Chapter 11
An Introduction to Computer Forensics

Chapter Outline

I. Evidence from cybercrimes rarely comes from examining the physical compartments of the computer.
II. Cases will normally require examining the various storage media inside the computer.
III. Once the computer is properly secured from the crime scene, it becomes time to begin preparing for the forensic examination of the computer.
IV. What is Computer Forensics?
    A. Forensic science has experienced growth in popularity in the last decade, likely as the results of several high-profile court cases and the growing number of prime time television dramas that focus on forensics.
    B. Forensic science—A term used to describe the application of scientific techniques to the investigation of criminal activities and the presentation of evidence at trial.
    C. Computer forensics—The application of computer science to the investigation of criminal activities involving computer-related technology
    D. Any activity on a computer or a computer network can potentially be tracked, and evidence that is deleted may still be recoverable.
        1. This new trail of electronic or digital evidence is termed the "digital trail."
        2. Computer forensics experts use their training with various software applications to recover evidence stored at various points along the digital trail.
        3. Recent advances in software technology have provided a recent growth in popularity for the field of specialization in the public sector of law enforcement and the private sector of large corporations.
            a. These private-sector companies have created an additional problem for law enforcement agencies, as officers may undergo the necessary training to investigate cybercrimes, only to later leave the agency and accept a higher-paying position with a private company.
    E. How Computers Store Data
        1. The internal storage device most commonly encountered in computer analysis is referred to as the hard drive or hard disk.
        2. When computers were first evolving, there was no internal storage device, and all information had to be stored on soft, flexible disks (known as floppy disks).
            a. The most common size of floppy disks was 5.25 inches.
            b. When starting up a computer—a process referred to as "booting up" the computer—it was required that an operating system be loaded into the computer's memory from one of these 5.25-inch floppy disks.
        3. The physical makeup of a disk consists of several tracks that run around the inside of the disk.
            a. Each of these tracks is made of smaller clusters that range in size

depending upon the operating system being used.

b. When a file is saved to one of these disks, the information is written to one of these clusters, and the entire cluster is used even if the cluster size is larger than the file size.

c. Later operating systems have worked to improve the storage capacity of these clusters, in an attempt to avoid "wasting" space.

4. When a file is saved to one of these clusters, the operating system will record the location of that file in the file table (a master record of all files stored on the computer).

a. Earlier computers used the File Allocation Table (FAT) whereby the location of the file is stored in the File Allocation Table.

b. Newer operating systems may use the New Technology File System (NTFS).

i. NTFS is very similar in function to FAT but also allows users to provide better security to the files that are stored in the Master File Table (MFT).

c. Many people believe that deleting a file results in the removal of the file from the disk.

i. Deleted files are stored in the recycle bin.

ii. The operating system merely takes note that the file has been selected for removal.

iii. When a person goes to the recycle bin and indicates that they wish to retrieve a deleted file, the operating system then takes note that the file is no longer deleted and can be accessed once again.

iv. When the recycle bin is emptied, the file may no longer be recovered through the Windows interface.

1. Once the file is deleted, the operating system will modify the file's name so that the operating system is aware that the space is no longer being used to store data.

2. The content will remain on the disk until the space is needed to save additional data.

3. The space where these files are stored, but are not being used, is referred to as unallocated space because the space is not set aside for use and is therefore unallocated.

a. This area of disk space could contain valuable evidence.

F. Internet Activity Stored on a Computer

1. Anytime a user logs onto the Internet and visits web pages on the World Wide Web, there remains a record of these pages on the computer.

2. When a user connects to the Internet and attempts to view a website, the computer will send a request to the server that stores the web pages associated with the website.

a. Upon approval of the user's request; the data is transferred to the user's computer.

i. To speed up access to web pages, web browsers began using what is known as cache memory.
    1. The cache folder was where the images from these websites were stored.
    2. Once the images were stored on the computer, then should the user attempt to return to a website that had been visited earlier, there would be no need to download the pictures (only new images and text would download).
ii. Cache folders is still used today to speed up a user's web browsing.
iii. Should a user realize that these images are being stored on the computer, then they may attempt to delete the folder's contents by using a command to instruct the operating system to empty the cache folder.
    1. The files may still be recoverable even if deleted.
    2. The file's name is changed to include the hexadecimal notation, and the operating system will consider the space where the image is stored to be unallocated space ready for future use.
    3. Until new data is stored, the user's viewing habits may be constructed.

b. A second, more controversial tool used to speed up web browsing is the use of cookies.
    i. Cookies—Small bits of data that websites use to recognize returning visitors.
    ii. Cookies are popularly used by commercial websites to track purchasing and searching habits and make recommendations of similar items for users.
    iii. Debates over the use of cookies have revolved around whether a company should have the ability to store and request information from users without their express permission.
        1. Newer versions of web browsers make the activation and deactivation of the cookies feature easier.

c. Internet activity is also stored on the computer when a web page is accessed.
    i. Data has to be stored on the computer for the user to see the image of the web page.
    ii. Upon shut-down, the data is deleted and the operating system recognizes that the space is no longer being actively used.
    iii. If a computer forensics analyst can get to the data before the space is used again, then the contents of the web page may be recovered and viewed.

1. Useful for recovering evidence from electronic communications that are not stored on either the computer or the Internet service provider's server.

G. The Computer Forensics Process

1. The first step in the computer forensics process will normally involve the duplication of the suspect's hard drive or other digital storage media.
   a. The original storage media should only be examined when it is impractical or impossible to use a duplicated copy of the media.
   b. Using the original computer may result in this potential evidence being damaged or destroyed.
   c. Another important issue involves the date-and-time–stamping features of most operating systems because each time a file is accessed the date-and-time stamp is modified.
   d. Ways duplication of storage media takes place.
      i. The first involves the disk duplication features built into some of the more powerful operating systems.
         1. Linux—Uses a serious of text-based commands.
            a. Through use of the Linux commands an analyst can make a bit copy of the storage media, meaning that all files are copied—both those that are active and those that are not active.
      ii. The second involves the use of commercial software programs.
         1. Programs: Safeback, FTK Imager by AccessData, and EnCase.
         2. Programs operated by analyzing the target storage media and transferring the data to a target disk in a process called imaging.
      iii. Hard disks can also be used to store the images of the suspect's media.
   e. The forensic imaging of a suspect's media may take place at the scene of the alleged crime or may take place in a controlled laboratory setting.
      i. Portable computer forensics kits allow analysts to image suspect computers and storage media at the scene of a search warrant's execution.
         a. The kit contains a variety of connection cables that allow for the imaging of various storage media that may be encountered by investigators.
         b. Analysts will connect the suspect's media to the portable kit and create an image that can be examined then or at a later time in the forensics lab.
         c. This kits can be useful in cases involving

multiple computers or computers used for legitimate business purposes to allow the investigator to determine whether or not to seize the computers.
                i. Not seizing the computer can reduce financial harms for businesses.
            2. Some forensics kits and programs can consist of network connections to image a suspect's computer.
            3. It is still recommended that, whenever possible, the analyst or investigator assigned to the investigation seize the computer to ensure that the best evidence is available should a criminal trial become necessary.

2. The second phase of the process involves verifying the files and file signatures.
    a. This is necessary to ensure that claims of planted or manipulated evidence are minimized.
    b. One method of verifying the file is to generate a hash value for the hard disk before beginning any analysis of the disk.
        i. A hash value is basically a 32-character value that is representative of a disk's contents.
        ii. If any value on the hard disk is modified, then the hash value will be drastically altered.
        iii. The MD5 algorithm is the most commonly accepted method of generating hash values; it works by:
            1. Encrypting the original disk or file and generating a hash value.
            2. Generating new hash values when any changes to the file occur.
    c. The next step is to verify the file signature.
        i. Whenever a file is created or saved, the file's header will contain a signature that informs the operating system what type of file is being created or saved.
            1. When the file is opened, the header will inform the operating system what software program is needed to open the file.
        ii. Once the file is saved, then the operating system will identify the file through the file's name extension (.doc for Microsoft Word).
            1. When the operating system sees this extension—.doc—an icon of the document with the "M" is placed next to the filename and becomes visible to users in the Windows Explorer window.
            2. Changing the file's icon association is no more complex than changing the file extension of the file
                i. This can be used by anyone wanting

to mislead the operating system into behaving as if a file is associated with a different program; for instance, one could attempt to misrepresent child pornography images as documents.

ii. Investigators may wish to consider this when drafting search warrants for digital evidence, as there is the potential for a suspect to claim that a search warrant for child pornography images could not include a search of document files.

d. File signature analysis involves the comparing of actual file signatures with the header of the file to determine whether there are any discrepancies.

i. Notations should be made if the header indicates that the file is an image and the file signature indicates that the file is a document.

ii. The forensic analyst will then physically locate each of the files and examine the files to determine what data the file in question contains.

iii. Newer computer forensic software packages include a feature that will provide a listing of these files and their locations.

3. The third phase in the process is the actual examination of the storage media (the forensic analysis).

a. Various computer forensic software programs come with a variety of tools and features that will examine the entire hard disk for images and document files, search the storage media for files that contain keywords, and search the unallocated space on the disk for data that has been deleted but not yet overwritten.

b. Many consider the forensic analysis to be the most tedious and unexciting part of the high-technology crime investigation, but it is one of the most important stages in the process.

i. The size of the storage media being examined will determine the amount of time necessary to conduct a full analysis.

ii. Another factor that will influence the amount of time it takes to conduct an examination is the type of evidence the forensic analyst is searching for.

1. For image-based evidence, one should search the entire drive's contents for the necessary images.

2. For text-based evidence, one should search the entire drive and can use the keyword search.

a. A keyword search will allow the analyst to

search the entire hard disk for certain words or phrases.

4. The fourth phase is the completion of the forensics report.
    a. Forensics report—A written report concerning the evidence that was uncovered during the examination.
    b. Some of the computer forensic software packages include a report feature that will allow the forensic analyst to save images, documents, and text segments during the actual forensic analysis.
        i. A narrative can be included to describe where the actual image was discovered.
    c. The final reports can then be formatted to provide an easy-to-read document including all evidence recovered throughout the investigation and analysis that can be used in the following ways:
        i. By prosecutors in decided whether to pursue charges.
        ii. By prosecutors and defense in the discussion of a plea bargain.
        iii. At trial in the prosecutor's case.
        iv. By the forensic analyst to refresh his or her memory before giving testimony in court.

V. The Computer Forensic Software Packages
    A. Two categories of forensic software packages:
        1. Graphical User Interface (GUI) —A user-friendly software with which a user can click on an icon to perform a function.
            a. Best choice for law enforcement agencies that do not have full-time computer scientists on staff to conduct forensic analyses.
            b. EnCase by Guidance Software is considered one of the best GUIs because of the software's various component programs.
                i. Users may preview the suspect disk to conduct any analysis that could be undertaken if an image of the disk was being examined.
                    1. Previewing can save hours of analysis of disks that do not contain evidence, because an investigator can then concentrate on imaging only those disks that are important to the case.
                ii. Users can image a disk through the use of a network patch cable or a serial cable.
                    1. Thus, the forensic analyst does not have to remove the physical disk from the computer.
                    2. The EnCase software allows users to create a boot disk that will prevent any data from being written to a suspect's disk during the computer's start-up process.
                        a. When the computer is running, a disk image can be created.
                        b. When the image is created, the EnCase software then allows the analyst to search

the hard drive by:

i. Examining the image files located on the hard drive through a gallery view.
ii. Examining the files through the use of a hex view (reading the hexadecimal file components).
iii. Searching the entire disk for keywords.

3. EnCase also has a reporting feature that allows a forensic analyst to save keyword hits, images and personal comments into an easy-to-format report that can be printed or e-mailed to attorneys involved in the case.

c. Forensic Took Kit by AccessData contains many of the same features as the EnCase software.
   i. Forensic Tool Kit (FTK) allows users to examine an imaged disk by:
      1. Examining image files through a gallery view.
      2. Examining files through a hexadecimal view.
      3. Searching the disk for keywords.
   ii. FTK contains an e-mail feature that will automatically search out e-mails stored on the disk and provide the information in an easy-to-read format.
   iii. It also has the ability to import image files that are created using a wide variety of imaging software formats and allows for users to import EnCase image files.
   iv. The FTK forensic software is the password-cracking component, Password Recovery Tool Kit (PRTK), that allows users to gain access to encrypted files

2. Non-Graphical User Interface-based Software Utilities
   a. Maresware has a suite of command line utilities available for use by computer forensics analysts.
      i. Advantage is cost.
      ii. Disadvantage is the complexity of their use; it requires an understanding of the DOS command line.
      iii. The Maresware utility suite contains, but is not limited to, t the following applications:
         1. DISCK_CRC—A hashing utility that will generate an MD5 value for a particular file or collection of files.
         2. HEX_SECT—A hex editor tool that allows the user to examine the hexadecimal and text values location within a suspect disk.
         3. STRSRCH—A keyword search program that will search a suspect disk for keywords and then provide

the user with information relating to how many times the keyword appears and where on the disk the information is located.

iv. The software is not often sold and in fact may be phased out.

3. Another suite of command lines tools is manufactured by NTI.
   a. Some of the utilities available are:
      i. CRCMD5—A utility used to calculate a cyclical redundancy check value or an MD5 hash value that can be used to verify that there have been no changes made to a file or group of files during an examination.
      ii. DiskSearch 32—A keyword search program designed to operate on Microsoft DOS systems.
      iii. Filter Intelligence and SafeBack—Filter Intelligence allows users to locate names, phrases, and word groups.
   b. According to NTI, this program is useful when handling investigations that involve e-mail addresses or electronic communications.

VI. Admissibility of Digital Evidence
   A. Once the forensics analysis has been completed, there is the possibility that the digital evidence may be needed at trial.
   B. At this stage there are several issues that investigators must consider:
      1. The legal standards associated with admitting digital evidence into a criminal trial.
         a. Two landmark decisions that must be considered.
            i. *Daubert v. Merrell Dow Pharmaceuticals* (1993) —Involves the admission of scientific evidence into the federal trial system.
               1. A civil case opinion that has been extended into the criminal court arena.
               2. The court ruled that scientific evidence was required only to be reliable and scientifically valid, and there was no requirement for scientific evidence to be generally accepted before it could be entered into evidence under Federal Rules of Evidence Rule 702.
            ii. *Frye v. United States* (1923) —Guides the admission of scientific evidence into state trial court proceedings.
               1. The court ruled that for scientific evidence to be entered into trial, it must be shown that the process of gathering and analyzing the evidence has crossed the line from experimental process to demonstrative process.
               2. The court decided that it was best to consider only the admission of scientific evidence when the evidence is generally accepted in the scientific community from which it is derived.

b. *State v. Hayden* (1998)—Involved the admission of digital evidence.
    i. The court ruled that digital enhancements, when conducted by trained experts who use appropriate software, were admissible.
c. *United States v. Scott-Emuakpor* (2000) —The court ruled that training in computer forensics was sufficient to warrant qualification as an expert witness.
d. *Williford v. State* (2004) —The court determined that EnCase satisfied the requirements for admission of scientific evidence.
    i. The general consensus being that digital evidence obtained from computer forensic examinations will be considered admissible so long as the evidence is obtained through the use of an authenticated computer forensics software package and the user has been trained in the standard methodology of computer forensics analysis.

2. The integrity of the digital evidence associated with the criminal charges against a defendant.
    a. Even if digital evidence in general is deemed to be admissible under the criteria necessary for scientific evidence, there is still the need to authenticate the digital evidence and establish the chain of custody for such evidence.
    b. To ensure a more thorough authentication process, some experts have recommended that analysts generate a hash value for the suspect's storage media before making an image of the media; then, after the image is made, another hash value should be generated from the storage media.
        i. The two hash values can be compared to show that no changes have been made to the suspect's storage media during the imaging process.
        ii. After examining the storage media, a third hash value could be generated to show that no changes were made to the suspect's media during the collection of evidence.
    c. The second consideration involves establishing the chain of custody for the digital evidence.
        i. "Chain of custody" is a legal term for the grouping of people who have handled evidence in any manner during an investigation, whether it is to examine the evidence or just to move the evidence to and from a given location.
        ii. When evidence is admitted at trial, the chain of custody is used to prove that the evidence presented at trial is the same evidence that was seized from the crime scene and that it was not modified.
        iii. If multiple individuals touch digital evidence, then there are multiple opportunities for the evidence to be dismissed.
        iv. Professionals in the field of computer forensics have recommended several techniques to assist in preventing

problems with the chain of custody.

1. The use of a professional labeling process for digital evidence.
2. The use of a special evidence transaction log.
    a. Using any spreadsheet program, a log can be created that includes the following information:
        i. Evidence Inventory Number
        ii. Date and Time
        iii. Location Remove and Taken
        iv. Reason the Evidence is Being Removed
        v. Who Removed the Evidence

VII. Conclusion

Key Terms

Bit copy: A copy of active and non-active files of a storage media.

Book disk: This prevents any data from being written to a suspect's disk during the computer's start-up process.

Booting: Starting up a computer.

Cache folder: A folder where images from websites are stored so that if a user accesses the website again, there would be no need to download the pictures.

Chain of custody: A legal term for the grouping of people who have handled evidence in any manner during an investigation.

Cookies: Small bits of data that websites use to recognize returning visitors.

Computer forensics: The application of computer science to the investigation of criminal activities involving computer-related technology.

Digital trail: The trail of electronic or digital evidence of any activity on a computer or a computer network, including even those that were deleted.

File signature: A signature that informs the operating system what software program is being created or saved.
File signature analysis: The comparing of actual file signatures with the header of the file to determine whether there are any discrepancies.

File table: The master record of all files stored on the computer.

Floppy disk: A soft, flexible disk that was used to store information on a computer.

Forensic science: The application of scientific techniques to the investigation of criminal activities and the presentation of evidence at trial.

Forensics report: A written report concerning the evidence that was uncovered during the examination.

Graphical User Interface (GUI): A computer forensic software package that allows a user to click on an icon to perform a function.

Hard drive or hard disk: The internal storage device on a computer.

Hash value: A 32-character value that is representative of a disk's content.

Imaging: The process of analyzing target storage media and transferring data to a target disk.

Keyword search: A search technique that allows the analyst to search the entire hard disk for certain words or phrases.

Non-Graphical User Interface (Non-GUI): A computer forensic software package based on command line utilities.

Portable computer forensic kits: Kits that allow analysts to image suspect computers and storage media at the scene of a search warrant's execution.

Recycle bin: Place on hard drive where files are moved when the delete command is given.

Slack space: The difference between the hard drive space needed to store the deleted original file and the new file saved.

Unallocated space: The hard drive space where files deleted from the recycle bin are stored, but are not being used.