UNIVERSITI PUTRA MALAYSIA

PRIVACY-PRESERVING COMPUTER FORENSICS FRAMEWORK

WALEED ABDULJABBAR HALBOOB

FSKTM 2015 20
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 Zinah 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.
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

yang diperlukan untuk mempunyai penyelesaian yang sah, menyediakan pengimejan dan analisis berkesan (pencarian dan penyahsulitan), dan, akhirnya, menyokong kedua-dua carian berdasarkan 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 Mahmod, Associate Professor Dr. Nur Izura Udzip, 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: ____________________  Signature: ____________________
Name of Chairman of Signature: ____________________
Supervisory Signature: ____________________
Committee: ____________________
Name of Member of Signature: ____________________
Supervisory Signature: ____________________
Committee: ____________________
Name of Member of Signature: ____________________
Supervisory Signature: ____________________
Committee: ____________________
Name of Member of Signature: ____________________
Supervisory Signature: ____________________
Committee: ____________________
TABLE OF CONTENTS

ABSTRACT i
ABSTRAK iii
ACKNOWLEDGEMENTS v
APPROVAL vi
DECLARATION viii
LIST OF TABLES xii
LIST OF FIGURES xiv

CHAPTER

1 INTRODUCTION 1
1.1 Background 1
1.2 Problem Statement 2
1.3 Research Objectives 3
1.4 Scope and Limitation 3
1.5 Thesis Organization 4

2 LITERATURE REVIEW 5
2.1 Computer Forensics Overview 5
2.2 Selective Imaging Concept and State-of-the-Art 6
2.3 Privacy Issue in Computer Forensics 8
2.4 Lawfulness Evaluation Criteria 9
2.4.1 Collection Limitation Principle 10
2.4.2 Data Quality Principle 10
2.4.3 Purpose Specification Principle 11
2.4.4 Use Limitation Principle 11
2.4.5 Security Safeguards Principle 12
2.4.6 Openness Principle 12
2.4.7 Individual Participation Principle 12
2.4.8 Accountability Principle 13
2.5 Lawfulness Evaluation Criteria: A Discussion 13
2.6 Current Research Directions 14
2.6.1 Forensics-enabled Privacy Enhancing Techniques 14
2.6.2 Anti-forensics Solutions 14
2.6.3 Privacy-preserving Digital Forensics Approaches 15
2.7 Privacy-preserving Computer Forensics Approaches 15
2.7.1 Policy-based Approaches 16
2.7.2 Cryptographic-based Approaches 17
2.8 Discussion and Evaluation 23
2.8.1 Lawfulness Summary 23
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
    Integration with Computer Forensics Analysis Tools 89
5.5 Comparison 90
<table>
<thead>
<tr>
<th>6</th>
<th>CONCLUSION</th>
<th>94</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Overview</td>
<td>94</td>
</tr>
<tr>
<td>6.2</td>
<td>Conclusion</td>
<td>94</td>
</tr>
<tr>
<td>6.3</td>
<td>Research Contributions</td>
<td>95</td>
</tr>
<tr>
<td>6.4</td>
<td>Future Works</td>
<td>95</td>
</tr>
</tbody>
</table>

REFERENCES 96
APPENDICES 102
BIODATA OF STUDENT 115
LIST OF PUBLICATIONS 116
### LIST OF TABLES

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

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


Analysis. Master, Concordia University, Master thesis.


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


