Innovation Trends of Digital Forensics – Cybercrime

Dr. G Angeline Prasanna¹, Keerthana S. ², AjithKumar K. ³, Princy Joseph K ⁴

Associate Professor & Head¹³, PG Students²,³,⁴

Department of Computer Science, AJK College of Arts and Science, Coimbatore, Tamilnadu

Abstract
Digital forensic has progressed maneuvering easy computer crimes to research of elaborate circumstances that can become effect that is worldwide significant. The success is studied by this paper for the Digital forensic; its origins, its present point also its tips being future. This paper sets the scene with checking out literature in which past methods are electronic may be forensic closely by the assessment and study of current condition of art both in professional and forensics which can be educational. The paper finally concludes with offering future trends in this arena.

Keywords- Digital Forensic, Mobile Device Forensic, Forensic Framework, Network Forensic, String Analysis

I. INTRODUCTION

Modern society increasingly depends on communication networks, mobile appliances, Internet of Things (IoT) solutions, cyber-physical system (CPS) technologies, and cloud-based services. By taking advantage of the most cutting-edge information and communication technology (ICT) features, commercial activities, business transactions, and government services have grown, transforming the lifestyle of nearly all individuals. The tight coupling of the physical world with ICT technologies has led to indubitable benefits. At the same time, it accounts for the proliferation of new threats and cyber security issues such as identity theft, cyber bullying, data leakage exploiting social engineering or information hiding, malicious software turning IoT nodes into “zombies,” distributed denial of service (DDoS) orchestrated through botnets, and malware targeting specific appliances such as those affecting VoIP devices and smart vehicles. As a consequence, cyber attacks can have a significant socio-economic impact on both global enterprises and individuals. Therefore, cybercriminals should be promptly identified, and high-quality evidence of the attacks should be made available in the
courtroom. Shockingly, seeking after cybercrimes frequently requires complex examinations that can traverse universal borders and be exposed to various wards and legitimate frameworks. This issue, together with the immense volume and wealth of data, exceptionally heterogeneous ICT advances, and complex present day equipment/programming systems, raises new difficulties, particularly in the field of computerized crime scene investigation. Truth be told, it's entirely expected to need to follow an assailant through the Internet, or to gather intimations to illicit advanced exercises covered in enormous volumes of information. In this manner, law implementation office’s (LEAs) and security experts must refresh and frequently reexamine their way to deal with advanced legal sciences. For example, legal specialists must be completely arranged to accumulate powerful and unambiguous advanced confirmations, while LEAs ought to be prepared to "handle" a multidisciplinary process that traverses a few fields, including law, software engineering, fund, organizing, information mining, and criminal equity.

II. DIGITAL FORENSICS: PAST

The formal beginning of academic community research in the area of digital forensic investigation was in 2002 with an article called “Network Forensics Analysis” authored by Corey who studied Network Forensic Analysis Tool (NFAT) and highlighted its benefits in regard to traffic capture, traffic analysis and security issues. Although digital forensics might seem to be a novel discipline, its roots date back to 1970, when engineers recovered the only copy of a database that had been inadvertently deleted. From such a starting point, digital forensics rapidly evolved. Nowadays, it’s possible to address a variety of aspects concerning digital investigations, for instance, undeleting data or dumping network traffic to reconstruct attacks offline. The standard toolbox used for digital forensics covers all the different aspects of the cyber investigation procedure. However, as hinted at earlier, evolutions in the ICT domain reduce its efficiency and pose hazards. More specifically, modern forensic procedures and their main limitations can be summarized as follows.

Stored Data and File system Analysis

Digital forensics experts are expected to maintain the integrity of seized evidence, such as hard drives. Thus, safe practice consists of accessing the device in read-only mode and using forensic tools to create working copies, or forensic images. For instance, forensic images can be used to analyze files and installed applications, or to inspect unused space to spot remnants of deleted contents. Unfortunately, the increasing number of devices and the volume of available data make creating and analyzing such working copies highly time consuming. A recent approach triages a live system, which allows forensics examiners to acquire
evidences potentially hidden in volatile digital artifacts (for example, the contents of RAM or the clipboard, the list of open files, running processes, active network connections, or mapped drives). This aspect is also vital to avoid losing evidence as a consequence of a reboot, which could cause encryption of the file system or deletion of temporary data. In fact, an increasing number of OSs make it easy to encrypt the file system including the swap area (see, for example, popular Unix-based OSs such as Linux and mac OS). Although this contributes to the security and privacy of end users and enterprises, it also represents a hurdle for forensic investigations. Due to the integration and almost per-click setup in some environments, the percentage of encrypted file systems can be expected to increase, eventually becoming the default scenario in the long-term future.

Network Forensics

Typically, network forensics requires collecting and analyzing the traffic produced by a host, an intermediate node, or an entire portion of a network. Forensic analysts rely on traffic captures as well as on the available logs of network services and security appliances such as intrusion detection systems. To this aim, different granularities are possible. The most fine-grained techniques perform per-packet inspections, where the threat/attack is traced back by analyzing header fields or IPv4/IPv6 addresses. Unfortunately, collecting and storing each packet of large network infrastructures is economically infeasible and could saturate resources of routers or dedicated security tools such as firewalls. Moreover, the amount of data could be impossible to evaluate using commodity hardware, making the forensics analysis too expensive, especially if real-time requirements must be satisfied. A more coarse-grained approach exploits information collected by ad hoc tools such as firewalls, intrusion detection systems, or honeypots. Also in this case, the huge traffic volume, the heterogeneity of the used technologies (for example, proprietary communication protocols deployed in automation environments mixed with the TCP/IP suite if data has to be transmitted through the Internet), the availability of information-hiding-capable cyber threats, traffic encryption, or obfuscating mechanisms such as onion routing can render network forensics efforts highly challenging.

Reverse Engineering

Reverse engineering is performed by inspecting the binary of a malware sample, the recordings of network traffic, or other execution traces, such as log files generated by the guest OS. However, the effectiveness of such approaches is limited for a subset of modern threats deploying anti-forensics countermeasures such as code obfuscation, or multistage loading architectures, in which each stage of the threat is hidden and encrypted. The same holds for information hiding capable malware that covertly communicates with a remote command and control facility.

"Selected Areas in Communications, IEEE

![Graph showing growth in articles from 2002 to 2008]

Figure 1: Number of articles in digital forensics in the period of 2002 – 2008 (No articles matched our search criteria in 2003)

III. NOW

From the very first days of digital forensic investigation’s life up to now, there were vast improvements in digital forensics techniques ranging from recovering deleted evidences and searching the megabytes storage devices to deal with petabytes storage devices [46], cloud based investigations [47], mobile device examinations [48], wireless network investigations [49], and database forensics [50]. Generally, the current perspective of digital forensic investigation can be categorized into four main types namely Computer Forensic, Smart Device Forensic, Network Forensic and Database Forensic. Among mentioned categories, computer forensics has attracted the most attention of academicians and professionals. On the other hand, digital criminals and
intruders are trying to minimize footprints of their actions utilizing anti-forensic techniques. Some of the common approaches of anti-forensics are using cryptography, steganography, metadata tempering, program packing, generic data hiding, and even disk sanitizing, we started. In spite of all studies in the field of digital forensic investigation, we are yet to have a comprehensive reliable study which offers analysis of related scientific research trends in the field. To address the aforementioned issue, this survey with searching in computer science research journals indexed by four main scientific database namely Science Direct, Springer, IEEE and ACM for papers published in the period of Jan 2008- Mar 2013 that include “forensic” as an author keyword or in the paper abstract.

Future Paradigms

Many of the digital forensics tools tailored to discovering evidence are expected to reside on the suspect’s device, but offer limited features for investigating unknown and complex environments, including big data-like sources. Consequently, the majority of forensic software is unsuitable for identifying anomalies in an automatic or unattended way. One of the major challenges to be addressed in the near future, therefore, is the creation of tools and techniques to analyze the bulk of data and report possible digital clues to the examiner for further investigation. Alas, such tools and techniques’ engineering, including proper visualization features to help the forensic examiner, is a complex task, particularly because of the lack of unified standards and the nontrivial computational requirements.

Fortunately, digital investigation can leverage the features of cloud computing, for instance, to offload the most demanding operations of digital forensics procedures, such as log analysis, data indexing, and multimedia processing. From this perspective, one of the most interesting aspects of the cloud is the opportunity to exploit a new paradigm in which forensics is provided as a utility, à la forensics as a service (FaaS). An additional benefit of pursuing an FaaS paradigm is the possibility of concentrating the software in a single point, which makes updates and improvements easier. This can also hide complexity from end users, allowing professionals to concentrate on the investigation. Similarly, digital investigations can leverage the proliferation of software-defined networking techniques, which offer additional layers of abstraction useful for analyzing attacks or infections without the need for resource-consuming traffic analysis campaigns.

Emerging Cloud Computing or Cloud Forensic Challenges

Usage of cloud services such as Amazon Cloud Drive, Office 365, Google Drive and Dropbox are now commonplace amongst the majority of Internet users. From a digital forensics point of view, these services present a number of unique challenges,
as has been reported in the 2014 National Institute of Standards and Technology’s draft report [NIST, 2014]. Typically, data in the cloud is distributed over a number of distinct nodes unlike more traditional forensic scenarios where data is stored on a single machine. Due to the distributed nature of cloud services, data can potentially reside in multiple legal jurisdictions, leading to investigators relying on local laws and regulations regarding the collection of evidence. This can potentially increase the time, cost and difficulty associated with a forensic investigation. From a technical standpoint, the fact that a single file can be split into a number of data blocks that are then stored on different remote nodes adds another layer of complexity thereby making traditional digital forensic tools redundant.

**DFaaS**

Digital Forensics as a Service (DFaaS) is a modern extension of the traditional digital forensic process. Since 2010, the Netherlands Forensic Institute (NFI) have implemented a DFaaS solution in order to combat the volume of backlogged cases This DFaaS solution takes care of much of the storage, automation, investigator enquiry in the cases it manages. Describe the advantages of the current system including efficient resource management, enabling detectives to directly query the data, improving the turn around time between forming a hypothesis in an investigation its confirmation based on the evidence, and facilitating easier collaboration between detectives working on the same case through annotation and shared knowledge. While the aforementioned DFaaS system is a significant step in the right direction, many improvements to the current model could greatly expedite and improve upon the current process. This includes improving the functionality available to the case detectives, improving its current indexing capabilities and on-the-fly identification of incriminating evidence during the acquisition process.

**Applying Complementary Cutting Edge Research to Forensics**

Current investigation practice involves the analysis of data on standalone workstations. As such, the sophistication of the techniques that can be practically employed are limited. Much research has been conducted in a variety of areas that has theoretical relevance to digital forensics, but has been impractical to apply to date. A movement towards DFaaS and high-performance computing, as discussed above, offers advantages beyond merely expediting the techniques currently used in forensics investigations, which remain reliant on manual input. It also promises a situation where this complementary research may practically be brought to bear on digital forensic investigations. One such research area is that
of Information Retrieval (IR). Traditionally, IR is concerned with identifying documents within a corpus that help to satisfy a user’s “information need”. Traditionally, IR researchers have been faced with the trade-off between the competing goals of precision (retrieving only relevant documents) and recall (retrieving all the relevant documents), whereby improving on one of these metrics typically results in a reduction in the other. In IR for legal purposes, recall has long been acknowledged as being the more important metric, given that a single missing relevant document could have serious consequences for the prosecution of a criminal case, the enforcement of a contract, etc. However, focusing on recall frequently results in an investigator being required to manually sift through a large quantity of non-relevant documents. This is in contrast to web search, for example, where users typically do not require all of the relevant documents to be retrieved, of which there may possibly be millions. Instead, a web searcher wishes to avoid wasting time on non-relevant material.

IR for digital forensics is often seen as a typical example of legal information retrieval. Although this is certainly true at the point a case is being built for court, it could be argued that the level of recall required at the triage stage can be sacrificed somewhat for greater precision, in order to allow investigators make speedy decisions about whether a given device should be investigated fully. Thus there is the potential for configurable IR systems to be utilised in forensics investigations, whose focus will change depending on the stage of the investigation.

The primary advantage of applying IR techniques to digital investigations is that once the initial preprocessing stage has been completed, searches can be conducted extremely quickly. Furnas has shown that less than 20% of searchers choose the same keywords for topics they are interested in. This suggests that many queries must be run to achieve full recall, and also suggests that standard IR techniques such as query expansion and synonym matching could also be applied to increase recall.

However, increasing recall typically reduces precision by also retrieving non-relevant documents as false positives. There are a number of ways in which this problem can be alleviated. The use of the aforementioned data reduplication techniques would eliminate standard system files from consideration (Beebe and Diet- rich [2007] note that the word “kill” appears as a command in many system files). Additionally, common visualization approaches such as ranking and clustering are likely to help investigators in their manual search of retrieved documents.
Another consideration is that event timeline reconstruction is extremely important in a crim- inal investigation [Chabot et al., 2014]. When constructing a timeline from digital evidence some temporal data is readily available (e.g. chat logs, file modification times, email times- tamps, etc.), although it should be acknowledged that even this is not without its own challenges. Within the IR community, much research has been conducted into the extraction of tempo- ral information from unstructured text .This can be used to dramatically reduce the manual load on investigators in this area. However, increasing recall typically reduces precision by also retrieving non-relevant documents as false positives. There are a number of ways in which this problem can be alleviated. The use of the aforementioned data reduplication techniques would eliminate standard system files from consideration (Beebe and Diet- rich [2007] note that the word “kill” appears as a command in many system files). Additionally, common visualization approaches such as ranking and clustering are likely to help investigators in their manual search of retrieved documents.

Another consideration is that event timeline reconstruction is extremely important in a crim- inal investigation [Chabot et al., 2014]. When constructing a timeline from digital evidence some temporal data is readily available (e.g. chat logs, file modification times, email times- tamps, etc.), although it should be acknowledged that even this is not without its own challenges. Within the IR community, much research has been conducted into the extraction of tempo- ral information from unstructured text .This can be used to dramatically reduce the manual load on investigators in this area.

IV. CONCLUSION

In this paper a number of current challenges in the field of digital forensics are discussed. Each of these challenges in isolation can hamper the discovery of pertinent information for digital investigators and detectives involved in a multitude of different cases requiring digital forensic analysis. Combined, the negative effect of these challenges can be greatly amplified. These is- sues alongside limited expertise and huge work-loads has resulted in the digital evidence back- log increasing to the order of years for many law enforcement agencies worldwide. The predicted ballooning of case volume in the near future will serve to further compound the backlog problem – particularly as the volume of evidence from non-traditional sources, such as cloud-based and Internet- of-Things sources, is also likely to increase.

In terms of research directions, practices already in place in many
Computer Science sub-disciplines hold promise for addressing these challenges including those in distributed, parallel, GPU and FPGA processing, and information retrieval. More intelligent reduplicated evidence data storage and analysis techniques can help eliminate the duplicated processing and duplicated expert analysis of previously content. These research directions can be applied to the traditional digital forensics process to help combat the aforementioned backlog through more efficient allocation of precious digital forensic expert time through the improvement and expedition of the process itself.

REFERENCES


