Network Forensics: Notions and Challenges

Ahmad Almulhem
Computer Engineering Department
King Fahd University of Petroleum and Minerals
Dhahran 31261, Saudi Arabia
Tel: +966-3-860-7554, Fax: +966-3-860-3059
E-mail: ahmadsm@kfupm.edu.sa

***Abstract*—Network forensics is an extension of the network security model which traditionally emphasizes prevention and detection of network attacks. It addresses the need for dedicated investigative capabilities in the current model to allow investigating malicious behavior in networks. It helps organizations in investigating outside and inside network attacks. It is also important for law enforcement investigations. In this paper, various aspects of network forensics are reviewed as well as related technologies and their limitations. Also, challenges in deploying a network forensics infrastructure are highlighted.**

***Index Terms*—Network Forensics, Network Security, Computer Forensics, Computer Security**

I. Introduction

When it comes to network security, organizations typically use various products [1]. Generally, these products addresses security from two main perspectives; namely *prevention* and *detection*. Examples of prevention products include firewalls and access control mechanisms. Similarly, examples of detection products include intrusion detection systems and antivirus tools.

While many attacks are foiled by the used products, novel attacks still circumvent prevention products without being detected. In these situations, investigating the attacks is a very challenging task. In many cases, serious attackers are skillful at hiding evidences. Therefore, firewall logs and intrusion detection alerts may completely miss these attacks or may prove inadequate for a comprehensive investigation, especially when the goal is to apprehend the perpetrator [2].

In computer security literature, *network forensics* has been proposed to introduce investigation capabilities in current networks [3], [4]. It refers to a dedicated investigation infrastructure that allows for the collection and analysis of network packets and events for investigative purposes. It is proposed to complement the mentioned network security model.

Network forensics is of a great importance for today's organizations. On one hand, it helps to learn the details of outside attacks ensuring similar future attacks are thwarted. Additionally, network forensics is essential for investigating insiders' abuses which constitute second costliest type of attack within organizations [1]. Finally, law enforcement requires network forensics for crimes in which a computer or digital system is either being the target of a crime or being used as a tool in carrying a crime.

The rest of the paper is organized as follows. First, network forensics terminologies and process model are presented in section II. Then, the current practice in network forensics is presented in section III. In section IV, related technologies are reviewed showing their connection to network forensics and their limitations. Then, in section V, main challenges in designing a network forensics infrastructure are highlighted.

II. Background

A. Terminology

The term network forensics was previously used in few contexts without an official definition [3]. However, it generally refers to the collection and analysis of network data such as network traffic, firewall logs, IDS logs, etc. Technically, it is a member of the already-existing and expanding field of *digital forensics* [4], [5]. In particular, it is concerned with digital forensics in networked environments.

In real life, forensic science refers to the use of scientifically proved techniques to answer questions related to criminal and civil litigation. Analogously, network forensics is defined as:

> "*Network Forensics: The use of scientifically proved techniques to collect, fuse, identify, examine, correlate, analyze, and document digital evidence from multiple, actively processing and transmitting digital sources for the purpose of uncovering facts related to the planned intent, or measured success of unauthorized activities meant to disrupt, corrupt, and or compromise system components as well as providing information to assist in response to or recovery from these activities [4].*"

B. Model

In theory, digital forensics (and hence network forensics) is not a protection product. In particular, it is not supposed to replace firewalls and intrusion detection systems. Instead, it is a complex *process* in which methodologies, tools and human intelligence combine for the purpose of investigation.

In the literature, few models have been proposed to model the digital forensics process [4], [5], [6], [7], [8]. There is no consensus about which model best (or even correctly) represents the process. However, the proposed models share a common foundation when fine details are ignored. In particular, they are based on standard investigation models that are applied in real-life crimes.

The *Integrated Digital Investigation Process* (IDIP) is a representative model of the digital forensic process [8]. It

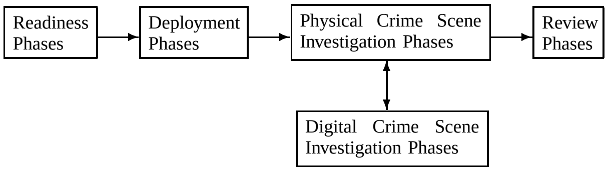
Readiness Phases
Deployment Phases
Physical Crime Scene Investigation Phases
Review Phases
Digital Crime Scene Investigation Phases

Fig. 1. The Integrated Digital Investigation Process (IDIP)

consists of various phases that are organized into five groups as shown in Figure 1. The following is a brief description of these groups:

- **Readiness phases**: The goal of these phases is to ensure that the personal and infrastructure are able to fully support an investigation when an incident occurs.
- **Deployment phases**: The goal of these phases is to provide a mechanism for an incident to be detected and confirmed.
- **Physical Crime Scene Investigation phases** : The goal of these phases is to collect and analyze the physical evidence and reconstruct the actions that took place during the incident.
- **Digital Crime Scene Investigation phases** : The goal of these phases is to analyze digital devices that were obtained from the physical investigation phases.
- **Review phases** : The goal of these phases is to review the whole investigation and identify areas of improvement.

III. State of the Art

Network forensics is currently a manual and time consuming process [9]. It is typically conducted by experienced system administrators. A typical investigation proceeds by analysing various types of logs. In a typical network setting, logs can be found in a number of places. For instance, a network is usually equipped with a dedicated auditing facility, such as *Syslogd* in Unix networks. Also, applications like web servers and network devices like routers and firewalls, maintain their own logs.

Various tools and homemade scripts are typically used for the investigation. For example in a Unix environment, an investigator may use free utilities like *tcpdump* [10], *grep*, *strings*, etc. Some investigators employ commercial tools known as *network forensic analysis tools* [11], [12], [13]. The architectures of these commercial tools are not disclosed. However, they provide functionalities similar to those free utilities. Although, they are generally more user-friendly and versatile.

Since network forensics is generally a manual and brute-force process, it is usually both time consuming and error-prone. Additionally, the mentioned logs are not meant for thorough investigation. The logs may lack enough details or contrarily have lots of unrelated details. They also come in different incompatible formats and levels of abstraction.

IV. Related Technologies

In this section, related technologies are reviewed showing their connection to network forensics and their limitations.

A. Intrusion Detection Systems

Overview: An *intrusion detection system (IDS)* refers to a system designed to detect computer and network attacks [14], [15]. It monitors computing resources (a single host or an entire network) and generates alerts when an attack is detected.

IDSs are deployed as *host-based* and/or *network-based*. Also, they employ two main approaches to detect attacks.

- *Signature-based*: An approach where detection is achieved by matching against a database of known attacks.
- *Anomaly-based*: In this approach, an IDS builds a model of “normal” activities of a system. It then alerts when a deviation is detected.

Network Forensics: An IDS is a valuable addition to a network forensics system. It can play the role of a sensor that triggers the forensics process. Additionally, the generated alerts constitute an important source of information which can be collected and analyzed later. These alerts also help analyzing data collected from other sources.

Limitations: There are a number of limitations pertaining to using IDS for network forensics:

- Detection Reliability: When relying on an IDS's output, there are a number of concerns. First, IDS suffers from false alarms; namely *false-positives* and *false-negatives*. A false-positive refers to the case when an IDS generates an alert for a nonexistent attack, while a false-negative refers to the case when an IDS misses an actual attack. A second concern is specific for network-based IDS. They can be a target for known classes of attacks; namely *evasion* and *insertion* attacks [16]. Also, they can not handle *encrypted* traffic.
- Data Details: In general, IDS's output lacks enough details for serious investigation. Typically, the output is a one-line text alert.

B. Honeypots

Overview: A honeypot refers to a set of services, an entire operating system or even an entire network that is built to lure and contain intruders [17], [18]. Although, honeypots are meant to be compromised, they are in reality a tightly sealed compartment that is well controlled and monitored.

Essentially, all honeypots share the same concept. They do not have any production value or any authorized activity. Thus, any attempt to interact with them is most likely malicious. Besides containing and studying attacks, they also can be setup to draw attention away from real targets [19].

Network Forensics: From an investigative perspective, a honeypot is an ideal tool to closely study attackers and capture their tools, keystrokes, etc. Few studies have been proposed to adopt honeypots for forensics purposes [20], [21]. A notable example is the *Honeynet Project*; a voluntary research organization dedicated to study the tools, tactics, and motives of attackers [18].

Limitation: From a legal point of view, honeypots can be problematic for at least two reasons. First, a honeypot has no value. It is solely setup to be compromised and attacked. Therefore, compromising it does not incur any damages. In other words, it is not possible to legally claim any damages.

Secondly, honeypots can be regarded as a borderline between keeping attackers out of a network and inviting them in [17]. Therefore, they may be challenged as an unfair *entrapment*.

C. Computer Forensics

Overview: Computer forensics is the oldest member in the digital forensics family. Traditionally, it refers to the forensics analysis of standalone computers found in crime scenes [22]. In particular, it involves analyzing their data storage devices, like hard disks. Typically, an investigator uses specialized software to recover deleted files, encryption keys, passwords, emails, etc.

Computer forensics has evolved over time following the standard methodologies used by law enforcement in investigating crimes in real life. Typically, the computer itself is not necessarily a victim of an attack, it is usually a tool used by a criminal. The forensics process usually follows well defined procedures to preserve, identify, extract, document, and interpret recovered data in the seized computer.

In general, computer forensics is not limited to personal computers. It also refers to investigating other digital devices that have some type of data storage medium. Examples of such devices include cellular phones, PDAs, digital cameras, etc. Like computers, these devices can also be found in crime scenes or with suspects.

Network Forensics: When performing a network forensics investigation, computer forensics techniques can be employed to investigate the computers as if they were not networked. In other words, a networked computer can be isolated and analyzed as a standalone computer. Accordingly, computer and network forensics actually complement each other.

Limitations: Computer forensics is solely for investigating standalone computers. It lacks in terms of investigating networked computers. In particular, it does not deal with issues that arise as a result of distributed sources of data. Such issues include data correlation, attack propagation, etc.

Additionally, computer forensics exclusively deals with persistent data stored on a local hard drive or other medium. In networked environment, however, there is a need to deal with volatile data such as network traffic. Accordingly, network forensics requires live data collection and analysis.

V. Challenges

A key challenge in network forensics is to first ensure that the network is forensically ready. For a successful network investigation, the network itself must be equipped with an infrastructure to fully support this investigation [5], [4], [6], [7], [8], [23]. The infrastructure should ensure that the needed data exists for a full investigation.

Designing a network forensic infrastructure is a challenging task because of the many possibilities in this design space. The following is a brief description of some of these challenges:

- Data sources: A typical network has several possible sources of data which includes raw network packets and logs of network devices and services. Although, it is desirable to collect data from all the possibles sources, this option is not always feasible especially for large networks. Therefore, an important decision is to select a subset of data sources which gives a good coverage of the network and makes the collection processes practical.
- Data granularity: A related issue to selecting data sources is to decide on how much details should be kept. For instance, when collecting network packets, one may collect whole packets, packets' headers, connection information (ip addresses, port numbers), etc. Similar to the above item, keeping extensive data details is not practical in large networks.
- Data integrity: It is critical to ensure the integrity of collected data. The outcome of the forensics process can be adversely affected if the collected data are altered either deliberately or accidentally. Therefore, measures have to implemented to ensure data integrity during and after data collection and analysis.
- Data as Legal Evidences: Using the collected data internally within an organization is quite different from presenting the data in a court of law. In the latter case, the collected data has to pass stringent legal procedures in order to qualify as evidences in a court of law. They have to pass an *admissibility* test; a screening process by the court [2], [24].
- Privacy Issues: Collected data is expected to include sensitive information such as personal emails and files. Therefore, proper handling of these data is crucial. The data has to be protected by access control measures, so only authorized personnel have access.
- Data Analysis: A major challenge is analyzing the collected data, in order to produce useful information that can be used in a decision making process. Such analysis process is in many respect challenging due to the complexity of a typical network environment and the amount and diversity of data involved. Innovative tools are needed to help human investigators to analyze data. These tools may apply techniques from fields like data mining [25], and information visualization [26].

VI. Conclusion

Nowadays, organizations use various products to protect their computer network. While many attacks are overcome by these products, novel attacks still circumvent prevention products without being detected. In these situations, investigating the attacks is a very challenging task.

In computer security literature, *network forensics* has been proposed to introduce investigation capabilities in current networks. It refers to a dedicated investigation infrastructure that allows for the collection and analysis of network packets and events for investigative purposes. In this paper, various aspects of network forensics were reviewed as well as related technologies and their limitations. Also, challenges in deploying network forensics infrastructure were highlighted.

References

[1] R. Richardson, "2008 csi/fbi computer crime & security survey," CSI, Tech. Rep., 2008.

[2] P. Sommer, "Intrusion detection systems as evidence," *Computer Networks*, vol. 31, 1999.

[3] M. Ranum, "Network forensics: Network traffic monitoring," Network Flight Recorder, Inc., Tech. Rep., 1997.

[4] G. Palmer, "A road map for digital forensic research," in *Digital Forensic Research Workshop*, Utica, New York, 2001.

[5] *Electronic Crime Scene Investigation: A Guide for First Responders*, U.S. Department of Justice: National Institute of Justice, July 2001.

[6] N. Beebe and J. G. Clark, "A hierarchical, objectives-based framework for the digital investigations process." *Digital Investigation*, vol. 2, no. 2, pp. 147–167, 2005.

[7] V. Baryamureeba and F. Tushabe, "The enhanced digital investigation process model," in *Digital Forensic Research Workshop*, Utica, New York, 2004.

[8] B. Carrier and E. H. Spafford, "Getting physical with the digital investigation process," *International Journal of Digital Evidence*, vol. 2, no. 2, 2003.

[9] C. Fennelly, "Analysis: The forensics of internet security," *SunWorld*, July 26, 2000.

[10] V. Jacobson, C. Leres, and S. McCanne, "Packet capture library," http://www.tcpdump.org/, (Last visited: May 26, 2007).

[11] N. King and E. Weiss, "Analyze this!" *Information Security Magazine*, Feb. 2002.

[12] N. Enterprise, "Netdetector: Proactive security surveillance solution," http://www.niksun.com/, (Last visited: May 26, 2007).

[13] S. Enterprises, "Netintercept: A network analysis and visibility tool," http://www.sandstorm.com/, (Last visited: May 26, 2007).

[14] T. Lunt, "Detecting intruders in computer systems," in *1993 Conference on Audit and Computer Technology*, 1993.

[15] M. Roesch and C. Green, *Snort Users Manual*, Apr. 2003.

[16] T. Ptacek and T. Newsham, "Insertion, evasion, and denial of service: Eluding network intrusion detection," *Secure Networks, Inc.*, Jan. 1998.

[17] B. Scottberg, W. Yurcik, and D. Doss, "Internet honeypots: Protection or entrapment?" in *Proceedings of the IEEE International Symposium on Technology and Society (ISTAS)*, 2002.

[18] L. Spitzner, "The honeynet project," http://www.honeynet.org, (Last visited: May 26, 2007).

[19] K. Takemori, K. Rikitake, Y. Miyake, and K. Nakao, "Intrusion trap system: an efficient platform for gathering intrusion-related information," in *10th International Conference on Telecommunications*, vol. 1, 2003, pp. 614–619.

[20] A. Yasinsac and Y. Manzano, "Honeytraps, a network forensic tool," in *Sixth Multi-Conference on Systemics, Cybernetics and Informatics*, 2002.

[21] B. Redmon, "Maintaining forensic evidence for law enforcement agencies from a federation of decoy networks: An extended abstract," *Mitretek Systems*, Fall 2002.

[22] J. H. W.G. Kruse, *Computer Forensics: Incident Response Essentials*. New York: Addison Wesley, 2001.

[23] A. Almulhem and I. Traore, "Experience with engineering a network forensics system," *Lecture Notes in Computer Science*, vol. 3391, pp. 62–71, Jan. 2005.

[24] D. Brezinski and T. Killalea, "Guidelines for evidence collection and archiving," RFC 3227, BCP 55, Feb. 2002.

[25] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*. Addison-Wesley, 2005.

[26] R. Marty, *Applied Security Visualization*. New York: Addison Wesley, 2008.