1

Introduction

André Årnes

Testimon Forensic Laboratory, Norwegian University of Science and Technology (NTNU), Gjøvik, Norway; and Telenor Group, Oslo, Norway

The world is becoming increasingly interconnected. We find connected devices in virtually every home, and computer networks are the nervous systems of corporate and government organizations everywhere. According to Internet Live Stats (2016), there are almost 3.5 billion Internet users in the world as of August 2016, covering close to 50% of the world's population. The Internet is, however, a network of networks consisting of competing and concurrent technologies with users from different organizations and countries. Unfortunately for the investigator, the Internet was designed for robustness and redundancy, rather than security and traceability. This increases the complexity and uncertainty of digital investigations and represents a formidable challenge for digital forensics practitioners.

Digital forensics is becoming increasingly important with the escalation of cybercrime and other network-related serious crimes. Understanding the laws and regulations governing electronic communications, cybercrimes, and data retention requires the continuous acquisition of new knowledge, methods, and tools. Digital evidence is everywhere and plays an important role in virtually any criminal investigation, from petty crimes to cybercrime, organized crime, and terrorism. It is therefore critically important that students of computer science and security acquire a fundamental understanding of digital forensics, in order to take part in the public debate and to act as experts in a legal context.

1.1 Forensic Science

Forensic science is a branch of science that is widely popularized in fiction and in contemporary media, ranging from Sir Arthur Conan Doyle's first Sherlock Holmes novel *A Study in Scarlet* published in 1887 to today's *CSI* and similar crime shows. It is commonly understood that forensic science is both highly inquisitive, requiring a creative mindset, and formalistic, requiring a strict adherence to established processes. An authoritative textbook in the field, *Criminalistics* (Saferstein, 2007), states that "forensic science in its broadest definition is the application of science to law." The terms *criminalistics* and *forensic science* are used interchangeably, although criminalistics has a

Digital Forensics, First Edition. Edited by André Årnes.

stronger flavor of the services of a crime laboratory. For the purpose of this book, we will only use the first term, as defined in Definition 1.1.

Definition 1.1: Forensic Science

The application of scientific methods to establish factual answers to legal problems.

A forensic scientist is responsible for the important task of establishing facts related to questions such as: what has happened, how did it happen, who has been involved, and when did it occur? To solve such problems, a forensic scientist draws on methods and tools from a wide range of theoretical and applied sciences, including biology, medicine, physics, geology, computer science, and electrical engineering. As it is often not possible to answer a problem with full certainty, a forensic scientist is also trained to apply statistics to express the results in terms of probabilities (for a comprehensive discussion, see Aitken & Taroni, 2004).

1.1.1 History of Forensic Science

Forensic science was established as a separate scientific domain during the 1800s and early 1900s. The contributions of this new area of science dramatically changed the effectiveness of law enforcement. A comprehensive overview of the contributions is available in Saferstein (2007), but some notable innovators and milestones are:

- Mathieu Orfila (1787–1853), considered the father of forensic toxicology, published the first scientific text on forensic toxicology in 1814.
- Alphonse Bertillon (1853–1914) developed a method for identification through body measurements and published a system on personal identification in 1879.
- Francis Galton (1822–1911) studied fingerprints as a means of identification and published the book *Finger Prints* in 1892.
- Hans Gross (1847–1915) established the principles for the application of science in investigations in several publications, the first one in 1893.
- Alberts S. Osborn (1858–1946) established scientific principles for document examination and published the book *Questioned Documents* in 1910.
- Leone Lattes (1887–1954) studied characteristics of blood types for identification and created a method for the analysis of blood groups in blood stains in 1915.
- Edmond Locard (1877–1966), recognized worldwide for promoting the scientific method in criminal investigation, established a police laboratory in Lyon in 1910.

1.1.2 Locard’s Exchange Principle

Edmond Locard formulated the famous Locard’s exchange principle, which has served as an important principle for subsequent research within forensic science. The principle states that “when a person or object comes in contact with another person or object, a cross-transfer of materials occurs” (Saferstein, 2007). In this way, every criminal can be connected to a crime through trace evidence. It should, however, be noted that the principle cannot necessarily be directly applied to digital forensics, as the dynamics of digital evidence is different from that of physical evidence. In this textbook, we will, nonetheless, adopt Definition 1.2.

Definition 1.2: Locard's Exchange Principle

Whenever two objects come into contact with one another, there is an exchange of materials between them.

1.1.3 Crime Reconstruction

Crime reconstruction (or crime scene reconstruction) is the process of determining the most likely hypothesis, or sequence of events, through the application of the scientific method. For the purpose of this textbook, we apply Definition 1.3, based on the book *Crime Reconstruction* by Chisum and Turvey (2008).

Definition 1.3: Crime Reconstruction

Crime reconstruction is the determination of the actions and events surrounding the commission of a crime.

A crime reconstruction can leverage a wide range of forensic methods, for example firearm ballistics tests, statistical simulations, and biological experiments. The objective is to establish a hypothesis about the event or sequence of events and then to test whether the hypothesis is possible or not. If the hypothesis is confirmed, then one possible explanation has been identified. If it is refuted, then the explanation is not possible and other hypotheses will have to be considered.

1.1.4 Investigations

An investigation is a systematic examination, typically with the purpose of identifying or verifying facts. A key objective during investigations is to identify key facts related to a crime or incident, and a common methodology used in this textbook is referred to as *5WH* (Stelfox, 2013; Tilstone *et al.*, 2013), as defined in Definition 1.4.

Definition 1.4: 5WH

5WH defines the objectives of an investigation as *who*, *where*, *what*, *when*, *why*, and *how*.

The 5WH formula sets the following objectives (Stelfox, 2013):

- *Who*: Persons involved in the investigation, including suspects, witnesses, and victims
- *Where*: The location of the crime and other relevant locations
- *What*: Description of the facts of the crime in question
- *When*: The time of the crime and other related events

- *Why*: The motivation for the crime and why it happened at a given time
- *How*: How the crime was committed.

1.1.5 Evidence Dynamics

Evidence dynamics is defined as “any influence that adds, changes, relocates, obscures, contaminates, or obliterates physical evidence, regardless of intent” (Chisum & Turvey, 2000). The concept is useful in understanding the actual behavior of evidence and plays an important role in crime scene reconstructions. Although the definition is originally intended for physical evidence, it is equally applicable to digital evidence. For example, evidence dynamics can describe the mechanisms for writing to a sector on a hard drive, or the operations for creating, changing, or deleting a file in a file system. For the purpose of this textbook, we will use Definition 1.5, based on Chisum and Turvey (2000).

Definition 1.5: Evidence Dynamics

Evidence dynamics refers to any influence that adds, changes, relocates, obscures, contaminates, or obliterates evidence, regardless of intent.

1.2 Digital Forensics

Digital forensics refers to forensic science applied to digital information, whereas a *digital investigation* refers to investigations in the digital domain. We will use the definition from the first Digital Forensics Research Workshop (Digital Forensics Research Workshop, 2001), as defined in Definition 1.6.

Definition 1.6: Digital Forensics

The use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation, and presentation of digital evidence derived from digital sources for the purpose of facilitating or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorized actions shown to be disruptive to planned operations.

Other terms, such as *network forensics*, *device forensics*, and *Internet forensics*, are often used to label specialized fields within digital forensics. As information technology has become an integral part of all aspects of society, digital forensics is growing in importance. Most legal cases today have an aspect of digital forensics, involving for example mobile phones, credit card transactions, email systems, Internet logs, and GPS systems. As many types of digital evidence can be volatile and easily manipulated, the trusted preservation of evidence through the use of standardized forensic tools and methods has become essential.

You may have previously seen references to digital archaeology and digital geology when discussing digital forensics. This refers to the analogy introduced in Farmer and Venema (2004), where *digital archaeology* refers to digital traces in computer systems created by human behavior, whereas *digital geology* refers to digital traces created by the computer systems themselves as part of their inherent processes. The goal of digital forensics is usually to gather facts about human behavior (i.e., digital archaeology), but it is a prerequisite to understand how the computer systems behave (i.e., digital geology) in order to interpret digital evidence.

As forensic scientists and forensic practitioners, our role in digital forensics is to establish factual answers to legal problems. This responsibility calls for strong standards for the processing of evidence and for the soundness of the analysis and its conclusions. This textbook will provide a comprehensive introduction to this process and its principles.

1.2.1 Crimes and Incidents

An important consideration for digital forensics is that it is commonly applied both in criminal law and in private law. Law enforcement increasingly depends upon digital forensics to process digital evidence in the context of a *crime* under investigation, whereas public and private companies and organizations depend upon digital forensics as a tool for supporting legal action in the case of an *incident* (e.g., contract or policy violations). Digital crimes and incidents consist of a *digital event* or a sequence of events, as defined in Carrier and Spafford (2004a). For simplicity, we refer to the event investigated as an *incident* in this textbook, unless we are specifically referring to criminal law. The location of the incident is referred to as the *scene of the incident*, or in the case of a crime, the *digital crime scene*.

This textbook addresses the field of digital forensics, with an emphasis on after-the-fact forensic analysis (often referred to as *post mortem*). Digital forensics can be initiated by real-time detection of cybersecurity incidents (e.g., intrusion detection) or as part of security incident handling processes. However, the topics of security monitoring and security incident management are not addressed in this textbook, as these are separate fields with different processes, contexts, and objectives.

1.2.2 Digital Devices, Media, and Objects

A central distinction in digital forensics is that between digital devices, digital media, and digital objects. A *digital device* is a physical object, such as a laptop, a smartphone, or a car. A digital device necessarily contains one or more storage media, such as a hard drive or memory, referred to as *digital media*. The digital media contain data, stored in binary format, referred to as *digital data*. Forensic analysts often work with discrete collections of digital data, referred to as *digital objects* in this textbook, based on the original definition in Carrier and Spafford (2004c).

1.2.3 Forensic Soundness and Fundamental Principles

The scientific method refers to the use of "a method or procedure . . . consisting of systematic observation, measurement, and experiment, and the formulation, testing, and modification of hypotheses" (Turvey, 2008). Ideally, this should also be the gold standard for digital forensics, with the implication that any investigation of digital evidence must be entirely reproducible by a third party. However, our ability to capture digital evidence in real life is far from perfect, and forensic soundness has therefore come to mean that the professionally recognized principles and standards of digital forensics are observed. Our definition of *forensically sound* is thus provided in Definition 1.7.

Definition 1.7: Forensically Sound

An investigation is forensically sound if it adheres to established digital forensics principles, standards, and processes.

The two fundamental principles discussed in this textbook are evidence integrity and chain of custody. *Evidence integrity* refers to the preservation of evidence in a complete form without any intentional or unintentional changes, as defined in Definition 1.8.

Definition 1.8: Evidence Integrity

Evidence integrity refers to the preservation of evidence in its original form.

While evidence integrity is an ideal in digital forensics, it is often not achievable, as data inevitably changes in live computer systems and networks during investigations. Due to this, documentation of all steps in the investigation is an important objective. This is referred to as the *chain of custody*, as defined in Definition 1.9.

Definition 1.9: Chain of Custody

Chain of custody refers to the documentation of acquisition, control, analysis, and disposition of physical and electronic evidence.

It should be noted that there are, of course, many other important principles in digital forensics, some of which are extensively discussed in this textbook. We have, nevertheless, chosen to adopt these two principles as fundamental principles that should be observed throughout all phases of the digital forensics process.

1.2.4 Crime Reconstruction in Digital Forensics

Crime reconstruction can help test hypotheses about a possible chain of events. It leverages the five-step process for event-based crime scene reconstruction as proposed by Carrier and Spafford (2004b).

1) *Evidence examination*: Identify and characterize evidence relevant to an incident.
2) *Role classification*: Examine the role of the evidence as a cause or effect of an event.
3) *Event construction and testing*: Identify events and assess whether they are possible.
4) *Event sequencing*: Combine events into event chains.
5) *Hypothesis testing*: The hypothesis is tested using the scientific method.

The method can also be applied in the case of digital forensics through the use of physical or virtual testbeds set up to perform simulated experiments, as discussed in Årnes *et al.* (2007).

1.3 Digital Evidence

Central to any digital investigation is the notion of digital evidence, which is defined in Definition 1.10, based on the definition by Carrier and Spafford (2004a).

Definition 1.10: Digital Evidence

Digital evidence is defined as any digital data that contains reliable information that can support or refute a hypothesis of an incident or crime.

In digital forensics, we aim to process and store digital evidence in a way that is consistent with the principles of evidence integrity and chain of custody. A number of digital evidence storage and exchange formats have been developed to support this (Flaglien *et al.*, 2011), but close attention to a manual process and detailed documentation of the chain of custody are nonetheless required.

1.3.1 Layers of Abstraction

As you will see in the remaining chapters in this book, it is useful to discuss digital evidence in the context of *layers of abstraction*. This refers to the practice, used in all areas of computing, of hiding implementation details of higher layers of abstraction in order to reduce complexity. A forensic analyst has to analyze and reconstruct data at all layers of abstraction to be able to extract and explain relevant digital evidence. For example, a forensic analyst may have to analyze data at the binary level of a disk drive to reconstruct a text file that contains an email with content relevant to the investigated case. A well-known example from computer networks is the Open Systems Interconnection (OSI) reference model, which divides network protocols into seven layers of abstraction (see also Section 7.3).

1.3.2 Metadata

Metadata is a valuable source of evidence in digital forensics that will be thoroughly discussed in this textbook. Metadata, or *data about data*, contains information about data objects. For example, the metadata associated with a digital photograph can contain the time of taking the photo, the geographical location, and the camera used. The analysis of metadata is an important activity throughout the forensic process, as metadata can contain information that is key to solving a case.

1.3.3 Error, Uncertainty, and Loss

Other aspects of digital evidence that have to be understood by a forensic scientist are error, uncertainty, and loss (Casey, 2002). Such uncertainties can impact the interpretation of timestamps, geographical location, and authorship or ownership of data. This must be understood in the context of evidence dynamics in order to accurately interpret and present digital evidence in a legal context. Failure to do so can lead to a weakened or lost case in court, or even the wrongful conviction of an innocent.

1.3.4 Online Bank Fraud – A Real-World Example

To better facilitate learning across the topics covered in this textbook, we will refer to the real-world online bank fraud example of SpyEye. This case is particularly relevant for us, as it is both legally and technically complex, with relevance to all the chapters in this textbook. The case is well documented, as it has been investigated by the US Federal Bureau of Investigation (FBI) and tried in public court with broad media coverage. The hackers behind the malware were recently convicted to a total of 24 years in prison (US Department of Justice, 2016).

1.3.4.1 Modus Operandi

Online bank fraud is a well-known crime pattern in which a large number of computers are compromised and infected with Trojan malware, allowing their computers to be monitored and remote controlled. The computers are part of a network of infected computers – a *botnet* – and typically controlled by one or more command-and-control centers. Once established, the botnet is continuously monitored to gather personal information and credentials, and to establish an overview of the online bank accounts accessed by the victims through the compromised computers.

At some point during a regular online bank session, when a victim is active, the command-and-control center issues an instruction to initiate a transaction of a specific amount of money from a victim's account. The Trojan malware, having circumvented the security of the online banking session, fools the unsuspecting victim to authenticate and authorize the transaction. When such an illegitimate transaction is completed, an amount of money is transferred to a complicit third party – a *money mule* – whose primary task is typically to withdraw the money and transfer it to a (possibly foreign) account in a jurisdiction that is more forgiving to financial fraud and cybercrime.

The crime pattern involves multiple actors and a long-term commitment, requiring a criminal organization with access to highly specialized competencies and tools. Central actors in such schemes are thus the malware programmers, the command-and-control center managers, the mules, as well as the organization that is recruiting the mules and profiting from the operation. Access to the required botnet, consisting of the command-and-control center and a network of compromised computers with active malware, can be purchased as a service on underground forums, as discussed in Namestnikov (2009). The victims are, of course, both the online bank customers and the financial institutions.

1.3.4.2 The SpyEye Case

The SpyEye case closely followed this pattern, infecting 50 million computers worldwide and compromising at least 10,000 bank accounts. The convicted programmers responsible for the malware and botnets, Aleksandr Andreevich Panin ("Gribodemon") of Russia and Hamza Bendelladj ("Bx1") of Algeria, were subsequently sentenced to a total of 24 years in prison for causing losses to financial institutions and individuals of close to a billion US dollars. One of Panin's clients reportedly made $3.2 million in 6 months (US Department of Justice, 2014).

Due to the scale and complexity of the case, law enforcement agencies from several countries were involved in the investigations, supported by cybersecurity experts from the private sector. From a technical perspective, the investigation ranged from the forensic analysis of a large number of infected computers, to online investigations of their distribution and sale in underground forums, and to tracing and identifying the perpetrators and malware where forensics is complicated by obfuscation and anti-forensic countermeasures.

1.4 Further Reading

We recommend that students of this textbook study supplementary literature to gain a broader understanding of digital forensics. For this purpose, here is a list of textbooks that have previously been utilized, in full or for inspiration, in our digital forensics curriculum:

- *Forensic Discovery* by Dan Farmer and Vietse Venema: used as the main textbook from 2007 to 2011 (Addison-Wesley Professional; Farmer & Venema, 2005).
- *Open Source Forensics* by Cory Altheide and Harlan Carvey: used as the main textbook in 2012 and 2013 (Syngress; Altheide & Harlan, 2011).
- *The Basics of Digital Forensics: The Primer of Getting Started in Digital Forensics* by John Sammons: used as the main textbook in 2014 (Syngress; Sammons, 2012).
- *Guide to Computer Forensics and Investigations – Processing Digital Evidence*, 5th ed., Bill Nelson, Amelia Phillips and Christopher Stewart: supplementary literature (Cengage Learning; Nelson *et al.*, 2015).
- *Incident Response and Computer Forensics*, 3rd ed., by Jason Luttgens, Matthew Pepe, and K. Mandia: supplementary literature (McGraw-Hill Osborne Media; Luttgens *et al.*, 2014).
- *Digital Archaeology: The Art and Science of Digital Forensics* by Michael W. Graves: supplementary literature (Addison-Wesley; Graves, 2014).

We have further utilized a wide range of specialized supplementary literature that is worth studying, including:

- *File System Forensic Analysis* by Brian Carrier (Addison-Wesley Professional; Carrier, 2010).
- *Windows Registry Forensics: Advanced Digital Forensic Analysis of the Windows Registry* by Harlan Carvey (Syngress; Carvey, 2016).
- *Network Forensics, Tracking Hackers through Cyberspace* by Sherri Davidoff and Jonathan Ham (Prentice Hall; Davidoff & Ham, 2012).
- *Practical Mobile Forensics* by Satish Bommisetty, Rohit Tamma, and Heather Mahalik (Packt Publishing; Bommisety *et al.*, 2014).
- *Practical Malware Analysis: The Hands-On Guide to Dissecting Malicious Software* by Michael Sikorski and Andrew Honig (No Starch Press; Sikorski & Honig, 2012).

1.5 Chapter Overview

This book is divided as follows:

- *Chapter 2, "The Digital Forensics Process"*: This chapter provides a comprehensive overview of the digital forensics process and its fundamental principles. A thorough understanding of these aspects is a prerequisite for any forensic practitioner.
- *Chapter 3, "Cybercrime Law"*: This chapter provides an introduction to the legal aspects of cybercrime investigations, with an emphasis on the Cybercrime Convention and criminal law with examples.
- *Chapter 4, "Forensic Readiness"*: With the process and legal considerations covered, it is time to look at the processes and capabilities required to perform a forensic investigation, whether in a government or corporate context. This chapter introduces the concept of forensic readiness and its requirements.
- *Chapter 5, "Computer Forensics"*: This chapter describes the most classic scenario in digital forensics – the forensic analysis of a regular computer. From a technical perspective, this chapter provides a fundamental insight into digital evidence, in preparation for the more specialized technical chapters.
- *Chapter 6, "Mobile and Embedded Forensics"*: This chapter covers digital forensics for mobile devices and embedded systems. As the world is becoming increasingly digital, with personal mobile devices and the "Internet of Things" everywhere, the science of digital forensics is facing a wide range of new challenges.
- *Chapter 7, "Internet Forensics"*: This chapter covers digital forensics on the Internet, with tracing, acquisition, and advanced analytics of networked evidence from various sources, such as open sources, personal services (e.g., email), and cloud services.
- *Chapter 8, "Challenges in Digital Forensics"*: This chapter provides an overview of research topics and open research questions within the field of digital forensics, with an emphasis on computational forensics. The purpose of the chapter is to provide inspiration for further research, and it is supported by examples from research at NTNU.
- *Chapter 9, "Educational Guide"*: This chapter provides guidelines and additional material for educators and students in digital forensics. The chapter includes an overview of available training programs, standards, and supporting literature.

1.6 Comments on Citation and Notation

For the benefit of the reader, the following standards have been adopted in the textbook:

- *Citations*: Citations to authoritative textbooks, research papers, and online sources are provided throughout. Students are encouraged to research the primary sources to better understand the subject matter.
- *Definitions, Examples, and Legal Provisions*: These are highlighted in separate gray boxes throughout the book. Examples can be either real-world case examples or illustrative scenarios.

- *Figures*: All photographs and illustrations are made by the chapter authors, unless otherwise specified.
- *Software*: All references to software and hardware tools are included as examples only. They do not represent a recommendation or preference regarding tool choices, and they should not be interpreted as guidelines or instructions on tool usage.