Scientific Research Publishing

Journal of Information Security, 2023, 14, 454-463
https://www.scirp.org/journal/jis
ISSN Online: 2153-1242
ISSN Print: 2153-1234

Forensics: Collection of Sound Digital Evidence

Richard Kabanda, Harihar Thapa, Luis Rivera

College of Engineering, University of New Haven, West Haven, USA
Email: rkabanda2015@gmail.com

How to cite this paper: Kabanda, R., Thapa, H. and Rivera, L. (2023) Forensics: Collection of Sound Digital Evidence. *Journal of Information Security*, **14**, 454-463.
https://doi.org/10.4236/jis.2023.144025

Received: September 18, 2023
Accepted: October 22, 2023
Published: October 25, 2023

Abstract

This summary paper will discuss the concept of forensic evidence and evidence collection methods. Emphasis will be placed on the techniques used to collect forensically sound digital evidence for the purpose of introduction to digital forensics. This discussion will thereafter result in identifying and categorizing the different types of digital forensics evidence and a clear procedure for how to collect forensically sound digital evidence. This paper will further discuss the creation of awareness and promote the idea that competent practice of computer forensics collection is important for admissibility in court.

Keywords

Electronic Discovery Reference Model, National Institute of Standards and Technology (NIST), Message-Digest Algorithm, Secure Hash Algorithm, Forensically Sound Digital Evidence

1. Introduction

Computer technology has become part and parcel of everyone's life and is one of the fastest-growing areas. It should be noted that computer systems store valuable company and personal identifiable information (PII) whereas the computer network system provides an appropriate means of accessing or retrieving and processing services. As a result, they have become major targets for attackers, hence increasing computer crimes as further discussed in this paper.

To make it possible for relevant crime investigations to be carried out and concluded, Forensic science is utilized. However, this paper will stick to Digital Forensics which is a subset of forensic science. It is the application of science to identify, collect, examine, and analyze data while preserving its integrity and maintaining a chain of custody for the data [1].

Digital forensics is used to identify and preserve digital evidence in its most-purest form.

Digital forensics is the field of forensic science that is concerned with retrieving, storing, and analyzing electronic data that can be useful in criminal investigations [2]. This may include information from phones, pads, computers, hard drives, memory disks, SD cards, and other data storage devices. With digital forensics, the investigator can know who did what and when on a given system.

According to Dutelle [3] "Digital Evidence is the information and data of value to an investigation that is stored on, received, or transmitted by an electronic device".

For any successful and meaningful forensics investigation or recovery from system damage, the collection of sufficient, significant, and forensically sound data is vital. Forensically sound means that, from the acquisition of digital evidence and throughout the investigation, the evidence must remain in its original state [4]. Ignoring digital forensics recommended practices may lead to destroying the evidence. Once evidence is damaged, then the likelihood is higher that it will be denied in courts of law.

2. Categories of Digital Crime

Digital crime takes many different forms, and these may be conducted over different types of digital devices such as network logs, computer memory locations, emails, websites, peripheral devices, printers, and so on. The major crimes are described below. There are 3 main categories of digital crimes, and these include crimes against people, crimes against property, and crimes against public [5].

Crimes against people are mostly personal. These are crimes that happen over various digital technologies and are directed at individuals using the internet and digital platform. Some of which include identify theft, hate crime using social media, defamation through emails or messaging, blackmailing, child exploitation, cyber stalking using fake identity among others.

Crimes against property are directed towards the property of individuals or groups. Criminals are looking for intellectual property, patents, money and even theft of property. These crimes do not care about any permission of authors or owners. Some of the noted ones are piracy of digital media such as movies, songs, software, and programs.

Some cybercrimes are committed against the public interest. Cybercrimes against the public include trafficking from small to large quantities, money laundering, crime against the government and terrorism that targets government and public.

All these crimes are dependent on digital technology such as mobile phones, computers and small electronic devices including IoTs that keep logs which can be analyzed if required. Because it's these devices and networks that are used to execute the crimes.

Examples of appropriate data sources for cybercrimes include log files, storage drives, firewall logs, and disk drives.

3. Types of Digital Forensics

There are different types of Digital Forensics, and these are discussed in detail as follows.

Email forensics is one type of digital forensics that consists of the examination and detailed analysis of emails and their content to determine their legitimacy in a forensically acceptable manner.

Mobile forensics involves examining and analyzing mobile devices where contacts, call logs, SMS, Videos, and images may be investigated.

Disk forensics involves the extraction of data from media stores which involves searching for files that are active, edited, or deleted.

Network Forensics deals with the examination and analysis of network traffic to collect evidence.

Database Forensics involves studying and examining the database and its metadata for possible evidence.

Wireless Forensics is when wireless network traffic is collected and analyzed, and the payload is studied for any malicious code is called malware forensics.

Memory forensics deals with the collection of data from the memory of systems. These may include registry and cached memory.

4. The Five (5) Rules for Sound Digital Evidence

Digital evidence may be used to prove the innocence of the suspect and in this case, it is categorized as exculpatory evidence whereas if it yields conviction in criminal cases, it is called inculpatory evidence [4].

There are 5 general rules that need to be followed for digital evidence to be admissible and these include:

1) Digital evidence must be admissible: Collection and preservation of digital evidence in a recommended way are vital to make it admissible by the Jury or elsewhere.

2) Digital evidence is authentic. It must correlate with the incident in a relevant way to prove something.

3) Digital evidence should be clear and complete. It should be well-aligned and connected to the whole story. Incomplete evidence should not be submitted as it might drive the judgment in a different direction.

4) It should be reliable. The tools and the methodology used for digital evidence should not cause any doubt about its accuracy and authenticity.

5) Digital evidence must be complete, reliable, easy to comprehend, believable, and acceptable [6]. A forensic examiner must be able to explain the methodology used and how integrity was preserved.

5. The Forensic Process

It is a scientific method that uses tools and analysis approach that pertains to the crime scene. In digital forensics, keeping the evidence secure and maintaining its integrity must be maintained for it to be admissible. The increase in the number

of people using Internet of Things (IoT) devices, smart phones, small scale devices and other networked devices has led to the establishment of incidences of crime which has necessitated forensic investigations in these areas as well. Having skilled forensic investigators with the availability of forensics tools that not only covers the digital computer and laptops with general operating systems, but also the ability to perform small scale forensics has made it possible to collect evidence from those devices.

Digital forensic is an application of forensic science applied to digital devices. It includes the study and analysis of four phases which are followed sequentially. They are Identification, Collection, Analysis and Presentation

5.1. Identification

This is the first phase and is where a digital forensic investigator is presented with a task to determine the progress of the action. It could be the electronics devices, IoTs, smart phones, computers, storage devices, logs from network devices that could potentially have been used during the crime. It is at this stage that the investigator establishes and understands the scope of the investigation. It is during the identification that the quantity of the data involved is established.

Identification of the digital device type, make model, service provider, and location in terms of storage of data evidence are carried out by the investigator. NIST SP 800-86 lists the most common sources of data as desktop computers, servers, network storage devices, and laptops [7]. At this stage, the tools needed to complete the examination are identified.

A decision on what steps to be taken to protect and preserve the integrity of digital data should be made. For example, turning on a computer in a crime scene leads to some of the logs being deleted, files being deleted. Careful consideration should be given when identifying digital devices for a crime scene.

It is at the identification stage that a chain of custody is also established by creating a record to maintain the integrity of the evidence [6].

5.2. Collection

This phase involves collecting the data in a forensically sound manner during the investigation. This is one of the most important steps in digital forensics. To collect evidence in a crime scene, investigators must have the knowledge to collect it in a forensically sound manner so that it must be submissible in the court with no modification to any of the digital evidence. If any modification is noted, its validation is questioned. So, to perform sound collection the steps are:

1) Identify the source media to collect. Know how to collect the type of data stored on it, and how to access it.

2) Determine the parameters under which proper data acquisition and imaging should be done. For digital data, trusted imaging applications can be used with the forensic machine in creating a copy of the evidence for subsequent analysis to prevent tampering with the original.

3) Digital forensic evidence should be preserved. The imaging process should not modify the data at any one time during the chain of custody and always have more than one copy to work with during the analysis phase.

4) Digital evidence can be authenticated by proving that the collected evidence is identical to the original data. One way to achieve this reliably is through strong cryptographic hashes such as SHA-5 [8].

5.3. Analysis

After collecting the digital data in a forensically acceptable manner, the analysis phase begins. In analyzing the available data, a methodical approach is used so that analysts can draw appropriate conclusions based on the available data. The evidence is reviewed, and all the information is related to a specific case to determine and understand criminal activities. Analysts should examine copies of files, not the original files [7]. This phase is crucial for any investigation as the success or failure of the case will depend on its outcome.

During this phase, the analyst is expected to tie together the different pieces of evidence gathered in the previous phases and reconstruct in a sound, factual, and complete way the criminal actions being investigated.

5.4. Production and Presentation

After completing the investigation, you need to organize the evidence uncovered, and then organize and present your findings in a clear, concise, simple, and understandable manner. This will be performed through reports, which may be addressed to your manager, lawyers, or judges. The presentation of the findings should be done convincingly and factually while remaining accessible and understandable by non-technical people.

From the discussion above, we can conclude that the most crucial task in any digital investigation process is evidence collection. The success or failure of any investigation is dependent on the nature and soundness of the evidence collected.

It is required that an original copy of the evidence from the suspect system be saved as much as possible from being modified.

All satisfactory steps must be taken and ensure that the digital data being analyzed is the same as the original copy acquired. This is because digital data is prone to modification. It is recommended to always keep a backup of the original data and never analyze the original copy of the data.

To ensure the integrity of the data, hash values for all data must be computed and checked regularly. There are different tools used to accomplish this and some of them may include the Secure Hash Algorithm (SHA-1, SHA-2, SHA-3). or MD5 (Message-Digest algorithm).

Conklin emphasizes the following questions during evidence collection to be put in mind for sound evidence collection.

- Who collected the evidence?

- How was it collected?
- Where was the evidence collected?
- Who has had possession of the evidence?
- How was it protected and stored?
- When was it removed from storage? Why? Who took possession? [6]

Once the evidence is collected, it must be properly controlled to prevent tampering. They go ahead to talk about the chain of custody accounts for all people who handled or accessed the evidence, when and where it was stored, and who stored it [6].

6. Data Acquisition Approaches

The investigator must identify the best possible method to acquire the evidence without altering the original data, record of all activities associated with the acquisition and handling of the original data and copies must be maintained, must not perform any activities that are beyond the capability of the investigator, must consider all the safety, legal rights and organizational policies and procedures [9]. Data acquisition of digital data per device can be done in live mode or in dead mode.

In a live acquisition, digital data is copied from a running system under investigation to a forensic system using write blocker and software such as FTK imager for copying operating systems into the forensics devices along with the hash verifications.

With the dead acquisition, the digital data under investigation is copied to the forensic system which may also be called a trusted server. The disk drive under investigation/suspect is moved and connected to the forensic system in a trusted environment and booted from a trusted bootable device which can be a USB drive, CD-ROM, etc.

Whereas dead acquisitions are preferred over live, there are situations whereby a live acquisition is the only option in some cases where the critical services cannot be shut down because the shutdown would alert the suspect his acts have been detected or would create a risk of losing data when the power supply is cut off to the system.

7. Categories of Digital Data

During digital forensics, digital data is categorized based on the order of their volatility. And investigator must approach data acquisition starting with most volatile data to least volatile data and finally nonvolatile data.

Any data that is lost when the system is shut down is called volatile data. For example, CPU registers, cache, RAM, ARP cache, process tables, etc. are some of the volatile data, and should be collected first for a live system.

Nonvolatile data is data that persists when the power is off, for example, the hard disk. Some of the nonvolatile data are data on hard disk, remote server logged data, backups. Successful data acquisition must start with the most vola-

tile data and lastly done for data that has less or no volatility decreases.

8. Data Acquisition Techniques

Three main data acquisition techniques can be used to copy data from the suspect system to a trusted one:

1) The network can be used to copy data from the suspect computer system to a designated server. This may be used for both live and dead acquisitions. Network communication tools such as Netcat may be used to perform the exercise. Alternatively, network drive mounting may be performed and the copy data to the server. Mounting storage devices like network shares, CD drives, and hard drives make the files accessible through the computer's file system.

2) The suspect storage device, for example, the hard drive may be removed from the suspect system and connected to the trusted server for the data to be accessed.

3) A new hard disk may be connected or installed in the suspect system. The suspect system may boot from a trusted USB or CD-ROM drive and finally imaging the suspect disk to the new disk.

During the collection of evidence from a device, the target system may include SATA, SCSI, or IDE storage device drives. The target may also have different peripheral hardware attached to it.

The good practice is that the system under investigation should be shut down and boot it into a trusted computer system that is used to investigate. Alternatively, the hard drive can be removed and attached to the system where the forensic investigation will take place.

However, in some cases, it is recommended that the machine under investigation is kept powered on to collect all information needed from the investigation from the active memory because shutting down the system may render all data lost from memory locations.

Below are the recommended seven steps to be followed while carrying out a dead acquisition:

Step 1: Here the suspect system is shut down to allow the investigator to note and record that they have established a timestamp upon which no other modifications will occur in the system [8].

Step 2: The next step is to remove the drive from the suspect system to ensure that chain of custody form is created and filled out. The chain of custody accounts for all people who handled the evidence. When, who, where the evidence was obtained, where it was stored, and who had control or possession of the evidence for the period since when it was obtained [6].

Step 3: Check for other media and remove any other remaining storage media, for example floppy drive and zip drive. A chain of custody should be filled out for each device [8].

Step 4: Record Unified Extensible Firmware Interface (UEFI) or Basic Input Output System (BIOS) information.

At this stage, the system can now be safely booted up to check the BIOS or the UEFI information. The UEFI or BIOS should be reported in the chain of custody form. The BIOS or UEFI information collected consists of the system time and date. The BIOS or system time is important because it can differ from the actual time and time zone set for the geographical area in which you are located [8].

Step 5: involves the imaging of the driver. This is the riskiest part of the evidence-collection process. Care must be taken to avoid writing to the original media, every time it is accessed. Before using a drive to store an image, you must always use some wiping software to clean the drive from any previous evidence [8].

Step 6: After creating images of the suspect media, you need to record the cryptographic hashes produced by the imaging programs. This can be message Digest 5 (MD5), SHA256, etc. The cryptographic hash ensures the integrity of the data [8].

Step 7: lastly bag and tag the evidence and clearly label the drive to which the forensic image has been written and store it in a safe place, with no access to unauthorized personnel [8].

9. Collection and Investigations of Remote Evidence

In some cases, the suspect machine may be a production machine, or it may be in a hostile environment, both of which make it difficult to power it off. In those cases, the data collection must be carried out remotely. In most cases, it is recommended to investigate the machine before carrying out the remote collection. This will allow for checking the presence of suspect artifacts on the remote machine before initiating the data collection [8].

Numerous tools are available for the investigator to carry out forensic analysis on a remote machine without any physical access. One should be able to carry out key forensics analysis tasks, including imaging, in-depth examination of the file signatures, file hashing and copying files to the workstation used in a forensics investigation, and generating reports of suspicious information.

In situations where the case and target environment are very sensitive, for example in a foreign location, the ability to keep the investigation concealed, without the subject knowing that he/she is being investigated, is critical. Failing to do so can jeopardize the evidence and have unintended consequences.

The following actions can help keep an investigation undercover:

Minimize the number of simultaneous operations to limit the system resource usage.

Name the remote investigative program a system-related name for example lmhosts.exe.

Firewalls should be configured to allow inbound connections from the investigator's machine.

Ensure that the program used to carry out the remote investigation does not leave any traces of events in the event logs.

Look for only the data that relates to the investigation.

The investigation should be done when the suspect expects a lot of regular activities on the hard drive antivirus scans.

One of the critical challenges with remote network collection is that the machine must be powered on and running, and the data on its hard drive is, for the most part, changing constantly [8]. Consequently, it will be difficult to rely on the MD5 hash to authenticate the acquired data. In that case, data validity will depend on how sound the process will be. A well-specified procedure must be followed.

In expectation of malicious activities, companies must be prepared forensically. Companies should create and implement adequate and standard data collection mechanisms with policies in place.

10. Challenges

Forensic evidence collection has faced many challenges over time. For example, collecting useful evidence from massive data sets in terabytes makes it difficult to search for exactly what is needed.

Law enforcement and court noted a lack of extensive knowledge and understanding by prosecutors with elements of digital evidence [10].

Recent technological upgrades have impacted forensic solutions because the tools used may require either an upgrade or a change which makes it expensive for forensic examiners.

11. Conclusion

Digital forensics can be challenging and difficult to start out. This paper provides a comprehensive guideline that covers all phases of collecting forensically sound digital evidence with different acquisition approaches and techniques for live data, dead data, and remote data collection. Challenges are discussed briefly.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Gerard Johansen (2020) Digital Forensics and Incident Response: Incident Response Techniques and Procedures to Respond to Modern Cyber Threats. 2nd Edition, Packt Publishing, Birmingham, USA.

[2] Salfati, E. and Pease, M. (2022) NIST Digital Forensics and Incident Response (DFIR) Framework for Operational Technology. https://doi.org/10.6028/NIST.IR.8428

[3] Dutelle, A.W. (2020) Intro to Crime Scene Investigation. 4th Edition, World Headquarters, Burlington, MA, USA.

[4] Dr. Hayes, D.R. (2021) A Practical Guide to Digital Forensics Investigations. 2nd Edition, Pearson, Hoboken, USA.

[5] Kanellis, P., Kiountouzis, E., Kolokotronis, N. and Martakos, D. (2006) Digital Crime and Forensic Science in Cyberspace. https://doi.org/10.4018/978-1-59140-872-7

[6] Conklin, A.W., White, G., Cothren, C. and Davis, R.L. (2021) Principles of Computer Security: CompTIA Security + and beyond. 6th Edition, McGraw Hill, New York, Chicago, San Francisco, USA.

[7] Grance, T., Chevalier, S., Scarfone, K.A. and Dang, H. (2006) Guide to Integrating Forensic Techniques into Incident Response. https://www.nist.gov/publications/guide-integrating-forensic-techniques-incident-response

[8] Phillips, A., Cowen, D. and Davis. C. (2009) Hacking Exposed Computer Forensics. 2nd Edition, McGraw Hill, New York, Chicago, San Francisco, USA.

[9] Adams, R., Hobbs, V. and Mann, G. (2013) The Advanced Data Acquisition Model (ADAM): A Process Model for Digital Forensic Practice. *Journal of Digital Forensics, Security and Law*, **8**, Article 2. https://doi.org/10.15394/jdfsl.2013.1154

[10] Sean, E.G., Robert, C.D. and Brian, A.J. (2015) Digital Evidence and the U.S. Criminal Justice System—Identifying Technology and Other Needs to More Effectively Acquire and Utilize Digital Evidence. https://www.ojp.gov/ncjrs/virtual-library/abstracts/digital-evidence-and-us-criminal-justice-system-identifying