Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 56 (2015) 370 – 375

Procedia Computer Science

International Workshop on Cyber Security and Digital Investigation (CSDI 2015)

Privacy Levels for Computer Forensics: Toward a More Efficient Privacy-preserving Investigation

Waleed Halboob [a,b,*], Ramlan Mahmod [a], Nur Izura Udzir [a], Mohd. Taufik Abdullah [a]

[a] Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
[b] Center of excellence in Information Assurance (CoEIA), King Suad University, Riyadh, Saudi Arabia

Abstract

Computer forensics and privacy protection fields are two conflicting directions in computer security. In the other words, computer forensics tools try to discover and extract digital evidences related to a specific crime, while privacy protection techniques aim at protecting the data owner's privacy. As a result, finding a balance between these two fields is a serious challenge. Existing privacy-preserving computer forensics solutions consider all data owner's data as private and, as a result, they collect and encrypt the entire data. This increases the investigation cost in terms of time and resources. So, there is a need for having privacy levels for computer forensics so that only relevant data are collected and then only private relevant data are encrypted. This research paper proposes privacy levels for computer forensics. It starts with classifying forensic data, and analyzing all data access possibilities in computer forensics. Then, it defines several privacy levels based on the found access possibilities. The defined privacy levels lead to more efficient privacy-preserving computer forensics solution.

Peer-review under responsibility of the Conference Program Chairs

Keywords: Computer Forensics; Privacy protection; Privacy Levels; Cryptography; Forensic imaging

1. Introduction

The currently widely used procedure for collecting 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

* Corresponding author. Tel.: +966-503792189; fax: +966-11-469 5237.
E-mail address: wmohammed.c@ksu.edu.sa

Peer-review under responsibility of the Conference Program Chairs
doi:10.1016/j.procs.2015.07.222

investigation cost in terms of the required time and resources [1,2]. 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.

In addition, 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 such as specifying privacy policies, collecting only relevant data, using cryptographic techniques, taking into account existing privacy act(s), and auditing the investigation process [3,4,5,6,7,8].

However, collecting the relevant data while ignoring irrelevant data is a key point for privacy preservation in computer forensics, as mentioned earlier, and then only private relevant data are encrypted. But, this needs having some privacy levels that which data have to be collected or not, and encrypted or not. Existing privacy-preserving computer forensics solutions [9,10,11,12,13,14] are not efficient because they require the collection and encryption of all the data. This is because collecting and encrypting all data is not an acceptable solution especially when dealing with shared storages or servers.

In this research paper, privacy levels are proposed for computer forensics. However, the privacy levels help for classifying the forensic data into relevant and non-relevant as well as classifying the relevant data into private and non-private. Hence, at the end, only relevant data are collected, and only private relevant data are encrypted during the collection process. This leads to improve the investigation efficiency by reducing the required costs (in terms of time and resources).

The rest of this paper is structured as follows. Section 2 presents a brief review of the related works on privacy preservation in computer forensics. Section 3 presents the proposed privacy levels for computer forensics. Section 4 discusses the proposed privacy levels. Finally, Section 5 concludes the paper with possible future work.

2. Related Work

Several researches [3,4,5,6,7,8] have studied the conflict between privacy preservation and computer forensics. These studied suggest specifying accountability and before-the-fact privacy policies, along with using cryptographic techniques for preserving the private data during the evidence acquisition phase. Furthermore, they suggest that existing computer forensics tools, such as EnCase [15], Forensic ToolKit [16], must consider existing privacy acts. In addition, cryptographic techniques can be used to provide a privacy protection.

However, existing solutions can be classified into two branches: policy-based and cryptographic approaches. The policy-based approaches can be further classified into two approaches namely policy statements and privacy policies [17,18]. In fact, the goal of using the policy-based approaches is to let the data owner know how his private data should be collected, used, and disclosed. The cryptographic approaches [9,10,11,12,13,14] protect the data owner's private data during the investigation process. They encrypt all relevant and irrelevant data (even private or not) using some cryptographic techniques such as searchable encryption technique. All data owner's data are considered relevant and private, means the entire data are collected and encrypted. This increases the investigation cost in terms of time and resources. As discussed earlier, there is a need for collecting only relevant data as well as encrypting only private relevant data.

In Halboob *et al.* [19], we proposed a computer forensics framework that includes a definition for some privacy levels for computer forensics. Nevertheless, it has been proven that one level is not practical. Though, in this research different and refined privacy levels are defined.

3. The Proposed Privacy Levels

The privacy levels are generally used to explain the levels of privacy protection that the data collector must provide. Recent studies on computer forensics consider all the data owner's data as private. Therefore, they encrypt and encrypt the entire data. In fact, considering all the forensic data to be private requires protecting (e.g., encrypting) all these data, which consumes more time for data encryption and decryption.