



UNIVERSITI PUTRA MALAYSIA

PRIVACY-PRESERVING COMPUTER FORENSICS FRAMEWORK

WALEED ABDULJABBAR HALBOOB

FSKTM 2015 20



UPM
UNIVERSITI PUTRA MALAYSIA
BERILMU BERBAKTI

PRIVACY-PRESERVING COMPUTER FORENSICS FRAMEWORK

By

WALEED ABDULJABBAR HALBOOB

**Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in
Fulfillment of the Requirement for the Degree of Doctor of Philosophy**

June 2015

COPYRIGHT

All material contained within the thesis, including without limitation text, logos, icons, photographs and all other artwork, is copyright material of Universiti Putra Malaysia unless otherwise stated. Use may be made of any material contained within the thesis for non-commercial purposes from the copyright holder. Commercial uses of material may only be made with the express, prior, written permission of Universiti Putra Malaysia.

Copyright©Universiti Putra Malaysia



DEDICATIONS

To my father Abduljabbar, mother Hamidah, father-in law Abdulsalam (may ALLAH rest his soul in peace) and mother-in-law Niayam for their patience, encouragement and support.

To my wife Zinab and daughter Salma for their patience, encouragement and time.

To all my family members for their support.

To my friends in Malaysia for their standing with me all the time.



Abstract of theses presented to the senate of University Putra Malaysia in fulfillment of the requirement for the Doctor of Philosophy

PRIVACY-PRESERVING COMPUTER FORENSICS FRAMEWORK

By

WALEED ABDULJABBAR HALBOOB

June 2015

Chairman: Professor Ramlan Mahmod, PhD

Faculty: Computer Science and Information Technology

Computer forensics and privacy preservation are conflicting fields in computer security. Computer forensics tools essentially image and analyze all the data found in a targeted suspect's storage, even if these data are private and irrelevant to the crime under investigation. In contrast, privacy preservation techniques are used to protect a data owner private identity, information, and/or activities from any unauthorized access, use, or disclosure. Thus, there is a need to balance these two conflicting fields. In other words, there is a tremendous need to find a lawful and fair computer forensics solution that images and analyzes a suspect's data while preserving the privacy. Over the past decade, the conflict between privacy preservation and computer forensics has been investigated in several studies. However, the solutions proposed by previous researchers are not efficient and lawful as well as they did not provide a sufficient analysis. The objective of this research is to propose a computer forensics framework to preserve the privacy of data owners in an efficient and lawful manner while providing sufficient digital evidence analysis. Computer forensics privacy levels and policies are specified to help improve the framework's efficiency and lawfulness, respectively. A selective imaging concept is used for providing an efficient imaging and analysis. The private data are encrypted using an advanced encryption system (AES). Advanced forensic format 4 (AFF4) is used as a container for the imaged relevant data. The framework is implemented to ensure that it is workable and measure its efficiency. A qualitative evaluation method was used to evaluate both the lawfulness of the framework and sufficiency of the analysis by observing these criteria. Moreover, other related work was implemented to compare with the proposed framework. The results obtained show that the proposed framework satisfies all the required features for having a lawful solution, provides efficient imaging and analysis as well as sufficient analysis. It can be concluded that the proposed

framework has several advantages compared to the other related works, namely an efficient and lawful method for selective imaging and analysis, and sufficient analysis. It also provides a forensics sound and flexible solution with a distributed analysis.



Abstrak tesis yang dibentangkan kepada senat Universiti Putra Malaysia dalam memenuhi keperluan untuk ijazah Master Sains

RANGKA KERJA FORENSIK KOMPUTER YANG MENGEKALKAN PRIVASI

Oleh

WALEED ABDULJABBAR HALBOOB

June 2015

Pengerusi : Prof. Ramlan Mahmod, PhD

Fakulti : Sains Komputer dan Teknologi Maklumat

Forensik komputer dan pemeliharaan privasi adalah bidang yang bercanggah dalam keselamatan komputer. Alat Forensik komputer pada asasnya ialah imej dan menganalisis semua data yang terdapat dalam media penyimpanan milik suspek sasaran, walaupun data ini adalah peribadi dan tidak berkaitan dengan jenayah yang sedang disiasat. Sebaliknya, teknik pemeliharaan privasi digunakan untuk melindungi identiti peribadi, maklumat, dan/atau aktiviti pengguna daripada mana-mana pihak yang tidak dibenarkan untuk mengakses, menggunakan, atau mendedahnya. Oleh itu, terdapat keperluan untuk mengimbangi kedua-dua bidang yang bercanggah. Dalam erti kata lain, terdapat keperluan yang besar untuk mencari penyelesaian forensik komputer yang sah dan adil untuk pengimejan dan mengkaji data milik suspek manakala memelihara privasi mereka juga. Sepanjang dekad yang lalu, konflik di antara pemeliharaan privasi dan komputer forensik telah disiasat dalam beberapa kajian. Walau bagaimanapun, penyelesaian yang dicadangkan oleh penyelidik sebelum ini tidak cekap dan sah serta mereka tidak menyediakan analisis yang mencukupi. Objektif kajian ini adalah untuk mencadangkan rangka kerja forensik komputer untuk memelihara privasi pengguna dengan cara yang cekap dan sah di samping menyediakan analisis bukti digital yang mencukupi. Tahap privasi dan polisi bagi forensik komputer dinyatakan untuk membantu meningkatkan kecekapan rangka kerja dan kesesuaian dengan undang-undang, masing-masing. Konsep pengimejan terpilih digunakan untuk pengimejan hanya pada data yang berkaitan, sehingga mampu menyediakan rangka kerja yang cekap. Data peribadi disulitkan menggunakan *Advanced Encryption System (AES)*. *Advanced forensic format 4 (AFF4)* digunakan sebagai bekas untuk data yang berkaitan dengan pengimejan. Rangka kerja ini dilaksanakan untuk memastikan bahawa ia boleh digunakan serta diukur kecekapannya. Satu kaedah penilaian kualitatif juga digunakan untuk menilai kedua-dua hal iaitu kesahihan rangka kerja dan kecukupan analisis. Walau bagaimanapun, kerja-kerja lain yang berkaitan juga dilaksanakan, dinilai dengan cara yang sama, dan dibandingkan dengan rangka kerja yang dicadangkan. Hasilnya menunjukkan bahawa rangka kerja pertama yang dicadangkan memenuhi semua ciri-ciri

yang diperlukan untuk mempunyai penyelesaian yang sah, menyediakan pengimejan dan analisis berkesan (pencarian dan penyahsulitan), dan, akhirnya, menyokong kedua-dua carian berasaskan kata kunci dan atribut untuk menganalisis sasaran data yang disimpan. Ia boleh disimpulkan bahawa rangka kerja yang dicadangkan mempunyai beberapa kelebihan dibanding dengan kerja lain yang berkaitan, iaitu satu kaedah yang cekap dan sah bagi pengimejan dan analisis terpilih, dan membolehkan analisis yang mencukupi. Ia juga menyediakan forensik yang kukuh dan penyelesaian yang fleksibel dengan analisis diedarkan.



AKNOWLEDGEMENT

It is a great opportunity to thank Professor Dr. Ramlan Mahmud, Associate Professor Dr. Nur Izura Udzir, and Dr. Mohd Taufik Abdullah for their great help on this thesis and for their supporting guidance, ideas, and materials.

I would also like to express my thanks to the Faculty of Computer Science and Information Technology, especially the ICT unit, for providing general help and assistance. Also, I'd like to thank the Library and the School of Graduate Studies for helpfully fulfilling my every request.

Last but not least, I would like to thank my family for giving me the motivation and moral support needed to complete this thesis. Only Allah can truly reward what they have done.

I certify that a Thesis Examination Committee has met on 17/06/2015 to conduct the final examination of WALEED ABDULJABBAR HALBOOB on his thesis entitled "PRIVACY-PRESERVING COMPUTER FORENSICS FRAMEWORK" in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Doctor of Philosophy.

Members of the Thesis Examination Committee were as follows:

Abu Bakar Md Sultan, PhD

Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Abdul Azim Abd Ghani, PhD

Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Internal Examiner)

Zuriati Ahmed Zukarnain, PhD

Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Internal Examiner)

Felix Freiling, PhD

Professor
Friedrich-Alexander-Universität Erlangen-Nürnberg
Germany
(External Examiner)

ZULKARNAIN ZAINAL, PhD

Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 17 June 2015

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfillment of the requirement for the degree Doctor of Philosophy. The members of the Supervisory Committee are as follows:

Ramlan Mahmod, PhD

Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Nur Izura Udzir, PhD

Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Mohd. Taufik Abdullah, PhD

Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

BUJANG BIN KIM HUAT, PhD

Professor/ Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: _____ Date: _____

Name and Matric No.: _____

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: _____
Name of
Chairman of
Supervisory
Committee: _____

Signature: _____
Name of
Member of
Supervisory
Committee: _____

Signature: _____
Name of
Member of
Supervisory
Committee: _____

Signature: _____
Name of
Member of
Supervisory
Committee: _____

TABLE OF CONTENTS

	Page
ABSTRACT	i
ABSTRAK	iii
ACKNOWLEDGEMENTS	v
APPROVAL	vi
DECLARATION	viii
LIST OF TABLES	xii
LIST OF FIGURES	xiv
CHAPTER	
1	INTRODUCTION
1.1	Background
1.2	Problem Statement
1.3	Research Objectives
1.4	Scope and Limitation
1.5	Thesis Organization
2	LITERATURE REVIEW
2.1	Computer Forensics Overview
2.2	Selective Imaging Concept and State-of-the-Art
2.3	Privacy Issue in Computer Forensics
2.4	Lawfulness Evaluation Criteria
2.4.1	Collection Limitation Principle
2.4.2	Data Quality Principle
2.4.3	Purpose Specification Principle
2.4.4	Use Limitation Principle
2.4.5	Security Safeguards Principle
2.4.6	Openness Principle
2.4.7	Individual Participation Principle
2.4.8	Accountability Principle
2.5	Lawfulness Evaluation Criteria: A Discussion
2.6	Current Research Directions
2.6.1	Forensics-enabled Privacy Enhancing Techniques
2.6.2	Anti-forensics Solutions
2.6.3	Privacy-preserving Digital Forensics Approaches
2.7	Privacy-preserving Computer Forensics Approaches
2.7.1	Policy-based Approaches
2.7.2	Cryptographic-based Approaches
2.8	Discussion and Evaluation
2.8.1	Lawfulness Summary

	2.8.2	Unaddressed Issues	25
3		RESEARCH METHODOLOGY	26
	3.1	Problem Identification	26
	3.2	Designing the Framework	27
	3.2.1	Imaging and Analysis Efficiency	27
	3.2.2	Imaging and Analysis Lawfulness	28
	3.2.3	Analysis Sufficiency	29
	3.3	Implementing the Framework	29
	3.4	Experimental Design	30
	3.5	Evaluating the Framework	31
	3.5.1	Evaluating the Imaging and Analysis Efficiency	31
	3.5.2	Evaluating the Imaging and Analysis Lawfulness	32
	3.5.3	Evaluating the Analysis Sufficiency	32
4		THE PROPOSED FRAMEWORK	34
	4.1	Overview	34
	4.2	Privacy Levels and Policies: Specification	34
	4.2.1	Computer Forensics Privacy Levels	34
	4.2.2	Computer Forensics Privacy Policies	38
	4.3	Architecture of the Framework	42
	4.4	Components of the Framework	45
	4.4.1	Selective Imaging Model (SIM)	45
	4.4.2	Selective Analysis Model (SAM)	55
	4.5	System Implementation	59
5		RESULTS AND DISCUSSIONS	63
	5.1	Evaluation Criteria	63
	5.2	Performance Analysis	63
	5.2.1	Offset Ordering Process Efficiency Impact	63
	5.2.2	The Imaging Efficiency	70
	5.2.3	The Analysis Efficiency	81
	5.3	Lawfulness Analysis	85
	5.3.1	Privacy Policies	86
	5.3.2	Relevant Data Collection	86
	5.3.3	Identical Copies	86
	5.3.4	Auditing	87
	5.3.5	Integrity	87
	5.3.6	Authenticity	87
	5.3.7	Non-repudiation	88
	5.3.8	Encryption-based Access Control	88
	5.4	Analysis Sufficiency	88
	5.4.1	Keywords-based Search	89
	5.4.2	Attributes-based Search	89
	5.4.3	Integration with Computer Forensics Analysis Tools	89
	5.5	Comparison	90

6	CONCLUSION	94
6.1	Overview	94
6.2	Conclusion	94
6.3	Research Contributions	95
6.4	Future Works	95
	REFERENCES	96
	APPENDICES	102
	BIODATA OF STUDENT	115
	LIST OF PUBLICATIONS	116



LIST OF TABLES

Table		Page
2.1	Summary of Results of Evaluating Lawfulness of Related Privacy-Preserving Computer Forensics Works	23
4.1	Forensic Data Access Possibilities in Computer Forensics	37
4.2	Privacy Levels for Computer Forensics	38
5.1	Evaluating Cases to Determine Effect of Offset Ordering Process	64
5.2	Efficiency Results of Hard Disk Selective Direct Imaging	65
5.3	Efficiency Results for Flash Device Direct Imaging	66
5.4	Efficiency Results for Hard Disk Selective AFF4 Imaging	68
5.5	Efficiency Results for Flash Device Selective AFF4 Imaging	69
5.6	Imaging Time in Proposed Framework	71
5.7	Distribution of Encrypted AFF4 Imaging Time with Compression	73
5.8	Distribution of Encrypted AFF4 Imaging Time without Compression	74
5.9	Distribution of Normal AFF4 Imaging Time with Compression	74
5.10	Distribution of Normal AFF4 Imaging Time without Compression	75
5.11	Imaging Time in Related Works and Proposed Framework	76
5.12	Required Searching Time of Related Works and Proposed Framework	81
5.13	Required Decryption Time for Related Works and Proposed Framework	83
5.14	Summary of Results of Evaluating Proposed Framework and Related Works	91

LIST OF FIGURES

Figure		Page
2.1	Searchable Encryption Concept	18
2.2	Steps of Law <i>et al.</i> (2011) Model	20
2.3	Execution Process of First Scheme by Hou <i>et al.</i>	22
3.1	Methodology Steps	26
4.1	Proposed Framework	35
4.2	General Architecture of Proposed Framework	43
4.3	Architecture of Selective Imaging Model (SIM)	45
4.4	Flowchart of Offset Ordering Process	47
4.5	Pseudo code of Offset Ordering Process	48
4.6	Pseudo Code of Ordering Items inside the ORDMR Report	49
4.7	Flowchart of the Selective Imaging Method	50
4.8	Pseudo Code of Selective Imaging Method	51
4.9	Flowchart of Integrity Verification Process	53
4.10	Pseudo Code of Integrity Verification Process	54
4.11	General Architecture of Selective Analysis Model (SAM)	56
4.12	Flowchart of Investigator Authentication and Partial AFF4 Image Integrity Verification	58
4.13	Pseudo Code of Investigator Authentication and Partial AFF4 Image Integrity Verification	59
4.14	Screenshot of Prototype Main Window	60
4.15	Screenshot of Selective Imaging Model (SIM) Prototype	61
4.16	Screenshot of Searching Window of Selective Analysis Model (SIM) Prototype	62
5.1	Efficiency Results of Hard Disk Direct Selective Imaging	66
5.2	Efficiency Results for Flash Device Direct Selective Imaging	67
5.3	Efficiency Results for Hard Disk Selective AFF4 Imaging	68
5.4	Efficiency Results for Flash Device Selective AFF4 Imaging	70
5.5	Imaging Time of Proposed Framework	72
5.6	Distribution of Encrypted AFF4 Imaging Time with Compression	73
5.7	Distribution of Encrypted AFF4 Imaging Time without Compression	74
5.8	Distribution of Normal AFF4 Imaging Time with Compression	75
5.9	Distribution of Normal AFF4 Imaging Time without Compression	75
5.10	Imaging Time in Related Works and Proposed Framework	77
5.11	Full Imaging Time of The Hou <i>et al.</i> Model	78
5.12	Image Size of Proposed Framework	79
5.13	Image Size in Related Works and Proposed Framework	80
5.14	Required Searching Time for the Related Works and the Proposed Framework	83
5.15	Required Decryption Time for Related Works and Proposed Framework	85

A.1	Initializing RSA Cryptosystem and Generating Public/Private Key Pair	102
A.2	Initializing AES Cipher and Generating an AES Key	103
A.3	Initializing and Using The SHA-1	103
A.4	Implementation Source Code of Offset Ordering Process	104
A.5	Implementation Source Code of Encrypted AFF4 Imaging	106
A.6	Implementation Source Code of Normal AFF4 Imaging	108
A.7	Implementation Source Code of Authenticating Investigator	109
A.8	Implementation Source Code of Searching And Decryption	110



CHAPTER 1

INTRODUCTION

1.1 Background

Digital forensics is a computer security discipline that focuses on identifying, collecting, preserving, analyzing, and presenting digital evidence from digital systems so that the presented digital evidence is acceptable in a court of law. According to Stephenson (2002), digital forensics has three branches:

- 1) Computer forensics: Deals with gathering digital evidence from computers and computer storage (e.g., hard disks, flash memories, DVDs, etc.) whether the computer storage is used in personal computers, mobile devices, or servers. This term is sometimes used to refer to all three branches.
- 2) Network forensics: Considers the capture of digital evidence from network traffic and devices. However, mobile forensics is sometimes considered under this branch, and some authors deal with it as a separate branch.
- 3) Software forensics: Aims to assist in discovering who wrote a particular code to trace malicious users.

This research falls into the computer forensics branch, in which the investigation process has five main steps namely identification, collection, preservation, analysis, and presentation. The widely used procedure for collecting and analyzing digital evidence in computer forensics involves the creation of a bit-by-bit image from the data owner's physical storage and then later analyzing the bit-by-bit image at a Computer Forensics Laboratory (CFL). Using this procedure, all of the data found in the storage of the data owner (suspect, victim, or any related party to the crime) are collected and analyzed. In fact, this procedure has been proven to be a non-practical solution because of increases in the quantities of storage and data commonly owned, which increase the investigation cost in term of the required time and resources (Stüttgen *et al.*, 2013). The problem becomes worse when dealing with a server's storage because of the huge amount of data involved and many users not related to the crime under investigation. Therefore, this procedure creates a significant problem when the data owner's privacy is a concern. Collecting only relevant data is a key point for privacy preservation. Recently, a selective imaging concept has been proposed to gather only data relevant to the crime, which would reduce the investigation cost. However, selectively imaging only the relevant data is still not a sufficient solution for privacy preservation in computer forensics, and many other requirements must be addressed as discussed below.

Privacy preservation in computer forensics is an essential issue for several reasons, including the following (Bui, 2003; Croft and Olivier, 2010; Law *et al.*, 2011; and Hou *et al.*, 2013):

- In some countries, privacy acts exist and should be taken into account throughout the investigation process.
- The targeted data storage(s) may contain irrelevant data belonging to other unrelated parties or users, or could belong to the private sector (e.g., banking system, Internet Service Provider, etc.) and contain very sensitive private data (such as trade secrets, banking information, and so on).

The computer forensics and privacy protection fields are two conflicting directions in computer security. The former tries to find digital evidence related to a specific crime, while privacy protection tries to protect the user's privacy. As a result, finding a balance between a computer forensic investigation and privacy protection is a serious challenge (Ryan and Shpantzer, 2004; and Hou *et al.*, 2013).

To find the balance between computer forensics and privacy preservation, existing privacy act(s) must be taken into account, which requires addressing several issues such as following (Burmester *et al.*, 2002, Bui, 2003; Saboohi, 2006; Adams, 2008; Croft and Olivier, 2010; and Hou *et al.*, 2013):

- Collecting only data relevant to the crime. The relevancy is determined based on the investigation's goal and scope.
- Ensuring the integrity and authenticity of the collected relevant data.
- Preserving the privacy of the relevant data. Although encryption can be used here, how can the forensic data be encrypted in a forensically sound manner (e.g., without altering its corresponding metadata), and how can the encrypted data be analyzed sufficiently?
- Auditing the investigation process so a court of law can check whether or not the investigator has exceeded the investigation's scope and goal.
- Controlling access to the collected data so that only authorized investigators can analyze the data. Also, in a case where the collected data are disclosed to the public or unauthorized parties, a court of law should be able to use audit trails and access control mechanisms to track the collected data flow from the crime scene to the court room to discover who disclosed it and how.
- Different countries have different privacy acts, and some countries have different acts for the private and public sectors.

Several research efforts have tried to address some of the above issues, but the field still needs more effort because the investigated issues have not yet been totally addressed, and some issues still have research gaps, as will be presented in the next section.

1.2 Problem Statement

Several works have studied the conflict between privacy preservation and computer forensics (Burmester *et al.*, 2002; Bui, 2003; Saboohi, 2006; and Adams, 2008). These studies have suggested several solutions such as specifying accountability and privacy policies, using cryptographic techniques, taking into account existing privacy act(s), collecting only relevant data, and auditing the investigation process.

Existing solutions can be either policy-based or cryptographic approaches. The policy-based approaches are used, in general, to point out how the data owner's data will be collected, used, managed, and disclosed. In computer forensics, Srinivasan (2006; 2007) proposed four policies just for the digital evidence collection step. Therefore, there is a need to cover the other investigation steps.

Regarding cryptographic approaches, Croft & Olivier (2006; 2010) proposed a mechanism for investigating *Call Data Records (CDRs)* stored in a mobile service provider's server. These CDRs are grouped into several levels, and each data group is

encrypted several times upon its level. Thus, this mechanism is not efficient because of encrypting all the data several times. Law *et al.* (2011) proposed a model in which the investigator makes a bit-by-bit image of all the data, the data owner builds and encrypts an index file for each file, and the investigator prepares and encrypts search keywords and searches for relevant data in the index files. This work has a huge collection cost for imaging and building index files. In Hou *et al.* (2011a), two searchable encryption schemes were proposed. In the first scheme, the data owner encrypts all the data, and the investigator prepares and encrypts a single search keyword and passes it to the data owner. The data owner searches for the relevant data and submit them to the investigator. Therefore, this scheme assumes that the data owner is trusted and will not hide any relevant data. The second scheme is proposed to address this issue using a *Third Trusted Party (TTP)*. The TTP is used to search for relevant data. However, the TTP can hide any data and trusting it is not a final solution. In Hou *et al.* (2011b), the first scheme is extended to support multiple search keywords. In Hou *et al.* (2013), the first scheme is also extended to ensure the integrity and authenticity of the collected data and support multiple investigators.

The above related works (Croft & Oliver, 2006; 2010; Law *et al.*, 2011; Hou *et al.*, 2011a; 2011b; and 2013) have several drawbacks, including the following: i) they are not efficient because they require the collection and encryption of all the data; ii) they are not lawful because they do not take into account all the privacy protection requirements (such as privacy policies, access control, and auditing) for enforcing existing privacy acts; and finally, they do not provide sufficient analysis because they rely only on prepared search keywords for selecting and analyzing the relevant data. As a result, they support only text-based documents, and there is no guarantee that the prepared keywords will cover all the relevant data. In addition, the collected encrypted data cannot be analyzed with the existing widely known and acceptable tools (e.g., EnCase, FTK, etc.).

The research problem of this research is to seek for a privacy-preserving computer forensics framework that covers both privacy-based and cryptographic-based approaches, preserves the privacy of data owners in an efficient and lawful manner, and provides a sufficient analysis.

1.3 Research Objectives

The objective of this research is to propose an efficient and lawful privacy-preserving computer forensics framework while providing a sufficient analysis and based on the selective imaging concept.

1.4 Scope and Limitation

The scope of this research is digital forensics, especially computer forensics. To be more specific, this research focuses on privacy preservation while investigating computers and computer storage devices, whether these storage devices are used in personal computers, mobile devices, or servers.

Digital evidence identification, in which the private and relevant forensic data files are identified, is outside the scope of this research. This research considers digital evidence

collection, preservation, and analysis steps based on the selective imaging concept. Digital evidence presentation is not considered too. However, for digital evidence identification existing computer forensics tools will be studied, and tools that are suitable for identifying relevant and private data forensic files, as well as suitable for integrating with our proposed framework, will be used.

In addition, the data granularity considered by this research is the file level. Each file will be treated as private or not and relevant or not. Thus, classifying the content of structured files (e.g., database files) as private or non-private and relevant or non-relevant will not be considered by this research.

1.5 Thesis Organization

The rest of this thesis is organized as follows:

Chapter 2 presents the literature review, starting with introducing overviews of the computer forensics concept, as well as privacy preservation issues in computer forensics. The current research directions are introduced, along with the evaluation criteria of the lawfulness used by this research. Finally, the existing privacy-preserving computer forensics-related works are reviewed and evaluated.

Chapter 3 presents the research methodology steps used for specifying the proposed privacy levels and policies as well as designing, implementing, and evaluating the proposed framework.

Chapter 4 presents the proposed privacy levels and policies. It also covers, in detail, the components of the proposed framework, namely the selective imaging module and selective analysis module. Finally, it presents the framework's implementation.

Chapter 5 presents the results of an evaluation of the proposed framework. A comparison between the proposed framework and other related works is also presented.

Chapter 6 presents the conclusion and contributions of this research, followed by topics for future work.

REFERENCES

- AccessData (2013). Forensic ToolKit (FTK 5.5). Retrieved 5 March, 2014, from <http://accessdata.com/solutions/digital-forensics/forensic-toolkit-ftk>
- Adams, C. W. (2008). *Legal Issues Pertaining to the Development of Digital Forensic Tools*. Paper presented at the Third International Workshop on Systematic Approaches to Digital Forensic Engineering (SADFE '08), Oakland, California, USA.
- AFF4 (2014). Advance Forensic File Format (AFF4) Java Open Source Tool. Retrieved 12 July, 2014, from <http://code.google.com/p/aff4/>
- Al Fahdi, M. A., Clarke, N. L., & Furnell, S. M. (2013). *Challenges to Digital Forensics: A Survey of Researchers & Practitioners Attitudes and Opinions*. Paper presented at the Information Security for South Africa, Johannesburg, South Africa.
- Antoniou, G., & Gritzalis, S. (2006). *RPINA- Network Forensics Protocol Embedding Privacy Enhancing Technologies*. Paper presented at the IEEE International Symposium on Communications and Information Technology, Bangkok, Thailand.
- Antoniou, G., Sterling, L., Gritzalis, S., & Udaya, P. (2008). Privacy and Forensics Investigation Process: The ERPINA protocol. *Computer Standards & Interfaces*, 30(4), 229–236.
- Antoniou, G., Wilson, C., & Geneiatakis, D. (2006). *PPINA - A Forensic Investigation Protocol for Privacy Enhancing Technologies*. Paper presented at the IFIP International Federation for Information Processing, Crete, Greece.
- APEC. (2005). APEC Privacy Framework. Retrieved 10 March, 2014, from http://www.apec.org/Groups/Committee-on-Trade-and-Investment/~/_media/Files/Groups/ECSG/05_ecsg_privacyframewk.ashx
- Beebe, N. (2009). Digital Forensics Research: *The Bad, The Good and the Unaddressed*. Paper presented at Advances in Digital Forensics V - IFIP International Conference on Digital Forensics. Orlando, Florida, USA.
- Beebe, N. L., & Clark, J. G. (2005). A Hierarchical, Objectives-based Framework for the Digital Investigations Process. *Digital Investigation*, 2(2), 147-167.
- Benredjem, D. (2007). *Contributions to Cyber Forensics: Processes and E-Mail*

Analysis. Master, Concordia University, Master thesis.

- Braid, M. (2001). Collecting Electronic Evidence after a System Compromise. Retrieved 21 July, 2014, from <http://www.giac.org/paper/gsec/659/collecting-electronic-evidence-system-compromise/101519>
- Bui, S., Enyeart, M., & Luong, J. (2003). Issues in Computer Forensics. Santa Clara University, Computer Engineering, USA.
- Burmester, M., Desmedt, Y., Wright, R., & Yasinsac, A. (2002). "Security or Privacy, Must We Choose?". Paper presented at the Symposium on Critical Infrastructure Protection and the Law.
- Caloyannides, M. A. (2004). *Privacy Protection and Computer Forensics*, Artech House, Second Edition
- Carrier, B., & Spafford, E. H. (2003). Getting Physical with the Digital Investigation Process. *International Journal of Digital Evidence*. 2(2), 1-20.
- Carrier, B. D., & Spafford, E. H. (2004). *An Event-based Digital Forensic Investigation framework*. Paper presented at the In Proceedings of the 2004 Digital Forensic Research Workshop, Baltimore, Maryland.
- Cohen, M., & Schatz, B. (2010). Hash Based Disk Imaging Using AFF4. *Digital Investigation*, 7(2010), 121.
- Croft, N. J., & Olivier, M. S. (2006). *Sequenced Release of Privacy Accurate Call Data Record Information in a GSM Forensic Investigation*. Paper presented at the Information Security South Africa, Pretoria, South Africa.
- Croft, N. J., & Olivier, M. S. (2010). Sequenced Release of Privacy-accurate Information in a Forensic Investigation. *Digital Investigation*, 7(1-2), 95-101.
- CSVReader. (2014). Java CSV Library. Retrieved 2 August, 2014, from <http://sourceforge.net/projects/javacsv/>
- DFF. (2014). Digital Forensics Framework (DFF). Retrieved 1 August, 2014, from <http://www.digital-forensic.org/>
- Garfinkel, S., Malan, D., Dubec, K.-A., Stevens, C., & Pham, C. (2006). *Advanced Forensic Format: an Open Extensible Format for Disk Imaging*. Paper presented at the Advances in Digital Forensics II, IFIP Advances in Information and Communication, Orlando, Florida, USA.

- Gupta, A. (2013). *Privacy Preserving Efficient Digital Forensic Investigation Framework*. Paper presented at the Sixth International Conference on Contemporary Computing (IC3), Noida, India.
- Hou, S., Uehara, T., Yiu, S. M., Hui, L. C. K., & Chow, K. P. (2011a). *Privacy Preserving Confidential Forensic Investigation for Shared or Remote Servers*. Paper presented at the Seventh International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), Dalian, China.
- Hou, S., Uehara, T., Yiu, S. M., Hui, L. C. K., & Chow, K. P. (2011b). *Privacy Preserving Multiple Keyword Search for Confidential Investigation of Remote Forensics*. Paper presented at the Third International Conference on Multimedia Information Networking and Security, Shanghai, China.
- Hou, S., Yiu, S.-M., Uehara, T., & Sasakix, R. (2013). A Privacy-Preserving Approach for Collecting Evidence in Forensic Investigation. *International Journal of Cyber-Security and Digital Forensics (IJCSDF)*, 2(1), 70-78.
- Justice, U. S. D. O. (2008). *Electronic Crime Scene Investigation: A Guide for First Responders*, Second Edition: National Institute of Justice.
- Kenneally, E. E., & Brown, C. L. T. (2005). Risk Sensitive Digital Evidence Collection. *Digital Investigation*, 2(2), 101-119.
- Law, F. Y. W., Chan, P. P. F., Yiu, S. M., Chow, K. P., Kwan, M. Y. K., Tse, H. K. S., & Lai, P. K. Y. (2011). *Protecting Digital Data Privacy in Computer Forensic Examination*. Paper presented at the IEEE Sixth International Workshop on Systematic Approaches to Digital Forensic Engineering (SADFE), Oakland, CA, USA.
- OECD. (2013). *Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data*. Retrieved 15 October, 2014, from <http://www.oecd.org/sti/ieconomy/2013-oecd-privacy-guidelines.pdf>
- Palmer, G. (2001). *Report from the First Digital Forensic Research Workshop (DFRWS)*. Utica, New York, USA.
- Perumal, S. (2009). Digital Forensic Model Based On Malaysian Investigation Process. *International Journal of Computer Science and Network Security*, 9(8), 38-44.
- Project, C. (2009). *Enron Email Dataset*. Retrieved 4 February, 2015, from <https://www.cs.cmu.edu/~enron/>

- Recovery, C. (2013). CnW Recovery Software. Retrieved 4 February, 2015, from <http://www.cnwrecovery.com/>
- Reith, M., Carr, C., & Gunsch, G. (2002). An Examination of Digital Forensic Models. *International Journal of Digital Evidence*, 1(3), 1-12.
- Richard, G., & Roussev, V. (2004). *Breaking the Performance Wall: The Case for Distributed Digital Forensics*. Paper presented at the Proceedings of the 2004 Digital Forensics Research Workshop (DFRWS'04), Baltimore, Maryland.
- Richard, G., & Roussev, V. (2006). *File System Support for Digital Evidence Bags*. Paper presented at the Advances in Digital Forensics II - IFIP Advances in Information and Communication Technology (Vol. 222/2006, pp. 29-40): Springer.
- Richter, J., Kuntze, N., & Rudolph, C. (2010). *Securing Digital Evidence*. Paper presented at the Fifth IEEE International Workshop on Systematic Approaches to Digital Forensic Engineering (SADFE'2010) The Claremont Resort, Oakland, CA, USA.
- Ryan, D. J., & Shpantzer, G. (2004). Legal Aspects of Digital Forensics. Retrieved 29 July, 2014, from <http://euro.ecom.cmu.edu/program/law/08-732/Evidence/RyanShpantzer.pdf>
- Saboochi, M. (2006). Collecting Digital Evidence of Cyber Crime. Retrieved 3 July, 2014, from <http://www.supremecourt.gov.pk/ijc/articles/10/2.pdf>
- Saleem, S., Popova, O., & Bagillib, I. (2014). Extended Abstract Digital Forensics Model with Preservation and Protection as Umbrella Principles. *Procedia Computer Science*, 35(2014), 812 – 821.
- SleuthKit. (2014). Sleuth Kit (TSK) & Autopsy 3.0.10. Retrieved 2 December, 2013, from <http://www.sleuthkit.org/>
- Software, G. (2014). EnCase Forensics V7. Retrieved 4 February, 2015, from <https://www.guidancesoftware.com/products/Pages/encase-forensic/overview.aspx>
- Song, D. X., Wagner, D., & Perrig, A. (2000). *Practical Techniques for Searches on Encrypted Data*. Paper presented at the IEEE Symposium on Security and Privacy, Berkeley, CA, USA.
- Spafford, E. (2006). *Some Challenges in Digital Forensics*. Paper presented at the Advances in Digital Forensics II - IFIP International Conference on Digital

Forensics, National Centre for Forensic Science, Orlando, Florida, USA.

Srinivasan, S. (2006). *Security and Privacy in the Computer Forensics Context*. Paper presented at the International Conference on Communication Technology (ICCT'6), Guilin, China.

Srinivasan, S. (2007). Security and Privacy vs. Computer Forensics Capabilities *Information Systems Control Journal*, 4, 1-3.

Stahlberg, P., Miklau, G., & Levine, B. N. (2007). *Threats to Privacy in the Forensic Analysis of Database Systems*. Paper presented at the ACM SIGMOD international conference on Management of data New York, NY, USA.

Stephenson, P. (2002). The Forensic Investigation Steps. *Computer Fraud & Security* (10), 17-19.

Stephenson, P. (2003). A Comprehensive Approach to Digital Incident Investigation. *Information Security Technical Report*, 8(8), 42-54.

Stephenson, P. (2003). The DFRWS Framework Classes. Retrieved 3 March, 2010, from http://people.emich.edu/pstephen/my_papers/DFRWS_Classes.PDF

Stüttgen, J. (2011). *Selective Imaging: Creating Efficient Forensic Images by Selecting Content First*. Diploma Friedrich Alexander Universität Erlangen Nürnberg.

Stüttgen, J., Dewald, A., & Freiling, F. C. (2013). *Selective Imaging Revisited*. Paper presented at the Seventh International Conference on IT Security Incident Management and IT Forensics, Nuremberg (Nürnberg), Germany.

Turner, P. (2005). Unification of Digital Evidence from Disparate Sources (Digital Evidence Bags). *Digital Investigation*, 2(3), 223-228.

Turner, P. (2006). Selective and Intelligent Imaging using Digital Evidence Bags. *Digital Investigation*, 3(1), 559-564.

Turner, P. (2007). Applying a Forensic Approach to Incident Response, Network Investigation and System Administration using Digital Evidence Bags. *Digital Investigation*, 4(1), 30-35.

Ucal, M. (2005). Searching on Encrypted Data: University of Southern California. Retrieved 2 July, 2014, from http://www.slidefinder.net/s/searching_encrypted_data_mehmet_ucal/32171906

Wikipedia. (2014). Data Privacy. Retrieved 23 July, 2014, from <http://en.wikipedia.org/wiki/Privacy>

