DIGITAL FORENSICS: OPERATIONAL, LEGAL AND RESEARCH ISSUES

M. Pollitt, M. Caloyannides, J. Novotny and S. Shenoi

Abstract With more than 93% of the world's data being computer generated [23], digital forensics offers significant opportunities and challenges. This paper discusses the operational and legal issues related to digital evidence. In addition, it highlights current needs and future research opportunities.

Keywords: Digital Evidence, Computer and Network Forensics, Forensic Processes, Legal Issues, Research Directions

1. Introduction

Digital forensics played a crucial role in the investigations of the 9-11 terrorists, the Enron, WorldCom and Martha Stewart scandals, and the DC sniper attacks. In the sniper case, electronic data extracted from a GPS device in the suspects' automobile provided information about their route across the country. Three unsolved murders along this route, in Louisiana, Alabama and Georgia, which involved weapons of the same caliber as the DC attacks, were connected to the sniper suspects via ballistics and fingerprint evidence. Indeed, due to network convergence and the ubiquity of embedded systems and devices, practically every crime – not just white-collar crime – has a digital forensic component [9,15,21]. Laptops, cell phones and PDAs seized from drug dealers and other violent criminals often yield vital evidence [15,21]. Meanwhile, electronic evidence is playing an increasingly important role in civil litigation. This paper discusses the operational and legal issues related to digital evidence. In addition, it highlights current needs and future research opportunities.

2. Operational Issues

Computers and other electronic devices invariably contain information relevant to investigations. This is the electronic analog of Locard's forensic principle of "exchange," where "every contact leaves a trace" [10]. This section discusses the nature of digital evidence and the forensic process, i.e., the extraction and analysis of digital evidence.

2.1. Digital Evidence

A crime scene contains observable evidence, e.g., blood, which has probative value. However, forensic science has developed techniques, e.g., DNA analysis, that yield information that is not physically observable. Computers, hard drives and flash memory cards are patent (observable) evidence that contain latent information, e.g., files, application metadata, file system debris and operating system debris. Extracting and analyzing latent evidence requires specialized tools and techniques. Moreover, the "forensic specialists" who perform these activities must be well trained and capable of providing expert testimony [18].

The Scientific Working Group on Digital Evidence (SWGDE) has defined digital evidence as: "information of probative value stored or transmitted in digital form" [6]. This definition is now widely accepted by other groups, e.g., the International Organization on Computer Evidence and the G-8 High Tech Crime Subcommittee.

2.2. Digital Forensics

Various terms, e.g., computer forensics, cyber forensics, media analysis and network forensics, are used to refer to the process of acquiring, preserving, examining, analyzing and presenting digital evidence. We prefer "digital evidence forensics," or simply, "digital forensics."

2.2.1 Acquisition. In the acquisition step, electronic devices at the scene are gathered, marked for identification and entered into an evidence control system. When evidence resides on devices that cannot – for legal or practical reasons – be physically collected, e.g., a hospital network containing "live" medical data, it is necessary to create a physical duplicate of the evidence. Dynamic evidence, e.g., network packets in transit, must be captured and recorded on physical media, both accurately and reliably.

2.2.2 Preservation. The preservation step requires the "original" piece of digital evidence to be maintained so that the opposing counsel and the court may be assured that it is both reliable and unaltered [12]. Failure to do so may result in the inadmissibility of evidence. Consequently, forensic examinations are almost always conducted on duplicate media [11].

Thus, preservation has two components: physical preservation of the original and the creation of duplicate media for forensic examination. The former is accomplished via "chain of custody" procedures; the latter by creating either a forensically accurate duplicate or a demonstrably accurate representation of the original. The latter process is often referred to as creating a "forensic image" or a "bit-stream copy."

2.2.3 Examination. Modern computers and data storage devices have huge, ever increasing capacities, making it extremely difficult to identify information of probative value. Examination is the process whereby evidence is subjected to review for its source, origin, and manner of creation, alteration or destruction.

The first step is to document all aspects of the evidence, e.g., physical and logical characteristics, active and deleted file structures, metadata and data contained in unallocated areas. This documentation process assures that both inculpatory and exculpatory information is preserved. Next, the actual examination may begin. Since examinations are usually conducted to support legal proceedings, the nature and specifics of the evidence that would be probative is unique to each examination [20]. Therefore, it is important to define a plan of action and to select forensic software tools based on the goals of an examination.

Numerous forensic tools, including many of dubious quality, are available. Sometimes software not designed for forensic use is pressed into service, but this may raise legal challenges to the evidence or, even worse, render the evidence useless. Collectively, tools seek to include data as potentially probative or to exclude data that would not be probative. Obviously, the more restrictive the filter, the less data there will be to sift through, but it is also more likely that probative information will be missed. The ability to construct a process that invokes the fewest and most appropriate tools in the most efficient way is the artistic part of forensic examination. As storage volume and network bandwidth grow exponentially, it will be increasingly difficult to create tools and processes that are both efficient and accurate.

2.2.4 Analysis. Analysis involves the review of evidence for its content, probative value and usefulness to the investigative or legal objectives. Thus, analysis goes beyond the technical aspects of the original evidence and its constituent parts. While examination attempts to produce information that is potentially probative, it is not until the evidence is evaluated in the context of all of the other evidence that its value can be established. Data recovered from the original evidence may support other information known in the case, it may contradict other information in the case, it may be new information, or it may prove to be non-pertinent. To make these determinations as accurately as possible, it is necessary to have a clear understanding of the investigative process and detailed knowledge of all the investigative material.

2.2.5 Presentation. The output of forensic examination and analysis comprises the data files extracted from the forensic copy of the original evidence and a report documenting the forensic process, pertinent metadata and conclusions. Obviously, the presentation must be complete, concise and clear.

3. Legal Constraints

This section discusses U.S. laws governing the acquisition and use of digital evidence in criminal, civil and intelligence investigations.

3.1. Criminal Investigations

U.S. federal law relating to the acquisition of digital evidence by law enforcement agencies is governed by the Fourth Amendment and statutory privacy laws codified in the Wiretap Statute (18 U.S.C. Sections 2510-22), the Electronic Communications Privacy Act (ECPA) of 1986 (18 U.S.C. Sections 2701-02) and the Pen/Trap Statute (18 U.S.C. Sections 3121-27) [22,27].

The Fourth Amendment limits the ability of government agents to obtain evidence without a warrant. A warrantless search does not violate the Fourth Amendment if: (i) the agents' conduct does not violate an individual's "reasonable expectation of privacy," or (ii) the search falls within an exception to the warrant requirement [5,27,28]. Therefore, agents must consider if a search violates the expectation of privacy. Even if a search does violate this expectation, it may still be reasonable if it falls within an exception to the warrant requirement [27].

Determining what constitutes a reasonable expectation of privacy is difficult in cases involving computers. To assist federal investigators, the U.S. Justice Department recommends that agents treat a computer as a closed container (e.g., a file cabinet) [27]. The Fourth Amendment prohibits law enforcement agents from accessing and viewing computer information without a warrant if they would be prohibited from opening a closed container and examining its contents in the same situation [27]. However, courts have reached differing conclusions on whether or not individual computer files should be treated as separate closed containers.

Warrantless searches that violate a reasonable expectation of privacy comply with the Fourth Amendment if they fall within an established exception to the warrant requirement. For example, agents may conduct a search without a warrant if an individual with authority has consented to the search, if there is an expectation that the evidence will be destroyed, or if the evidence is in plain view [16,27]. Other exceptions occur during searches incident to a lawful arrest, inventory searches and border searches [27].

Because courts may or may not apply an exception, it is advisable to obtain a warrant before searching and seizing evidence. As with any search pursuant to a warrant, law enforcement agents must establish probable cause, and describe the place to be searched and the items to be seized [27]. However, because computer files may be encrypted, misleadingly labeled, stored in unusual formats or commingled with innocuous files, agents cannot simply establish probable cause, describe the files they need, and then go and retrieve the data. Instead,

they must understand the technical limits of different search techniques, plan the search carefully and then draft the warrant in a manner that authorizes them to take the steps necessary to obtain the evidence.

Agents should obtain multiple warrants if they have reason to believe that a network search will retrieve data stored in multiple locations. Also, agents should obtain a second warrant to search a computer seized pursuant to a valid warrant if the property targeted by the search is different from that specified in the first warrant.

The incidental seizure of certain materials may implicate the Privacy Protection Act (PPA) or the ECPA [27]. Under the PPA, law enforcement agents must take special steps when planning a search that may result in the seizure of First Amendment materials (e.g., materials posted on the web). The ECPA regulates how agents can obtain the contents of electronic communications stored by third-party service providers. To minimize liability under these statutes, agents should attempt to determine in advance the type of information that may be stored on the computer systems to be searched [27].

Law enforcement agents can conduct "no-knock" warrants if they suspect that announcing their presence would lead to the destruction of evidence [27]. Courts may also authorize "sneak-and-peek" warrants, which do not require agents to notify the person whose premises are searched. The USA PATRIOT Act amended 18 U.S.C. Section 3103a to allow a court to grant a delay of notice associated with the execution of a search warrant if it believes that providing immediate notification will have adverse effects [26]. However, warrants issued under the amended statute still prohibit the seizure of electronic information unless specifically authorized by a court.

In addition to physically seizing computers, agents frequently perform electronic surveillance. In these situations, it is important for agents to act in accord with the Wiretap Statute and the Pen/Trap Statute. The Wiretap Statute (Title III) governs real-time electronic surveillance in federal criminal investigations [7,24]. Title III broadly prohibits the interception of oral, wire and electronic communications. However, Title III contains dozens of exceptions and agents must be familiar with them before performing surveillance. The Pen/Trap Statute governs the use of pen registers and trap and trace devices [27]. The statute authorizes a government attorney to apply for a court order authorizing the installation of a pen register and/or trap and trace device when the information sought is relevant to an investigation. As modified by the USA PATRIOT Act, it now applies to communications on computer networks in addition to telephone communications and gives nationwide effect to pen/trap orders [26].

In addition to the applicable federal laws, state laws must also be considered. In particular, many states have laws regarding privacy, especially with regard to stored electronic information. Unfortunately, laws vary considerably from state to state, and a detailed discussion is beyond the scope of this paper. It is

important to note that when there is conflict, federal laws supercede state laws. However, in some instances, state laws are more specific and are, therefore, applicable.

3.2. Civil Investigations

Private corporations investigating their own employees may not adhere to most of the legal constraints discussed above. Frequently, private corporations require employees to sign computer use policies that include stipulations regarding the monitoring of computer activities. In these cases, employees typically resign their privacy rights.

The Fourth Amendment also does not apply to searches conducted by private parties who are not acting as agents of the government [27]. Therefore, no violation of the Fourth Amendment occurs when a private individual acting on his own accord conducts a search and makes the results available to law enforcement. However, law enforcement agents must limit themselves to the scope of the private search. Otherwise, the agents may violate the Fourth Amendment and any evidence gathered may be suppressed.

3.3. Intelligence Investigations

Searching, seizing or otherwise obtaining digital evidence in intelligence investigations can raise difficult questions of both law and policy. For example, the bounds of the Fourth Amendment are less clear than they are for ordinary criminal investigations [27]. Furthermore, the Foreign Intelligence Surveillance Act (FISA) creates a secret court and a legal regime for counterintelligence surveillance inside the United States [25]. When operating outside the United States, agents must be aware of the many laws in other nations that may affect an investigation. The U.S. Department of Justice Computer Crime and Intellectual Property Section (CCIPS) provides guidance on these matters.

4. Current Needs and Research Issues

This section discusses some of the significant areas of opportunity for research in digital forensics.

4.1. Data Storage and Analysis

The amount of data generated and stored in our daily activities is increasing rapidly. A recent University of California at Berkeley study estimated that almost 800 megabytes of data are produced annually for each person in the world; moreover, the per person figure is growing at about 30% per year [2,17]. In this environment, forensic investigators are faced with massive amounts of

evidence on digital media and networks. Single cases involving more than one hundred terabytes of data have already been encountered.

Storage area networks (SANs) are an attractive solution [3]. These high-speed special-purpose networks interconnect various data storage devices with massive data servers. SANs permit data from multiple cases to be loaded on networks and accessed, possibly remotely, by investigators. They allow for both physical and logical separation of forensic case files; this prevents evidence from being compromised, as with blood evidence in the well-known O.J. Simpson case. The increased storage capacity and scalability of SANs supports efficient information access, maximized hardware utilization, freedom from vendor dependence and the automation of forensic processes. SANs may also help solve current problems with long-term data storage by providing data redundancy (via RAID) and integrity. Promising research avenues include SAN architectures for large-scale, distributed operations, workflow modeling and analysis, validation and verification of SAN implementations, and the study of the legal ramifications of SAN use.

Another strategy for dealing with massive case files is to apply novel information manipulation techniques that can home in on pertinent data both rapidly and reliably. The main information manipulation techniques are data reduction and data mining. Data reduction methods involving known file filters and hash sets are currently used by forensic investigators, but these are limited in both scope and performance. Data mining employs a combination of machine learning, statistical analysis and modeling techniques to extract relevant information from large data sets. Data mining techniques are just beginning to find applications in digital forensic investigations, and numerous research opportunities exist in this area.

4.2. Specialized Devices

Digital investigations have largely concentrated on computers and networks. However, the ubiquity of hand-held electronic devices and new network appliances requires research on extracting digital evidence from non-traditional, and possibly damaged, equipment. Examples include fax machines, cell phones, smart cards, GPS navigational aids, digital cameras and various consumer electronics devices.

Fax machines store phone numbers of senders and recipients; upscale machines often maintain the actual faxed pages in memory. Cell phones store substantial amounts of data: dialed, received and missed calls as well as contact lists, photographs and MMS files. Information stored in smart cards (and not-so-smart cards) may include toll road access data, ostensibly anonymous prepaid phone cards and supermarket purchases. GPS devices store detailed path information that has been extracted and used successfully in criminal and

terrorism investigations. Digital cameras store data in memory chips, which are just as amenable to forensic techniques as traditional computer media, i.e., deleted files are not really deleted and slack space (between end-of-file and end-of-sector) will contain data. Like digital cameras, a multitude of other consumer electronics devices, such as MP3 players and camcorders, have digital memories and may contain useful information.

Embedded devices are an emerging trend. For example, automobiles now incorporate devices that store operational data, e.g., speed, brake status and throttle position, that may be relevant to investigations. The recent fatality accident investigation involving U.S. Rep. William Janklow (R-SD) exemplifies the use of digital evidence extracted from embedded devices in an automobile [1].

Research efforts should seek to identify electronic devices that may contain digital evidence and develop tools and techniques for extracting the evidence. In addition, it is necessary to develop best practices that meet the legal requirements for evidence admissibility.

4.3. Network Forensics

Forensic investigations involving networks languish for practical, jurisdictional and political reasons. Network administrators have no incentive to retain data that may be relevant to investigations outside their networks. This problem is even greater in situations where network nodes are located in foreign countries, especially non-friendly nations or those with different legal systems.

A potential solution is to develop and implement techniques that would make anonymity and pseudoanonymity harder to achieve on global networks like the Internet; this would limit the need for cooperation of foreign parties. IPsec [13], for example, would have resulted in increased sender authentication, but the pressure for its implementation has dissipated due to the popularity of NAT, which removes the urgency for new IP addresses [4], and SSL, which removes the urgency for a broadly accepted means to encrypt select traffic [8].

Clearly, it is difficult to solve the cultural, economic, political and jurisdictional problems that stifle network forensic investigations. However, it may be easier to develop technologies that ride on current IPv4 – and future IPv6 – networks, which could assist in tracing attacks across global networks. These technologies could then be proposed to cognizant organizations, e.g., IETF and IEEE, for possible codification as standards.

Peer-to-peer networks [19] and grid computing [14] are two emerging research areas in network forensics. Peer-to-peer applications create transient Internet-based networks, which are increasingly used to trade copyrighted and other illegal materials. Grid computing networks, on the other hand, utilize transient pools of computing nodes to perform tasks, each node contributing

cycles to the collaborative effort. Special techniques and tools must be developed to deal with the scale, jurisdictional and dynamic participation issues in these two types of networks.

4.4. Forensic Tools and Processes

Forensic tools have advanced considerably during the past few years. Nevertheless, the development of highly efficient software that can be utilized on a variety of platforms on a wide range of target systems in a modular fashion could narrow the gap between demand and functionality.

It is equally important to study the forensic process to determine how to conduct examinations most efficiently. Experienced examiners are very efficient, but no efforts have been undertaken to analyze and model their approaches with a view to automating the forensic examination process.

Regardless of the tools and techniques used, it is imperative to ensure the efficiency and effectiveness of forensic examinations. Furthermore, as data, media, forensic tools and investigations become more complex, it is increasingly difficult to demonstrate the "truth" of a forensic examination. To address this issue, research is needed in several directions, including software engineering approaches for trusted forensic software, automated processes and tools, and the validation of software as it applies to a particular platform, data set and forensic procedure.

4.5. Legal Considerations

To ensure that digital evidence will withstand judicial scrutiny, it is necessary to research the reliability and legality of forensic tools, techniques and procedures.

As court proceedings increasingly incorporate digital evidence, judges, juries and attorneys are more likely to test the merit of digital forensic tools and methodologies. In the United States, judicial evaluations of digital evidence – as with other forensic disciplines – will apply the Frye, Daubert and Kumho Tire tests [18]. Under these doctrines, the reliability of the proffered evidence must be demonstrated. This demands comprehensive validation and verification studies of digital forensic tools, techniques and procedures. Unfortunately, with the massive, dynamic electronic landscape, the digital forensics community has had limited success in developing timely and efficient means for validation, let alone verification.

Research on the privacy implications of digital forensic tools and procedures is critical. Even if tools and procedures are reliable and verifiable, investigators cannot wantonly seize digital devices and extract evidence. Electronic devices may store information protected by various privacy laws. Violating privacy laws

at any stage during the digital forensic process not only renders the collected evidence useless, but also exposes the investigator to civil action.

5. Conclusions

Sophisticated tools and procedures are sorely needed to acquire, preserve, examine, analyze and present digital evidence in a dynamic electronic landscape. But it is equally important that the tools and procedures be used properly and legally. Digital forensics is indeed a wide open area for research. Its technical, operational and legal dimensions pose myriad challenges and research opportunities.

References

[1] Cable News Network. Prosecutor: Janklow ran stop sign before deadly crash. www.cnn.com, August 20, 2003.

[2] Cable News Network. Scientists report data storage explosion. www.cnn.com, October 29, 2003.

[3] T. Clark. *Designing Storage Area Networks: A Practical Reference for Implementing Fiber Channel and IP SANs.* Addison-Wesley, Reading, Massachusetts, 2003.

[4] K. Egevang and P. Francis. RFC 1631: The IP network address translator. www.faqs.org/rfcs/rfc 1631.html, 1994.

[5] M. Elmore. Big brother where art thou? Electronic surveillance and the Internet: Carving away Fourth Amendment privacy protections. *Texas Tech Law Review,* 32:1053-1083,2001.

[6] Federal Bureau of Investigation. Digital evidence: Standards and principles. *Forensic Science Communications,* 2(2), April 2000.

[7] Federal Bureau of Investigation. *Congressional Statement on Limited Expansion of the Predicate Offenses for Title III Electronic Surveillance.* www.fbi.gov, 2001.

[8] A. Freier, P. Karlton and P. Kocher. The SSL Protocol Version 3.0. IETF (draft-ietf-tls-ssl-version3-00.txt), November 1996.

[9] S. Hinde. The law, cybercrime, risk assessment and cyber protection. *Computers & Security,* 22(2):90-95, 2003.

[10] K. Inman and N. Rudin. *Principles and Practices of Criminalistics: The Profession of Forensic Science.* CRC Press, Boca Raton, Florida, 2001.

[11] International Association of Computer Investigative Specialists. *Forensic Procedures.* www.cops.org, 2001.

[12] International Organization on Computer Evidence. *G8 Proposed Principles for the Procedures Relating to Digital Evidence.* www.ioce.org, 2000.

[13] Internet Engineering Task Force. IP security protocol (IPsec). www.ietf.org /html.charters/ipsec-charter.html.

[14] J. Joseph and C. Fellenstein. *Grid Computing.* Prentice Hall, Upper Saddle River, New Jersey, 2003.

[15] C. Kirby. Cyber sleuths: Computer forensics booms as importance of electronic evidence grows. *San Francisco Chronicle,* February 26, 2001.

[16] J. LaDue. Electronic surveillance and conversations in plain view: Admitting intercepted communications relating to crimes not specified in the surveillance order. *Notre Dame Law Review,* 65:490-533, 1990.

[17] P. Lyman and H. Varian. How much information? www.sims.berkeley.edu/how-much-info-2003, 2003.

[18] A. Moenssens, editor. Amendments to the Federal Rules of Evidence. www.forensic-evidence.com, 2003.

[19] D. Moore and J. Hebeler. *Peer-to-Peer: Building Secure, Scalable and Manageable Networks.* McGraw-Hill/Osborne, Emeryville, California, 2001.

[20] M. Noblett, M. Pollitt and L. Presley. Recovering and examining computer forensic evidence. *Forensic Science Communications,* 2(4), October 2000.

[21] M. Pollitt. Digital evidence. *Proceedings of the Thirteenth Interpol International Forensic Science Symposium,* Lyons, France, 2002.

[22] E. Sinrod, *el al.* Cyber-crimes: A practical approach to the application of federal computer crime laws. *Santa Clara Computer and High Technology Law Journal,* 16:177-232, 2000.

[23] State Bar of Texas. Electronic discovery and computer forensics collection. Computer and Technology Section, www.sbot.org/discovery_library.htm.

[24] R. Strang. Recognizing and meeting Title III concerns in computer investigations. *United States Attorneys' Bulletin,* 49(2):8-13, 2001.

[25] L. Tien. Foreign Intelligence Surveillance Act. Electronic Frontier Foundation, www.eff.org, 2001.

[26] U.S. Department of Justice. *Field Guidance on New Authorities that Relate to Computer Crime and Electronic Evidence Enacted in the USA Patriot Act of 2001.* www.cybercrime.gov, 2001.

[27] U.S. Department of Justice. *Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations.* www.cybercrime.gov, 2002.

[28] R. Winick. Searches and seizures of computers and computer data. *Harvard Journal of Law & Technology,* 8:75-128, 1994.