Digital Forensic Techniques for Investigation of Cybercrimes in Cloud Computing Environment

FINAL SYNOPSIS SUBMITTED
TO
MANGALORE UNIVERSITY
FOR THE AWARD OF THE DEGREE OF DOCTOR OF PHILOSOPHY IN
COMPUTER SCIENCE

By
EZZ ELDIN BADAWY GAD ALRAB HEMDAN
Research Scholar
Department of Post-Graduate Studies and Research in Computer Science,
Mangalore University, Mangalagangothri-574199
Mangalore, Karnataka, India

Under the supervision of
Dr. MANJAIAH D.H
Professor
Department of Post-Graduate Studies and Research in Computer Science,
Mangalore University, Mangalagangothri-574199
Mangalore, Karnataka, India

November 2017
FINAL SYNOPSIS SUBMITTED
TO
MANGALORE UNIVERSITY
FOR THE AWARD OF THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
COMPUTER SCIENCE

Title: Digital Forensic Techniques for Investigation of Cybercrimes in Cloud Computing Environment

Name of Candidate: Ezz Eldin Badawy Gad Alrab Hemdan

Name of Supervisor: Dr. Manjaiah D.H

Date of Registration: 02-11-2014

Place of Work: Department of Post-Graduate Studies and Research in Computer Science, Mangalore University, Mangalagangothri-574199, Mangalore-Karnataka-India
Digital Forensic Techniques for Investigation of Cybercrimes in Cloud Computing Environment

Preface

Security of cloud computing environment and its services and infrastructure has become one of the significant research directions especially digital forensics. The digital forensics is a highly specialized and interdisciplinary field, which requires a deep understanding of the underlying technical, regulatory, legal and other aspects, besides intimate knowledge of temporal trends-historical, modern and emerging trends. Applying digital forensic techniques and principles in a cloud computing environment which known as ‘Cloud Forensic’ is facing complex obstacles where little research has done to cover this critical area. There are numerous multifaceted challenges that obfuscate cloud forensic. These challenges such as crime scene reconstruction, isolating cloud instance, data provenance, evidence segregation, and dynamic nature of cloud computing. This makes the cloud forensic investigation as a tough mission for digital investigators to cope with cybercrimes in the cloud environment in order to reconstruct crime events about an incident that occurred and introduce a digital evidence to a court of law as admissible proof for the committed cloud based-crime.

Currently, cloud computing services and infrastructures suffer from the absence of the support for digital forensics. Furthermore, the current digital forensic techniques, methods and tools cannot contend with the dynamic nature of the cloud environment. Therefore, this work is to identify and explore cloud forensic challenges and opportunities as well as provides a number of proposed solutions for recent critical forensic issues related to the investigation of cloud-based crimes in forensically sound and timely fashion manner.

The rest of this pre-synopsis is organized as follows: Section 1 provides an introduction to cloud forensic area while the research motivation is presented in section 2. The literature review is presented in section 3 while the research statement is introduced in section 4. The research objectives are discussed in section 5 while the research contributions are provided in section 6. The conclusions and future works in the subject are introduced in section 7 while the thesis organization is presented in section 8.

1. Introduction

In recent times, cybercrime has become one of the critical threats to countries, organizations, and individuals around the world because of the problems that are caused breaching security
of their digital systems such as computers, servers and storage systems…etc. Therefore, these countries, organizations, and individuals take various strategies and procedures to protect and defend against these threats by securing their systems using firewalls, anti-viruses, intrusion detection, and prevention system and countless other protection mechanisms. Infrequently, these protection mechanisms perhaps not prevent attackers from breaching systems security, so as to, there is a serious need for methods or procedures that are able to discover and trace attackers and criminals who attacked the systems as well as determine the weakness of these systems. The science which concerns with discovering and tracing the attackers called digital forensics. The digital forensics is used to extract and analyze digital evidence which collected and acquired from a crime scene to reconstruct the crime events that occurred and provide it as an admissible proof about crime in a court of law.

Today, conventional view to cybercrimes has changed because the appearance of new technologies such as cloud computing that depend on a theory of distributed systems. Cloud computing infrastructures are distributing around the world to provide services for organizations and individuals with a method of on-demand manner or cost-per-use. Thus, traditional cybercrime techniques are changed to adapt to a dynamic landscape of the cloud computing environment. In the same time, digital investigators have to change and expand their tools and techniques to deal with the new types of cybercrimes in the cloud.

To understand the prominence of the digital forensic investigation in the cloud, two hypothetical criminal scenarios of cybercrimes that may occur in a cloud environment will explain. These scenarios are online banking service and malicious user in the cloud respectively.

1. **Online Banking Service:** A criminal scenario that can be occurred within an online banking service is explained as follows [1]: *Alex is a customer on online banking service from Bank XYZ. The banking service using datacenter which installed VMware ESXi server 5 to host virtual machines and the affected virtual guest resource running windows server 2012 R2 Web Server Edition. The customer believes that hackers hacked and exploited an application-level flaw in the guest operating system to steal his information such as name, username, password, credit card information and other critical information about the customer account. The bank has identified the flaw and can patch the system, but is first interested in preserving forensic evidence to support a criminal investigation or civil suit. Technical staff determined that the flaw would not have enabled persistent access to the guest server. The bank has already disabled inbound web access*
to the virtual guest, but the system needs to be kept up and online to support other critical business services.

2. Malicious User in the Cloud: A malicious activity that can be occurred in the cloud by a malicious user is described as shown in Figure 1 as follows [2]: Bucky is a cloud user who runs a Virtual Machine (VM) in a cloud. John is a malicious user. He rented also another VM in the cloud to enable him to access cloud infrastructure. John decided to use Bucky's VM to launch many types of attacks to other VMs running in the cloud to steal their data. One of the attacked VM's owner stores important data which are stolen by the malicious attacker. Consequently, the owner asked a forensic investigator to investigate the case. The investigator found the Owner's VM records each visiting activity information like IP address. Analyzing the visiting customer records, the investigator found that the attacking has done through Bucky's VM. Eventually, the investigator issued a subpoena to the corresponding cloud provider to provide him the network logs for this particular IP address. On the other hand, there are two scenarios may occur. Firstly, John managed to collude with the cloud provider after the attack. Therefore, while providing the logs to the investigator, the cloud provider supplied a tampered log to the investigator, who had no way to verify the correctness of the logs. Under this circumstance, John will remain undetected. Secondly, Even if the cloud provider was honest, John could terminate her rented VM and leave no trace of the attack. Hence, the cloud provider could not give any useful logs to the investigator. Finally, if the investigator can't find the truth, Bucky will be responsible for the malicious attack. To identify the actual attacker in the above malicious attack scenario, there is a need to introduce new techniques that are able to execute digital forensics procedures successfully in the cloud.

![Figure 1: Hypothetical Scenario for Malicious Activity in the Cloud.](image)
2. Research Motivation

Recently, cloud computing has become one of the most popular and significant computing paradigms. The cloud computing is an emerging revolutionary technology that has started changing the ways people live and work. A recent research by Market Research Media states that the global cloud computing market is expected to grow at a 30% Compound Annual Growth Rate (CAGR) reaching $270 billion in 2020 [3]. With the growth of cloud and related services, security and privacy in the cloud have become very critical issues in cybersecurity field where hackers can use cloud infrastructure with exceptional bandwidth, storage, and computing power to launch their illegal activities. According to a recent IDC IT Cloud Services User Survey, 74% of IT executives and CIOs have referred security as the main reason to prevent their migration to the cloud services model [4].

There are numerous complex challenges such as crime scene reconstruction, isolating cloud instance, data provenance, evidence segregation, and dynamic nature of cloud computing. These challenges are making forensic investigation process more challenging task for digital investigators to investigate cloud-based crimes, extract a digital evidence and reconstruct cybercrime events and finally send a forensic report to a court of law as admissible evidence about the committed crime that occurred in the cloud environment. Little work has been done to explore and introduce new theoretical and practical methods in cloud forensic area to support security professionals, cloud investigators and law enforcement officers to investigate incidents in the cloud in forensically sound and timely fashion manner.

Cloud forensic is challenging at best but can be performed in a manner consistent with law using new techniques, procedures, and tools. Several researchers have explored security issues relating to data stored in clouds which make cloud cannot be used to store data for many sectors such as business, healthcare, banks or national security agencies, which require an audit and regulatory compliance. Though, some researchers have clarified and identified challenges and opportunities along with starting design and develop techniques and tools to easily accomplish cybercrimes investigation in cloud computing environment. In addition to there are little case studies of cloud investigations of cybercrimes to illustrate and explain the issues and to educate practitioners in this domain.

The major challenges of cloud forensics originate from the very characteristics with which the cloud computing platform is identified. A number of cloud forensics challenges that obfuscate cloud forensic are such as each cloud server contains documents from many users. Hence, it is not possible to seize servers from a data center without violating the privacy of
various other customers, the storage system is no longer local. Therefore, even with a subpoena, law enforcement agents cannot seize the suspect’s machine and get access to the suspect’s documents and although the data belonging to a particular suspect is recognized, detaching it from other consumers’ data is problematic. Moreover, except the cloud provider’s word, there is usually no evidence that links a given data file to a specific suspect.

In the other side, scores of opportunities are available through utilizing massive cloud resources for supporting cloud forensic such as Forensic as a Service (FaaS) which introduce one powerful option for digital investigators through utilizing the vast capabilities of cloud computing [6]. This makes digital forensics as an “on-demand” service for allowing for as much storage and processor power as needed to conduct an investigation. In this case, a forensic server will reside on the cloud side, offline, until the need arises for them. Documents could be backed up to the cloud for investigators to use without having to disrupt normal business. Naturally, the cloud resources could be used for sorting, searching, and hashing the evidence data.

3. Literature Review
Digital forensic techniques for investigation of cybercrimes in cloud computing (cloud forensic) is comprising various areas such as digital forensics and cloud computing. The aim of this section is to gain a deeper understanding and identify challenges and opportunities in forensic science concerns in cloud computing as well as current research review which can help to develop strategies for measurements, standards and technology research to mitigate these challenges that cannot be handled with present technology and approaches. Despite significant research in digital forensics, little has been written about the applicability of digital forensic to cloud computing environment. This section provides State-of-the-Art of research work related to cloud forensic to understand current challenges and proposed solutions.

3.1 Current Research Problems in Cloud Forensics
In this division, a study of open problems and impediments in cloud forensics will discuss as follows: Collection and Acquisition of Forensic Data, Log Information, Service Level Agreement (SLA), Virtual Machine Introspection (VMI), Data Provenance in Cloud, Trusted Platform Module (TPM), Isolating a Cloud Instance and Forensic Analysis for Cloud Storage Services.
• **Collection and Acquisition of Forensic Data**

Data collection and acquisition from cloud infrastructure is a challenging step in cloud forensics. Cloud Service Providers (CSPs) can play a vital role in this step by providing a web-based management console like AWS management console. Dykstra et al. [7] recommended a cloud management plane for use in IaaS service model. From the console panel, customers, as well as investigators, can collect VM image, network, process, database logs, and other digital evidence, which cannot be collected in other ways. The only problem with this solution is that it requires an extra level of trust – trust in the management plane. In traditional evidence collection procedure, where we have physical access to the system, this level of trust is not required. Dykstra et al. [7] proposed a trust model with six layers: Guest application/data, Guest OS, Virtualization, Host OS, Physical Hardware, and Network. The further down the stack is, the less cumulative trust is required. There are many issues that make the data acquisition in cloud forensic is challenging. These issues such as physical inaccessibility of digital evidence, dependence on the CSP, volatile data in the cloud, less control in cloud infrastructure, legal and trust issues, multi-tenancy, and large bandwidth.

• **Log Information**

One of the significant digital evidence for forensic investigation is log information. Some researchers have identified logging related to the cloud. There are many challenges related to log information in cloud forensics such as decentralization, volatility of logs, multiple tiers and layers, accessibility of logs, dependence on the CSP and absence of critical information in logs.

Marty et al. [8] proposed a log management solution, which can solve many logging challenges. The proposed approach consists of three steps. In the first step, logging must be enabled on all infrastructure components to collect logs. The second step is for establishing a synchronized, reliable, bandwidth efficient, and encrypted transport layer to transfer log from the source to a central log collector. The final step deals with ensuring the presence of the desired information in the logs.

To facilitate logging in the cloud, Zafarullah et al. [9] proposed logging provided by OS and the security logs. In order to investigate digital forensics in the cloud, they set up cloud computing environment by using Eucalyptus. Using Snort, Syslog, and
Log Analyzer (e.g., Sawmill), they were able to monitor the Eucalyptus behavior and log all internal and external interaction of Eucalyptus components. For their experiment, they launched a Distributed Denial of Service (DDoS) attack from two virtual machines and analyzed bandwidth usage log and processor usage log to detect the DDoS attack. From the logs in /var/eucalyptus/jetty-request-05-09-xx file on Cloud Controller (CC) machine, it is possible to identify the attacking machine IP, browser type, and content requested. From these logs, it is also possible to determine the total number of VMs, controlled by single Eucalyptus user and VMs communication patterns. Their experiment shows that if the CSPs come forward to provide better logging mechanism, cloud forensics will be benefited greatly.

To get essential logs from all the three cloud service models, Birk et al. [10] proposed that the CSP could provide process, network and access logs to the user by read-only Application Programming Interface (APIs). By using these APIs, the user can provide appreciated information to digital investigators. In PaaS, users have full control on their application and can log a variety of access information. They proposed a central log server, where the user can store the log information. In order to protect log data from possible eavesdropping and modifying the action, the users can encrypt and sign the log data before sending it to the central server.

- **Robust Service Level Agreement (SLA)**
  In the recent time, there is an immense gap in the existing Service Level Agreement (SLA), which neither defines the responsibility of CSPs at the time of some malicious crimes nor their role in the digital forensic investigation. Researchers gave emphasis on sound and robust SLA between cloud service providers and users [11]-[12]. To resolve the transparency issues, the CSP should build a long-term trust relationship with users or customers. A robust SLA should state how the providers deal with the cloud-based crimes, i.e., how and to which extent they help in the forensic investigation process. In this context, another question can come – how we can be sure of the robustness of an SLA. To ensure the quality of an SLA, it can by taking assistance from a Trusted Third Party.

- **Virtual Machine Introspection (VMI)**
  Virtual Machine Introspection (VMI) is the process of externally monitoring and auditing the runtime state of the virtual machine (VM) from either the Virtual Machine
Monitor (VMM) or from some virtual machine other than the one being examined. By runtime state, it is referring to memory, registers, disk, processor, network, and other hardware-level events. Through this process, can execute a live forensic analysis of the system, while keeping the target system unaffected. Hay et al. [13] presented that if a VM instance is compromised by installing some rootkit to hide the malicious events, it is still possible to identify those malicious events by performing the VMI process. They used an open source VMI library, Xen (VIX) suite to perform their experiment. However, this tool is no longer maintained under this name, it is now known as LibVMI [14].

- **Data Provenance in the Cloud**

Provenance plays an important role in the forensic investigation in the cloud because provenance provides the history of an object. By implementing secure provenance, we can get some important forensic information, such as, who owns the data at a given time, who accesses the data, and when. Some researchers have applied the principles of provenance to cloud forensics. Secure provenance can ensure the chain of custody in cloud forensics as it can provide the chronological access history of evidence, how it was analyzed, and preserved. There have been several projects for secure provenance in cloud computing [15]-[16], but no CSP has practically implemented any of the mechanisms yet.

- **Trusted Platform Module (TPM)**

To preserve the integrity and confidentiality of the data, several researchers proposed Trusted Platform Module (TPM) as a solution. The TPM for cloud computing proposed by several researchers for ensuring trust in cloud computing [17]-[18]. By using TPM, can get machine authentication, signing, hardware encryption, secure key storage, and attestation. It can provide the integrity of the running virtual instance, trusted log files, and trusted deletion of data to customers. However, Dykstra et al. [7] declared that TPM is not totally secure and it is possible to change a running process without being discovered by the TPM. Furthermore, at current, CSPs have heterogeneous hardware and few of them have TPM. Hence, the CSPs cannot ensure a homogeneous hardware environment with TPM in the future.
• **Isolating a Cloud Instance**

A cloud instance must be isolated if any incident takes place in that instance. Isolation is necessary because it helps to protect evidence from contamination. However, as multiple instances can be located in one node, this task becomes challenging. Delport et al. [19] provided some possible techniques for cloud isolation such as Instance Relocation, Server Farming, Failover, Address Relocation, Sandboxing, and Man in the Middle (MITM). These techniques are as follows:

- **Instance Relocation Technique**: To move an instance, data on the secondary storage, the content of the virtual memory, (e.g., swap memory), and the running processes must be moved. Relocation can be done in two ways – manual and automatic. In the manual mode, the administrator has all the power to move the instance. In automatic mode, CSP moves the instance from one node to another. While moving, the challenge is to ensure confidentiality, integrity, and availability of other users’ data.

- **Server Farming Technique**: In which can be used to re-routing request between user and node.

- **Failover Technique**: Where there is at least one server that is replicating another. There are three ways of failover – Client-based failover, DNS-based failover, and IP-address take over. Address relocation is another technique, which is actually a special case of DNS-based failover. When it is detected that the main computer has failed, the traffic is rerouted to the backup server. However, this technique depends on the success of replication.

- **Sandbox Technique**: It can also isolate an instance by placing it in a sandbox. One approach of creating a sandbox is installing a sandboxing application in cloud operating system. Another approach is creating a virtual box around an instance and observes all the communication channel.

- **Man in the Middle (MITM) Technique**: It can be by placing a MITM between cloud instance and hardware. In that manner, can get log information from the network, RAM, CPU, and hard drive. To get advantage from this technique, the CSP should embrace this technique for implementation in its cloud infrastructure.

• **Forensic Analysis for Cloud Storage Services**

Criminal can keep their secret files (e.g., child pornography, terrorist documents) in cloud storage and can destroy all evidence in local storage to remain clean. Cloud
storage services are part of cloud computing services that are subject to exploit by attackers to steal or modify data in cloud storage. Store data in the cloud storage that are remotely distributed on cloud servers in overseas jurisdictions rather than in local machines make a new challenge for forensic practitioners and law enforcement agencies to acquire digital evidence for analysis and examination in a forensically manner to be admissible in the court of law.

Quick and Choo [20]-[22] studied data remnants on client devices and found that there is information in cloud storage accounts(i.e. Dropbox, Microsoft SkyDrive and Google Drive) which is not available on user machine which may either accessed an account through web browser or is synchronized to an account using the client software[23]. This information includes previous and historical versions of files and information that identify the cloud storage user such as computer name, IP address, times and dates associated with the modification made in his account’s contents. Quick and Choo also explore methods to preserve cloud-stored information because there is critical information in the cloud account which may not available in user machines. This information may help investigators to collect vital evidential data to reconstruct crime event related to cloud storage.

Shams and Hassan [5], they summarize some of the challenges in three service models of cloud computing for publicly deployed cloud in Table 1.

<table>
<thead>
<tr>
<th>TABLE 1: SOME OF THE CHALLENGES IN CLOUD FORENSICS [5]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Challenges of Cloud Forensics</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Physical inaccessibility</td>
</tr>
<tr>
<td>Dependence on CSP</td>
</tr>
<tr>
<td>Volatile Data</td>
</tr>
<tr>
<td>Trust Issue</td>
</tr>
<tr>
<td>Large bandwidth</td>
</tr>
<tr>
<td>Multi-tenancy</td>
</tr>
<tr>
<td>Decentralization of Logs</td>
</tr>
<tr>
<td>Volatility of logs</td>
</tr>
<tr>
<td>Logs in multiple tiers and layers</td>
</tr>
<tr>
<td>Accessibility of logs and Depending on CSP for logs</td>
</tr>
<tr>
<td>Chain of Custody</td>
</tr>
<tr>
<td>Problem of current forensic tools</td>
</tr>
<tr>
<td>Crime scene reconstruction</td>
</tr>
<tr>
<td>Cross-border law</td>
</tr>
</tbody>
</table>
3.2 Existing Systems and Current Solutions
Shams Zawoad and Ragib Hasan [5], they mentioned and discussed some existing proposed solutions related to Cloud Forensics, which can mitigate some of the challenges of cloud forensics as follow in Table 2.

<table>
<thead>
<tr>
<th>Challenges</th>
<th>Proposed Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trust issue for depending on CSP</td>
<td>Trust Model</td>
</tr>
<tr>
<td>Preserving integrity</td>
<td>Distributed signature detection framework</td>
</tr>
<tr>
<td>Decentralization of logs, logs in multiple tiers and layers, absence of critical information in logs, Volatility of logs</td>
<td>Log management solution</td>
</tr>
<tr>
<td>Depending on CSP for logs</td>
<td>API provide by CSP for logs</td>
</tr>
<tr>
<td>Dependability on CSP for data acquisition</td>
<td>Cloud management plane</td>
</tr>
<tr>
<td>Compliance issue, dependability on CSP</td>
<td>Robust SLA</td>
</tr>
<tr>
<td>Compliance issue, Developing a robust SLA</td>
<td>Trusted third party</td>
</tr>
<tr>
<td>Cross-border law</td>
<td>Global unity</td>
</tr>
<tr>
<td>Live forensics issue</td>
<td>Virtual machine introspection</td>
</tr>
<tr>
<td>Volatile Data</td>
<td>Continuous synchronization</td>
</tr>
<tr>
<td>Trust issues of cloud computing</td>
<td>Trusted platform module (TPM)</td>
</tr>
<tr>
<td>Multi-tenancy issue</td>
<td>Isolating a cloud instance</td>
</tr>
<tr>
<td>Chain of custody</td>
<td>Data provenance in cloud</td>
</tr>
</tbody>
</table>

3.3 Research Tools and Test Environments
Researchers and investigators who work in digital forensic investigation field, they used many forensic tools and test environments to perform digital forensic investigation process. There are many forensic tools that are used in digital forensics such as Encase and FTK. There is research required to either modify existing tools or to introduce new tools tailored to meet cloud forensics needs. For instance, outsourcing infrastructure in most of the cloud service models increases the need for tools that are capable of performing analysis and examination using a secure remote connection. Also, it requires digital forensics tools to acquire and process memory and network dumps. A summary of most of the tools used to perform digital forensics evidence extraction or analysis are listed in Table 3.

To test the theoretical approaches, researchers need to identify a suitable cloud test environment for their experiment testbed. Table 4 summarizes the test environments that were used that could suit a cloud computing projects in general and cloud forensics scenarios in
particular. These environments are designed with the cloud computing structure and characteristics in mind.

**TABLE 3: SUMMARY OF DIGITAL FORENSIC TOOLS [24]**

<table>
<thead>
<tr>
<th>Used Tool(s)</th>
<th>Possible Usability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wireshark [26]</td>
<td>Captures network traffic between VM and the CSP</td>
</tr>
<tr>
<td>Microsoft Expression Encoder4 [27]</td>
<td>VM windows video recorder</td>
</tr>
<tr>
<td>FTK Imager [28], EnCase [29]</td>
<td>Acquisition of memory and disk images.</td>
</tr>
<tr>
<td>FTK Remote Agent [30]</td>
<td>Acquisition of evidence remotely</td>
</tr>
<tr>
<td>Encase Remote Agent [30]</td>
<td></td>
</tr>
<tr>
<td>X-Ways [31]</td>
<td>Acquisition of Windows and Linux live system</td>
</tr>
<tr>
<td>Sleuth Kit Hadoop [32]</td>
<td>Faster processing of video files for forensic acquisition and analysis (initial stage framework)</td>
</tr>
<tr>
<td>FROST [30]</td>
<td>Digital forensics tools for the OpenStack cloud platform</td>
</tr>
<tr>
<td>XenAccess [33]</td>
<td>Xen VM introspection library (Hypervisor level)</td>
</tr>
<tr>
<td>VMWatcher [34]</td>
<td>VMware VM introspection (Hypervisor level)</td>
</tr>
<tr>
<td>VMwall [35]</td>
<td>VMware VM introspection (VM level)</td>
</tr>
</tbody>
</table>

**TABLE 4: SUMMARY OF TEST ENVIRONMENTS USED FOR CLOUD FORENSICS [24]**

<table>
<thead>
<tr>
<th>Cloud Test Environment</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BonFire [36]</td>
<td>An EU project enables operating a multi-site cloud-based facility on top of different infrastructure testbeds such as Emulab.</td>
</tr>
<tr>
<td>Eucalyptus [37]</td>
<td>A software used to build Amazon Work Station (AWS) private and public cloud.</td>
</tr>
<tr>
<td>OpenNebulla [38]</td>
<td>An industry standard used to provide virtual data centres and IaaS.</td>
</tr>
<tr>
<td>CloudSim [39]</td>
<td>A solution to create large-scale cloud computing data center, virtual hosts, and capability of analysis for network traffic.</td>
</tr>
<tr>
<td>Emulab [40]</td>
<td>Public facility available for researchers to develop, debug and evaluate their systems.</td>
</tr>
<tr>
<td>OpenStack [41]</td>
<td>A project used to create various IaaS architectures such as storage, compute and network.</td>
</tr>
<tr>
<td>Rackspace [42]</td>
<td>Based on OpenStack and provides IaaS.</td>
</tr>
<tr>
<td>Amazon[43],[44],[45]</td>
<td>An appropriate solution provides a various flavour of IaaS, Amazon Simple Store Service (S3) as storage, Amazon Elastic Comput Cloud (EC2) as a computation required for AWS. Amazon Elastic Block Store (EBS) used for backups.</td>
</tr>
</tbody>
</table>
3.4 Computer Forensics via Cloud Forensics

Computer forensics conventionally depend on having physical access to systems, providing digital investigators with the ability to acquire and interact with hardware resources such as memory and disks. For instance, extracting data from magnetic drives has been a principal of computer forensics where digital investigators and examiners establish a chain of custody, create a forensically sound image of a drive, and interact with it in a non-volatile state.

In the Cloud, digital investigators and users do not have access to the physical hardware and resources are shared, classical computer forensic techniques are not easily applied. When a file has deleted the mapping to that file is destroyed instantly and space can then be overwritten and can happen rapidly. If an image itself is shut down it disappears unless put in a suspended mode where you are still charged for it. Backups of data and images can be made, but given the variables, the digital investigators would need to be able to have snapshots of particular instances in time.

A comparison study is tabulated in Table 5 to clarify the differences between computer forensics and cloud forensics.

<table>
<thead>
<tr>
<th>TABLE 5: COMPARISON BETWEEN COMPUTER FORENSICS AND CLOUD FORENSICS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Computer Forensic</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Access Control</strong></td>
</tr>
<tr>
<td><strong>Application</strong></td>
</tr>
<tr>
<td><strong>Database</strong></td>
</tr>
<tr>
<td><strong>Operating System</strong></td>
</tr>
<tr>
<td><strong>Compute</strong></td>
</tr>
<tr>
<td><strong>Storage</strong></td>
</tr>
<tr>
<td><strong>Network</strong></td>
</tr>
</tbody>
</table>

Page | 13
4. Research Statement

The research statement is adopted as the following:

“Cloud Forensic cannot be performed as computer and network forensic in traditional IT systems due to distributed and dynamic nature of cloud computing but can be performed in a manner that is consistent with law by developing and introducing practical digital forensic procedures, approaches and techniques to help Cloud Investigators for investigation of cloud-based crimes in legal and easy manner. For instance, cloud forensic storage require understand and identify data remnants that are available after accessing, using and conducting client systems, moreover analyzing large generated log data from cloud services and infrastructures in order to reconstruct timeline of crime events as well as investigate private cloud environment plus investigate crimes in clustered VMware hypervisors along with the need for building cloud architecture support digital forensics besides utilizing immense cloud capabilities to provide forensic as a service for countries, organizations and individuals”.

This statement reflects several points. First, cloud storage services need more understanding from a digital forensics point of view to facilitate investigation of crimes related to cloud storage services. Second, Cloud infrastructures and services generating an immense amount of log data which need high and fast processing platforms to reconstruct timeline about cloud-based events that may assist in tracking criminals in an effective manner. Third, VMware provides hypervisors for deploying and serving virtual machines. A cluster of ESXi hypervisors can be subject for criminals so that there was little work to provide tools for extract and acquire digital evidence from them. Fourth, One of the drawbacks of cloud systems is it is no support digital forensics therefore, there is a serious need to build a cloud forensic model for being ready for further investigation of cloud-based crimes. Finally, utilizing the huge capabilities of storage and computing resources in building cloud-based forensics laboratories for training students, law enforcement officers as well as building a central laboratory at a country level and other countries to enhance collaboration between their forensics agencies in fighting and investigation severe cyber-attacks.

5. Research Objectives

The central theme of this research work concerns on applying digital forensic techniques in cloud computing environment for the investigation of cybercrimes. Applying digital forensic techniques and principles in a cloud environment facing complex challenges and obstacles which are required to be resolved to perform a convenient cloud forensics. Thus, this thesis
concerns with developing proficient digital forensic techniques for investigation of cybercrimes in cloud computing environment in forensically sound and timely manner. The aim of this research is to study and resolve the following issues:

- Explore and identify challenges and opportunities for performing digital forensics investigation in cloud computing environment.
- Secure cloud storage system with understand and identify data remnants that are available after accessing, using and conducting cloud storage as well as provide a forensic investigation method for investigating cloud storage crimes.
- Perform forensic analysis of large volume of cloud log data to help in reconstructing a timeline of cybercrime events in the cloud.
- Provide a method for investigation of cybercrimes in private cloud environment.
- Develop an application to help in the extraction of digital evidence from VMware clustered ESXi hypervisors.
- Design and develop a cloud computing architecture model to support digital forensics.
- Provide a way to utilize massive cloud capabilities for building a cloud forensic laboratory to investigate cybercrimes as well as provide a learning environment for law enforcement officers, researchers, and students in universities.

6. Research Contributions

This research work concerns with developing proficient digital forensic techniques for investigation of cybercrimes in cloud computing environment in forensically sound and timely manner. It is introduced decent different contributions to the field of cloud forensics. The research contributions are proposed as follows:

- **Explore and Identify Challenges and Opportunities of Cloud Forensics**

  A literature study is done to explore and identify challenges and opportunities for performing digital forensics investigation in the cloud computing environment which is known as ‘cloud forensic’. The identification of cloud forensic challenges and opportunities helped us to accomplish and complete this research.

---

This research work identifies some shortcomings and complicates in cybercrimes investigation in the cloud through studying technical issues and offering particular proposed solutions. Cloud computing poses numerous challenges that are specific to the forensic investigation. From these challenges which addressed in this research are as follows:

- Secure cloud storage services with identifying data remnants that are available after accessing, using and conducting Box cloud storage as a case study and the need for a method to the forensic investigation of this cloud storage services in forensically sound manner.
- Performing forensic analysis of large volume of log data to reconstruct a timeline of cybercrime events in the cloud environment.
- How to apply forensics investigation of cybercrimes in private cloud environment as well as provide a forensic methodology or a framework to assist the digital investigators in analyzing and examine virtual machines and its snapshots.
- Acquire and extract digital evidence from VMware clustered ESXi hypervisors.
- Offer cloud services with ensuring that Cloud Service Providers (CSPs) support the digital investigation process (i.e. building cloud architecture support forensics).
- The need for the design and the guidance for building a cloud forensic laboratory at the country level to investigate cybercrimes as well as provide a learning environment for law enforcement officers, digital investigators, researchers and students through using massive cloud capabilities.

- **Secure Cloud Storage System based on combining Checking Data Integrity and Digital Forensic**

A cloud forensic approach based on data integrity checking and digital forensic for assisting and helping digital investigators for performing digital forensics for cloud storage services such as Box cloud storage service. This approach introduced a secure cloud storage system

---


combining checking data integrity and digital forensics. This approach works as first-hand preventive check-up in the integrity of uploaded data in cloud storage, then, if there is change or modifications and tampering with the data that was already stored in cloud storage provider, the investigators will start performing the digital investigation process. This approach can help to save the investigators time. The system is able to provide notice while some kinds of change or manipulation of data, because of the notice, it is easy to identify the criminal of the data. So that, the user can start trace attackers who performed the criminal activity.

In conclusion, the purpose of this approach to fill gaps of previous work that not take in consideration automatic check of data integrity that can speed up taking action against attackers as well as provide a forensic analysis of data artifacts and remnants that remain in user machine when using Box cloud storage as a case study.

- **Log Data Analysis for Supporting Cloud Forensics**

A forensic analysis approach is presented using Apache Hadoop and Apache Spark to analysis large size of log data and extracts knowledge from them which can assist the digital investigators during investigating cloud-based crimes that occurred through a particular time. As log data can use to study vulnerabilities and weaknesses in systems such as networks to fix them. Similarly, it can help to accomplish timeline reconstruction for investigation purpose that helps to draw a conclusion about what happened during a crime and assist in identifying attackers who hacked the systems. The investigators use log data analysis to construct a traceback for past sequence of events and analyze criminal actions in detail. The proposed approach can use to obtain past sequence of events in cloud computing environment during a specific interval of time then visualized using existing visualization tools. In addition to, develop an automatic timeline application for analyzing and monitoring of cloud log data for a forensic analysis purpose.

The proposed approach can be built on top of cloud infrastructures to perform automated log analysis for each part of the cloud system. Building such type of application

---


within the cloud can reduce time and cost of performing cloud forensic process through determining occurred events in the cloud. Primary steps to perform timeline reconstruction in cloud environment from available log data can be as follows:

1. Process available access log data, which are recorded in cloud storage.
2. Perform pre-processing and cleaning operations on log data.
3. Detect and select specific attributes such as IP address and, date and times for timeline analysis purpose that helps to collect the evidence of criminal actions.
4. Apply specific statistics operations on these attributes such as count number of access to cloud server by a specific user per specific date and time during the criminal actions.
5. Visualize output of the previous step to conclude what has been happened during a specific period of time.
6. Generate a forensic report as the evidence that summary timeline events that support the trial to persuade the court against the suspect user.

- **Investigation of Cybercrimes in Private Cloud Environment**

A digital forensic approach is provided for investigation of cybercrimes in a private cloud environment. The proposed approach can help digital investigators and practitioners in acquisition and collection of digital evidence from the private cloud infrastructures especially virtual machine which is considered the core element of virtualized cloud systems. The proposed approach can help the digital investigators to determine attackers of virtual machines using live and dead forensics as well as collect and extract information from the client device, hypervisor, and hypervisor management system to facilitate the investigation process in an effective manner.

In addition to this, a forensic methodology is suggested for investigating and analyzing virtual machine and its snapshots for helping in the reconstruction of criminal activities which done using or against the virtual machine in virtualized environments such as cloud computing.

---


• **Acquisition and Extraction of Digital Evidence from Clustered VMware ESXi Hypervisors** ⁵

An efficient forensic acquisition application called Hypervisor Forensic Acquisition Application (HFAA) is proposed for acquisition and extraction digital evidence from clustered VMware ESXi servers. The proposed system is built on top of open source Virtual Infrastructure (VI) Java Application Programming Interface (API), which is built on top of the Web Services API. Using this API instead of Web Services can give much shorter, faster, and more readable code. This application utilizes this feature to provide full control to manage and handle vCenter/ESXi for performing incident response and digital investigation.

This application can provide an initial step and guide to digital investigators and researchers to develop new applications and tools for real-time forensic acquisition and analysis of virtualized environments like cloud environment. With some additional development and evaluation through adding new features to the proposed application, this can become a good tool for using in cloud forensic community.

• **Cloud Forensics Investigation Model** ⁶

A Cloud Forensics Investigation Model (CFIM) is proposed to investigate cybercrimes in the cloud environment. The proposed system is a smart system that is able to take a snapshot of the state of running virtual machine in virtual datacenter and send to Trusted Center Server (TCS) that monitor the status of the VMs as well as store snapshots of the virtual machines for sending them for Forensic Server (FS) up to require for performing forensics process. The proposed model support a concept of Forensic as a Service (FaaS) that provide various benefits of conducting digital forensics through using Forensic Server on the cloud side. This model can help tracking malicious users in the cloud environment as

---


well as determine weaknesses of the running virtual machines for future use and finally provide support and help in the digital investigation process in the cloud.

The proposed model can provide the following features to support cloud forensics:

- Prevent loss of volatile data through recording snapshots of running VMs in a persistent storage.
- Help to issue a search warrant in respect of cloud environment through determining the location of TCS servers.
- No need to resume a suspend VM before the acquisition, which may potentially change the evidence through use VM snapshots.
- No need for large bandwidth to download an image of VM instance because the proposed model used the FS on the cloud side.
- Provide proactive strategy by preserving regular snapshots of VMs that can significantly help incident response and handling.
- Provide the Forensic as a Service (FaaS) concept through utilization of huge and massive computing and storage resources of cloud computing for performing the investigation process in a timely fashion manner.

- **Cloud Forensic Laboratory for Investigation of Cybercrimes**

Here, we provide the idea of designing and establishing a Cloud Forensic Laboratory (CFL) to face the sophisticated yet advanced level of cybercrimes taking place currently comprises of various layers requiring levels of analysis consequently in classical IT and cloud systems. The proposed system will benefit from enormous storage and processing capabilities which are providing by cloud computing to perform acquisition, extracting, analysis, examination and reporting for the large size of digital evidence for both of crimes in IT systems and cloud. The proposed system will help and assist the digital investigators and practitioners to perform the digital investigation process in forensically sound and timely fashion. Design a cloud forensic laboratory to investigate cybercrimes as well as provide a learning environment for law enforcement officers, researchers, and students in universities.

---


7. Conclusions and Future Directions

This research presented a study on applying digital forensics techniques in cloud computing environment for the investigation of cybercrimes in forensically sound and timely fashion. The work conclusions are introduced in section 7.1 and the future research direction is presented in section 7.2.

7.1 Conclusions

This work concerned with developing proficient digital forensic techniques for investigation of cybercrimes in cloud computing environment in forensically sound and timely manner. It is introduced research contributions to the field of cloud forensics. They can be summarized as follows:

- A literature review is done to explore and identify challenges and opportunities for performing digital forensics investigation in the cloud computing environment. The identification of cloud forensic challenges and opportunities such as secure and forensic analysis of cloud storage services, log data analysis, design cloud computing model to support digital forensics, design cloud-based forensic laboratory which helped us to accomplish and complete this work.

- A cloud forensic approach based on data integrity checking for assisting and helping digital investigators is proposed for performing digital forensics for box cloud storage as a case study. The experiment results showed that there are data artifacts that remain in the user machine which uses Windows 7 about using box cloud storage such as IP address, and user account information like the username. The proposed approach can potentially useful tool for performing cybercrimes investigation related to cloud storage.

- A forensic approach is introduced for analysis large size of log files to extract knowledge which can assist digital investigators and examiners during the investigation of cloud-based crimes that occurred through a particular time. In this approach, we used Apache Hadoop and Apache Spark for analysis web log data. Apache Hadoop for analysis of log data is used while an Apache Spark is used to provide batch and real-time analysis of web server log data. In each approach, three different programs are implemented and tested on three different log files in size. Each program extracts the different type of information that can help digital investigators in reconstructing timeline related crimes that are occurred. The results show that Hadoop and Spark can be used as fast platforms for processing various large size of log files and extract beneficial information that can support digital investigators.
in analysis massive amount of cloud log data in a given frame time as well as reconstructed timeline related to incidents. Furthermore, the results can provision to reconstruct and generate a timeline related to historical past sequence events occurred during a crime as well as identify the malicious user’s IP address, date and time, with the number of access.

- A forensic approach is provided for investigation of cybercrimes in a private cloud environment. The proposed approach can help digital investigators and practitioners in acquisition and collection of digital evidence from the private cloud infrastructures especially virtual machine which is considered the essential element of virtualized cloud systems. In addition to can collect and extract information from the client device, hypervisor, and hypervisor management system to facilitate the investigation process in an effective manner. Also, presented a forensic methodology which can use for investigating and analyzing virtual machine and its snapshots for assisting in the reconstruction of criminal activities which done using a virtual machine.

- An efficient forensic acquisition application called Hypervisor Forensic Acquisition Application (HFAA) is proposed for the acquisition of clustered VMware ESXi servers. The design of proposed application can support to scale in a dynamic cloud computing environment where distributed clusters of hypervisors that share virtual machines and storage for providing better and on-demand services. This application aims to be utilized for extracting digital evidence from clustered VMware ESXi hypervisors for assisting the digital investigators in performing the digital investigation process. This application is an initial step for the digital investigators, practitioners, and researchers to develop new approaches and methods for evidence acquisition and extracting from clustered VMware ESXi hypervisors and hosted virtual machines in a forensically sound and timely fashion manner. With some supplementary development through adding new features to the proposed application, this can make it as a good tool to be used in cloud forensic area.

- A Cloud Forensics Investigation Model (CFIM) to investigate cybercrimes in the cloud environment is proposed. The proposed system is a smart system that is able to take a snapshot of the state of running virtual machine in virtual datacenter and send to Trusted Center Server (TCS) that monitor the status of the VMs as well as store snapshots of the virtual machines for sending them for Forensic Server (FS) up to require for performing forensics process. The proposed model supports a concept of Forensics as a Service (FaaS) that provide various benefits of conducting digital forensics through using Forensic Server on the cloud side. The implementation of the proposed model within cloud architecture can
increase the probability of tracking malicious users in the cloud environment, determine weaknesses in cloud services such as virtual machines for future use as well as support cloud forensics investigations.

- The idea of building a Cloud Forensic Laboratory (CFL) is introduced which is based on using cloud computing capabilities for investigation of cybercrimes that are a cloud or classical IT based. The proposed CFL can decrease the time that required to investigate cybercrimes in the cloud through using enormous capabilities of cloud computing. Finally, a forensic report can be generated about the committed crime for presenting it in a court of law as admissible proof. The proposed system can improve the investigation process by providing an environment for the digital investigators and experts to access various forensic tools and test environments remotely. Also, the proposed system can help the digital investigators and practitioners to perform the digital investigation process in forensically sound and timely fashion manner.

7.2 Open Problems and Future Works

The work carried out in this research leaves a number of open questions which are worth addressing in the future as follows:

- Develop and design procedures and tools to segregate forensic data between multiple tenants in the cloud.

- Collaboration between international law enforcement agencies in cloud forensics.

- Providing proactive measures can significantly facilitate cloud forensic investigations.

- Changes in Service-Level Agreement (SLA) to support and facilitate forensic readiness in the cloud.

- Design and implement new methods to recover the deleted data, identify the ownership of the deleted data, and use the deleted data for event reconstruction in the cloud.

- Extending and modifying current digital forensics tools to use them in the cloud computing environment.

- Develop real-time cloud data analysis system for support online attacks detection and tracking.

- Developing new versions to perform more sophisticated forensic acquisition process on clustered VMware ESXi hypervisors as well as fixed the limitations mentioned above by adding more useful features.
Implement the proposed cloud forensics model with OpenStack, which support different types of hypervisors such as Microsoft Hyper-V, Citrix XenServer, and Oracle VM. In addition, integrate an intrusion detection and prevention system to reduce digital forensics process cost and time.

Implement the proposed CFL as a real system for training and investigation of real crime cases.

8. Thesis Organization
The thesis of this work is organized as follows:

Chapter 1: Provides an overview of the research area, the motivation for working in this research work, challenges and shortcomings in the research area, research objectives and contributions, finally the thesis organization.

Chapter 2: Provides basic concepts of cloud computing, digital forensics, cybercrimes and cloud forensics in addition to literature review about challenges and opportunities of performing digital forensic techniques in cloud computing environment.

Chapter 3: Provides a cloud forensic approach based on data integrity checking for assisting and helping digital investigators for performing digital forensics for cloud storage services such as Box cloud storage service.

Chapter 4: Provides a forensic approach using Apache Hadoop and Apache Spark to analysis large size of log data and extract knowledge from them which can assist digital investigators during investigating cloud-based crimes that occurred through a particular time.

Chapter 5: Provides a forensic approach for investigation of cybercrimes in a private cloud environment using VMware virtualization technology. In addition to, introduce a forensic methodology for investigating and analyzing virtual machine and its snapshots for helping in the reconstruction of criminal activities which done using a virtual machine.

Chapter 6: Provides an efficient forensic application for extracting digital evidence from clustered VMware ESXi hypervisors called Hypervisor Forensic Acquisition Application (HFAA) for helping digital investigators and experts in performing the digital investigation process where the virtual data center is used in the cloud computing environment.
Chapter 7: Provides a Cloud Forensic Investigation Model (CFIM) to investigate of cybercrimes in the cloud. The proposed system is an intelligent system that able to take a snapshot periodically for the state of each Virtual Machine (VM) which runs in the cloud and send it automatically to Trusted Center Server (TCS) that works as a datastore for VMs. This TCS server is responsible for monitoring and recording the states of the VMs and send it to Forensic Server (FS) upon request from Cloud Service Provider (CSP) to deal with the FS.

Chapter 8: Provides the idea of designing and establishing a Cloud Forensic Laboratory (CFL) which is based on cloud computing platform for facing the sophisticated yet advanced level of cybercrimes taking place currently comprises of various layers requiring levels of analysis consequently in classical IT and cloud systems.

Chapter 9: Gives conclusions and directions for future work on the subject.
List of Publications and Submissions

**Book Chapters**


**International/National Conferences/Journals**


Bibliography


