

The feature extraction method uses the following quantities

$$P_{i,j}\left(\vec{k}\right) = \left|M_i\left(\vec{x}\right) - \gamma.M_j\left(\vec{x}\right)\right| \tag{9}$$

$$M_i\left(\vec{x}\right) = \left(2^{-i}\psi(2^{-i}.\vec{x}\right) * A \tag{10}$$

where $M_i\overrightarrow{(x)}$ represents the response of the Mexican Hat Wavelet Filter. $P_{i,j}\overrightarrow{(k)}$ is the interaction between scales *i* and *j* with γ as the scaling factor. The scheme was designed for both color and grey scale images. The Wavelet filtering was adopted using FFT. Marr wavelet allows for different degrees of robustness against distortion, cropping affects only few feature points, and is band limited reducing noise sensitivity problem in feature extraction.

The embedding process is outlined as follows

- The feature extraction method generated reference centers of disks.
- Image normalization was done to select the location for the watermark.
- Coordinate transformation coefficients between original normalized images were generated.
- Location of blocks in original image for watermarking was determined from the normalized image.
- The coordinates were transformed from normalized image to original image.
- A 2D FFT was applied to each disk and the watermark was embedded in Transform domain.
- 2D IFFT was performed on the watermarked blocks to replace the original image blocks.

A secret key was used in the watermark detection process. It was concluded that the scheme can be further improved if the feature points were more robust under severe geometric distortions.

Liu and Chou [49] utilized the color features of the HVS to design the watermarking scheme. Dejey and Rajesh [24] adopted Discrete Wavelet Transform-Fan Beam Transform (DWT-FBT) and proposed two nonblind color image watermarking schemes. The first was a wavelet fan beam watermarking on luminance and chrominance and the other, wavelet fan beam watermarking on chrominance only. The proposed schemes provided for a trade off on capacity, robustness and imperceptibility. The schemes were robust to blurring, sharpening and histogram equalization attacks.

Bi *et al.*[42] used Mallet's Multiband Discrete Wavelet Transform and Empirical Mode Decomposition and proposed a blind watermarking scheme. Selecting a dilation factor $M\geq 2$, 1D scaling and a wavelet filter $H_l(\xi)$, $0\leq l \leq M-1$, the original image was decomposed. The watermark bits were embedded in suitable sub images. The robustness of the watermark was tested against JPEG compression, Black and White Noise, Gaussian Noise, ConvFilter and rotation scaling effects. Multiwavelet watermarking was robust against compression attacks, cropping and scaling. Wavelets are suitable for attacks on the mid frequencies against multiwavelets for low and high frequency ranges.

Bhatnagar et *al.*[62] applied fractional wavelet packet transform(FRPWT) in digital watermarking. The host image was decomposed using fractional wavelet packet transform and a grey scale image was used as the watermark compared to the previous randomly generated Guassian Noise. A secret key known only to the creator was used to change the frequency bands at all sublevels. The reference image was obtained by inverse fractional wavelet transform. Since FRPWT depends on the transform order all along the axis to decompose and reconstruct the image, this method was more robust and secure to copyright protection. This method avoided ambiguity problem faced by SVD methods. This method was tested on host images of size 256 x 256 and the watermark size was 64 x 64.

Sparse representation of an image could be achieved by DCT or DWT. Wavelets performed well for 1D piecewise smooth functions. Higher order wavelets could not see the smoothness along the edges. Contourlets represent image edges sparsely, employ iterative filter banks and allow for different number of directions at each scale. [56] – [60] proposed contourlets based watermarking schemes. These schemes adopted directional information of images edges. Akhaee et al. [61] proposed a multiplicative watermarking approach in contourlet domain. The proposed blind watermarking scheme was robust to AWGN and compression attacks.

Ridgelets [17]-[18] deal with line singularity issues faced in 2D wavelets. Kalantari et al. [19] proposed a robust watermarking method in the ridgelet domain. To adapt to curved edges, the image is partitioned into blocks so that the curved edge formed a straight edge. The watermark is inserted into the blocks with high entropy by modifying the amplitude of the ridgelet coefficients. The distribution of the ridgelet coefficients is unknown. Due to this a host distribution independent decoder working near the optimal point is used. To achieve maximum robustness the decoder is optimized by taking into consideration the Guassian noise attack.

Curvelets is another multiscale transform developed by Candues *et al.* [64]. This transform could efficiently represent edges and singularities along curves. Curvelets also have the capability of better recovery under noisy circumstances. Leung *et al.*[63] adopted curvelets and proposed a watermarking scheme leading to the addition of a HVS adopting the orientation parameter of the curvelets. The proposed scheme is robust against image processing methods. Zhang *et al.* [65] proposed a multipurpose blind watermarking scheme adopting multiscale curvelet transform. Here, single level watermark was embedded onto the significant coefficients by quantization in individual frequency scales. The coefficient selection is done on the concept, that, the coefficient energy was proportional to its sensitivity. This scheme provided for image authentication and copyright protection and was robust to Gaussian low pass filtering, contrast enhancement and Gaussian noise contamination.

Besides watermark embedding, optimum recovery of the watermark is important for all applications. Several techniques have been proposed to recover the watermark. Barni *et al.* [52] addressed the problem of watermark recovery. The architecture of an optimum decoder for an additive/ multiplicative watermark embedded in the DFT domain was derived by relying on Baeyes Decision theory. A statistical analysis to model the DFT coefficients was used to derive the actual structure of the decoder and the decoder was further simplified. Bian and Liang [53] proposed a watermark detector that applied Bessel K PDF which performed well even on weak watermarks. Most of the schemes described above adopted inverse transform to extract the watermark efficiently.

5. EVALUATING WATERMARKS

Initially most of the watermarking techniques adopted a different test series, different images, and different methodologies. It was highly difficult to obtain a comparative description without reimplementing and testing them separately. But the implementation would be quiet different due to the change in test benches, which would further suggest sometimes weaker implementation or mismatches in the results. Hence, evaluation methodologies were required with common benchmarks. This would result in a less detailed table of results along with a reliable summary of the proposed schemes.

5.1. Performance Evaluation Metrics

Petitcolas and Fabien [51] stated that the first step of the evaluation procedure is to identify the target of evaluation algorithms. A full scheme evaluation is the collection of functionality services to which a level of assurance is globally applied and for each of which a specific level of strength is selected. Six to seven levels of assurances are globally accepted. Levels of perceptibility ranges from not perceptible to slightly perceptible to completely perceptible. Assurance levels to access perceptibility are through human perceptual models. Another metric is to consider geometric distortions. Detection probabilities of the watermark and bit error rate measurements are the accessing parameters for robustness. The level of robustness ranges from no robustness to provable robustness. Robustness is also tested with a random payload of a given size, if the application size is fixed. Speed is a parameter that varies from hardware to software implementations. Difference in statistical properties of the original and the watermarked image leads to detection attacks. Very few watermarking schemes consider this criterion.

In brief, the following are some quality measures to evaluate the performance of watermarked images. The two most important metrics are PSNR and BER. Others include MSE and SSIM.

1) **PSNR**- Peak Signal to Noise Ratio determines the quality of the recovered watermark.

$$\text{PSNR} = 10 \log_{10} \left(\frac{N \times N}{MSE} \right) \qquad (11)$$

where N=peak signal value of the original signal

2) **BER** – Bit Error Ratio is the number of error bits in the overall bits received. This metric compares the host image and the watermarked image.

BER = $\frac{C}{H \times W}$ (12)

where H and W represent the height and width of the watermarked image. C indicates the number of bits received in error.

Transform Adopted	Concept & Details	Results & Summary
SS	Watermark is distributed over a wide range of frequencies	Larger bandwidth is required to transmit a narrow bandwidth signal. Not robust against high amplitude noise.
DCT	Watermark is spread over a range of frequencies and hence not detectable	Not robust to misalignments and compression attacks.
DFT	Watermark is added in the mid frequency ranges	Robust to filtering, noise addition, scaling compression, rotation and cropping attacks.
DWT	Watermark is embedded in mid frequency ranges	Orthogonal and symmetrical, reconstruction. Perform well for 1D functions. Comparatively robust to misalignments and compression attacks, geometric distortions, signal processing attacks, histogram equalisation attacks.
Multiwavelets	Watermark is embedded into low and high frequency ranges.	Robust against compression, cropping and scaling.
Contourlets	Watermark can be embedded along different direction of the curved edges	Robust to AWGN and compression attacks.
Ridgelets	Deal with line singularity issues faced in 2D wavelets	Maximum robustness against Gaussian noise attacks
Curvelets	Can represent edges and singularities along curves	Better recovery under noise conditions. Robust to Gaussian low pass filtering, contrast enhancement and noise contamination.
Shearlets	Multiscale geometric analysis, high frequency ranges	Performed well compared to DCT and DWT

TABLE 1: Comparative Study of Frequency Domain Watermarking Schemes.

3) **SSIM** – Used to measure the similarity between two images as an improvement on PSNR and MSE. This metric is calculated on various windows of an image. Consider two windows *p,q* of size N X N, then

SSIM = $\frac{(2\mu_p\mu_q + c_1)(2\sigma_{pq} + c_2)}{(\mu_p^2 + \mu_q^2 + c_1)(\sigma_p^2 + \sigma_q^2 + c_2)}$ (13)

where

- μ_p , μ_q is the average of *p* and *q respectively*
- σ_p^2 is the variance of *p*
- σ_q^2 is the variance of q

- σ_{pq}^{2} is the covariance of *p* and *q*
- c_1 and c_2 are variables used to stabilize the denominator

4) **MSE** – Mean Square Error is defined as average squared difference between the reference image and the distorted image.

MSE = $\frac{1}{pq}\left[\sum_{i=1}^{p}\sum_{j=1}^{q}\left(N(i,j)-W(i,j)\right)^{2}\right]$ (14)

Where,
p and *q* = height and width of the image respectively,
$N(i,j)$ = pixel values of the original image
and $W(i,j)$ = pixel values of the watermarked image.

5.2. Watermarking Attacks

The watermarked image is subjected to various attacks. Boato *et al.* [48] proposed the first benchmarking tool to evaluate watermarking robustness based on GA (Genetic Algorithms). Robustness was evaluated in terms of perceptual quality measured by WPSNR (Weighted Peak Signal to Noise Ratio). The watermarked image was subjected to JPEG2000 compression, AWGN, resizing and amplitude scaling. The attacks to which a watermarked image is subjected can be broadly classified as intentional and non-intentional attacks. Hartung et al.[66] classified attacks as simple attacks(noise addition, cropping, compression), disabling attacks(geometric distortion, rotation, cropping), ambiguity attacks and removal attacks. In this section we describe intentional and non-intentional attacks in brief.

1) Compression Attacks- During compression using JPEG and JPEG2000 standards, lossy compression techniques produce irreversible changes to the watermarked images wherein the watermark may become fragile whilst lossless compression techniques are more robust in recovering the watermark.

2) Interference Attacks- These attacks add noise to the watermarked image. Salt and pepper noise, denoising, averaging, AGWN etc.

3) Signal Processing Attacks- These attacks include lossy compression, linear, nonlinear and adaptive filtering, denoising and noise addition.

4) Geometric Attacks- Most of the watermarking techniques is not robust to geometric attacks. These attacks are rotation, cropping, flipping, scaling, row-column removal and resizing.

5) Cryptographic Attacks- The security of the system can be determined by detecting the weakness in the code, cipher, protocols and the management entrapping the system. These attacks include cipher text only, plain text only and chosen text that can be chosen plaintext or chosen cipher text.

6) Collusion Attacks- Such attacks are common where the attacker has access to more than one copy of the watermarked image. The attacker can predict the watermark by colluding them. These attacks can be linear or nonlinear. They are powerful due to their capability of achieving their objective and the degradation is very low.

7) Active Attacks – These attacks are those that try to break the system. They tend to alter the watermark or remove the watermark. These are done by viruses.

8) Passive Attacks- These attacks look at sensitive information which can be subjected to other attacks.

9) Histogram Equalization attacks- These attacks normally tend to enhance image intensities. These include brightness, contrast adjustments.

A brief summary of the watermarking schemes is given in Table 1. A comparative analysis on watermarking attacks using Fractional Wave Packet Transform is depicted in Table 2. The image considered is Lena (256 X256) and the watermark used is the Logo.

Attack	**No Attack**	**Median Filtering**	**Salt and Pepper Noise**	**Gaussian Filter**
PSNR	50.491	22.256	35.7823	28.181

TABLE 2: PSNR for various attacks

6. CONCLUSION

Embedding the watermark in low frequency components is robust to low pass filtering, compression and geometric attacks while embedding the watermark in high frequency components is robust to histogram equalization and geometric attacks. A multitransform approach, wherein the watermark can be embedded in both the high frequency and low frequency components dealing with line singularities is left as an open area for research to achieve maximum robustness. An attacker having multiple copies of the watermarked image can remove the watermark by collusion attack. A watermarking scheme which is transparent, robust to geometric distortions and collusion attacks can be designed by adapting multiwavelets.

A watermarking scheme that is robust to desynchronisation attacks is still a challenging issue. It is noted that most of the watermarking schemes could resist rotation, scaling, translation and other affine transforms but very few were resilient to cropping attacks. There exists a trade-off between imperceptibility and robustness, and imperceptibility and capacity. In the process of designing a robust watermark, it is necessary to consider collusion attacks which are still an open area for researchers. Hence, digital watermarking is an interesting area which provides an open space for research.

7. REFERENCES

[1] Craver, S.; Memon, N.; Boon-Lock Yeo Yeung, M .M., "Resolving rightful ownerships with invisible watermarking techniques: limitations, attacks, and implications," IEEE Journal on Selected Areas in Communications, vol.16, no.4, pp.573-586, May 1998.

[2] Podilchuk, C.I.;Wenjun, Zeng, "Image-adaptive watermarking using visual models," IEEE Journal on Selected Areas in Communications, vol.16, no.4, pp.525-539, May1998.

[3] Watson A. B., “DCT quantization matrices visually optimized for individual images,” in Proc. SPIE Conf. Human Vision, Visual Processing, and Digital Display IV, vol. 1913, pp. 202– 216, Feb. 1993.

[4] Chiou-Ting Hsu; Ja-Ling Wu, "Multiresolution watermarking for digital images," IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, vol.45, no.8, pp.1097-1101, Aug 1998.

[5] Wei, Z.H.; Qin, P.; Fu, Y. Q., "Perceptual digital watermark of images using wavelet transform," IEEE Transactions on Consumer Electronics, vol.44, no.4, pp.1267-1272, Nov 1998.

[6] Cox I. J.; Kilian J.; Leighton F. T.; and Shamoon T., “Secure spread spectrum watermarking for multimedia,” IEEE Transactions on Image Processing, vol. 6, pp. 1673–1687, 1997.

[7] Chun-Shien Lu; Shih-Kun Huang; Chwen-Jye Sze; Hong-Yuan Mark Liao, "Cocktail watermarking for digital image protection," IEEE Transactions on Multimedia, vol.2, no.4, pp.209-224, Dec 2000.

[8] Lin, S.D.; Chin-Feng Chen, "A robust DCT-based watermarking for copyright protection," IEEE Transactions on Consumer Electronics, vol.46, no.3, pp.415-421, Aug 2000.

[9] Ming-Shing Hsieh; Din-Chang Tseng; Yong-Huai Huang, "Hiding digital watermarks using multiresolution wavelet transform," IEEE Transactions on Industrial Electronics, vol.48, no.5, pp.875-882, Oct 2001.

[10] Chun-Shien Lu; Liao, H.Y.M., "Multipurpose watermarking for image authentication and protection," IEEE Transactions on Image Processing, vol.10, no.10, pp.1579-1592, Oct 2001.

[11] Solachidis, V.; Pitas, I., "Circularly symmetric watermark embedding in 2-D DFT domain," IEEE Transactions on Image Processing, vol.10, no.11, pp.1741-1753, Nov 2001.

[12] Barni, M.; Bartolini, F.; Piva, A., "Multichannel watermarking of color images," IEEE Transactions on Circuits and Systems for Video Technology, vol.12, no.3, pp.142-156, Mar 2002.

[13] Ruizhen Liu; Tieniu Tan, "An SVD-based watermarking scheme for protecting rightful ownership", IEEE Transactions on Multimedia, vol.4, no.1, pp.121-128, Mar 2002.

[14] Chen, Tao; Wang, Jingchun, "Image Watermarking method using integer-to-integer wavelet transforms", Tsinghua Science and Technology , vol.7, no.5, pp.508-512, Oct. 2002.

[15] Jengnan Tzeng; Wen-Liang Hwang; I-Liang Chern, "Enhancing image watermarking methods with/without reference images by optimization on second-order statistics," IEEE Transactions on Image Processing, vol.11, no.7, pp.771-782, Jul 2002.

[16] Yiwei Wang; Doherty, J.F.; Van Dyck, R.E.,"A wavelet-based watermarking algorithm for ownership verification of digital images,” IEEE Transactions on Image Processing, vol.11, no.2, pp.77-88, Feb 2002.

[17] Do, M.N.; Vetterli, M., "The finite ridgelet transform for image representation," IEEE Transactions on Image Processing, vol.12, no.1, pp.16-28, Jan 2003.

[18] E. J. Candès and D. L. Donoho, “Ridgelets:Akey to higher-dimensional intermittency?,” Phil. Trans. R. Soc. Lond. A., pp. 2495–2509, 1999.

[19] Kalantari, N.K.; Ahadi, S.M.; Vafadust, M., "A Robust Image Watermarking in the Ridgelet Domain Using Universally Optimum Decoder," IEEE Transactions on Circuits and Systems for Video Technology, vol.20, no.3, pp.396-406, March 2010.

[20] Furon, T.; Duhamel, P., "An asymmetric watermarking method," IEEE Transactions on Signal Processing, vol.51, no.4, pp.981-995, April 2003.

[21] Izquierdo, E.; Guerra, V., "An ill-posed operator for secure image authentication," IEEE Transactions on Circuits and Systems for Video Technology, vol.13, no.8, pp.842-852, Aug 2003.

[22] Wong, P.H.; Au, O.C.; Yeung, Y. M., "Novel blind multiple watermarking technique for images," IEEE Transactions on Circuits and Systems for Video Technology, vol.13, no.8, pp.813-830, Aug. 2003.

[23] Chu, W.C., "DCT-based image watermarking using subsampling," IEEE Transactions on Multimedia, vol.5, no.1, pp.34-38, March 2003.

[24] Dejey. R. S.; Rajesh, D.; "Robust discrete wavelet fan beam Transforms based color image watermarking", IET Transactions on Image Processing, vol 5, no 4, pp315–322.

[25] Chih-Wei Tang; Hsueh-Ming Hang, "A feature-based robust digital image watermarking scheme," IEEE Transactions on Signal Processing, vol.51, no.4, pp.950-959, Apr 2003.

[26] P. Bas, J. M. Chassery, and B. Macq, "Robust watermarking based on the warping of pre-defined triangular patterns," Proc. SPIE Security and Watermarking of Multimedia Contents- II, vol. 3971, pp. 99–109, 2000.

[27] J.P. Antoine and P. Vandergheynst, "Two-dimensional directional wavelets in image processing," International. J. Image. Syst. Technol., vol. 7, pp. 152–165, 1996.

[28] D. Marr, Vision. San Francisco, CA: Freeman, pp. 54–61, 1982.

[29] Suhail, M.A.; Obaidat, M.S., "Digital watermarking-based DCT and JPEG model," IEEE Transactions on Instrumentation and Measurement, vol.52, no.5, pp.1640-1647, Oct. 2003.

[30] Yang Zhao; Campisi, P.; Kundur, D., "Dual domain watermarking for authentication and compression of cultural heritage images," IEEE Transactions on Image Processing, vol.13, no.3, pp.430-448, March 2004.

[31] Tae Young Kim; Hyuk Choi; Lee, Kiryung; Kim, Taejeong, "An asymmetric watermarking system with many embedding watermarks corresponding to one detection watermark," IEEE Signal Processing Letters, vol.11, no.3, pp.375-377, March 2004.

[32] Bao, P.; Xiaohu Ma, "Image adaptive watermarking using wavelet domain singular value decomposition," IEEE Transactions on Circuits and Systems for Video Technology, vol.15, no.1, pp.96-102, Jan. 2005.

[33] Briassouli, A.; Tsakalides, P.; Stouraitis, A., "Hidden messages in heavy-tails: DCT-domain watermark detection using alpha-stable models," IEEE Transactions on Multimedia, vol.7, no.4, pp.700-715, Aug. 2005.

[34] Siyue Chen; Leung, H., "Ergodic chaotic parameter modulation with application to digital image watermarking," IEEE Transactions on Image Processing, vol.14, no.10, pp.1590-1602, Oct. 2005.

[35] Langelaar, G.C.; Lagendijk, R.L., "Optimal differential energy watermarking of DCT encoded images and video," IEEE Transactions on Image Processing, vol.10, no.1, pp.148-158, Jan 2001.

[36] Das, T.K.; Maitra, S.; Mitra, J., "Cryptanalysis of optimal differential energy watermarking (DEW) and a modified robust scheme," IEEE Transactions on Signal Processing, vol.53, no.2, pp.768-775, Feb. 2005.

[37] Piva A.; Barni M.; Bartolini F.; and Cappellini V., "DCT-based water- mark recovering without resorting to the uncorrupted original image," in Proc. IEEE Int. Conf. Image Processing, vol. 1, pp. 520–523, Oct. 1997.

[38] Kumsawat, P.; Attakitmongcol, K.; Srikaew, A., "A new approach for optimization in image watermarking by using genetic algorithms," IEEE Transactions on Signal Processing, vol.53, no.12, pp.4707-4719, Dec. 2005.

[39] Mobasseri, B.G.; Berger, R.J., "A foundation for watermarking in compressed domain," IEEE Signal Processing Letters, vol.12, no.5, pp.399-402, May 2005.

[40] Biao-Bing Huang; Shao-Xian Tang, "A contrast-sensitive visible watermarking scheme," IEEE Transactions on MultiMedia, vol.13, no.2, pp.60-66, April-June 2006.

[41] Xiao-Ping Zhang; Kan Li, "Comments on "An SVD-based watermarking scheme for protecting rightful Ownership", IEEE Transactions on Multimedia, vol.7, no.3, pp.593-594, June 2005.

[42] Ning Bi; Qiyu Sun; Daren Huang; Zhihua Yang; Jiwu Huang, "Robust Image Watermarking Based on Multiband Wavelets and Empirical Mode Decomposition," IEEE Transactions on Image Processing, vol.16, no.8, pp.1956-1966, Aug. 2007.

[43] Coltuc, D.; Chassery, J. M., "Very Fast Watermarking by Reversible Contrast Mapping", IEEE Signal Processing Letters, vol.14, no.4, pp.255-258, April 2007.

[44] Celik, M.U.; Lemma, A.N.; Katzenbeisser, S.; Van der Veen, M., "Lookup-Table-Based Secure Client-Side Embedding for Spread-Spectrum Watermarks", IEEE Transactions on Information Forensics and Security, vol.3, no.3, pp.475-487, Sept. 2008.

[45] Wei-Hung Lin; Shi-Jinn Horng; Tzong-Wann Kao; Pingzhi Fan; Cheng-Ling Lee; Yi Pan, "An Efficient Watermarking Method Based on Significant Difference of Wavelet Coefficient Quantization", IEEE Transactions on Multimedia, vol.10, no.5, pp.746-757, Aug. 2008.

[46] Tsz Kin Tsui; Xiao-Ping Zhang; Androutsos, D., "Color Image Watermarking Using Multidimensional Fourier Transforms", IEEE Transactions on Information Forensics and Security, vol.3, no.1, pp.16-28, March 2008.

[47] Altun, H.O.; Orsdemir, A.; Sharma, G.; Bocko, M.F., "Optimal Spread Spectrum Watermark Embedding via a Multistep Feasibility Formulation", IEEE Transactions on Image Processing, vol.18, no.2, pp.371-387, Feb. 2009.

[48] Boato, G.; Conotter, V.; De Natale, F. G B; Fontanari, C., "Watermarking Robustness Evaluation Based on Perceptual Quality via Genetic Algorithms", IEEE Transactions on Information Forensics and Security, vol.4, no.2, pp.207-216, June 2009.

[49] Liu, K. C.; Chou, K. H.; “Robust and transparent watermarking scheme for color images”, IET transactions on Image Processing, vol 3, no. 4, pp.228–242.

[50] Meerwald, P.; Koidl, C.; Uhl, A., "Attack on “Watermarking Method Based on Significant Difference of Wavelet Coefficient Quantization”, IEEE Transactions on Multimedia, vol.11, no.5, pp.1037-1041, Aug. 2009.

[51] Petitcolas; Fabien A. P., "Watermarking schemes evaluation", IEEE Signal Processing Magazine, vol.17, no.5, pp.58,64, Sep 2000.

[52] Barni, M.; Bartolini, F.; De Rosa, A.; Piva, A., "A new decoder for the optimum recovery of nonadditive watermarks", IEEE Transactions on Image Processing, vol.10, no.5, pp.755-766, May 2001.

[53] Bian, Y.; & Liang, S., “Locally Optimal Detection of Image Watermarks in the Wavelet Domain Using Bessel K Form Distribution”, IEEE Transactions on Image Processing, vol.22, no.6, pp.2372–2384, June 2013.

[54] Jiwu Huang; Shi, Y.Q.; Yi Shi, "Embedding image watermarks in dc components", IEEE Transactions on Circuits and Systems for Video Technology, vol.10, no.6, pp.974-979, Sep 2000.

[55] Wong, P.W.; Memon, N., "Secret and public key image watermarking schemes for image authentication and ownership verification", IEEE Transactions on Image Processing, vol.10, no.10, pp.1593-1601, Oct 2001.

[56] M. Jayalakshmi; S. N. Merchant; and U. B. Desai, “Digital watermarking in contourlet domain”, in Proc. 18th Int. Conf. Pattern Recognition, 2006, vol. 3, pp. 861–864.

[57] M. Jayalakshmi; S. N. Merchant; and U. B. Desai, “Blind watermarking in contourlet domain with improved detection,” Int. Conf. Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP’06).

[58] Xueqiang, L.; Xinghao, D.; and Donghui, G., “Digital watermarking based on non-sampled contourlet transform,” in Proc. IEEE Int.Workshop Anti-counterfeiting, Security, Identification, pp. 138–141, 2007.

[59] Li, H.; W, Song; and Wang, S., “A novel blind watermarking algorithm in contourlet domain,” in Proc. 18th Int. Conf. Pattern Recognition, vol. 3, pp. 639–642, 2006.

[60] Xiao S.; Ling H.; Zou F.; and. Lu, Z., “Adaptive image watermarking algorithm in contourlet domain,” Proc. Japan-China Joint Workshop on Frontier of Computer Science and Technology, pp. 125–130, 2007.

[61] Akhaee, M.A.; Sahraeian, S.M.E.; Marvasti, F., "Contourlet-Based Image Watermarking Using Optimum Detector in a Noisy Environment," IEEE Transactions on Image Processing, vol.19, no.4, pp.967- 980, April 2010.

[62] Bhatnagar,G.; Raman,B.; Wu,Q.M.J., “Robust watermarking using fractional wavelet packet transform”, IET Transactions on Image Processing, vol 6, no.4, pp.386-397, June 2012.

[63] Leung, H.Y.; Cheng, L.M.; Cheng, L.L., “Digital Watermarking Schemes using Multiresolution and Curvelet and HVS Model”, Proceedings of 8th International Workshop, IWDW 2009, Guildford, UK, vol 5703, pp 4-13, August 24-26, 2009.

[64] Candès E. J.; Demanet, L.; Donoho, D. L.; and Ying, L., “Fast discrete curvelet transforms,” Tech Rep., Appl. Comput. Math., California Inst. Technol., 2005. [On line]. Available: http://www.curvelet.org.

[65] Chune Zhang; Cheng, L. L.; Zhengding Qiu; Cheng, L.M., "Multipurpose Watermarking Based on Multiscale Curvelet Transform," IEEE Transactions on Information Forensics and Security, vol.3, no.4, pp.611-619, Dec. 2008.

[66] F. Hartung; J. K. Su; B. Girod, “Spread spectrum watermarking: Malicious attacks and counterattacks,” pp. 147-158, 1999.

[67] Khalifa, O.O.; Binti Yusof, Y.; Abdalla, A.H.; Olanrewaju, R.F., "State-of-the-art digital watermarking attacks," 2012 International Conference on Computer and Communication Engineering (ICCCE), pp.744-750, July 2012.

Computer Forensic: A Reactive Strategy for Fighting Computer Crime

Abdullahi Mohammed Abdulmohammedabdul@yahoo.com
Faculty of Physical Science
Department of Computer Science
University of Port Harcourt
Port Harcourt, 500001, Nigeria

Enoch O. Nwachukwu Enoch.nwachukwu@uniport.edu.ng
Faculty of Physical Science
Department of Computer Science
University of Port Harcourt
Port Harcourt, 500001, Nigeria

Abstract

Computer Forensics is the science of obtaining, preserving, documenting and presenting digital evidence, stored in the form of encoded information, from digital electronic storage devices, such as computers, Personal Digital Assistance (PDA), digital cameras, mobile phones and various memory storage devices. All must be done in a manner designed to preserve the probative value of the evidence and to assure its admissibility in legal proceeding. The word forensics means “to bring to the court”. Forensics deals primarily with the recovery and analysis of latent evidence. Latent evidence can take many forms, from fingerprints left on a window to deoxyribonucleic acid (DNA) evidence recovered from blood stains to the files on a hard drive. This paper provides a high-level overview on computer forensics investigation phases for both technical and non-technical audience. Although the term “computer” is used, the concept applies to any device capable of storing digital information.

Keywords: Computer Forensic, Digital Evidence, Digital Forensic, Time Stamp, Computer Crime.

1. INTRODUCTION

In a perfect world the need for determining the activity conducted within a computer would not be necessary; however, this is not a perfect world and there are times when it is imperative that the activity of a computer be investigated. There should be a way for an individual to analyze a computer, in times when possible misconduct has occurred. For this reason, computer forensics, a newly developed area of computer science, becomes an increasingly more important aspect daily and will be widely used in the twenty-first century.

The widespread use of computers has caused computer crimes to increase at an alarming rate. Computers have given criminals a new approach to carrying out their misdeeds. After a crime or a questionable act is suspected on a computer, a digital investigation must follow. The investigation is used to determine the scope of the problem. The computers investigated will typically be either those used to commit the crime or those which are the targets of the crime.

Computer security is a vast field that touches all aspects of data confidentiality, integrity and availability for suitably controlling access to data. Access control is only one of the ten domains of Information Systems Security categorized by the International Information Systems Security Certification Consortium (ISC) ^{2} which is responsible for certifying Information Systems Security professionals globally.

Out of the two aspects of security, the proactive comprises of detective, preventive and deterrent measures while the reactive deals with corrective, investigative, recovery and compensatory measures taken to guarantee a certain degree of data assurance. Most of what is studied today in computer security only emphasizes the proactive components. Owing to many factors, investigating root cause analysis and studying computer usages or file structures to determine exploitable trends have never been the norm in most environments.

This paper introduces the reactive part of computer security otherwise called computer forensics. It attempts to serve as an introduction into the vast field of computer forensics. While defining forensic science holistically and introducing such terms as Computer Forensic Investigation process, digital evidence, chain of custody, event reconstruction and digital forensic process.

2. COMPUTER FORENSIC DEFINED

The term forensics derives from the Latin ‘forensis’, which meant in open court or public, “which itself comes from the term forum, referring to an actual location, public square or marketplace used for judicial and other business."

Contemporary use of the word forensics, therefore, generally continues to relate to law, and has come to mean \scientific tests or techniques used with the detection of crime." Thus, computer forensics implies a connection between computers, the scientific method, and crime detection. Digital forensics is largely used interchangeably with computer forensics, but implies the inclusion of devices other than general-purpose computer systems, such as network devices, cell phones, and other devices with embedded systems [1]. However, largely everyone except academic computer science researchers use the term in connection with the law. Many computer scientists have simply been using the word “forensics" as a process of logging, collecting, and auditing or analyzing data in a post hoc investigation." Digital forensics is a branch of forensic science encompassing the recovery and investigation of material found in digital devices, often in relation to computer crime [1] [2]. The term digital forensics was originally used as a synonym for computer forensics but has expanded to cover investigation of all devices capable of storing digital data. With roots in the personal computing revolution of the late 1970s and early '80s, the discipline evolved in a haphazard manner during the 1990s, and it was not until the early 21st century that national policies emerged [1].

Digital forensics investigations have a variety of applications. The most common is to support or refute a hypothesis before criminal or civil courts. Forensics may also feature in the private sector; such as during internal corporate investigations or intrusion investigation.

The technical aspect of an investigation is divided into several sub-branches, relating to the type of digital devices involved; computer forensics, network forensics, database forensics and mobile device forensics. The typical forensic process encompasses the seizure, forensic imaging (acquisition) and analysis of digital media and the production of a report into collected evidence.
Carrier [2] points out that as well as identifying direct evidence of a crime, digital forensics can be used to attribute evidence to specific suspects, confirm alibis or statements, determine intent, identify sources, or authenticate documents.

Beckett [3] asserts that the term computer forensics was in informal use in academic publications from at least 1992; however the term remained informally defined for many years. A commonly cited definition of the field in Australian literature is McKemmish’s [4] definition of forensic computing:

“The process of identifying, preserving, analyzing and presenting digital evidence in a manner that is legally acceptable” [4]

The American Academy of Forensic Sciences defines forensics as follows:

"The word forensic comes from the Latin word forensic: public; to the forum or public discussion; argumentative, rhetorical, belonging to debate or discussion. From there it is a small step to the modern definition of forensic as belonging to, used in or suitable to courts of judicature, or to public discussion or debate. Forensic science is a science used in public, in a court or in the justice system. Any science, used for the purpose of the law, is a forensic science." [5]

This broad definition of forensics and McKemmish's earlier definition inform the definition of computer forensics given by the Scientific Working Group on Digital Evidence (SWGDE), whose definition is [6]:

"The scientific, examination, analysis, and/or evaluation of digital evidence in legal matters."

Researchers attending the first Digital Forensic Research Workshop, 2001 defined Digital Forensic Science as [7]:

"The use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation and presentation of digital evidence derived from digital sources for the purpose of facilitating or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorized actions shown to be disruptive to planned operations"

This broad definition reflects a change in forums in which the techniques of computer forensics are increasingly being applied. While traditionally, computer forensics was exclusively targeted in legal forum, computer forensics is increasingly practiced in non-legal context such as corporate investigations, intelligence and military.

The terms digital forensics, forensic computing and computer forensics are today arguably used interchangeably. Historically, computer forensics and forensic computing related to the interpretation of computer related evidence in courts of law. Technology however does not stand still, nor does language, and the meaning of the term has remained consistently under negotiation. Two factors have been at play underlying this process: the changing state of uptake of digital technologies, and with it moves within organizations to consider governing and regulating the use of information technology [8].

However, many experts feel that a precise definition is not yet possible because digital evidence is recovered from devices that are not traditionally considered to be computers. Some researchers prefer to expand the definition such as definition by Palmer [7] to include the collection and examination of all forms of digital data, including that found in cell phones, PDAs, iPods, and other electronic devices.

It is not just the content of emails, documents and other files which may be of interest to investigators but also the 'metadata' associated with those files [9]. A computer forensic examination may reveal when a document first appeared on a computer, when it was last edited, when it was last saved or printed and which user carried out these actions.

More recently, commercial organizations have used computer forensics to their benefit in a variety of cases such as:

≈ Intellectual Property theft
≈ Industrial espionage
≈ Child exploitation/abuse
≈ Employment disputes
≈ Economic Fraud investigations
≈ Forgeries

- ≈ Matrimonial issues
- ≈ Bankruptcy investigations
- ≈ Inappropriate email and internet use in the work place
- ≈ Regulatory compliance

3. COMPUTER FORENSIC PROCESS

From a technical standpoint, the main goal of computer forensics is to identify, collect, preserve, and analyse data in a way that preserves the integrity of the evidence collected so it can be used effectively in a legal case. Forensic process comprises the following phases [13]:

- Collection
- Examination
- Analysis
- Reporting

3.1 Data Collection Phase

The first step in the forensic process is to identify potential sources of data and acquire data from them. Common data sources include desktop computers, servers, network storage devices, and laptops with internal drives that accept media, such as CDs and DVDs, and also have several types of ports (e.g., Universal Serial Bus (USB), Personal Computer Memory Card International Association (PCMCIA) to which external data storage media and devices can be attached. External storage examples include [14]:

- o Thumb drives
- o Memory and flash cards
- o Optical discs
- o Magnetic disks.

Standard computer systems also contain volatile data that is available until the system is shut down or rebooted.

In addition to computer-related devices, many types of portable digital devices (e.g., PDAs, cell phones, digital cameras, digital recorders, and audio players) may also contain data.

Analysts should be able to survey a physical area, such as an office, and recognize the possible sources of data.

Once exhibits have been seized a forensic duplicate of the media is created, usually via a write blocking device, a process referred to as Imaging or Acquisition [10]. The duplicate is created using a hard-drive duplicator or software imaging tools such as DCFLdd, IXimager, Guymager, TrueBack, EnCase, FTK Imager or FDAS. The original drive is then returned to secure storage to prevent tampering.

The acquired image is verified by using the SHA-1 or MD5 hash functions. At critical points throughout the analysis, the media is verified again, known as "hashing", to ensure that the evidence is still in its original state [16].

Before the analyst begins to collect any data, a decision should be made by the analyst or management on the need to collect and preserve evidence in a way that supports its use in future legal or internal disciplinary proceedings. Furthermore, a clearly defined chain of custody should be followed by keeping a log of every person who had physical custody of the evidence, documenting the actions that they performed on the evidence and at what time, storing the evidence in a secure location when it is not being used, making a copy of the evidence and performing examination and analysis using only the copied evidence, and verifying the integrity of the original and copied evidence [17].

Some proactive measures taken by organizations to collect data for forensic purposes include:

- Configuration of most operating systems (OSs) to audit and record certain event types, such as authentication attempts and security policy changes, as part of normal operations.
- Implementation of centralized logging. Certain systems and applications forward copies of their logs to secure central log servers.
- Performing regular system backups allows analysts to view the contents of the system as they were at a particular time.

In addition, security monitoring controls such as intrusion detection software, antivirus software, and spyware detection and removal utilities can generate logs that show when and how an attack or intrusion took place.

3.2 Examination

The examination process helps make the evidence visible and explain its origin and significance. This process should accomplish several things. First, it should document the content and state of the evidence in its totality. Such documentation allows all parties to discover what is contained in the evidence [16]. Included in this process is the search for information that may be hidden or obscured. Once all the information is visible, the process of data reduction can begin, there by separating the "what" from the "chaff". Giving the tremendous amount of information that can be stored on computer storage media, this part of the examination is critical.

3.3 Analysis

After acquisition the contents of image files are analysed to identify evidence that either supports or contradicts a hypothesis or for signs of tampering (to hide data) [11].

During the analysis an investigator usually recovers evidence material using a number of different methodologies (and tools), often beginning with recovery of deleted material. Examiners use specialist tools (EnCase, ILOOKIX, FTK, etc.) to aid with viewing and recovering data. The type of data recovered varies depending on the investigation; but examples include email, chat logs, images, internet history or documents. The data can be recovered from accessible disk space, deleted (unallocated) space or from within operating system cache files [9].

Once evidence is recovered the information is analysed to reconstruct events or actions and to reach conclusions, work that can often be performed by less specialist staff. Digital investigators, particularly in criminal investigations, have to ensure that conclusions are based upon data and their own expert knowledge **[9]**. In the US, for example, Federal Rules of Evidence state that a qualified expert may testify "in the form of an opinion or otherwise" so long as [12]:

(1) The testimony is based upon sufficient facts or data.

(2) The testimony is the product of reliable principles and methods.

(3) The witness has applied the principles and methods reliably to the facts of the case.

3.4 Reporting

Once the computer forensic analysis is complete, presenting an understandable, defendable and complete report is key. The evidentiary packages produced must be complete, easy to understand and always explained in precise detail. The addition of relationship charts, entity explanations, timelines, histories and mail-thread analysis gives a clear understanding of the issue, as well as the players [15]. The analyst shall be able to defend the process and testify to the methodologies used relating to the facts in a case, when questions start getting tough during expert witness testimony.

When completed reports are usually passed to those commissioning the investigation, such as law enforcement (for criminal cases) or the employing company (in civil cases), who will then decide whether to use the evidence in court. Generally, for a criminal court, the report package will consist of a written expert conclusion of the evidence as well as the evidence itself [9].

4. CONCLUSION

Given the enormity of task in cybercrime control and policing, the absence of dearth of trained and qualified computer forensics law enforcement officers, there is urgent need for the Federal Government to pay attention to the training of adequate EFCC and police officers in the computer forensic sciences to enhance effective policing of the ever increasing cyber criminals. The problem is serious, particularly now that the Federal Government has passed the information Technology Bill and Electronic Evidence Act for this purpose. A law made but cannot be enforced is no law. Cyber criminals will be forced to retreat if a large percentage of fraudsters are arrested, prosecuted and punished at first attempt. It is strongly recommended that Polytechnics and Universities should establish Computer Forensics certificate, diploma and degree courses to meet the ever-increasing demand for this type of urgently needed personnel. The provision of adequately qualified experts will beef up their deployment in the police and military. This may well be antidote to the fast eroding confidence in e-commerce and international trade in Nigeria.

In this paper, we have reviewed the literatures in computer forensics and identified the four (4) main phases of computer forensics investigation process: *Collection*, *Analysis*, *Examination* and *Reporting*.

Our future research will focus on Computer forensic Investigation Process Models, where we shall apply a risk based approach in computer forensic investigation. The legal aspect on computer forensics is an interesting area that should be further investigated [14]. Probable research area might be a way to categorize and approve computer forensic tools for certain investigations and situations? How cross-country investigations are handled, and how are differences between the countries rules and regulations managed?

5. REFERENCES

[1] M Reith, C Carr and G Gunsch, (2002).“Anexamination of Digital Forensic models”. International Journal of Digital Evidence [online]. Available at: www.acm.org [Accessed on 15/10/2012].

[2] B. Carrier, (2001b). “Defining Digital Forensic examination and Analysis Tools” Digital Research Workshop II. Available at: www.acm.org [Accessed on 15/10/2012].

[3] J. Beckett, J., “Digital Forensics: Validation and Verification in a Dynamic Work Environment”. 40th Annual Hawaii International Conference on System Science. 2007: Hawaii.

[4] R. McKemmish, 1999, “What is Forensic Computing?” Trends and Issues in Crime and Criminal Justice, Australian Institute of Criminology. Available at: http://aic.gov.au/documents/9/C/A/%7B9CA41AE8-EADB-4BBF-9894-64E0DF87BDF7%7Dti118.pdf [Accessed on 20/10/2012]

[5] AAFS. “So you want to be a forensic scientist?” American Academy of Forensic Science, Available at: http://www.aafs.org/default.asp?section_id=resources&page_id=choosing_a_career

[6] N.L Beebe and J.G Clark. “A Hierarchical, Objectives-Based Framework for Digital Investigations Process”, 4th Digital Forensics Research Workshop. 2004: Baltimore, MD.

[7] G. Palmer, G. “A road Map for Digital Forensic Research, in First Digital Forensic Research Workshop”, G. Palmer, Editor. 2001: Ucita, New York. www.acm.org.

[8] B. Schatz. “Digital Evidence: Representation and Assurance”. Doctorate Thesis, Submitted to Information Security Institute, Faculty of Information Technology, Queensland University of Technology, 2010.

[9] C. Eoghan (2004). "Digital Evidence and Computer Crime, Second Edition". Elsevier. ISBN 0-12-163104-4.

[10] A. Richard, H. Val and M. Graham (2012). "'The Advanced Data Acquisition Model (ADAM): A process model for digital forensic practice" Journal of Digital Forensics, Security and Law, Vol. 8(4).

[11] B. Carrier (2001) "Defining digital forensic examination and analysis tools". Digital Research Workshop II. CiteSeerX: 10.1.1.14.8953.

[12] Federal Rule of Evidence. Available at; http://federalevidence.com/rules-of-evidence#Rule702 Retrieved on 2nd May, 2014.

[13] A. J. Marcella, Jr. and D. Menendez. "Cyber Forensics: A Field Manual for Collecting, Examining, and Preserving Evidence of Computer Crimes," (2nd Edition): Taylor & Francis Group, LLC. 2008.

[14] E.O. Nwachukwu, A Mohammed. D.C. Igweze and V.O. Ewulonu "Microsoft Windows Based Computer Forensic" International Journal of Information Technology and Business Management; 2927(1):2012-2014.

[15] National Institute of Justice. "Electronic Crime Scene Investigation: A Guide for First Responder". Washington, D.C.: U.S. Department of Justice, National Institute of Justice, 2004. NCJ 187736. http://www.ojp.usdoj.gov/nij.

[16] S. Garfinkel, "Forensic Feature Extraction and Cross-Drive Analysis". The 6th Annual Digital Forensic Research Workshop Lafayette, Indiana, August 14-16, 2006.

[17] J. E. Regan. "The Forensic Potential of Flash Memory". Master's Thesis. Naval Postgraduate School, Monterey, CA, 2009.

[18] F. Adelstein. "MFP: The Mobile Forensic Platform". International Journal Of Digital Evidence, Spring 2003, Volume 2. Issue 1.

Outliers In Data Envelopment Analysis

Shaik Khaleel Ahamed *khaleelska@gmail.com*
Research Scholar, C.S.E.Dept
S.V.U.College of Engineering
S.V. University
Tirupati, A.P, 517501, India

Prof.MM Naidu *mmnaidu@yahoo.com*
Professor, C.S.E. Dept
S.V.U.College of Engineering
S.V. University
Tirupati, A.P, 517501, India

Prof.C.Subba Rami Reddy *csruma@yahoo.com*
Professor, Statistics.Dept
S.V. University
Tirupati, A.P, 517501, India

Abstract

Data Envelopment Analysis is a linear programming technique that assigns efficiency scores to firms engaged in producing similar outputs employing similar inputs. Extremely efficient firms are potential Outliers. The method developed detects Outliers, implementing Stochastic Threshold Value, with computational ease. It is useful in data filtering in BIG DATA problems.

Keywords: Constant Return to Scale, Data Envelopment Analysis, Super Efficiency, Threshold Value.

1. INTRODUCTION

An ‘outlier’ is an observation that is radically dissimilar with majority of observations. It falls outside a cloud of normal observations. The presence of an outlier may be due to reporting errors. Such observations shall be corrected or removed for a valid empirical analysis and consequent conclusions. If an outlier arrives from the same probability distribution as others, they do occur with small probability. Such observations shall be carefully examined since they carry special information that cannot be retrieved from the normal observations. Outliers do not possess any item in a neighborhood of a specified radius. Detection of outliers is constituted by two sub problems.

(i) Define inconsistency in a data set and
(ii) To provide an efficient method to identify the inconsistent observations (outliers).

2. DATA ENVELOPMENT ANALYSIS

Data Envelopment Analysis is a linear programming technique that measures efficiency of decision making units. In efficiency evaluation production plans are projected onto the envelopment frontier determined by the most efficient observations that are potential outliers. Outliers elevate the frontier leading to the under estimation of efficiency scores of inefficient decision making units. Charnes, Cooper and Rhodes (1978) proposed a technology set that is based on the axioms of inclusion, free disposability and minimum extrapolation, whose boundary serves as envelopment frontier that admits constant returns to scale. The efficiency scores of interior production units are under estimated in the presence of outliers in the CCR (1978) model. Banker, Charnes and Cooper (BCC, 1984) extended the CCR model, whose production possibility set is based on the axioms of inclusion, convexity, free disposability and minimum extrapolation. The extremely efficient decision making units are potential Outliers.

3. DEA – Outliers

a) Timmer (1971) was the first one to recognize high sensitivity of DEA scores when outliers are present, in linear programming problems. By suitably finding the threshold value, a specified percent of firms were removed from the reference set to arrive at output elasticities with respect to inputs, in the frame work of Cobb-Douglass production function (1928), with acceptable magnitudes. The deleted input and output plans are viewed as Outliers**. The percentage of firms removed from the data is subjective.**

b) In DEA all efficient decision making units are flagged as potential outliers. The efficiency score of efficient firms is 100%. Andersen and Petersen (1993) suitably tailored the DEA constraints to assess super efficiency scores of efficient firms. Such production unit with larger efficiency score (input approach) is ranked better. The input super efficiency score is larger than or equal to unity, for such production plans. In their approach firm's input and output vector, whose efficiency is under evaluation, is removed from the reference set, the assessed DMU being efficient. Consequently, the input vector falls below the input efficient frontier and the deletion pushes the frontier upwards, toward inefficient units all producing a given level or more of an output. Deletion of an efficient production plan from the reference set leads to the contraction of input sets. Such input efficient decision making unit whose deletion from the reference set resulted in maximum contraction of input set is the most influential observation, possibly an outlier (refer to the figure). The property of frontier displacement refers to efficient decision making units. If the input and output combination of efficient firm is removed from the reference set, for the same firm its production plan is projected on to the constrained frontier. If input orientation is pursued this score emerges to be one or more than one. Suppose the input efficiency score is 1.5, then this score is interpreted as, that this firm will continue to be efficient in the presence of input expansion up to a factor 1.5. This approach can be extended in a straight forward manner to output and graph orientation. **The super efficiency measurement above gives a single measurement of irregular polyhedron. The threshold value to identify outliers is due to subjective choice.**

c) Wilson (1995) identified outliers following leave-one-out approach, and the search was in relation to efficient frontier, under exclusively input perspective and output perspective. **Wilson's method requires more computational labour while his threshold value is subjective.**

d) Simar (2003) suggests that a production plan shall be treated as an outlier if it is sufficiently influential under both orientations (input and output). **His threshold values to identify outliers are subjective.**

e) Tran et.al (2008) proposed a new method for detecting outliers in Data Envelopment Analysis. They consider the CCR-DEA formulation and the observed plans which determined the CCR frontier as potential outliers. Their approach depended on the intensity parameters of efficient firms arrived at construction of the DEA hull. With reference to CCR-DEA hull the intensity parameters are non-negative. If a firm is inefficient, its intensity parameter is assigned with a zero value by every firm, including itself. An efficient firm evaluated relatively efficient by itself may participate in the construction of DEA frontier for the evaluation of inefficient decision making units, there by possess positive intensity parameters. An efficient firm that appears the most with positive intensity parameter values while inefficient firms are evaluated may be viewed as an influential observation. For identification of outlier not only the count of positive intensity parameter values is important as metric but their sum can also be used as another metric. Stosic and Sampario de souza (2003) proposed a method which is based on a combination of a boot strap and resampling schemes for automatic detection of outliers, which takes into consideration the concept of leverage. The leverage metric measures the effect produced on the efficiency scores of all others DMUs, when a particular firm is removed from the data set. Outliers are expected to display leverage much above the mean leverage and hence should be selected with lower probability than the other DMUs when resampling is performed.

f) Sampario de Souza et.al (2005) defined the leverage of jth DMU as,

$$l_j = \sqrt{\frac{\sum_{k=1,k+j}^{n} \left(\theta_{kj}^{*} - \theta_k\right)^2}{n-1}}$$

where θ_{kj}^{*} is the efficiency score of kth DMU based on the data set from which jth DMU's production plan is removed, and θ_k is efficiency score of kth DMU. **Based on unaltered data set, one can compute mean leverage, in boot strap samples choice of threshold value being subjective.**

g) Johnson et.al (2008) believed outliers are found not only among extremely efficient but also inefficient observations. The leverage of an input and output observation to displace the frontier is chosen as a metric to identify an outlier both in efficiency and inefficiency perspectives. The leverage estimate is provided by super efficiency and super inefficiency score. **For this purpose the efficient and inefficient frontiers are used, which bind the production possibility set from above and below, the choice of threshold value is subjective.**

h) Chen and Johnson (2010) formulated an alternative to the above approach. They consider Hull that satisfies the axioms of inclusion and convexity. The axiom of free disposability is withdrawn, on which the convex Hull is built. The methodology developed to identify outliers is similar to the super efficiency evaluation proposed by Andersen and Petersen (1993). The leverage of a DMU to contract the production possibility set while its input vector and output vector are removed from the reference technology determines if the DMU under evaluation is outlier or not. Removal of free disposability axiom, removes the weak efficient subset of the DEA production possibility set from the reference technology, overall boundary shift attributed to an efficient decision making unit serves as a metric to classify it as an outlier or not**. The threshold value is subjective and the method involves greater computational labour.**

4. NEW METHOD- ITS MERITS OVER OTHER METHODS

The proposed study is an attempt to identify outliers in a scenario that there are n production units combining m similar inputs to produce s similar outputs. The production units may be profitable or non-profitable organizations. The input and output vectors of the production units spin a production possibility set under the axioms of inclusion, free disposability, closure under ray expansion and contraction and minimum extrapolation. The production units can be decomposed into four disjoint sets constituted by, (i) extremely efficient, (ii) efficient, (iii) weakely efficient and (iv) inefficient. The surface of the pp set is spun by the extremely efficient ones. All the extremely efficient firms constitute the reference technology of production process. If the input and output vectors of an extremely efficient firm is deleted from the reference technology then the production possibility set experiences contraction. The new pp set is a subset of the original pp set. An inefficient firm's input and output vectors deletion leaves the pp set intact. The potential outliers are the extremely efficient firms. An important direction in the attempt to identify outliers is suggested by Andersen and Petersen (1993) through their super efficiency measurement problem. Their approach reveals such extremely efficient firm with the largest (smallest) super efficiency score under input (output) orientation is certainly an outlier. In this method for identification of outliers, a threshold value needs to be specified which is subjective. Further, super efficiency score provides one measurement of an irregular polyhedran that accounts for contracted region. When an extremely efficient firm's input and output vectors are deleted from the reference technology, for some inefficient firms, their efficiency scores will increase and for the remaining inefficient firms, their efficiency scores would be intact. The increments of efficiency-scores of inefficient firms provide additional measurements of contracted region embedded in an irregular polyhedron. These additional measurements combined with the difference between the super efficiency score and unity provides a means to obtain statistically based threshold value that facilitates outliers identification. The various methods of outlier identification outlined in the review suffer from subjective threshold value and heavy computational labour. **The merits of the new method are that the threshold value is statistically determined, requires least computational labour. This method is of immence use in data filtering in problems that constitute inputs and outputs with a monotonic relationship between inputs and outputs, particularly useful in BIG DATA problems.**

4.1 Data Envelopment Analysis-Constant Return To Scale-Outliers

Charnes, Cooper and Rhodes (1978) proposed a fractional programming problem to measure technical efficiency of decision making units. Applying Charnes and Cooper transformation, this problem can be transformed into a linear programming problem. Under input perspective the optimal solution not only assigns a technical efficiency score to each decision making unit, but provides such scores to its peer DMUs that are based upon the input and output weights of the decision making unit for which the CCR-DEA problem is solved.

Let $x_{ij,} i \in I; y_{rj,} r \in S$ be the inputs and outputs of the decision making unit $j \in J$. For j=0, the following CCR problem is solved:

$$\delta_0^1 = \max \sum_{r=1}^{s} v_r y_{r0}$$

$$\text{s.t} \sum_{i=1}^{m} u_i x_{i0} = 1 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

$$\sum_{r=1}^{s} v_r y_{rj} - \sum_{i=1}^{m} u_i x_{ij} \le 0, \forall j \in J$$

$$v_r \ge 0, r \in S; u_i \ge 0, i \in I$$

For efficient decision making units δ_0^1 =1 and the corresponding slack is zero for $j = 0 \in J$.

The potential decision making units are the efficient ones. Solving the above problem for each decision making unit, efficient firms can be identified. These firms are potential super efficient. To assess super efficiency of extremely efficient decision making units. Andersen and Petersen (1993) formulated an input oriented envelopment problem.

$$\delta_0^2 = \min \lambda$$

$$\text{s.t} \sum_{\substack{j=1 \\ j \ne 0}}^{n} \lambda_j x_{ij} \le \lambda x_{i0}, i \in I \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

$$\sum_{\substack{j=1 \\ j \ne 0}}^{n} \lambda_j y_{rj} \ge y_{r0}, r \in S$$

$$\lambda_j \ge 0, \forall j \in J - \{0\}$$

i) The super efficiency problem is solved for the extremely efficient decision making units.
ii) Super efficiency score measures the ability of an extremely efficient decision making unit to remain efficient in the event of further radial augmentation of inputs upto some degree.

iii) Under constant return to scale frame work the super efficiency problem is always feasible if input and output values are positive.
iv) Super efficiency score reveals the ability of the firm to contract the production possibility set.
v) The dual of the above envelopment problem is,

$$\delta_0^2 = \max \sum_{r=1}^{s} v_r y_{r0}$$

$$\text{s.t} \sum_{i=1}^{m} u_i x_{i0} = 1 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ (3)$$

$$\sum_{r=1}^{s} v_r y_{rj} - \sum_{i=1}^{m} u_i x_{ij} \le 0, j \in j - \{0\}$$

$$v_r \ge 0, r \in S$$

$$u_i \ge 0, i \in I$$

The optimal solution of (1) is a feasible solution of (2). Therefore,

$$\delta_0^2 \ge \delta_0^1$$

For extremely efficient firm, $\delta_0^1 = 1 \Rightarrow \delta_0^2 \ge 1$.

Problem (1) and (3) can be equivalently expressed as,

$$\delta_0^1 = \max \frac{\sum_{r=1}^{s} v_r y_{ro}}{\sum_{i=1}^{m} u_i x_{io}}$$

$$\text{s.t} \ \frac{\sum_{r=1}^{s} v_r y_{rj}}{\sum_{i=1}^{m} u_i x_{ij}} \le 1, j \in J \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ (4)$$

$$v_r \ge 0, r \in S; u_i \ge 0, i \in I$$

$$\delta_0^2 = \max \frac{\sum_{r=1}^{s} v_r y_{ro}}{\sum_{i=1}^{m} u_i x_{io}}$$

$$\text{s.t} \ \frac{\sum_{r=1}^{s} v_r y_{rj}}{\sum_{i=1}^{m} u_i x_{ij}} \le 1, j \in J - \{0\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ (5)$$

$$v_r \ge 0, r \in S; u_i \ge 0, i \in I$$

Applying Charnes and Cooper transformation problem (4) and (5) can be reduced to (1) and (3) respectively.

Every feasible solution of program (4) is a feasible solution of (5). If $\left(\overline{v},\overline{u}\right)$ and $\left(\overline{\overline{v}},\overline{\overline{u}}\right)$ are optimal solutions of (4) and (5) respectively, then we have,

$$\frac{\sum_{r=1}^{s}\overline{v}_r y_{rj}}{\sum_{i=1}^{m}\overline{u}_i x_{ij}} \le \frac{\sum_{r=1}^{s}\overline{\overline{v}}_r y_{rj}}{\sum_{i=1}^{m}\overline{\overline{u}}_i x_{ij}} \le 1,\ j\in J$$

$$\Rightarrow \frac{OD^{'}}{OD} \le \frac{OD^{''}}{OD}$$

$$\frac{OE^{'}}{OE} \le \frac{OE^{''}}{OE}$$

$$\frac{OF^{'}}{OF} \le \frac{OF^{''}}{OF}$$

For j=0, $\dfrac{\sum_{r=1}^{s}\overline{v}_r y_{ro}}{\sum_{i=1}^{m}\overline{u}_i x_{io}} \le \dfrac{\sum_{r=1}^{s}\overline{\overline{v}}_r y_{ro}}{\sum_{i=1}^{m}\overline{\overline{u}}_i x_{io}}$

since this firm is efficient, $\dfrac{\sum_{r=1}^{s}\overline{v}_r y_{ro}}{\sum_{i=1}^{m}\overline{u}_i x_{io}} = 1$

$$\frac{\sum_{r=1}^{s}\overline{\overline{v}}_r y_{ro}}{\sum_{i=1}^{m}\overline{\overline{u}}_i x_{io}} \ge 1$$

$$\frac{OB^{'}}{OB} \ge 1$$

$$\frac{OB^{'}}{OB} - d_B = 1$$

$$d_B = \frac{OB^{'}}{OB} - 1$$

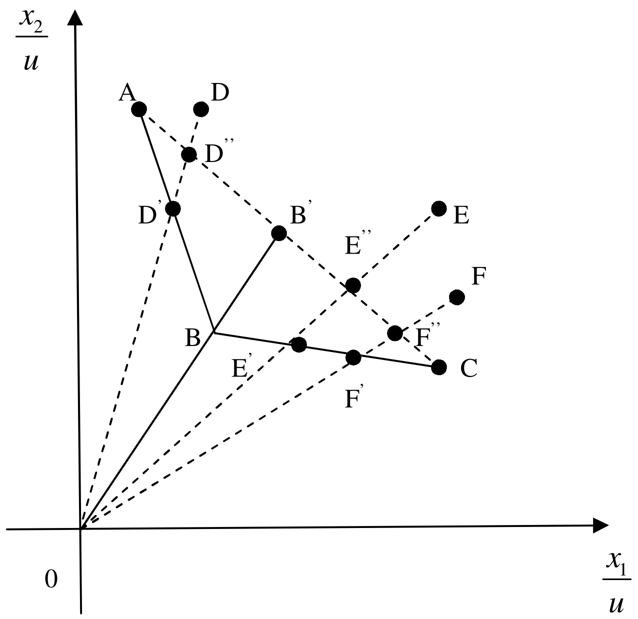


FIGURE 1: Unit Output Isoquant.

In the figure above first and second input requirements to produce unit output are measured along horizontal and vertical axes respectively. The input isoquant is determined by the extremely efficient firms A,B and C. the firms D,E and F are inefficient for which the firm B is an efficient peer, solving problem(1) for firm B, its standard efficiency score and cross efficiency scores for the remaining decision making units can be obtained. The cross efficiency scores are as follows: $\frac{OD^{'}}{OD},\frac{OE^{'}}{OE},\frac{OF^{'}}{OF}$. Such efficiency scores of a firm evaluated with other firm's efficiency scores are called cross efficiency scores.

Solving the super efficiency problem (3), super efficiency scores for firm B and cross efficiency scores for other firms can be obtained. The cross efficiency scores of other firms are,

$$\frac{OD^{''}}{OD},\frac{OE^{''}}{OE},\frac{OF^{''}}{OF}$$

$$\frac{OD^{''}}{OD}\geq\frac{OD^{'}}{OD}$$

$$\frac{OE^{''}}{OE}\geq\frac{OE^{'}}{OE}$$

$$\frac{OF^{''}}{OF}\geq\frac{OF^{'}}{OF}$$

The area of the triangle ABC measures the contraction of the production possibility set. The super efficiency score of B, provides one measurement of contracted production possibility set,

$d_B=\frac{OB^{'}}{OB}-1$ that lies between zero and one.

d_B gives a measurement of production possibility set contraction.

Define

$$d_D = \frac{OD''}{OD} - \frac{OD'}{OD}$$

$$d_E = \frac{OE''}{OE} - \frac{OE'}{OE}$$

$$d_F = \frac{OF''}{OF} - \frac{OF'}{OF}$$

$d_{D,}d_E$ and d_F are also measurements of contraction of the production possibility set. We take average of all these measurements to arrive at a more meaning full measure of contraction.

$$\overline{d_B} = \frac{d_B + d_D + d_E + d_F}{\eta_B}, \text{ where } \eta_B = 4$$

The above arithmetic mean gives rise to a Student t-test, in which $\overline{d}$ is tested against zero, if sample size is small

$$t_B = \frac{\overline{d_B}}{s/\sqrt{\eta_B}}$$ follows Student's t-distribution with $\eta_B - 1$ degrees of freedom.

If $\overline{d_B} \geq t_\alpha \frac{s}{\sqrt{\eta_B}}$, then firm B is an outlier, where α is the level of significance.

If there are other decision making units that are inefficient and for which firm B is not an efficient peer, for such firms problems (1) and (3) assign the same efficiency scores, so that their deviations vanish.

(i) For outlier determination a threshold value is needed, whose choice often subjective. This method provides a threshold value $t_\alpha \frac{s}{\sqrt{\eta_B}}$ that is statistically determined which depends upon the level of significance.

(ii) Further, this method need not choose every extremely efficient decision making unit as an outlier.

(iii) It is a common practice to identify large super efficient firms as outliers, 'how large' is a subjective matter.

(iv) For the identification of an outlier this method uses not only the super efficiency scores, but also the potential improvements of efficiency of inefficient decision making units.

5. FUTURE RESEARCH DIRECTION

Economic data often are subjected to returns to scale. Returns to scale may be constant, increasing or decreasing. The present study assumes constant returns to scale. The super efficiency problems are always feasible, if input and output values are positive and returns to scale are constant. However, if return to scale are either increasing or decreasing it is likely that for some extremely efficient firms their super efficiency problems are infeasible. A natural extension of the present study is identification of outliers, suitably fine tuning the super efficiency problems to be free from infeasibility, in the presence of non-constant returns to scale.

6. REFERENCES

[1] Andersen, P. and N. C. Petersen. (1993). "A Procedure for Ranking Efficient Units in Data Envelopment Analysis." *Management Science*, 39:1261-1264.

[2] Charnes, W.W. Cooper, Z.M. Huang and D.B. Sun, Polyhedral cone-ratio DEA models with an illustrative application to large commercial bank, Journal of Economics 46 (1990) 73-91.

[3] Banker, Charnes and Cooper (1984).”Estimating Most Productive Scale Size Using Data Envelopment Analysis.” *European Journal Of Operations Research* 35-44

[4] Charnes, A., Cooper W.W., and Rhodes, E., (1978), “Measuring the Efficiency of Decision-Making Units”, European Journal of Operations Research, 2, 429-444.

[5] Chen and Johnson (2010) ; “A Unified model for detecting Outliers in DEA, Computers and Operations Research, Vol. 37. 417-425.

[6] Daraio, C. and L. Simar (2003), Introducing environmental variables in nonparametric frontier models: a probabilistic approach, Discussion paper 0313, Institute de Statistique, Universities Catholique de Louvain, Belgium.

[7] Johnson, A.L., Chen W.C., McGinnis, L.F., (2008).,” Internet-based benchmarking for warehouse operations”. Working Paper, 2008.

[8] J.R. Doyle and R.H. Green, Efficiency and cross-efficiency in DEA: derivations, meanings and uses, Journal of Operational Research Society 45 (1994) 567-578.

[9] J.H. Dula and B.L. Hickman, Effects of excluding the column being scored from the DEA envelopment LP technology matrix, Journal of Operational Research Society 48 (1997) 1001-1012.

[10] J. Zhu, Robustness of the efficient DMUs in data envelopment analysis, European Journal of Operational Research 90 (1996) 451-460.

[11] J. Zhu, Super-efficiency and DEA sensitivity analysis, European Journal of Operational Research 129 (2001) 443-455.

[12] M. Halme and P. Korhonen, Restrciting weights in value efficiency analysis, European Journal of Operational Research 126 (2000) 175-188.

[13] P.C. Pendharkar, “A Data Envelopment Analysis-Based Approach for Data Preprocessing,” IEEE Transactions on Knowledge & Data Engineering, Vol. 17, No. 10, 2005, pp. 1379-1388.

[14] R.G. Dyson and E. Thanassoulis, Reducing weight flexibility in data envelopment analysis, Journal of Operational Research Society 39 (1988) 563-576.

[15] R. Green, J.R. Doyle and W.D. Cook, preference voting and project ranking using DEA and cross-evaluation, European Journal of the Operational Research 90 (1996) 461-472.

[16] Stosic, B. and Sampaio de Sousa, M.C. (2003) “Jackstrapping Dea Scores For Robust Efficiency Measurement.” Series Texto para Discussão N° 291, Universidade de Brasília.

[17] S. Talluri and J. Sarkis, Extensions in efficiency measurement of alternate machine component grouping solutions via data envelopment analysis, IEEE Transactions on Engineering Management 44 (1997) 27-31.

[18] Tran,N.M., Sheverly,G., and Preckel,P., (2008) “ A New Method for detecting Outliers in DEA”, Applied Economic Letters, 1-4.

[19] T. R. Anderson, A. Uslu, and K. B. Hollingsworth, "Revisiting extensions in efficiency measurement of alternate machine component grouping solutions via data envelopment analysis," Working paper 1998.

[20] Timmer, C. Peter,(1971)," Using a probabilistic frontier production function to measure technical efficiency", Journal of Political Economy 79, 776-794.

[21] Wilson, P. W. (1995) "Protecting Influential Observations in Data Envelopment Analysis." Journal of Productivity Analysis, 4:27–45.

Fast Mobile IPv6 Handover Using Link and Location Information

Mahmud Mansour *mansour30@hotmail.com*
Faculty of Information Technology
Department of Network Engineering
Tripoli University
Tripoli, Libya

Mohamed Alnas *M.J.R.Alnas@bradford.ac.uk*
School of Informatics, Mobile Computing
Network and Security Research Group
University of Bradford
Bradford, UK

Abstract

There are two causes of latency in mobile handover: the move detection latency and registration latency. This delay is inherent in the round-trip incurred by Mobile IP as the registration request is sent to the home agent and the response sent back to the foreign agent. Throughout the time between Mobile Node (MN) leaving the old foreign network (oFN) and Home Agent (HA) receiving the MN registration message, HA does not know MN's latest Care of Address (CoA), and therefore it still forwards the packets destined for MN to the old foreign network. These packets will be discarded and lost.

This paper present an improved link layer mechanism with Location information Provider. Global position systems GPS used to assist FMIPv6 for fast handovers and reduced packet loss during handover. We introduce a new link layer combined with Location information Provider signalling in this algorithm accordingly. Further, we report the implementation details performed through simulations.

Therefore, link layer information and Location information Provider allows an MN and FAs to predict the loss of connectivity more quickly than the L3 advertisement based algorithm. The simulations evidence performance improvements in terms of latency and packet loss. It is also shown that by enabling Location information Provider inside the FA discovery method and improving link layer event services, an MN can be well prepared for handover and perform faster movements.

Keywords- Mobile IP, Link Layer Information, Global Position Systems, Fast Handover, Handover Latency, Packet Loss.

1. INTRODUCTION

The primary purpose of IP is to keep data packets delivering between hosts in the Internet. Mobile IP is an Internet standards protocol, proposed by the Internet Engineering Task Force (IETF), which enhances the existing IP to accommodate mobility [1, 2].

The most important functions in mobile IP is the addressing. Because host in the Internet must has a unique IP address, which species its location. Such an address consists of a network address and a host address. Mobile IP allows a MN to communicate with other nodes after changing its link-layer point of attachment from one Access point to another without changing the MN’s address [2,3]. The MN perform handovers between access Points while still using the preserving IP Address. Therefore, packets may be routed to it using this address regardless of the MN’s current point of attachment to the Internet [1,4].

Mobile IPv6 [1] is a protocol proposed to develop as a subset of IPv6 to support mobile connections; Mobile IPv6 allows mobile nodes to change their point of attachment whilst not breaking existing application sessions. Each MN is always identified by its home address, regardless of its current point of attachment to the Internet. While situated away from its home, a MN is also associated with a Care-of-Address (CoA), which provides information about the MN's current location. Packets destined to the MN's home address are transparently routed to its CoA. When the MN changes its point of attachment to the Internet, a handover occurs. The handover mechanism provided in the Mobile IPv6 causes latency, which makes the MN unreachable for a period of time [16, 17].

The protocol allows IPv6 nodes to cache the binding of a MN’s home address with its CoA, and then to send any packets destined for the MN directly to it at this CoA. MIPv6 offers a solution to solve the IP mobility, but due to intolerable high data lost rate and long handover latency. A new protocol, called Hierarchical Mobile IPv6 has been proposed by the RFC 4140 [3] document, which spread out Mobile IPv6 to allow for both micro mobility and macro mobility handling. Hierarchical Mobile IPv6 (HMIPv6) proposal suggests a Mobile Anchor Point (MAP) to act as a local HA to reduce signalling delays in handovers. However, the handover delays still remain unacceptable for some applications.

When a MN changes its point of attachment to the network, it moves from one network (Old network) to another new network and this process is known as handover. During the handover process, the MN usually has disconnected from the old network before connecting to the new network and thus there is a time when the MN has lost connectivity to the Internet. During this period time, it cannot send or receive IP packets to maintain existing application sessions, because of the link switching delay and this time period known as handover latency, it is the primary cause of packet loss.

The latency due to a handover using basic MIPv6 is directly proportional to the minimum round-trip time necessary for a binding update (BU) to reach either the home agent (HA), the correspondent node (CN) or old location in case forwarding from old location is allowed. The interruption time starts in the moment that the mobile node (MN) does not listen to the old location anymore and finishes when the first packet arrives via the new location either from the HA, CN or old location [2].Therefore, these packets may be lost and need to be retransmitted [3, 4].

There are two causes of latency in mobile handover: the move detection latency and registration latency. This delay is inherent in the round-trip incurred by Mobile IP as the registration request is sent to the home agent and the response sent back to the foreign agent. Moreover, there is a high Mobile IP handover delay because of the agent discovery and registration periods, eventually Mobile IP handover can cause significant performance degradation, especially in large-scale mobility environments. Mobile IP can use link layer information to force a handover to a new access network before any mobility at the network layer detected [2]. The Handover decisions based on movement calculation eliminate the need to wait for beacon signals from other FAs. In this paper, we propose the use of link-layer information combined with the global position systems (GPS) in every FA (Location Information of FAs) that can detect the direction of the MN to the new network agent. The link-layer trigger and the Location information of FAs enhance the overall performance of the Mobile IP handover [9].

2. MOVEMENT DETECTION

Movement detection is one of the most important operations performed by the MN, because it is used to discover the handover. To achieve this goal the MN will use any mechanism to detect its movement from one link to another. The Standard movement detection mechanism defined in Mobile IPv6 uses services defined in IPv6 Neighbor Discovery. Additional information provided by other mechanisms can be used besides the one provided by the standard mechanism in order to facilitate the movement detection [2]. The movement of the MN can be detected by using Location information that install inside all FAs. Therefore, FAs can discover the direction of the mobile node and the address of the new foreign agent that MN will move. This will reduce the delay of the registration [10].

3. HANDOVER LATENCY

The Handover latency is the most important issue in mobility network. It refers to the ability of the network to allow a call in progress to continue as the MN continues to travel and change its point of attachment. The handover refer to the time between the reception of the last packet through the old FA (oFA) and reception of the first packet through the new FA (nFA) [5, 6, 7]. During this time the MN, start disconnecting from the old network and start new registration with the new network while packet still forwards to the old network; these packets will be discarded and lost. The packet losses could cause impossible disruptions for real-time services, degrade the QoS and lead to severe performance deteriorations of upper layer protocols, especially when the handover is frequent and the distance between MN and the HA is great [8,9,10].

In general, handover can be classified as either proactive or reactive. Proactive handovers utilize link layer triggers to support the MN in determining that a handover is about to happen and establish packet flow to the target access point prior to the handover event, i.e. requires link layer coupling. This covers a hybrid of mobile assisted and mobile controlled handover types. Reactive schemes only follow the base mobile IP movement methods [13].

4. LINK LAYER INFORMATION AND LOCATION INFORMATION PROVIDER

The main reason to use the link layer to improve the handover delay and packet lost. This can be achieved by providing the information of the link layer; the MN can predict its connectivity more quickly than Network Layer advertisement-based algorithms. Therefore, it used this information to predict the breakdown of the link layer before is broken. This facilitates the execution of the handover, and the elimination of the time to detect handover. [11,12].

The Location information Provider built inside of FAs is now being used in most of the mobile networks to determine the location of any FAs address. GPS is used in different areas and is becoming more commonly used because it is integrated in various devices. Moreover, it can also be used in MN and other wireless access devices to facilitate good handover due to its accurate location-trace [18].

Building Location information Provider in the MN and FAs means the MN and FAs are able to track the position constantly. By using L2 and Location information Provider in MN and FAs, it is possible to decrease latency and packet loss. Handover decisions based on movement calculation remove the need to wait for beacon signals from other FAs and to discover handover target areas in advance. The link layer and GPS information Provider used in this paper to reduce the delay and Packet lost [20].

5. RELATED WORKS

In the past few years, different proposals have been presented to minimize the handover delay in Mobile-IPv6 networks. Many of the proposed methods require modification of the Access Routers (ARs). Two slightly different handover solutions using multicast routing are presented in [7] and [8].

The Post-Registration proposal involve Link layer [L2 triggers]. [15] The handover method is based on a network-initiated model of a handover, which does not require any MN involvement until the actual Layer 2 (L2) connection with the new Foreign Agent (nFA) is completed. Such a trigger is a signal related to the L2 handover process. Two types of triggers can be received: a source trigger at the oFA (L2-ST) and a target trigger at the nFA (LS-TT). The first trigger that is used is an early notice of an upcoming change in the L2 point of attachment of the MN, referred to as anticipation trigger. A second trigger, the Link Down trigger (L2-LD), indicates that the L2 link between the MN and the oFA is lost. The Link Up trigger (L2-LU) occurs when the L2 link between the MN and the new FA is established. A trigger initiated at the old FA is referred as a source trigger and a trigger initiated at the new FA is referred as a target trigger.

This approach uses Bi-directional (BET) edge tunnels to perform low latency change in the L2 without the MN involvement. A handover occurs when the MN moves from the oFA, Where the MN performed a Mobile IP registration to nFA. The MN delays its registration with the nFA, while maintaining connectivity using the BET between the oFA and nFA.

The other proposal is Pre-Registration [15], realizes an anticipated L3 handover. This handover method allows the MN to communicate with the new Foreign Agent (nFA) while still connected to the Old Foreign Agent (oFA). This way, the MN is able to pre-build its registration state on the nFA prior to an underlying L2 handover.

The network assists the MN in performing an L3 handover before the L2 handover is completed. Both the MN (mobile-initiated) and the FAs (network-initiated) can initiate a handover. A mobil-initiated handover occurs when the L2 anticipation trigger is received at the MN informing it that it will shortly move to the nFA. The L2 trigger contains information such as the nFA’s IP address [14].

The standard Mobile IPv6 procedures have to deal with the same handover latency problem as Mobile IPv4. In [3], Koodli species a protocol to improve handover latency in Mobile IPv6 as [2] does for Mobile IPv4.The Fast Handover method is an extension proposed for Mobile IPv6 and resembles a combination of Pre-Registration and Post-Registration. The Fast handovers for Mobile IPv6 [FMIPv6] Handover can be either Network-Initiated or Mobile-Initiated, depending on whether one of the ARs or the MN initiates the handover. The two main possibilities are router discovery performed by MN on Layer 3 and a link-specific event (L2 trigger) occurring in the MN or in the network. In [8], HMIPv6 a proposal suggest to an extension to Mobile IPv6, which aims to reduce the amount of signalling between the MN and its CNs during a handover, and to improve the performance in terms of handover speed.

In an IETF draft, which expired in April 2006 [19], Jung et al. propose a combination of the Fast Handovers and Hierarchical Mobile IP extensions to Mobile IPv6. The scheme is called Fast Handover for Hierarchical Mobile IPv6" (F-HMIPv6). The MN enters a new MAP domain, it first performs the HMIPv6 registration procedures with HA and MAP. Later, when the MN moves from a PAR to a NAR within the MAP domain, it will follow the local Binding Update (BU) Procedure of F-HMIPv6. During the handover, data packets sent by CNs will be tunneled by the MAP toward the NAR via a bi-directional tunnel, similarly to the FMIPv6 procedure. Optionally, the MAP may start bi-casting packets to PAR and NAR simultaneously. It should be noted that no bi-directional tunnel is established between PAR and NAR.

6. PROPOSED ALGORITHM

The Predictive handover for FMIPv6 (P-HMIPv6) provides a different approach for resolving the timing ambiguity problem. Link layer information such as signal strength is continuously available, providing important information about the availability of new links, and the FAs will use the location information’s of MNs and all neighbours (FAs) to facilitate handover. Therefore, the handover can be predict in advance before MN moves out of the coverage area of the oFA. The main idea behind the proposal is to apply link layer information and Location information of the FAs to predict a breakdown wireless link before the link is broken. The use of Proactive will significantly reduce handover latency and reduces packet loss in handover.

The proposal will consider the handover to start when it is predict that the link layer association to the oFA will lost. The handover will completed when the registration reply message received from the HA to the MN. Figure 1 describes the overall P-Mobile IP protocol message flow.

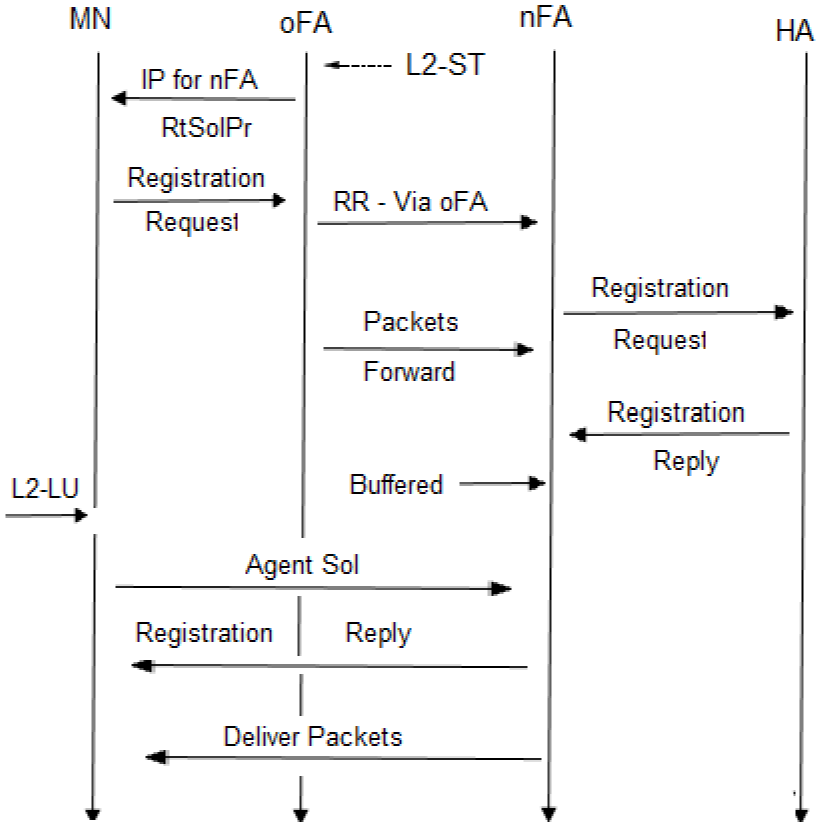


FIGURE 1: P-Mobile IP Protocol Message Flow.

1. The handover process starts when the MN become leavening the oFA and entering the overlap area between nFA and oFA.
2. The Foreign Agent receiving an L2 trigger informing that MN is about to move from oFA to nFA.
3. The oFA provided location information for nFA include MN home Agent address and that include by oFA, which include the direction that based on Layer 2 information.
4. The oFA sends a handover Initiation (HI) message containing the MNs home Agent address.
5. The nFA sends a handover Acknowledgement (H-Ack) message to oFA.
6. Tunnelling then establish between oFA and nFA.
7. The oFA will forward all packets received through the tunnel to the nFA.
8. MN then will receive fast acknowledgment (FB-Ack) via nFA.
9. The nFA will buffer packets that received from the oFA, and this packets will forwarding after the MN send FNA message to ask nFA to forward all buffered packets,

7. SIMULATION SCENARIO and CONFIGURATION

In this section, we evaluate the performance proposed for the FMIPv6 using the link and location information algorithm. We compare our algorithm against a Mobile IPv6 and Mobile IP We assume that the MN on area (A) and start handover to (B or E) the area for the overlapping is 25m, the Handover decisions based on movement calculation eliminate the need to wait for beacon signals from other FAs and also to discover handover target areas in advance.

We use network simulator CIMS NS-2 version ns-allinone-2.31 as a simulation tools in order to simulate FMIPv6 handover [21, 22]. The simulator is modified to emulate IEEE 802.11 infrastructured behaviours with multiple disjoint channels. This modification forces L2 handover operations, where stations only receive data packets via one FA at a time.

The network features three MNs connected to it; the first will move sequentially from oFA to nFA, starting at overlap of the nFA1, performing handovers at a rate of a 30 handovers/min. In each test, the MN1 will be the receiver of a CBR or FTP traffic source, generating either UDP or TCP packets. This traffic originates from the CN1 outside the network, or inside the domain from CN2. All presented results are taken as the average of multiple independent runs, coupled with a 95% confidence interval. The best possible handover point occurs at position A, as shown in Figure 2.

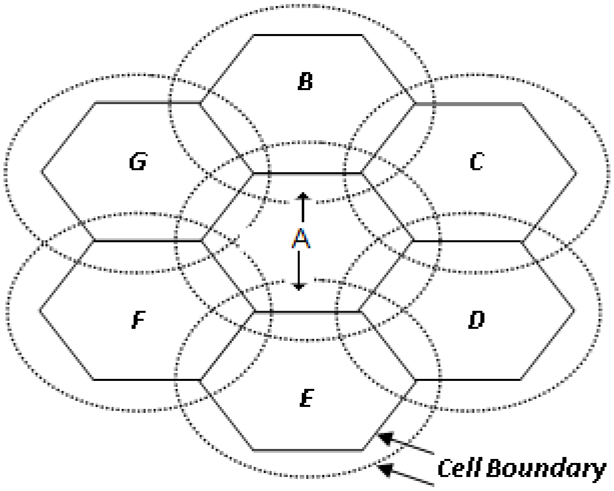


FIGURE 2: Overlapping Coverage Area.

8. PERFORMANCE ANALYSIS and EVALUATION

In our simulation, we use a 500m × 500m and a 1000m × 1000m area with a 3 to 7 MNs [5, 11]. The network bandwidth is 2 Mbps and the medium access control (MAC) layer protocol is IEEE 802.11 [19]. The packet size is 10p/s which will generate enough traffic when we increase the number of connections for example at 40 connections of source-destination pairs, it will generate 400 packets per second for whole scenario. Other simulation parameters are shown in Table1. These parameters have been widely used.

TABLE 1: Simulation Parameters.

Simulation parameter	Value
Simulator	Ns-allinone-2.31
Network range	600m×600m and 1000m×1000m
Transmission range	25m
Mobile nodes	3 and 5
Traffic generator	Constant bit rate
Bandwidth	2Mbps
Packet size	512 bytes
Packet rate	10 packet per second
Simulation time	750s and 1100s

Figure 3 and 4 showing the relation between the handover latency and packet loss, as we observe that the P-FMIPv6 performs better in terms of handover latency and packet loss compared to the others, although the fast handover protocol is proposed and design to minimize the latency and the packet loss during a handover while the worst case observed Mobile IP and Mobile IPv6.

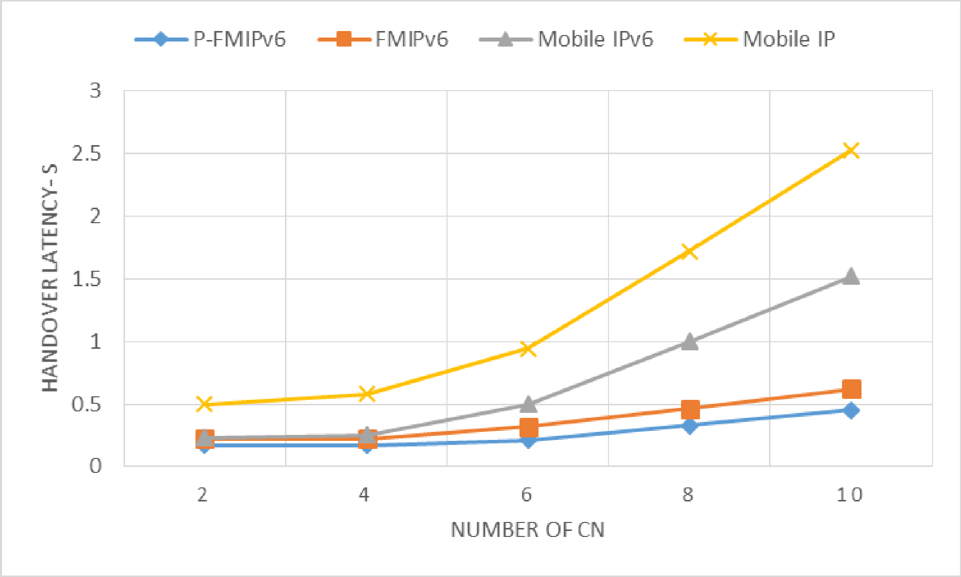


FIGURE 3: Impact of Handover Latency.

We observe in figure 4, that the number of packet loss increase with increase of CN, the P-FMIPv6 showing better performance comparing to the wars case of MIP and MIPv6.

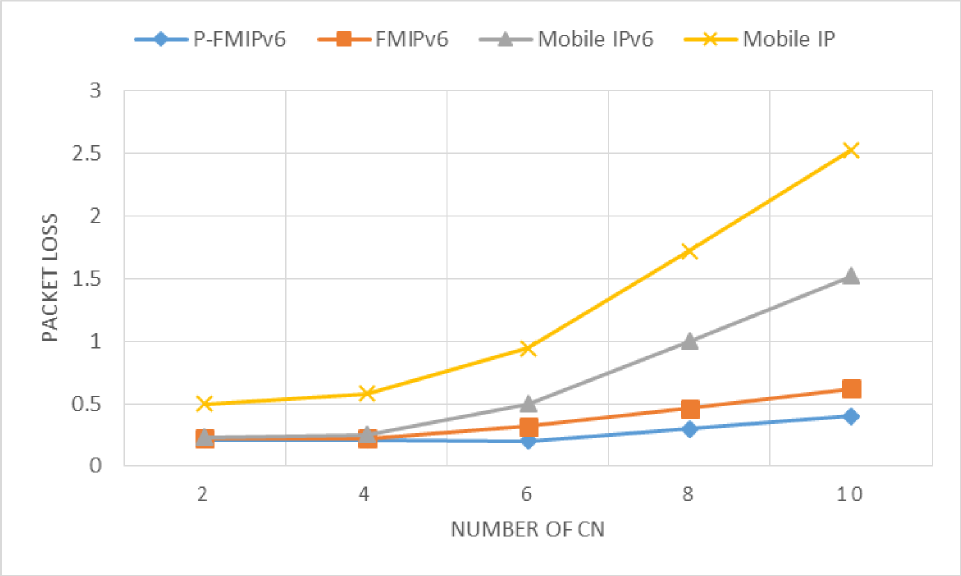


FIGURE 4: Impact of Packet Loss.

The overall throughput graph showing in Figure 5. The figure shows that as the number of sending rate increase the throughput increase. The P-FMIPv6 performs better than all other proposal. The reason for the throughput increase is that more packets are sent overall, although the number of packets lost increase as the sending rate increase. The P-FMIPv6 slightly performs well compared to the other three proposals. We can see that at rate of 10 the MIPv6, FMIPv6 and P-FMIv6 very close.

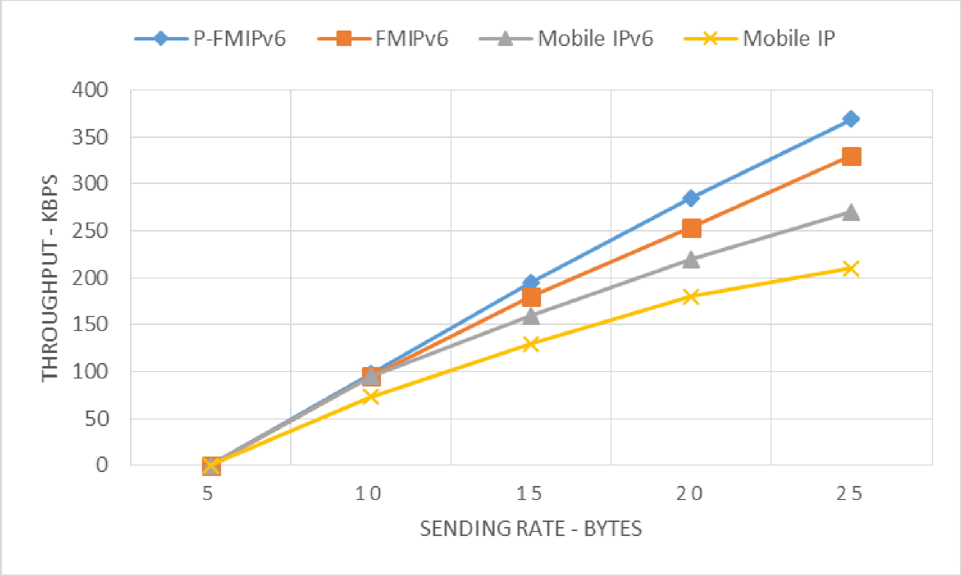


FIGURE 5: Throughput Versus Rate.

Obviously, the loss in the buffer increases when the buffer size is increased. The number of packets lost depends both on the size of the buffer used to store packets for potential handovers and the sending rate as seen in figure 6. The number of packets lost increases for Mobile IP since no buffer is used and increases as the sending rate increase since more packets are sent, while MN is unable to receive them during handover. While on the other hand, the number of packets lost decreases as buffer size increase for P-MIPv6 and FMIPv6.

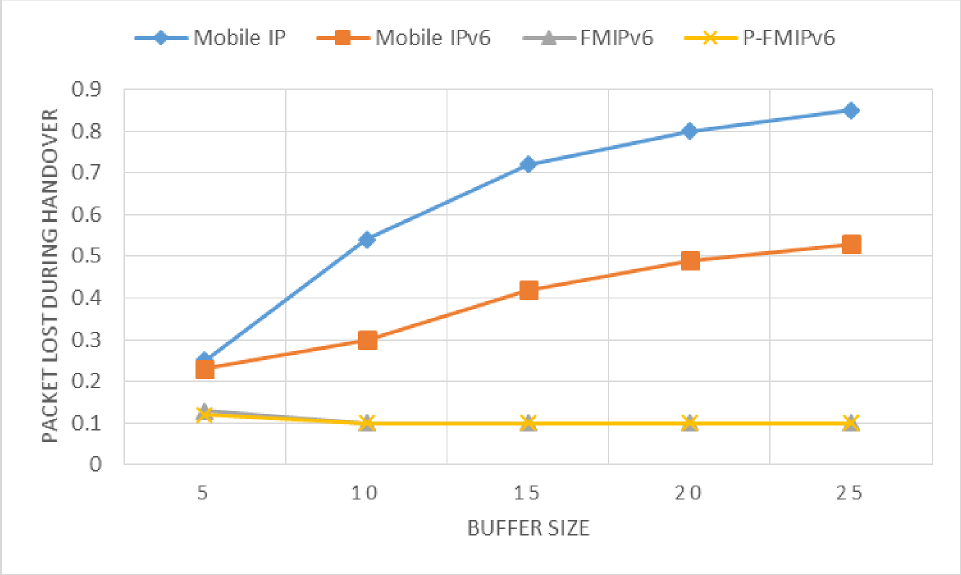


FIGURE 6: Packet Loss vs Buffer Size.

The result graph shows the uplink of MN to CN transmission behaviour with sixe handover in unit time of all four schemes. The result graph shows the transmission bit rate of each handover protocol. The MIPv6 and Mobile IP receive less data than other schemes because their time period take to finish the registration, while the FMIPv6 and P-FMIPv6 shows the highest transmission rate.

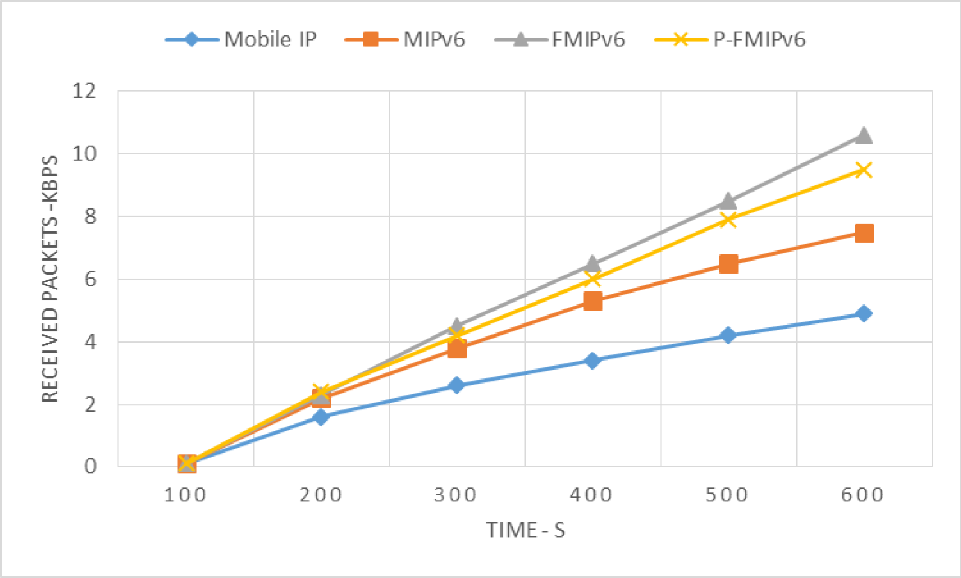


FIGURE 7: Handoff Behaviour.

9. CONCLUSION

In this paper, we developed and analyzed the proposed scheme of the P-Mobile IPv6 handover using link layer and location information scheme. The performance study in this paper indicate that the use of link layer and location information helps to minimize packet loss and improver the throughput. In our scheme, we analysed the performance by simulating the proposed scheme in ns-2 to get fast mobile handover for FMIPv6.

We then compared the experimental results with the results of the Mobile IP and Mobile IPv6 and FMIPv6. The performance study in this paper indicates that the use of link layer information with location information helps to minimize packet loss and improve the throughput of Mobile IP handover. We have seen that the starting point for packet loss could happen in two ways: first, packets may get lost in the oFA when the forwarding buffer overflows and secondly, packets may get lost in the nFA when, upon their arrival, the *ReRep* from the HA has not arrived in the nFA. The first reason for loss may be avoided by appropriately dimensioning the forwarding buffer. This buffer should be able to store arriving packets at least during a time equal to the delay on the nFA and oFA path. The second loss is more difficult to deal with. It is determined by the difference between the delays of the paths oFA, nFA and nFA, HA.

In addition, we evaluated the impact of L2 setup on different performance measures of Mobile IP, together with handover latency, packet loss and throughput. The simulation results show that P-Mobile IPv6 handover latency is not too sensitive to L2 setup latency and beacon periods compared to the other schemes of Mobile IP. Moreover, P-Mobile IPv6 can achieve a fast and seamless handover if MN’s moving speed is not too high, but is within reasonable limits.

10. REFERENCES

[1] D. Johnson, C. Perkins, and J. Arkko. ”Mobility Supportin IPv6”. Internet Engineering Task Force (IETF), RFC-3775, June 2004.

[2] C. E. Perkins, “Mobile Networking through Mobile IP”, *Internet Computing, IEEE,* vol. 2, pp. 58–69, 1998.

[3] H. Soliman, C. Castelluccia, K. E. Malki, and L. Bellier.”Hierarchical Mobile IPv6 Mobility Management (HMIPv6)”. *Internet Engineering Task Force* (*IETF*), *RFC-4140*, August 2005.

[4] P. Bhagwat, C. Perkins and S. Tripathi, “Network Layer Mobility: an Architecture and Survey”, *IEEE Personal Communications,* vol. 3, pp. 54–64, 1996.

[5] H. Balakrishnan, V. N. Padmanabhan, S. Seshan and R. H. Katz, “A Comparison of Mechanisms for Improving TCP Performance over Wireless Links”, *IEEE/ACM Transactions on Networking,* vol. 5, pp. 756–769, 1997.

[6] S. Mohanty and I. F. Akyildiz, “A Cross-layer (Layer 2 + 3) Handoff Management Protocol for Next-Generation Wireless Systems”, *Transactions on Mobile Computing,* vol. 5, pp. 1347–1360, 2006.

[7] I. F. Akyildiz, X. Jiang and S. Mohanty, “A Survey of Mobility Management in Next-generation All-IP-Based Wireless Systems”, *IEEE Wireless Communications*, vol. 11, pp. 16–28, 2004.

[8] I. F. Akyildiz, “Mobility Management for Next Generation Wireless Systems”, *Proceedings of the IEEE*, vol. 87, no. 8, pp. 1347–84, August 1999.

[9] J. Puttonen, “Using Link Layer Information for Improving Vertical Handovers”, *16th International Symposium on Personal, Indoor and Mobile Radio Communications IEEE*, 2005.

[10] H. Chung-Ming, C. Meng-Shu and L. Jin-Wei, “A Link Layer Assisted Fast Handoff Scheme Using the Alternative Path Approach”, *20th International Conference on Advanced Information Networking and Applications*, 2006.

[11] F. Fang and D. S. Reeves, “Explicit Proactive Handoff with Motion Prediction for Mobile IP”, *2004 IEEE Wireless Communications and Networking Conference, WCNC 2004*, vol. 2, pp. 855–860, 2004.

[12] S. Oh, H. Song and Y. Kim, “Seamless Fast Handover in Mobile IPv4 Using Layer-2 Triggers”, *2nd International Conference on Systems and Networks Communications, ICSNC 2007*, pp. 16–16. 2007.

[13] S. Thalanany, “Low Latency Handoffs in Mobile IPv4”, *draft-ietf-mobileip-lowlatency-handoffs-v4-04.txt*, June 2002.

[14] K. El-Malki and H. Soliman, “Fast Handoffs in Mobile IPv4”, Internet draft, *draft-emalki-mobileip-fast-handoffs-03.txt*, September 2000.

[15] S. Oh, H. Song and Y. Kim, “Seamless Fast Handover in Mobile IPv4 Using Layer-2 Triggers,” in *Systems and Networks Communications, ICSNC 2007, 2nd International Conference*, pp. 16-16, 2007.

[16] R. Koodli and C. E. Perkins, “Mobile IPv4 Fast Handovers”, Internet draft, *Internet Engineering Task Force*, *draft-ietf-mip4-fmip*, February 2006.

[17] R. Hsiehet, “S-MIP: a Seamless Handoff Architecture for Mobile IP”, *Proceedings of INFOCOM 2003*, March 2003.

[18] G. Dommety and T. Ye, “Local and Indirect Registration for Anchoring Handoffs”, draft-dommety-mobileip-anchorhandoff-01.txt, July 2000.

[19] Columbia University, *Columbia IP Micro-Mobility Software*, http://www.comet.columbia.edu/micromobility/.index.html.

[20] G. Pollini, “Trends in Handover Design”, *IEEE Communications Magazine*, 34, 3, 80–90, March 1996.

[21] S. Goswami, “Simultaneous Handoff of Mobile-IPv4 and 802.11”, Internet Draft, *IETF, draft-goswami-mobileip-simultaneous-handoff-v4- 02.txt*, February 2003

[22] H. Chung-Ming, C. Meng-Shu and L. Jin-Wei, “A link layer assisted fast handoff scheme using the alternative path approach,” in *Advanced Information Networking and Applications: 20th International Conference*, pp. 5, 2006.

INSTRUCTIONS TO CONTRIBUTORS

The *International Journal of Computer Science and Security (IJCSS)* is a refereed online journal which is a forum for publication of current research in computer science and computer security technologies. It considers any material dealing primarily with the technological aspects of computer science and computer security. The journal is targeted to be read by academics, scholars, advanced students, practitioners, and those seeking an update on current experience and future prospects in relation to all aspects computer science in general but specific to computer security themes. Subjects covered include: access control, computer security, cryptography, communications and data security, databases, electronic commerce, multimedia, bioinformatics, signal processing and image processing etc.

To build its International reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJCSS.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Started with Volume 9, 2015, IJCSS is appearing with more focused issues. Besides normal publications, IJCSS intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

We are open to contributions, proposals for any topic as well as for editors and reviewers. We understand that it is through the effort of volunteers that CSC Journals continues to grow and flourish.

IJCSS LIST OF TOPICS

The realm of International Journal of Computer Science and Security (IJCSS) extends, but not limited, to the following:

- Authentication and authorization models
- Computer Engineering
- Computer Networks
- Cryptography
- Databases
- Image processing
- Operating systems
- Programming languages
- Signal processing
- Theory
- Communications and data security
- Bioinformatics
- Computer graphics
- Computer security
- Data mining
- Electronic commerce
- Object Orientation
- Parallel and distributed processing
- Robotics
- Software engineering

CALL FOR PAPERS

Volume: 9 - **Issue: 5**

i. Submission Deadline : September 30, 2015 **ii. Author Notification:** October 31, 2015

iii. Issue Publication: November 2015

CONTACT INFORMATION

Computer Science Journals Sdn BhD

B-5-8 Plaza Mont Kiara, Mont Kiara

50480, Kuala Lumpur, MALAYSIA

Phone: 006 03 6204 5627
Fax: 006 03 6204 5628
Email: cscpress@cscjournals.org

CSC PUBLISHERS © 2015
COMPUTER SCIENCE JOURNALS SDN BHD
B-5-8 PLAZA MONT KIARA
MONT KIARA
50480, KUALA LUMPUR
MALAYSIA

PHONE: 006 03 6204 5627
FAX: 006 03 6204 5628
EMAIL: cscpress@cscjournals.org