Research on virtualized cloud forensics model and workflow

Guangxuan Chen[1, a *], Guomin Zhou[1,b], Guangxiao Chen[2,c], Qiang Liu[1,d]

[1]Zhejiang Police College, Hangzhou, China

[2] Universidad Carlos III de Madrid, Madrid, Spain

[3]College of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, China

[a]ericcgx@163.com, [b]zhouguomin@zjjcxy.cn, [c]rex1220@foxmail.com, [d]liuqiang@zjjcxy.cn

Keywords: digital forensics, cloud computing, cloud forensics, evidence acquisition

Abstract. As the cyber crimes are becoming increasingly rampant and indictable, the traditional digital forensics is no longer qualified in the more complex cloud computing environment. Due to the security issues and specific features of cloud computing, it is easy for the hackers to exploit the weakness of the cloud infrastructure for criminal activities. Meanwhile, the traditional computer crimes are shifting to the cloud computing architecture in which it is more difficult detected. This paper analyzed the challenge of digital forensics under cloud computing environment compared to that in traditional information systems and proposed a cloud forensics workflow model. The model is of referential value in the digital forensics under cloud computing environment.

1. Introduction

With the rapid development of information technology, computer crimes or cyber crimes are becoming increasingly rampant and indictable. Meanwhile, the cloud security issues are also becoming increasingly prominent, combined with various cloud related cyber crimes. The cyber criminals usually harness the various cloud providers' services directly or take the services as a bridge to commit other attacks and criminal activities. It is the uncertainty of the storage site of the cloud data and the real-time character of the data stream that make the traditional computer forensics increasingly incompetent when facing with cloud related cyber crimes.

It is critical for the justice department to establish an effective cyber crime investigation and digital forensics method to meet the sensitive and urgent needs. However, compared to the traditional information systems, cloud computing architecture is more complex with huge scale data amount and more diverse types of applications [1-2]. The now available way to acquire evidence is extremely difficult and is inefficient in dealing with large number of logs. The simple, discrete, non-system based clues and afterwards-forensics way of traditional digital forensics is no longer adapt to the new cloud computing environment [3]. How to discover and locate the digital forensics timely and accurately is besetting the investigators.

2. Challenge of cloud forensics

2.1 Security issues of cloud computing. According to Cloud Security Alliance (CSA) [4], there are mainly nine categories of risk in cloud computing environment, namely data breaches, data loss, account or service traffic hijacking, insecure interfaces and APIs, denial of service, malicious insiders, abuse of cloud services, insufficient due diligence and shared technology vulnerabilities. These weakness of the cloud service has gave the hackers, internal employees of cloud service providers, cyber-grifters and other cyber criminals the opportunity to commit cyber crimes.

2.2 Differences between traditional digital forensics and cloud forensics. There're a lot of differences of digital forensics between cloud computing environment and traditional information systems, which brings great challenge to the traditional digital forensics in the cloud computing era. These main differences can be reflected in the following aspects (shown in Table 1):

(1) Complexity of forensics scene and forensics environment

Single machine forensics and network related forensics are the two typical digital forensics methods in the traditional information environment. The two methods are direct with relatively clear forensics targets (the computer, medium and other digital devices seized from the suspects). While cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction [1]. This cloud model is composed of five essential characteristics, three service models (SaaS, PaaS, IaaS), and four deployment models (private cloud, public cloud, hybrid cloud and community cloud). Thus, it can be seen, investigators are facing with more complex forensics scene in the cloud computing era. Traditional computer architecture based forensics rules are no longer competent in the cloud computing environment.

(2) Evidence acquisition

Firstly, according to international digital forensics practice, digital forensics consists of four phases, i.e., evidence seizure, evidence acquisition, evidence analysis, and evidence presentation. As for the traditional digital forensics environment, it is clear to ascertain the devices to be seized. These devices, such as, PC, laptop, cellphone, storage medium, switch, firewall, server, usually in small quantities and easily moved. While in the cloud computing environment, cloud model involves a large number of server clusters, storage devices and network devices that deployed at different places. It is unpractical for the investigators to seize the complicated cloud system.

Secondly, it is practical to adopt the mature method to get the valid information or mirror the memory and disk in the traditional environment. While, in the cloud computing environment, virtual machines are the major roles in implementing various applications. It is difficult for the investigators to ascertain the location of the virtual machine at a specific time.

(3) Evidence analysis

Firstly, as for the traditional digital forensics environment, the mediums involved in the case are limited. The investigators are able to obtain a lot of valid evidence or clues from the mediums. In contrary, it is nearly impossible for the investigators to deal with the vast amount of data in the cloud to find the valid evidence and clues.

Secondly, there are a lot of mature digital forensics technology and products aimed to traditional digital forensics scene, such as ENCASE, X-WAYS, CELLdek. Meanwhile, there are related digital forensics rules and manuals circulated among investigators. While in the cloud environment, there aren’t any practicable forensics technologies and products available, not to mention the official manuals. Online forensics is the only way available and the evidence acquired usually doesn’t possess integrity and reproducibility.

Table 1. Main differences of digital forensics between cloud computing architecture and traditional information system

	Cloud computing architecture	Traditional information system
Forensics environment	Public cloud, Private cloud, hybrid cloud; SaaS, PaaS, IaaS	Single machine, computer networks
Evidence acquisition	Difficult to carry out physical seizure and online forensics is possible; Difficult to locate the evidence of a specific time point	Easy to carry out physical seizure; Easy to locate the evidence
Evidence analysis	Limited valid evidence from massive data; Lack of mature rules, technology and tools; Difficult to achieve the integrity and reproducibility of the evidence	Plenty of valid evidence from limited medium; Mature rules, technology and tools; Ensured integrity and reproducibility of the evidence

3. Cloud forensics model and workflow

The main thought of cloud forensics model is to put the collaborative technology into traditional computer forensics [5]. The hosts, distributed storage units, networks and RDB in the infrastructure layer are virtualized into multiple virtual machines in the server clusters. All the virtual machines are formatted into an independent virtual layer that can greatly cut down the purchase cost and maintenance cost. The forensics unit for the cloud is mounted in this virtual layer that makes data acquisition and log analysis efficiently with minor cost. The cloud forensics workflow is shown in Fig.1.

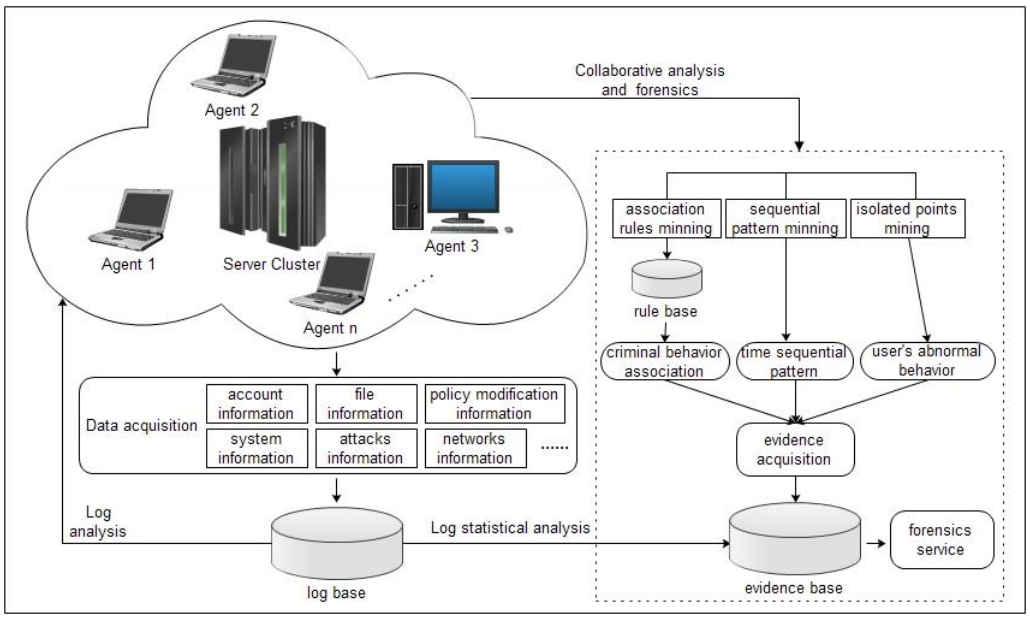

Fig.1 Basic cloud forensics workflow

The basic cloud forensics workflow can be described as follows. Log information acquisition will be conducted through data acquisition agents in the cloud and then the information will be stored in the log base after format unification. Then, collaborative forensics and analysis between the hosts and

server clusters in the cloud will be carried out. Such as, a large number of association rules in the log that according to a specific distribution rule will be mined and abstracted through association rule mining procedure; and the time sequence of the intrusions in which the time and event sequence features of the cyber crimes will be extracted can be found through sequential pattern mining procedure; and the isolated point analysis procedure can help us find the abnormal data pattern and further analysis can be conducted on this basis in order to get the valid evidence. The evidence acquired through above procedures and log base statistical analysis will be merged into the evidence base used for providing forensics service.

4. Conclusion

In order to better combat cyber crimes, especially that related to cloud services, it's necessary to develop cloud forensics method to assist the traditional digital forensics which is no longer qualified in the cloud computing environment. The cloud forensics model and workflow proposed in this paper is competent in massive data acquisition, log analysis, evidence acquisition and collaborative forensics analysis through procedures as association rules mining, sequential pattern mining and isolated points mining. And then, further analysis on user abnormal behaviors and criminal behaviors are conduct through the evidence base and collaborative forensics mechanism. The model is of referential value in the digital forensics under cloud computing environment both for the cloud service providers and law enforcement.

5. Acknowledgement

This study is supported by Scientific Research Program of Zhejiang Police College (No. 20140629), Scientific Research Project of Zhejiang Educational Department (No.Y201329872), Public Security Theory and Soft Science Research Project of Ministry of Public Security (No. 2015LLYJZJST035), and Fund of Key Laboratory of Public Security Information Application Based on Big-data Architecture, Ministry of Public Security, P.R. China.

References

[1] Information on http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf

[2] G.X. Chen, Y.H. Du, P.K. Qin, L. Zhang and J. Du, A New Single Sign-on Solution in Cloud, Lecture Notes in Electrical Engineering. 277 (2013) 755-761.

[3] G.X. Chen, Y.H. Du, P.K. Qin and J. Du, Suggestions to digital forensics in Cloud computing ERA, 2012 3rd IEEE International Conference on Network Infrastructure and Digital Content (IC-NIDC), (2012)540-544.

[4] Information on https://cloudsecurityalliance.org/

[5] W. G, Research on sequential pattern mining algorithm and its application for cloud forensics, Shandong Normal University, 2012.