COMPUTER CRIME AND DIGITAL EVIDENCE

E Casey, Stroz Friedberg LLC, Washington, DC, USA

Introduction

Computers can be involved in a crime in three general ways: (1) as a target; (2) as an instrument; (3) and as a source of evidence. Unauthorized access, theft of services, and other offenses defined in many computer crime laws focus on computers as targets. In the past, the term computer crime was associated with this limited set of offenses, but as computer use becomes more prevalent in society, computers are becoming involved in a wider range of crimes and the term computer crime is assuming a broader meaning.

Computers can be directly involved in many types of criminal activity, including terrorism, organized crime, stalking, and child exploitation. For example, sex offenders and obsessional harassers use computers to threaten and control victims, making the computer an instrument of the crime. In addition, due to the nature of digital data and a computer's storage of it, computers can contain evidence relating to crimes. For instance, serial killer Maury Roy Travis sent a letter to a newspaper with a map that showed where a victims' body could be found. Travis was tracked down via a 'cybertrail' comprised of a unique number printed on the map, associated web server access logs, and internet dial-up records that were generated when he connected to the internet to download the map from an online travel website. Investigators searched Travis' home and found incriminating evidence, including victims' blood and videotapes showing a number of the victims being tortured and killed.

The scope of computer crime becomes even broader with the proliferation of mobile devices and equipment with built-in computers such as personal digital assistants, mobile telephones, and computers embedded in cars. A personal digital assistant can contain significant details about a victim or an offender's life and activities. These data are potentially retrievable by others. A mobile telephone can reveal which telephone numbers an individual called or was called from at particular times. Additionally, it may be possible to ascertain the locations of a victim and likely suspect, leading up to a violent crime based on the locations of their mobile telephones. Sensing and diagnostic modules in cars – analogous to the black box on an airliner, recording data such as vehicle speed, brake status, and throttle position during the last five seconds before an impact – are used to investigate automobile accidents.

For forensic purposes, it is generally not computers themselves that are of primary interest but rather the data they contain. Additionally, related data on networks such as the internet and mobile telephone systems can be useful in an investigation. The term digital evidence is used to refer to any data stored or transmitted using a computer that may have probative value. It may support or refute a hypothesis of how an offense occurred or address critical elements of an offense, such as intent or alibi. Given the current ubiquity of digital evidence, it is a rare crime that does not have some associated data stored on or transmitted using computers. It is becoming routine for law enforcement agencies to devote resources to forensic examination of computers in most types of criminal investigation to seek related evidence on computers and networks.

Digital Crime Scene Investigation

When computers are an instrument of a crime or a source of digital evidence, it is useful to think of them as secondary crime scenes. Like a physical crime scene, digital crime scenes can contain many pieces of evidence and it is necessary to apply the same processes to preserve, document, and search the scene. In addition, Locard's exchange principle applies to the digital realm, helping investigators establish continuity of offense and track down criminals. According to Locard's exchange principle, when two entities (e.g., objects, people, locations) come into contact during the commission of a crime, an exchange of evidence occurs. Despite their best efforts to conceal or destroy incriminating digital evidence, criminals who use computers often leave behind digital traces that are useful in an investigation. In past homicide cases, victims' computers contained evidence that led to the murders and evidence on offenders' computers revealed their intent to kill.

In the UK case involving Dr. Harold Shipman, changes he made to computerized medical records on his practice computer system were instrumental in convicting him of killing hundreds of patients. Following Shipman's arrest, police made an exact copy of the hard drive from his computer, thus preserving a complete and accurate duplicate of the digital evidence. By analyzing the computer application Shipman used to maintain patient records, investigators found that the program kept an audit trail, recording changes made to patient records. This audit trail indicated that Shipman had lied about patients' symptoms and made backdated modifications to records to conceal the murders. During his trial Shipman claimed that he was familiar with this audit trail feature and was sufficiently knowledgeable about computers to falsify the audit trail if he had actually been trying to hide these activities. However, the court was convinced that Shipman had altered the records to conceal his crimes and sentenced him to life in prison.

Attributing computer activities to a particular individual can be challenging. Digital evidence can provide a circumstantial link between a person and activities on a computer, but it can be difficult to prove beyond a reasonable doubt that the defendant committed the crime. For instance, logs showing that a particular internet account was used to commit a crime do not prove that the owner of that account

was responsible since someone else could have used the individual's account. Attributing a crime to an individual becomes even more difficult when a crime is committed from a publicly accessible computer, such as at an internet cafe or public-library terminal.

Using evidence from multiple independent sources to corroborate each other and develop an accurate picture of events can help develop a strong association between an individual and computer activities. In one stalking case, investigators did not have sufficient evidence to prove that their prime suspect sent a threatening e-mail from a public-library terminal. Therefore, they had to interview witnesses, compare the e-mail with letters that were mailed to the victims by the suspect years earlier, and use other traditional investigative techniques to build a solid case.

As another example, a man accused of possessing child pornography argued that all evidence found in his home should be suppressed because investigators had not provided sufficient probable cause in their search warrant to conclude that it was in fact he, and not an imposter, who was using his internet account to traffic in child pornography (*US* v. *Grant*, US Court of Appeals, 1st Cir. 2000, available online at http://laws.lp.findlaw.com/1st/992332.html). During their investigation into an online child exploitation group, investigators determined that one member of the group had connected to the internet using a dial-up account registered to Grant. Upon further investigation, they found that Grant also had a high-speed internet connection from his home that was used as a FTP (File Transfer Protocol) server – the type of file-transfer server required for membership in the child exploitation group. Coincidentally, while tapping a telephone not associated with Grant in relation to another child pornography case, investigators observed that one of the participants in a secret online chat room was connected via Grant's dial-up account. Contemporaneous surveillance of the defendant's home revealed that his and his wife's cars were both parked outside their residence at the time. The court felt that there was enough corroborating evidence to establish a solid circumstantial connection between the defendant and the crime to support probable cause for the search warrant. Hence, using multiple independent sources of evidence, it is possible to establish a solid circumstantial link between online activities and an individual.

A Developing Forensic Discipline

From a forensic standpoint, it is necessary to process digital evidence in such a way that it will hold up under scrutiny in court. To address this need, formal principles and methodologies have been developed for processing evidence from a wide range of technologies.

Early approaches to processing digital evidence were developed primarily by law enforcement, with assistance from computer professionals, during the late 1980s and early 1990s in the USA. A number of new terms were created to describe this practice, including computer forensics, forensic computer analysis, and forensic computing. Also, new terms like network forensics, internet forensics, and incident forensics were created to accommodate other technologies. Although certain fundamental evidence-handling principles were applied to computers at this stage, such as maintaining chain of custody, no formalized methodology was developed. The lack of standards for how computers and networks were handled as a source of evidence resulted in a lack of consistency, making it more difficult for the practice to develop into the generally accepted norms of a forensic science discipline.

Several groups were formed to develop standards and a more scientific approach to processing evidence on computers and networks was established. The International Organization of Computer Evidence (IOCE) was established in the mid-1990s "to ensure the harmonization of methods and practices among nations and guarantee the ability to use digital evidence collected by one state in the courts of another state." In 1998, the Scientific Working Group on Digital Evidence (SWGDE) was established to "promulgate accepted forensic guidelines and definitions for the handling of digital evidence." As a result of these efforts, the American Society of Crime Laboratory Directors/Laboratory Accreditation Board (ASCLD/LAB) updated its accreditation manual in 2003 to include standards and criteria for "digital evidence examiners" in US crime laboratories. There are similar efforts to develop digital evidence examination into an accredited discipline under international standards (ISO 17025).

Although the SWGDE developed guidelines for training and best practices, it did not provide a solid methodology that would enable the field to develop into a science. In 2001, the first Digital Forensics Research Work Shop (DFRWS) was held, bringing together knowledgeable individuals from academia, the military, and the private sector to advance the field as a science. One outcome of this workshop was a framework for processing digital evidence and a suggested title for the field: "digital forensic science." Digital forensic science was defined as "the

use of scientifically derived and proven methods toward 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." The DFRWS workshop also led to the creation of the *International Journal of Digital Evidence* (IJDE), and later the *Journal of Digital Investigation*.

The approach to processing digital evidence developed at the DFRWS was based on traditional methods of the forensic sciences and has focused attention on forensic issues. For instance, more formalized methods of processing computers and networks are being developed that are modeled on physical crime-scene investigation. Researchers and practitioners are developing new techniques and tools for detecting, tracking, and attributing computer crime. Specifications for digital evidence processing tools are being developed to ensure that they address the needs of the forensic science community. In addition, related areas of expertise are emerging to deal with digital evidence from different technologies (e.g., networks, mobile telephones) and to perform specialized tasks (e.g., recovery of deleted or encrypted data, analysis of computer programs).

Digital Evidence

Because any crime can involve a computer, it is important to have a basic understanding of the kinds of digital evidence that might be available. Additionally, a familiarity with the fragility and limitations of digital evidence will minimize the risk of mishandling computers and damaging or misinterpreting the evidence they contain.

At its basic level, digital evidence exists in a physical medium such as a magnetic disk, a copper wire, or a radio signal in the air. Forensic examiners rarely scrutinize the physical medium and instead use computers to translate the data into a form that humans can interpret, such as text, audio, or video. Therefore, examiners rarely see the actual data but only a representation, and each layer of abstraction can lose information and introduce errors. For instance, analyzing the magnetic properties of a hard drive may reveal additional information useful for some investigations (e.g., overwritten data, the cause of damage to the disk). The risk of examining media at this low level is that the act of observing will cause changes that could destroy or undermine the evidence.

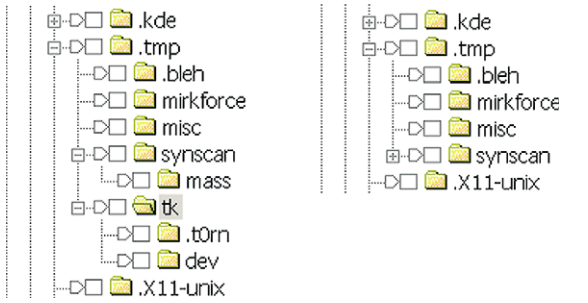

Figure 1 A folder named "tk" contained important evidence related to a computer intrusion investigation. The "tk" folder is visible using a newer version of a digital evidence examination tool (left), but not an older version containing a bug (right).

In fact, it is considered best practice to examine an exact replica of digital evidence to avoid altering the original. However, it can be difficult to obtain an exact and complete copy of a magnetic disk, random access memory (RAM), a copper wire, or a radio signal. For instance, programmatic mistakes (a.k.a. bugs) have been found in tools for collecting digital evidence from hard drives, resulting in only a portion of the data being copied. Bugs have also been found in tools for examining digital evidence on storage media, resulting in an inaccurate representation of the underlying data, as shown in Figure 1.

There are many other potential sources of error in digital evidence between the time data are created by a system and the time of preservation and analysis of the evidence. For instance, system malfunction can result in erroneous or missing log entries. In addition, as with other forms of evidence, poor training or lack of experience can lead forensic examiners to mishandle or incorrectly interpret digital evidence.

The multiple layers of separation between a forensic examiner and the original evidence can be problematic from a forensic standpoint. The possibility that important evidence was overlooked, misrepresented, or misinterpreted leaves the door open for criticism and reasonable doubt. To mitigate these risks, experienced digital evidence examiners do not base their conclusions on the results of one tool. For example, making copies of digital evidence with more than one tool reduces the chance that portions will not be collected. In addition, comparing results in multiple tools and validating important findings at a low level reduces the risks of misrepresentation and misinterpretation.

The mutability of digital evidence is another forensic concern. To demonstrate that digital evidence is an exact replica of the original and has not been altered since it was collected, it is common practice to

calculate a cryptographic hash (e.g., MD5, SHA1) of the evidence prior to collection. For instance, consider a letter found on a computer containing the sentence "Jane, I want to kill you" that has an MD5 value of 95a2592365b98fcac8c940de3d136943, as shown here:

```
C:\>type letter
Jane, I want to kiss you
C:\>md5sum letter
95a2592365b98fcac8c940de3d136943 *letter
```

By taking such precautions to document the original state of the evidence, any changes in the evidence can be detected quickly using the hash value. For instance, if the aforementioned letter was altered to contain the work "kill" instead of "kiss" this would be reflected in the MD5 value as shown here:

```
C:\>type letter
Jane, I want to kill you
C:\>md5sum letter
ccdcd9ac77345491a6c10609dc3ad338 *letter
```

Although the ephemeral nature of digital evidence has been mentioned, one benefit of digital evidence is that it can persist despite efforts to destroy it. For instance, deletion and formatting often involve removing higher-level "logical" references to the data (e.g., file names, locations of data on disk) but leave data on the physical medium. For instance, when a hard drive containing a Microsoft FAT file system is formatted, the file allocation table (FAT) is obliterated but the data from files remain on the disk and can be recovered. In addition to deleted files, other remnants of data can be found on disks in the form of RAM contents saved to disk by the operating system and temporary files created by some applications. For instance, Microsoft Word creates temporary files while a document is being edited, creating fragments such as the one shown in uninterpreted form (Figure 2).

Among other things, the data on line 4 in this fragment show where the associated document was located and what it was called (C:\Private\Secret1.doc.pgp). The ".pgp" file extension suggests that the original document was encrypted using a program called Pretty Good Privacy (PGP). This example demonstrates the valuable lesson that, even if a document was encrypted, it may be possible to recover portions of its contents in unencrypted form.

The file fragment in Figure 2 also demonstrates that a Microsoft Word document contains data that are not visible when the file is printed or viewed using Microsoft Word (a.k.a. metadata). In addition to the original file location and name, these metadata can include date–time stamps showing when the file was created and last modified, and can even indicate which computer was used to create the file, and the last ten authors of the document. For instance, the global unique identifier (GUID) value on the last line in Figure 2 suggests that the document was created on a computer containing a 3COM Ethernet network interface card with address "00-10-4B-DE-FC-E9." Because each network interface card is assigned a unique address, this information is very useful for identifying the source computer. Notably, newer versions of Microsoft Word do not include this address in the GUID value and knowledgeable computer users can alter this address on their computers. If the metadata in a suicide note indicate that it was created after the victim died on a computer other than the victim's, this could indicate that someone else wrote the note.

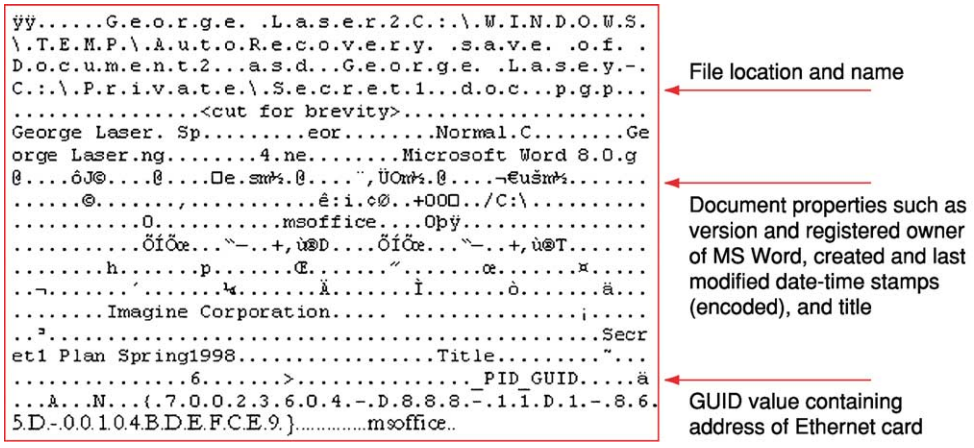

Figure 2 Fragment of a Microsoft Word document viewed in uninterpreted form showing metadata not visible when file is printed or viewed normally.

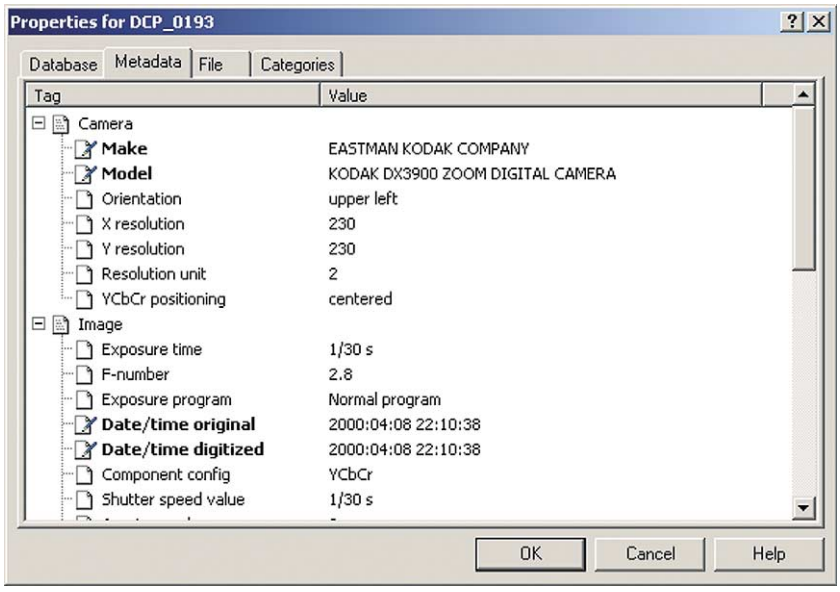

Figure 3 Metadata in photograph from a digital camera extracted using a tool called ACDSee. The metadata includes a time–date stamp created when the photograph was taken. In this instance, the date and time are inaccurate because the camera's clock was not set correctly – this photograph was actually taken in April 2003.

Evaluation of Source and Class Characteristics

Increasingly, offenders are becoming aware of the risks associated with using computers and are taking precautions to conceal their identities and destroy incriminating digital evidence. For instance, some offenders use anonymous internet services to make it more difficult for investigators to find them. Consider a harassment case in which the offender sends the victim threatening e-mail via an intermediate server. Normally the e-mail message would contain information about the computer used to send the message. Specifically, the e-mail header would contain the internet protocol (IP) address of the sender's computer (every computer on the internet is assigned an IP address to enable delivery of data). However, because the harasser sent the message via an intermediate server, the e-mail header will contain the IP address of that server, thus concealing the actual source. For example, headers in the following e-mail sent from a Yahoo! account indicate that the message was sent from an IP address in Japan (210.249.120.210):

> To: Count Rugen
> From: "Inigo Montoya" <inigo_montoya@yahoo.com>
> X-Originating-IP: 210.249.120.210
> Date: Wed, 04 Jun 2003 03:51:45-0000
> Subject: Prepare to die!

However, the sender merely connected to Yahoo! via this computer in Japan. Therefore, additional investigation would be required to determine the actual source of the message. Log files from the intermediate computer, such as those shown below, might contain the IP address of the actual sender's computer (172.16.34.14 in this example):

> **172.16.34.14**, anonymous, 6/4/03, 03:43:24, 210.249.120.210, GET, http://mailsrv.yahoo.com/login.html, 200
>
> **172.16.34.14**, anonymous, 6/4/03, 03:44:02, 210.249.120.210, GET, http://mailsrv.yahoo.com/inigo_montoya/inbox.html, 200
>
> **172.16.34.14**, anonymous, 6/4/03, 03:45:27, 210.249.120.210, GET, http://mailsrv.yahoo.com/inigo_montoya/compose.html, 200
>
> **172.16.34.14**, anonymous, 6/4/03, 03:51:36, 210.249.120.210, GET, http://mailsrv.yahoo.com/inigo_montoya/sent.html, 200

Similarly, it is not safe to assume that a file originated on the computer that it is found on, since it could have been created elsewhere and transferred via a network or cable. Additional investigation is required to determine the actual source of a network connection or piece of digital evidence. In the case of an IP address, the continuity of offense must be established, linking the offender to the crime. In the case of

a file, class and individuating characteristics can be used to assess the source. As an example, Figure 3 shows metadata extracted from a photograph taken with a Kodak DX3900 digital camera.

These metadata could be used to demonstrate that a photograph was likely taken using a suspect's camera, disproving a claim that he downloaded the file from the internet.

If these kinds of metadata are not available in a digital photograph, it may be possible to use other characteristics of a photograph to determine its source. For instance, Europol's Excalibur system uses image recognition technology to search a database of photographs from past investigations for similarities with a given image. If two photographs contain a common component such as a piece of fabric with a distinct design, this may indicate that they were taken in the same place, providing investigators with a lead.

Summary

As computers become more integrated in people's daily lives, investigators are encountering an increasing amount of evidence of witness, victim, and criminal activity in digital form. Even traditional crimes such as homicide and rape can involve digital evidence either directly or incidentally. Something as simple as a murder victim's personal diary on her computer can influence victimology, providing deep insight into her life and the people she interacted with, including the perpetrator and other victims of a serial homicide. In addition, digital evidence on computers and networks has helped identify and apprehend offenders in murder cases. Although some offenders take precautions to conceal, manipulate, or destroy digital evidence, sources may exist of which the offender was not aware or had no control over, particularly when networks are involved. An awareness of the kinds of data that may exist (e.g., deleted files, logs, metadata) and the inherent limitations (e.g., abstraction, mutable, evaluation of source) can help investigators make use of digital evidence.

There is a growing need for reliable methods and trained experts to process digital evidence as it is used in more criminal investigations. Efforts are being made to craft standards of practice and develop this field into a fully-fledged forensic discipline. One such endeavor is to develop a generally accepted training and certification process to help ensure that a crime scene expert who collects digital evidence, a forensic examiner who processes the recovered evidence, and an investigator who analyzes the evidence would all be applying the same principles and standards in their activities, written reports, and in court testimony. Additionally, tools for processing digital evidence are being tested to identify bugs that could introduce errors in the collection or examination stages. Training, standards development, and tool-testing initiatives must keep pace with advances in computer technology, making digital forensics an exciting and rapidly evolving field.

See Also

Children: Sexual Abuse, Overview; Sexual Abuse, Epidemiology; **Internet:** Forensic Medicine; Toxicology

Further Reading

Carrier B (2003) Defining digital forensic examination and analysis tool using abstraction layers. *International Journal of Digital Evidence* 1.

Carrier B, Spafford G (2003) Getting physical with the digital investigation process. *International Journal of Digital Evidence* 2.

Casey E (2002) Error, uncertainty and loss in digital evidence. *International Journal of Digital Evidence* 1.

Casey E (ed.) (2002) *Handbook of Computer Crime Investigation: Forensic Tools and Technologies*. London: Academic Press.

Casey E (2002) Cyberpatterns: criminal behavior on the internet. In: Turvey B (ed.) *Criminal Profiling: An Introduction to Behavioral Evidence Analysis*, 2nd edn. London: Academic Press.

Casey E (2004) *Digital Evidence and Computer Crime: Forensic Science, Computers, and the Internet*, 2nd edn. London: Academic Press.

Hollinger RC, Lanza-Kaduce L (1988) The process of criminalization: the case of computer crime law. *Criminology* 26: 101–126.