Editors: James A. Whittaker, jw@se.fit.edu
Michael Howard, mikehow@microsoft.com

# Computer Forensics

William H. Allen
*Florida Institute of Technology*

Increasingly, when the news media report on investigations into financial fraud, suspected terrorism, and other modern crimes, they mention the importance of evidence gathered from computers. Forensic examination of computer and other digital data has become an indispensable tool for law enforcement, corporate security, and intelligence gathering.

One definition of computer forensics is "acquiring, preserving, retrieving, and presenting data that has been processed electronically and stored on computer media."[1] However, many experts feel that a precise definition isn't yet possible because, increasingly, digital evidence is recovered from devices that aren't traditionally considered to be "computers." Some prefer to expand the definition to include the collection and examination of all forms of digital data, including that found in cell phones, PDAs, iPods, and other electronic devices. Even narrowing our definition to include only traditional computing systems (PCs, servers, and such), we find that computers can be involved in criminal activities in many different ways because of their ubiquitous nature. Among other things, a computer can

- be the target of a crime, including information theft, financial fraud, denial of service, or other direct attack;
- be used to commit crimes against other computers;
- be used to commit non-computer crimes, such as creating false documents or counterfeiting currency;
- be used to illegally copy or distribute copyrighted materials such as music or movies, or to store illegal documents, such as child pornography or information stolen from government agencies or corporations; or
- contain information such as contact lists, copies of falsified documents, or email that documents a conspiracy, which investigators could use to prevent or solve crimes.

In crimes that involve information stored in or manipulated by computers, forensic techniques are needed to extract and analyze that data. More controversially, a computer might hold information that could be used to determine an individual's intention to commit a crime. I ignore such investigations here, however, because they raise legal and social issues that are peripheral to this article's focus. Instead, I present an overview of the processes and problems related to computer forensics.

## Not as easy as it looks

Forensic specialists who work with physical evidence, such as blood, DNA, poisons, or firearms, have developed court-tested investigative procedures for collecting and processing it without contamination. These standards are generally based on decades of scientific research and empirical analysis. Equivalent standards for computer forensics are still emerging, hampered by the limited amount of academic research conducted in this area to date. Collecting and examining digital data thus presents several unique problems compared to those that arise with other types of evidence.

### Processing at a secure lab or in situ

Investigators can remove some types of physical evidence (hair or fiber samples, for instance) from crime scenes to analyze in labs; other types (such as fingerprints on a wall or skid marks on pavement) must be examined on site and preserved photographically or by other means. When collecting digital data, investigators can make exact copies using software tools that clone disks onto removable storage or other physical drives. This can be very useful when, for example, evidence related to an investigation that focuses on a single employee exists on a company's server; although investigators wouldn't be permitted to seize the hardware because of the harm it would inflict on the company, copying the data would allow the investigation to proceed.

### Ensuring that digital evidence remains unaltered

Many common software applications modify data when a file is opened—even if the user doesn't change the contents. Simply booting a modern operating system

causes changes to log and configuration files that could alter critical evidence. The system log, for example, is a vital piece of evidence that can show a suspect's login and logout times. When an investigator logs in to begin an examination, the system also records that activity and alters the file's contents and timestamps.

**In many ways, digital evidence has proven more difficult to analyze than physical evidence.**

To minimize the problems associated with such activities and verify the evidence's integrity, investigators often use cryptographic hash functions (commonly MD5 or SHA) to "fingerprint" individual files, or even entire disks, before copying data to another drive. This process, known as *imaging*, allows us to examine a copy of the files without contaminating the original data. If we discover evidence in the copy, we can use the hash function again to verify that the copy is a duplicate of the original data, thus proving that the same evidence exists on the original disk.

### Sifting through thousands of files

Several vendors have developed software specifically for forensic examination. Applications such as EnCase (www.encase.com), Forensic Toolkit (www.accessdata.com), and SleuthKit (www.sleuthkit.org) provide filtering and logging features that can greatly reduce investigators' workloads. For example, EnCase includes analytical tools that can find password-protected Word documents, extract relevant Windows registry entries, or recognize file types by their internal structures rather than their filename extensions. More complex searches can employ Boolean logic or regular expressions to locate specific text strings or binary patterns. Regardless of the tools employed, however, we must still follow well-established procedures, such as keeping a journal of all steps taken to discover evidence, documenting the chain-of-custody for all seized hardware and data, and verifying that evidence was not contaminated by the investigative process, to assure that discovered evidence is useful for courtroom presentation in criminal or civil cases.

### Examining more than existing files

As Simson Garfinkel and Abhi Shelat discussed in the first issue of *IEEE Security & Privacy*, digital data is more persistent than the average user expects.[2] Experienced criminals have learned to exploit this by hiding data in deleted files or other areas of a hard disk that conventional software can't access. Forensic tools such as those I mentioned can examine disks sector by sector to search for evidence contained in erased or partially overwritten files, but examiners must be careful not to accidentally manufacture evidence by piecing together sections of unrelated files. Savvy criminals use the more technically sophisticated technique of storing data in *file slack*—the space between a file's logical and physical ends. With a disk-editing tool such as the one included in Norton Utilities (www.symantec.com), users can insert information into unused bytes after a file's logical end where they're not visible when the file is opened in its usual application software. Fortunately, the forensic tools used for examining deleted files can also reveal the contents of file slack.

### Password-protected or encrypted content

The distinction between simple password protection and encryption is noteworthy here. The (somewhat questionable) security provided by some application-based password protection schemes can prevent naïve users who lack a password from gaining access to data. However, experienced investigators know that a text or hex editor (such as the one in Norton Utilities) will often provide unrestricted access.

Even with encrypted data, many users write down their passwords and keep them near their computers or use easily remembered (and easily discovered) passwords based on personal information. Investigators who have access to a suspect's computer usually also have access to detailed information on the suspect's family, friends, and pets (a common source of passwords). They often have access to the seized computer's physical location as well, which lets them search paper documents that might contain a recorded password.

Alternatively, password-cracking software can often discover the correct password via dictionary or brute-force searches. Encryption can sometimes block investigators, but errors in the design, implementation, or use of the encryption software make it less frequent than you might expect.[3]

### Reliability of digital evidence

In many ways, digital evidence has proven more difficult to analyze than physical evidence. Unlike fingerprints or DNA, bits are not unique. They have value only when grouped in patterns that represent information. Discovering certain patterns and reaching conclusions regarding their meaning is often subjective; forensic experts sometimes disagree about the evidence they represent or the way they were reconstructed.

Establishing the identity of a person who committed a specific crime is another issue that clouds computer forensic evidence's reliability. Although we might discover clear evidence of a crime, proving that a particular individual had sole access to the computer or email account

## Further resources on computer forensics

Many sources provide information on computer forensic tools, techniques, and training. Universities, private consultants, and IT training organizations provide a range of opportunities from short seminars to certification programs to graduate degree programs.

Several organizations provide online resources and practical training for those interested in computer forensics:

- International Association of Computer Investigative Specialists (www.cops.org)
- The SANS Institute (www.giac.org)
- International Information Systems Forensics Association (www.iisfa.org)

In addition to several recent articles in *IEEE Security & Privacy*, the following books and periodicals offer more information on digital evidence collection and examination:

- Eoghan Casey's *Digital Evidence and Computer Crime*, 2nd ed. (Elsevier Academic Press, 2004), takes an in-depth look at both legal and technical issues.
- *Computer Forensics: Incident Response Essentials* (Addison-Wesley, 2002), by Warren G. Kruse and Jay G. Heiser, provides a good tutorial on tools and techniques.
- John R. Vacca's *Computer Forensics: Computer Crime Scene Investigation* (Charles River Media, 2002) includes many case studies and guidelines for examiners.
- *Forensic Science Communications*, a quarterly peer-reviewed journal published by personnel at the US Federal Bureau of Investigation's forensics laboratory, often includes articles on digital data collection and analysis.

Various conferences and scholarly journals provide forums for research in computer forensics and related topics:

- International Journal of Digital Evidence (www.ijde.org)
- Digital Forensic Research Workshop (www.dfrws.org)
- *Digital Investigation: The International Journal of Digital Forensics & Incident Response*, published by Elsevier Advanced Technology (www.digitalinvestigation.net)

As these lists indicate, the majority of the information that's available is focused on training, tools, and techniques for computer forensic investigators.

A clear need exists for more active research in the area to bring computer forensics to the same level of maturity as other areas of forensic science.

during the period when the crime occurred is difficult if the machine is shared or improperly secured. This issue can be complicated by the presence of active viral or Trojan software on the suspect's computer, raising the question of whether unknown persons could have placed evidence on the suspect's hard drive.

### Privacy rights

The US Constitution protects citizens from "unreasonable search and seizure," but the courts have yet to agree on the restrictions those rights place on collecting digital evidence.

Based on privacy laws regarding telephone communications,[4] courts have held that email messages in transit are protected from monitoring without a warrant. Yet, recent rulings have also held that email stored, however briefly, on servers isn't as strongly protected. Thus, it might be possible to monitor email messages without a suspect's knowledge. Although privacy advocates are dismayed, law enforcement agencies see this as a new source of evidence.

### Legal issues

Modern computer systems' portability and connectivity lead to questions about jurisdiction. Is the crime civil or criminal? Does it violate local, state, national, or international law, or some combination of those? Which laws apply if a person commits a crime online—those in place where the suspect was physically located, those where the stolen or destroyed data was stored, or both?

Standards and procedures for the forensic examination of physical evidence, such as fingerprints or firearms, are far more broadly established than those for digital evidence. Requirements for warrants, the authority of agencies in charge of investigations, and standards for evidence's admissibility and presentation are shaped by the jurisdictions in which cases are presented. Investigators must be aware of these constraints until greater legal precedent establishes a clearer path for dealing with computer-related investigations.

Although computer forensics researchers and practitioners have made significant progress toward creating useful standards and practices for collecting and examining digital evidence, many challenges remain. The difficulty of extracting accurate, admissible evidence from computers and other devices will only increase as those with criminal intent adopt new methods. Forensic specialists will soon face a range of new issues, including:

- inevitable increases in criminals' technical skills,
- the widespread use of strong encryption,
- the potential use of Trusted Computing platforms[5] for illegal purposes,

- increased volume of data that can be stored on disk drives, and
- difficulties in examining data from legacy software applications or hardware that use proprietary formats, incompatible disk drives, or obsolete operating systems.

New tools and techniques have increased the reliability and speed with which investigators can conduct examinations, but new technologies will continue to challenge computer forensic specialists and researchers. We can only hope that increased awareness of digital evidence's importance in detecting and solving crime will lead to further research and development in this area. □

### References

1. M.G. Noblett, M.M. Pollitt, and L.A. Presley, "Recovering and Examining Computer Forensic Evidence," *Forensic Science Comm.*, vol. 2, no. 4, 2000; www.fbi.gov/hq/lab/fsc.
2. S.L. Garfinkel and A. Shelat, "Remembrance of Data Passed: A Study of Disk Sanitization Practices," *IEEE Security & Privacy*, vol. 1, no. 1, 2003, pp. 17–27.
3. E. Casey, "Practical Approaches to Recovering Encrypted Digital Evidence," *Int'l J. Digital Evidence*, vol. 1, no. 3, 2002; www.ijde.org/docs/02_fall_art4.pdf.
4. *Electronic Comm. Privacy Act (ECPA)*, *US Code*, title 18, part I, chapter 119, 1986.
5. R. Oppliger and R. Rytz, "Does Trusted Computing Remedy Computer Security Problems?" *IEEE Security & Privacy*, vol. 3, no. 2, Mar./Apr., 2005, pp. 16–19.

***William H. Allen*** *is an assistant professor of computer science at Florida Institute of Technology. His research interests include computer and network security, modeling and simulation of network-based attacks, and computer forensics. Allen received a PhD in computer science from the University of Central Florida. He is a member of the IEEE Computer Society, ACM, and Usenix. Contact him at wallen@fit.edu.*