Chapter 3

TOWARDS A FORMALIZATION OF DIGITAL FORENSICS

Jill Slay, Yi-Chi Lin, Benjamin Turnbull, Jason Beckett and Paul Lin

Abstract While some individuals have referred to digital forensics as an art, the literature of the discipline suggests a trend toward the formalization of digital forensics as a forensic science. Questions about the quality of digital evidence and forensic soundness continue to be raised by researchers and practitioners in order to ensure the trustworthiness of digital evidence and its value to the courts. This paper reviews the development of digital forensic models, procedures and standards to lay a foundation for the discipline. It also points to new work that provides validation models through a complete mapping of the discipline.

Keywords: Digital forensic models, standards, validation

1. Introduction

Many digital forensic researchers and practitioners have been active in the field for several years. However, it is difficult for a new researcher, particularly one with a narrow technical background, to have a holistic view of the discipline, the tasks involved and the competencies required to carry them out. Similarly, for a new practitioner, the scope and depth of the discipline along with the risks and opportunities are very unclear.

Several issues are in need of discussion. One is that of definition or terminology. What, if any, are the differences between "computer forensics," "forensic computing" and "digital forensics?" In this paper, we use "digital forensics" as an overarching notion that subsumes these terms.

While questions of terminology may be troubling, Pollitt [17] raises the more pressing issue of the quality of digital forensic examinations, reports and testimony in the light of errors in cases brought before U.S. courts over the years. He asks whether different policies, quality manuals, validated tools, laboratory accreditations and professional certifications would have made a difference in these cases. He calls for practitioners to examine all their methods and to expose them to external review to ensure that they are trustworthy.

This paper reviews the development of models, procedures and standards underlying digital forensics to provide a foundation for the discipline. The foundation will help ensure the soundness and reliability of digital forensic processes and the veracity of evidence presented in court.

2. Digital Forensics

Pollitt [15] provides one of the earliest definitions of digital forensics:

> "[Digital] forensics is the application of science and engineering to the legal problem of digital evidence. It is a synthesis of science and law. At one extreme is the pure science of ones and zeros. At this level, the laws of physics and mathematics rule. At the other extreme, is the courtroom."

Pollitt's definition is a foundational one in that it encompasses the digital forensic process and the possible outcomes. The analysis of computer systems is a clear goal, but the results must be legally acceptable.

McKemmish [10] describes digital forensics in the following manner:

> "[Digital forensics] is the process of identifying, preserving, analyzing and presenting digital evidence in a manner that is legally acceptable."

McKemmish suggests that digital forensics is a multi-disciplinary domain. On the other hand, Palmer [13] defines digital forensics as:

> "The use of scientifically derived and proven methods for 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."

Palmer's definition was developed at the Digital Forensic Research Workshop (DFRWS). In fact, it represented the consensus of the academic researchers in attendance. This definition of digital forensics also implies the notions of reliability and trustworthiness.

Digital forensics is concerned with investigations into misuse, and its outcomes must be acceptable in a court of law or arbitration proceedings. Therefore, there is a heavy reliance on the non-technical areas of the field, especially given its multidisciplinary nature. This issue is emphasized by Yasinsac and colleagues [22]:

> "[Digital] forensics is multidisciplinary by nature, due to its foundation in two otherwise technologically separate fields (computing and law)."

Several other authors have attempted to define digital forensics. Some explore the scientific and/or legal validity of digital evidence. For example, Bates [1] states:

> "The purpose of forensic investigation is to enable observations and conclusions to be presented in court."

Forte [6] writes:

> "The simple guiding principle universally accepted both in technical and judicial spheres, is that all operations have to be carried out as if everything will one day have to be presented before a judge."

Meyers and Rogers [11] draw attention to the validity of forensic methods and the use of digital evidence in courtroom proceedings:

> "[Digital] forensics is in the early stages of development and as a result, problems are emerging that bring into question the validity of computer forensics usage in the United States federal and state court systems."

Solon and Harper [19] discuss the fragility of digital evidence and the importance of handling evidence properly:

> "Computer-based digital evidence is very fragile; it can be easily altered, damaged or destroyed by improper handling. If the data has not been dealt with correctly, the judge will not allow it to be used in legal proceedings."

Pan and Batten [14] emphasize the use of systematic and sound procedures for evidence extraction:

> "In order to be usable as evidence in a court of law, [information] needs to be captured in a systematic way without altering it in so doing. Thus, the process of identification and handling of the evidence is of prime concern in a forensic investigation."

All these definitions state that the ultimate goal of digital forensics is to provide legitimate and correct digital evidence in a court of law instead of merely examining computer equipment or analyzing digital data. Thus, digital forensics is not a discipline that is focused entirely on technical issues – as some who enter the field from computer science have been taught to believe. Rather, it is a discipline that embraces both computer techniques and legal issues.

3. Digital Forensic Procedures

Operating procedures are an important issue in the field of digital forensics. The quality, validity and credibility of digital evidence are greatly affected by the forensic procedures applied to obtain and analyze evidence.

Reith and colleagues [18] emphasize the benefits of general procedures in digital forensics:

> “This allows a consistent methodology for dealing with past, present or future digital devices in a well-understood and widely accepted manner. For example, this methodology can be applied to a range of digital devices from calculators to desktop computers, or even unrealized digital devices of the future.”

General procedures for digital forensics should be flexible rather than being limited to a particular process or system. Reith and colleagues also identify a number of reasons why standard operating procedures (SOPs) are lacking in many operational laboratories. The reasons include the uniqueness of cases, changing technologies and differing legislation. Many of these issues can be addressed by having flexible SOPs that permit changes within a framework but with clear overall outcomes.

3.1 Research-Based General Procedures

Several researchers (see, e.g., [2, 9]) have focused on general digital forensic procedures. These research-based procedures form the foundation for the development of practitioner-based general procedures. Indeed the efforts undertaken in the context of research-based procedures make it much easier to develop practitioner-based general procedures.

Reith and colleagues [18] are certainly not alone in discussing general procedures and methods for partitioning the various stages in a digital forensic investigation. McKemmish [10] lists four major steps in digital forensic investigations: (i) identification of digital evidence; (ii) preservation of digital evidence; (iii) analysis of digital evidence; and (iv) presentation of digital evidence. The generalized procedures described by McKemmish focus on more than just the technical elements; they specifically cover outcomes applicable to judicial personnel (judges, lawyers and juries) who may have limited technical backgrounds.

McKemmish subsequently refined and extended his work with input from other authors, developing the Computer Forensic - Secure, Analyze, Present (CFSAP) model [12]. There are two principal differences between the CFSAP model and McKemmish’s earlier work. First, the evidence identification and preservation phases in McKemmish’s earlier model are combined to create the secure phase in the CFSAP model. This modification was deemed necessary because the boundaries between the two phases are sometimes not clear.

The second difference is the flowchart used to describe the CFSAP model. This modification gives users a better understanding of the procedures involved in digital forensic investigations. Moreover, a feedback loop permits movement from the analyze phase back to the secure phase to ensure that digital evidence is not overlooked.

Palmer [13] proposed an alternative model at the 2001 Digital Forensic Research Workshop (DFRWS), which we refer to as the DFRWS model. The DFRWS model incorporates six processes, each with its own candidate methods and techniques:

- **Identification:** Event/crime detection, signature resolution, profile detection, anomaly detection, complaint resolution, system monitoring, audit analysis.

- **Preservation:** Case management, imaging, chain of custody, time synchronization.

- **Collection:** Preservation, approved methods, approved software, approved hardware, legal authority, lossless compression, sampling, data reduction, recovery.

- **Examination:** Preservation, traceability, validation, filtering, pattern matching, hidden data discovery, hidden data extraction.

- **Analysis:** Preservation, traceability, statistical methods, protocols, data mining, timeline, link, special.

- **Presentation:** Documentation, expert testimony, clarification, mission impact, recommended countermeasures, statistical interpretation.

Stephenson [20] also describes a model with six phases. His model expands McKemmish's preservation phase into three phases: preservation, collection and examination. Stephenson's phases are listed below. Note that they provide specific information related to preserving digital evidence.

- **Identification:** Determine items, components and data possibly associated with the allegation or incident; employ triage techniques.

- **Preservation:** Ensure evidence integrity or state.

- **Collection:** Extract or harvest individual items or groupings.

- **Examination:** Scrutinize items and their attributes (characteristics).

- **Analysis:** Fuse, correlate and assimilate material to produce reasoned conclusions.

- **Presentation:** Report facts in an organized, clear, concise and objective manner.

In addition to highlighting the unique characteristics of the DFRWS model, Reith and colleagues [18] have proposed a processing model called the Abstract Digital Forensics Model (ADFM). ADFM is an extension of the DFRWS model, but it also draws from other sources such as the FBI crime scene search protocol [21]. According to ADFM, a digital forensic investigation has nine key phases:

- **Identification:** Identify and determine the type of the incident.
- **Preparation:** Organize necessary tools, required techniques and search warrants.
- **Approach Strategy:** Dynamically build an approach to maximize the collection of evidence and minimize victim impact.
- **Preservation:** Protect and maintain the current state of evidence.
- **Collection:** Record the physical crime scene and produce a duplicated image of digital evidence via qualified procedures.
- **Examination:** Perform an advanced search for relevant evidence of the incident.
- **Analysis:** Provide an interpretation of the evidence to construct the investigative hypothesis and to offer conclusions based on the evidence.
- **Presentation:** Provide explanations of conclusions.
- **Returning Evidence:** Ensure that the physical and digital assets are returned to their owners.

Carrier and Spafford [3] have proposed the Integrated Digital Investigation Process (IDIP) model, which is based on theories and techniques derived from physical investigative models. The IDIP model is based on the concept of a "digital crime scene." Instead of treating a computer as a substance that needs to be identified, it is treated as a secondary crime scene and the digital evidence is analyzed to produce similar characteristics as physical evidence.

The IDIP model has five phases: (i) readiness; (ii) deployment; (iii) physical crime scene investigation; (iv) digital crime scene investigation; and (iv) review. The purpose of the digital crime scene investigation phase is to collect and analyze digital evidence left at the physical crime scene. This digital evidence must be connected to the incident under investigation. The IDIP model is further broken down into seventeen sub-phases. Interested readers are referred to [3] for additional details.

3.2 Practitioner-Oriented Operating Procedures

Standard operating procedures (SOPs) are usually the ultimate goal of practitioner-based computer forensic models. Proper SOPs are essential for digital forensic practitioners to perform investigations that ensure the validity, legitimacy and reliability of digital evidence [4, 7–9, 16].

Definitions of SOPs have been discussed for several years. Pollitt [16] notes that standards serve to limit the liability for actions by examiners and their organizations. Lin and colleagues [8] emphasize that law enforcement agencies must define SOPs to enable personnel to conduct searches and process cases in a proper manner. Pollitt [16] examines SOPs from a scientific perspective while Lin and colleagues [8] focus more on the legal consequences of using improper SOPs. Several other authors (e.g., [9, 11]) agree that SOPs have profound significance, but emphasize that they should be flexible enough to accommodate the changing digital forensic environment.

Creating a permanent set of SOPs is infeasible. Pollitt [16] states that standards can impede progress and limit creativity. As new problems and tools become available, new methods for solving forensic problems will be created. Therefore, regular updates to SOPs will be necessary to deal with changing technologies and legal environments. Permanent SOPs will eventually become outdated and useless.

In general, digital forensic practitioners prefer to have a universally accepted set of SOPs. However, Palmer [13] argues that it is difficult to create such a set of SOPs because analytical procedures and protocols are not standardized and practitioners do not use common terminology. Reith and colleagues [18] reinforce this position by pointing out that forensic procedures are neither consistent nor standardized.

4. Accreditation Standards

Many organizations have sought to maintain high quality in their digital forensic processes by pursuing ISO 17025 laboratory accreditations. This international standard encompasses testing and calibration performed using standard, non-standard and laboratory-developed methods. A laboratory complying with ISO 17025 also meets the quality management system requirements of ISO 9001.

The high workloads and dynamic environments encountered in digital forensic laboratories can make it difficult to meet accreditation requirements. A Scientific Working Group meeting in March 2006 at Australia's National Institute of Forensic Science [5] addressed the principal issues regarding accreditation. While earlier discussions had concentrated on the trustworthiness of digital evidence, the meeting participants strongly believed that the Australian approach should focus on the validation of digital forensic tools and processes.

Given the high cost and time involved in validating tools and the lack of verifiable, repeatable testing protocols, a new sustainable model is required that meets the need for reliable, timely and extensible validation and verification. Also, a new paradigm should be adopted that treats a tool or process independently of the mechanism used to validate it.

If the domain of forensic functions is known and the domain of expected results is known, then the process of validating a tool can be as simple as providing a set of references with known results. When a tool is tested, a set of metrics can also be derived to determine the fundamental scientific measurements of accuracy and precision.

Mapping the digital forensics discipline in terms of discrete functions is the first component in establishing a new paradigm. The individual specification of each identified function provides a measure against which a tool can be validated. This allows a validation and verification regime to be established that meets the requirements of extensibility (i.e., the test regime can be extended when new issues are identified), tool neutrality (i.e., the test regime is independent of the original intention of the tool or the type of tool used), and dynamically reactive (i.e., testing can be conducted quickly and as needed).

The Scientific Working Group agreed to describe the digital forensic component in terms of two testable classes, data preservation and data analysis. These two classes in their broadest sense describe the science in sufficient detail to help produce a model that is useful for accreditation purposes, not only for validation and verification, but also for proficiency testing, training (competency) and procedure development.

Figure 1 presents the validation and verification model. The model covers data preservation and data analysis. Data preservation has four categories while data analysis has eight categories.

A breakdown of the forensic copy category can be used to illustrate the depth of the categorization of functions. Static data is data that remains constant; thus, if static data is preserved by two people one after the other, the result should remain constant. An example is a file copy or a forensic copy (bit stream image) of a hard disk drive. Dynamic data is data that is in a constant state of flux. If dynamic data is preserved once and then preserved again, the results of the second preservation could be different from those of the original preservation. For example, a

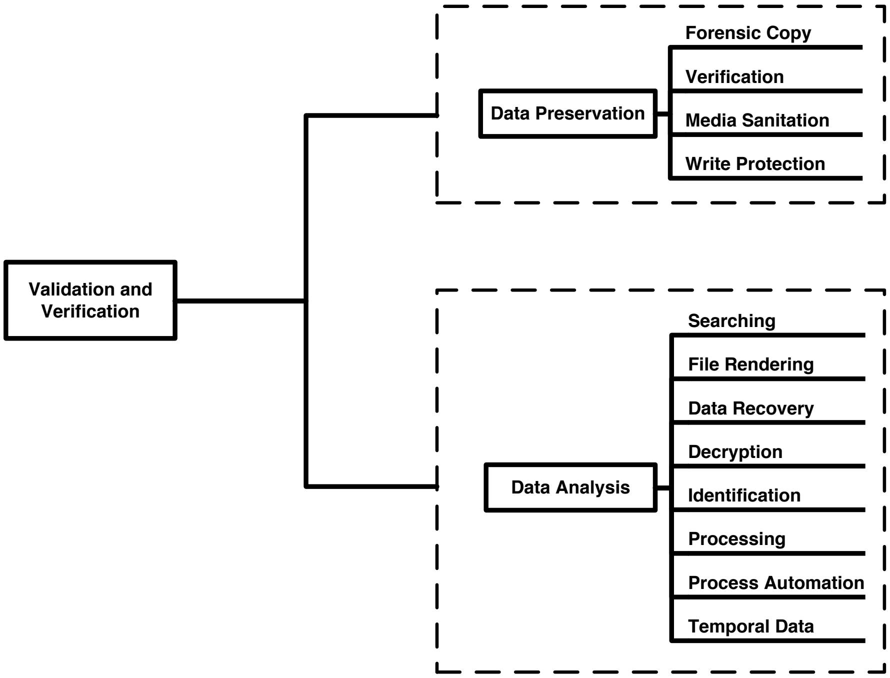

Figure 1. Validation and verification model.

TCP/IP traffic intercept or a memory dump is a snapshot at the instant of collection. A preservation of TCP/IP traffic or computer memory cannot be repeated with consistent results.

5. Conclusions

Experience has shown that the quality of digital forensic investigations is enhanced by the application of validated procedures and tools by certified professionals in accredited laboratories. Our review of digital forensic models, procedures and standards is intended to provide a foundation for the discipline. We hope that the foundation will help ensure the soundness and reliability of digital forensic processes and the veracity of the evidence presented in court.

References

[1] J. Bates, Fundamentals of computer forensics, *Information Security Technical Report*, vol. 3(4), pp. 75–78, 1998.

[2] J. Beckett and J. Slay, Digital forensics: Validation and verification in a dynamic work environment, *Proceedings of the Fortieth Annual Hawaii International Conference on System Sciences*, p. 266, 2007.

[3] B. Carrier and E. Spafford, Getting physical with the digital investigation process, *International Journal of Digital Evidence*, vol. 2(2), 2003.

[4] V. Civie and R. Civie, Future technologies from trends in computer forensic science, *Proceedings of the IEEE Information Technology Conference*, pp.105–108, 1998.

[5] Electronic Evidence Specialist Advisory Group, Electronic Evidence Specialist Advisory Group Workshop, National Institute of Forensic Science, Melbourne, Australia, 2006.

[6] D. Forte, Principles of digital evidence collection, *Network Security*, no. 12, pp. 6–7, 2003.

[7] A. Lin, I. Lin, T. Lan and T. Wu, Establishment of the standard operating procedure for gathering digital evidence, *Proceedings of the First International Workshop on Systematic Approaches to Digital Forensic Engineering*, pp. 56–65, 2005.

[8] I. Lin, T. Lan and J. Wu, A research of information and communication security forensic mechanisms in Taiwan, *Proceedings of the Thirty-Seventh International Carnahan Conference on Security Technology*, pp. 23–29, 2003.

[9] I. Lin, H. Yang, G. Gu and A. Lin, A study of information and communication security forensic technology capability in Taiwan, *Proceedings of the Thirty-Seventh International Carnahan Conference on Security Technology*, pp. 386–393, 2003.

[10] R. McKemmish, What is forensic computing? *Trends and Issues in Crime and Criminal Justice*, no. 118 (www.aic.gov.au/publications /tandi/ti118.pdf), 2002.

[11] M. Meyers and M. Rogers, Computer forensics: The need for standardization and certification, *International Journal of Digital Evidence*, vol. 3(2), 2004.

[12] G. Mohay, A. Anderson, B. Collie, O. de Vel and R. McKemmish, *Computer and Intrusion Forensics*, Artech House, Norwood, Massachusetts, 2003.

[13] G. Palmer, A road map for digital forensic research, *Proceedings of the 2001 Digital Forensic Research Workshop*, 2001.

[14] L. Pan and L. Batten, Reproducibility of digital evidence in forensic investigations, *Proceedings of the 2005 Digital Forensic Research Workshop*, 2005.

[15] M. Pollitt, Computer forensics: An approach to evidence in cyberspace, *Proceedings of the Eighteenth National Information Systems Security Conference*, pp. 487–491, 1995.

[16] M. Pollitt, Principles, practices and procedures: An approach to standards in computer forensics, *Proceedings of the Second International Conference on Computer Evidence*, pp. 10–15, 1995.

[17] M. Pollitt, Digital orange juice, *Journal of Digital Forensic Practice*, vol. 2(1), pp. 54–56, 2008.

[18] M. Reith, C. Carr and G. Gunsch, An examination of digital forensic models, *International Journal of Digital Evidence*, vol. 1(3), 2002.

[19] M. Solon and P. Harper, Preparing evidence for court, *Digital Investigation*, vol. 1(4), pp. 279–283, 2004.

[20] P. Stephenson, Modeling of post-incident root cause analysis, *International Journal of Digital Evidence*, vol. 2(2), 2003.

[21] K. Waggoner (Ed.), Crime scene search, in *Handbook of Forensic Services*, Federal Bureau of Investigation, Quantico, Virginia, pp. 171–184, 2007.

[22] A. Yasinsac, R. Erbacher, D. Marks, M. Pollitt and P. Sommer, Computer forensics education, *IEEE Security and Privacy*, vol. 1(4), pp. 15–23, 2003.