Overview of Digital Forensics and Anti-Forensics Techniques

Hussein Majed
Department of Computer Studies, Arab Open University
Beirut, Lebanon

Hassan N. Noura
Department of Computer Studies, Arab Open University
Beirut, Lebanon

Ali Chehab
Electrical and Computer Engineering, American University of Beirut
Beirut, Lebanon

***Abstract*—Digital forensics is very essential in any investigation where data is involved after a security breach. Data contents might be personal, business-related, or strictly confidential. The aim of digital forensics is to legally acquire and analyze the examined data, while Anti-forensics techniques aim to hide, manipulate, and even wipe the data, or to target the credibility of the acquired evidence. This paper presents the current anti-forensics techniques, the methods applied, and the available countermeasures.**

***Index Terms*—Digital forensics, anti-forensics, forensics techniques**

I. Introduction

We are witnessing a digital era, where almost everything relies on digital data. The presence of technology in various domain is reflected by the presence of digital data within such domains. Individuals, businesses, and governments are relying on digital data in every aspect, and the growth in this domain was augmented by the emergence of digital forensics due to many reasons such as cybercrimes, terrorism, and ordinary crimes where digital data might reveal some useful information [1]-[2]. On the other hand, the existence of digital forensics tools also triggered an alarm for hackers and threat actors, as well as the parties who are concerned about privacy to develop anti-forensics tools preserve the ability of Forensics tools (FT) useful in retrieving useful and relevant information. The computer forensics tools (CFT) and mobile forensics tools (MFT) are used to help the forensic examiner in collecting information from a device, make a true copy of that information to be analyzed and to extract reliable evidence that are legally accepted in courts. The information can be collected from 2 sources; it depends if the device is on and if the examiner has access to the device to access the volatile memory, and the second source is the non-volatile memory such as HDD, SSD, and flash. The evidence must be accurate and some measures such as hashing must be put in place to ensure that the source or evidence has not been subject to alteration.

The anti-forensics tools (AFT) and techniques are used to target the availability, usefulness of digital evidence by performing some actions such as altering, disrupting, and destroying the scientific validity of evidences [3], [4]. AFT takes several forms depending on the usage and purpose, artifact wiping, data hiding, trail obfuscation, and Attacks against FT.

This paper we describe the various types of AFT in the description section. The analysis section covers the variety of tools used by attackers to conduct different anti-forensic attacks. The countermeasures and limitations section explains the precautions that should be taken into consideration to minimize the effect of anti-forensics tools as well as the methods of detecting the usage of AFT on a system or network. the main goal of the study is to try and find suitable solutions against AFT to ensure data availability and integrity with minimum overhead on a system.

II. Description

There are many types and techniques used in anti-forensics; an intensive survey on the available tools were conducted to demonstrate the categories and subcategories [5], and the study showed that there are hundreds of AFT available.

AFT is divided into multiple categories according to the purpose of usage and type of attacks performed.

A. Artifact wiping

When deleting a file or folder from a storage device, the actual data remains on the storage device until it is overwritten by new data. the artifact wiping is considered an AF technique designed to completely erase and destroy data. Artifact wiping or secure wiping can be applied on files, entire disk, or partition. Available tools that serve the purpose are Eraser, External Examiner, Free Wipe Wizard, File Shredder, and Registry cleaner (see Figure 1).

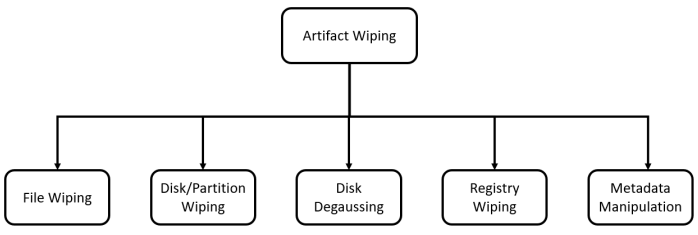

Figure 1. Artifact wiping

- File Wiping: used to destroy data associated with a specific file by filling randomly generated data on the

978-1-7281-6939-2/20/$31.00 ©2020 IEEE

occupied space on the storage device by the specified file.

- Disk/Partition Wiping: used to destroy data residing on all the sectors of a specified disk or partition.
- Disk Degaussing: is a process of exposing an electronic storage device to a magnetic field to neutralize and entirely erase any previously stored data.
- Registry Wiping: Registry is used by Microsoft Windows operating systems to store settings and data for the operating system and applications. The forensic examiners usually find useful artifact inside the registry during investigations. Registry wipers try to destroy such information permanently.
- Metadata Manipulation: Operating systems and software tools usually create metadata of each file created and saved on the storage device. The metadata is considered as information about information. They vary according to the file type but typically contain creation date, modified date, last accessed, last modified, and information about the author. Because this type of data is very sensitive and essential for drawing a timeline of events, manipulating and changing metadata is used to divert the forensic examiner into the wrong places.

B. Data Hiding

It is used to target the existence of data on storage, which makes the analysis and examination of digital evidence, by forensics examiners, difficult or impossible to conduct.
The taxonomy of data hiding includes steganography [6], data contraception, file system manipulation, hard disk manipulation, and network-based data hiding (see Figure 2).

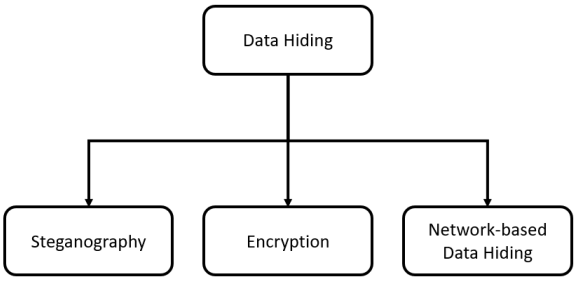

Figure 2. Data Hiding

- Steganography: is the art of hiding a covert message within a regular message, but not the fact that two parties are communicating with each other [7]. The secret message might be embedded in multiple types of files such as pdf, video, audio, images, and text.
- Encryption: Data Encryption is used to prevent access to stored data. Encryption can be applied to an individual file, database, email, or entire disk using several encryption algorithms.
- Network-based Data Hiding: It is the same concept of data encryption but on the fly instead of at rest. Network-based data hiding is used by multiple platforms to ensure privacy such as end-to-end encryption [8]. Normal Internet users may use VPN services to bypass geographical limitations if they have concerns about privacy. Attackers are using lately network-based data hiding or encryption to encrypt the communication traffic with bots or zombies as well as transferring data from attacked machines securely.

C. Trail obfuscation

This technique is more popularly known as "counterfeiting". Therefore trail obfuscation or evidence counterfeiting techniques are practiced to confuse and disorientate the investigation (see Figure 3). This could be achieved in a number of ways [9] .

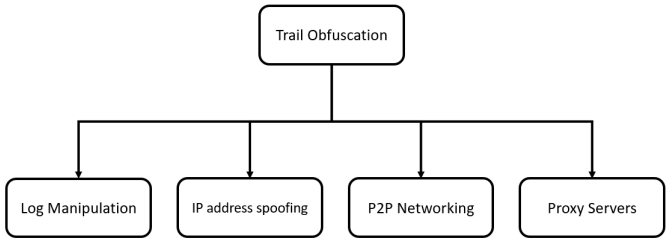

Figure 3. Trail obfuscation

- Log Manipulation: logs are usually saved on the attacked machine and external Syslog servers. if the attacker gains access to the machine, it will give him the ability to manipulate the log to divert the investigation and provide misleading information [10].
- IP address spoofing: attackers usually try to hide their identity or fake the source of the attack. IP spoofing is a process of spoofing the source of the packets to avoid being detected or to reflect and amplify attacks. It is usually used when performing a distributed denial-of-service attack.
- P2P Networking: provides a method of sharing data that does not require a central host or server. Instead, the data is shared among nodes that are connected to the network very quickly, each connected device with a small chunk of requested file can offer it to those who request it. The transmitted data might be illegal or breaking the law of copyright.
- Proxy Servers: by nature, a proxy server has been implemented to control and restrict the access to a predefined URL, some ISP providers use a transparent proxy server to provide the frequently requested URLs from cache engines. Attackers generally use a proxy server to hide their identity and the web requests on the destination web-server where the IP of the proxy server is shown.

D. Attacks against forensics tools & processes

The attacks against forensics tools are used to mislead or to divert the examiner of reaching the correct information. It can be conducted by many tools, some of which are used

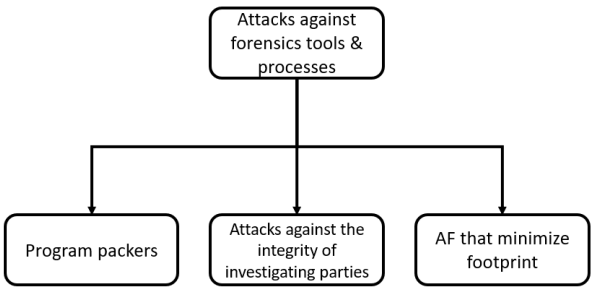

Figure 4. Attacks against forensics tools & processes

to minimize the footprint or to make reverse engineering impossible. In some cases, the criminals might go even further by trying to attack the credibility of the investigating parties (see Figure 4).

- Program packers are used against FT prevent the forensic examiner from applying reverse engineering on AFT. Another usage of Program packers is to avoid detection of AFT during forensic examination or scanning.
- Attacks against the integrity of investigating parties criminals use this method to attack the integrity of the investigating parties by performing malicious acts.
- AF that minimize footprint Installing AFT on a system might leave some traces and hence, attackers may adopt other approaches to minimize the footprint or data that was left behind [11] by using one of the following techniques:
  - Memory injection and syscall proxying: Buffer overflow and heap overflow allow an attacker to exploit a weakness in programs to inject a malicious code and to run it on the targeted system. It is considered as an entry point to a system and then, the attacker may upload tools and maintain access to the system.
  - The usage of live CD/DVD, bootable USBs, and virtual machines: Many operating systems can run using Live CD/DVD and Bootable USB. The attackers use these techniques to boot their own operating systems in read-only mode with special embedded tools to perform illegal activities on a system without leaving a trace. Virtual machines are used to run a guest operating system on a host machine; attackers might use this technology to perform attacks and securely erase the files associated with VM after performing the attack.
  - The usage of anonymous identities and storage.

III. Analysis

A. Artifact wiping

Multiple standards and techniques have been used in artifact wiping, and they are listed in various research studies [12]:

1) DoD 5220.11 M It is developed and supported by the US National Industrial Security Program. It exists in 2 main forms: 3 or 7 phases. The algorithm is based on the following steps: writing 0 and verify; writing 1 and verify; writing a random character and then verify.
2) NCSC-TG-025 It is developed and supported by the US National security Agency. it is based on the wiping techniques DoD 5220.22 M with multiple overwriting processes.
3) AFSSI-5020 It is similar to DoD 5220.11 M technique with just a difference in the verification process.
4) AR380-19 It is developed by the US Army; the concept is similar to the previous methods. The algorithm is based on the following steps: writing random character; writing a specified character; writing the complement of the specified character and then verify.
5) NAVSO P-5239-26 It is developed and supported by the US Navy. The algorithm is based on the following steps: writing a specific character; writing the complement of the specific character; writing a random character and then verify.
6) Gutmann It is developed by Peter Gutmann and Colin Plumb. It involves writing a series of 35 patterns over the region to be erased.
7) Schneir It is developed by Bruce Schneier; the process of wiping the data consists of 7 phases. The Schneier method overwrites the data on a storage device with a one, and then a zero, and finally with several passes of random characters.
8) Disk Degausser It is a physical device that contains a controlled magnetic field, measured in units of gauss. It erases data on Hard disk by generating a magnetic field so powerful that it permanently removes the magnetic properties from the iron oxide or chromium dioxide coatings. This type of AFT cannot be used on solid-state drives and flash media because they do not use magnetic technology to store data.
9) Registry Wiping Registry cleaner software is designed to locate specific hives within the registry which might leave clues about the attacker. Tools like Stellar BitRaser, RegSupreme, and RegSeeker are known to delete registry keys in Windows.
10) Metadata Manipulation The attacker can wipe the contents of an entire storage media, but this may lead to attracting attention. By modifying or overwriting the metadata of accessed files, the attacker ensures that the timeline could not be reliable. AFT like Timestomp and Defiler's Toolkit can be used to manipulate the metadata on Windows and Linux platforms. Attackers might also use the mount as a read-only technique to avoid creating metadata of their access depending on the situation and flexibility they have on a system[11].

B. Data Hiding

The purpose is to hide the existence of secret information (evidence) in a file or network traffic.

1) Steganography Techniques:

- The least significant bit substitution or overwriting is the most common steganography method used to hide data

in audio and image files by overwriting legitimate RGB color encodings or palette pointers in GIF and BMP files, coefficients in JPEG files, and pulse code modulation levels in audio files.
- Adding secret information to some specific parts of a file header.
- Adding secret information after the EOF Tag of an image.
- Spread-spectrum takes advantage of the fact that little distortions to image and sound files are hardly detectable.

2) Encryption: Many operating systems support entire disk encryption; Microsoft windows has Bit-locker which provides hard drive encryption, Device Encryption, and Used Disk Space Only encryption. Although users may use VeraCrypt and Truecrypt to encrypt specific data or to create a virtual encrypted volume, which can be mounted whenever needed. Some applications such as Microsoft office provide encryption of created files; they implement password protection for opening a file and for editing a file, database servers like Microsoft SQL server provide database encryption.

Macintosh also provides FileVault to encrypt the entire hard drive; Linux comes with different flavors but the main encryption software is DM-Crypt.

Attackers are using recently encryption in Ransomware attacks to encrypt the user data and ask for ransom.

3) Network-based Data Hiding: There are many VPN providers on the Internet that provide free VPN services for computers and mobiles in order to encrypt traffic on the fly; OPENVPN, for example, is a famous VPN provider that uses robust encryption.

Recently, botnets started encrypting traffic using either symmetric or asymmetric encryption between bots and command & control to avoid detection.

C. Trail obfuscation

Attackers may use a text editor to manipulate logs saved locally on compromised machines.

Crafting IP packets to spoof a source IP can be done by several well-known tools such as hping3, scapy, Libcrafter, and Yersinia.

Connecting to a P2P network is done using bit-torrent, ytorrent, limewire, and emule. these programs provide users with the ability to conenct to P2P networks to share contents.

Open Proxy servers are available online; Google is a good source to search for such servers. They are found in multiple types, some of them are http only while others support https. Also note that some old-style proxy servers that use SOCKS are available online.

D. Attacks against Forensics Tools & Processes

Program packers are designed to detect if the software is attached to a debugger that refuses to run; they can also encrypt the code to avoid being reversed. some tools like shiva, PECompact, and Burneye are used to target reverse engineering or detection scanning. Some packers require a password to decrypt and run a packed software.

There are many operating systems specially designed for penetration tests and might be used for hacking. Kali Linux, backtrack, Parrot, BlackArch, and Bugtrack are examples of operating systems that serve this purpose. Some of them come also with Forensic Mode which allows the attacker to boot up the system from a live cd/dvd/usb, and the storage devices of system will be mounted in read-only mode to leave no trace behind. These operating systems can be run on virtual machines within the host systems using hyper-v, Vmware workstation, Vmware ESXi, and virtual box.

IV. Limitations & Countermeasures

A. Artifact wiping

Artifact wiping has a big impact on digital forensics examination; the only known countermeasure to this type of attacks is frequent and incremental backup to either network-attached storage or to the cloud. The higher the frequency of backup the lesser the impact of wiping. However, these solutions have some limitations such as storage and resource consumption if applied on a NAS, in addition to being expensive and bandwidth-consuming when applied on cloud, depending on the amount of data stored.

B. Data Hiding

Data Hiding is a critical topic in the forensic world. Forensic examiners use many software tools to detect hidden data on the inspected devices. Below is a list of software used for each type of data hiding:

- Steganography: using Gargoyle, Stegdetect, and Stego Watch. All these software tools require the forensic examiners to know the type of steganography to detect the hidden information.
- Encryption: if the forensic examiners have access to the system while it is on, memory dumping is the best option to try and find the plaintext password in memory, otherwise, some sort of password attacks such as brute-force or dictionary attacks should be conducted to decrypt the contents. These types of attacks are resource and time consuming depending on the type of encryption and password complexity. They may also exploit some weakness in the old encryption algorithm to decrypt the data, but if the data is encrypted using modern ciphers, it will be difficult or impossible to retrieve the contents without the key.
- Network-based Data Hiding: Companies may restrict the usage of VPN on the firewall. This prevents employees from using VPN to leak sensitive data. Although the case is different at the level of Internet service providers due to privacy limitations, ISPs cannot stop service because it is encrypted and its main task is to provide connectivity. the national Internet gateway in each country may decrypt asymmetric traffic for security reasons such as the case in Turkey, since the country has a root certificate for this purpose. However, this solution is not effective against symmetric or end-to-end encryption.

C. Trail obfuscation

Keeping track of actions is usually done by securely saving logs on Syslog machines that are not exposed to compromising. Endpoint software is used to transmit the events, logs, Netflow generated by the routers, from machines to safe machines where the log is correlated and analyzed at a later stage. If the logs are saved on another machine, this allows the forensic examiners to retrieve them, even if they were wiped from the attacked machine by the attacker. Proxy servers logs are also essential to track users' activities. Netflow provides information about each packet and with the proper analysis, the forensic examiners may detect if P2P has been used on the network.

The detection of IP spoofing from external sources is crucial; avoiding IP spoofing requires some regulations and enforcement from concerned parties such as RIPE to force Internet service providers to never allow spoofing sources.

As previously stated, attackers may use AFT for many purposes, which prevents forensics examiners from finding useful information. Forensic examiners need to know if AFT has been used on the examined system to recover the affected information, as well as detecting the traces of AFT [13]. Many AFT tools fail to completely remove the relevant data, and they leave traces and footprints in the system registry, event logs, and storage about the usage of the tool on the system.

D. Attacks against Forensics Tools & Processes

There is no well-known software which runs against program packers, but some tools like Burndump can automatically detect if Burneye is running.

Keeping systems patched and up-to-date help in reducing the usage of buffer/heap overflow because patching the system provides immunity against known vulnerabilities but not against unknown Zero-day vulnerabilities. Some hardening actions are required to avoid the use of footprint minimization: the BIOS should be modified to stop booting from external devices, denying a change to the boot order, and protecting the entry to BIOS by a password. Although it can be bypassed by taking out the BIOS battery but it requires disassembling the computer case.

V. Recommendation

This section is dedicated to finding the best suitable solution with minimum overhead on the system. A software that can investigate the traces of AFT on a system with a check-list to detect the following:

- The BIOS is to be secured by a password to avoid modification.
- Virtualization should be disabled in BIOS
- The external boot should be disabled in BIOS
- Encrypted backup is to be adopted and scheduled
- Systems should always be up-to-date
- VPN access is to be blocked if not needed
- logs and events should be securely saved on servers

Developing software for this purpose is very useful and it is considered as a guideline to be followed by organizations that want to protect their data of any anti-forensic act that might target their systems.

VI. Conclusion

The goal of this work is to describe the various types of Anti-forensics attacks and techniques, and the problems that might face the forensics examiners in the area of digital forensics. Another goal is to show the countermeasures against these types of attacks taking into consideration their limitations. In conclusion, attackers are continuously enhancing the effectiveness of AFT, and they are developing new tools and techniques to make the digital forensic examination harder. It would be helpful if individuals, organizations, and governments conform and apply security policies that target data backups as well as the backup of logs. Encryption and data hiding are challenging in any investigation because encryption algorithms are designed to be hard to decrypt without the existence of super-computing. They are measured by a work function to show the effectiveness of the algorithm. Data hiding techniques are also designed to hide the detection of secret messages, especially if combined with encryption. Lastly, studying the traces of an AFT on a system would help the forensic examiners to detect that such techniques have been used on the examined system, which might reduce the time of the investigation.

References

[1] M. K. Rogers and K. Seigfried, "The future of computer forensics: a needs analysis survey," *Computers & Security*, vol. 23, no. 1, pp. 12–16, 2004.

[2] M. Al Fahdi, N. L. Clarke, and S. M. Furnell, "Challenges to digital forensics: A survey of researchers & practitioners attitudes and opinions," in *2013 Information Security for South Africa*, pp. 1–8, IEEE, 2013.

[3] M. Gül and E. Kugu, "A survey on anti-forensics techniques," in *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)*, pp. 1–6, IEEE, 2017.

[4] K. Conlan, I. Baggili, and F. Breitinger, "Anti-forensics: Furthering digital forensic science through a new extended, granular taxonomy," *Digital investigation*, vol. 18, pp. S66–S75, 2016.

[5] K. Conlan, I. Baggili, and F. Breitinger, "Anti-forensics: Furthering digital forensic science through a new extended, granular taxonomy," *Digital Investigation*, vol. 18, 08 2016.

[6] H.-M. Sun, C.-Y. Weng, C.-F. Lee, and C.-H. Yang, "Anti-forensics with steganographic data embedding in digital images," *IEEE Journal on selected areas in Communications*, vol. 29, no. 7, pp. 1392–1403, 2011.

[7] G. C. Kessler, "An overview of steganography for the computer forensics examiner," *Forensic science communications*, vol. 6, no. 3, pp. 1–27, 2004.

[8] A. Aminnezhad, A. Dehghantanha, and M. T. Abdullah, "A survey on privacy issues in digital forensics," *International Journal of Cyber-Security and Digital Forensics*, vol. 1, no. 4, pp. 311–324, 2012.

[9] G. Chhabra, "Anti-forensics techniques: An analytical review," 08 2014.

[10] H. N. Noura, O. Salman, A. Chehab, and R. Couturier, "Distlog: A distributed logging scheme for iot forensics," *Ad Hoc Networks*, vol. 98, p. 102061, 2020.

[11] S. Garfinkel, "Anti-forensics: Techniques, detection and countermeasures," *2nd International Conference on i-Warfare and Security*, 01 2007.

[12] M. Ölvecký and D. Gabriska, "Wiping techniques and anti-forensics methods," pp. 000127–000132, 09 2018.

[13] K. Park, J.-M. Park, E.-j. Kim, C. Cheon, and J. I. James, "Anti-forensic trace detection in digital forensic triage investigations," *Journal of Digital Forensics, Security and Law*, 01 2017.