7 Understanding Computer Forensics

Learning Objectives

After reading this chapter, you will able to:

- Understand the fundamental concepts in cyberforensics.
- Understand the meaning of the term "cyberforensics and the need for cyberforensics."
- Learn what "digital evidence" means along with the base term "forensics science."
- Get an overview of cardinal rules of computer forensics.
- Learn how cyberforensics is used in cybercrime investigations.
- Understand the legal requirements for cyberforensics and compliance aspects of cyberforensics.
- Get an overview of the role of forensics experts.
- Understand the "data privacy issues" involved in cyberforensics.
- Learn about forensic auditing.
- Learn about cyberforensic tool available in the market.
- Understand the challenges faced in cyberforensics.

7.1 Introduction

The purpose of this chapter is to address the other side of crime, that is, use of forensic techniques in the investigation of cybercrimes. "Cyberforensics" is a very large domain and addressing it in a single chapter is indeed a challenge. Complex technical aspects involved in digital forensics/computer forensics are not possible to cover in a single chapter. Therefore, this chapter is aimed at only providing a broad understanding about cyberforensics.

The term "chain of custody" has a recurring mention in this chapter because it is a central concept in forensics. We have provided a large number of information resources including several video clips on digital forensics investigation (some demonstrations too), reviews of forensics tools as well as interviews with experts. We recommend readers to visit those links in Video Clips, Further Reading. The discussion in this chapter will serve as background for Chapter 8 where hand-held forensics is addressed. The terms "cyberforensics," "digital forensics" and "computer forensics" are used interchangeably. Definitions of these terms are provided.

Cyberforensics plays a key role in investigation of cybercrime. "Evidence" in the case of "cyberoffenses" is extremely important from legal perspective. There are legal aspects involved in the investigation as well as handling of the digital forensics evidence. Only the technically trained and experienced experts should be involved in the forensics activities.

The requirements for setting up a digital forensics laboratory are explained in Section 7.11. Toward the end, a ready reckoner of cyberforensics tools is provided in a tabular form for readers' convenience (Tables 7.9, 7.10 and 7.11). Considering the widespread use of hand-held devices [personal digital assistants (PDAs), mobile phones and all its varieties as well as the iPods, etc.], we have addressed the forensics of hand-held devices in the next chapter. Some special topics such as "use of data mining in cyberforensics," "forensics auditing" and "antiforensics" are also discussed in this chapter. Case studies on digital forensics investigations are presented in Chapter 11 (in CD). With this background, let us proceed to understand the historical background.

7.2 Historical Background of Cyberforensics

The different types of cybercrimes are explained in Chapter 1. Computer is either the subject or the object of cybercrimes or is used as a tool to commit a cybercrime. The earliest recorded computer crimes occurred in 1969 and 1970 when student protestors burned computers at various universities. Around the same time, people were discovering methods for gaining unauthorized access to large-time shared computers. Computer intrusion and fraud committed with the help of computers were the first crimes to be widely recognized as a new type of crime.

The Florida Computer Crimes Act was the first computer crime law to address computer fraud and intrusion. It was enacted in Florida in 1978. [Mentioned in Chapter 6, (Box 6.6 in Chapter 6).]

The application of computer for investigating computer-based crime has led to development of a new field called *computer forensics*. Sometimes, computer forensics is also referred to as "digital forensics." Computer forensics/digital forensics has existed for as long as people have stored data inside computers.

"Forensics evidence" is important in the investigation of cybercrimes.

Discussion on the legal side of cybercrime (see Chapter 6) serves as a link to this chapter through the term "evidence," "digital evidence" in particular. Basically, computer forensics experts need digital evidence in cases involving data acquisition, preservation, recovery, analysis and reporting, intellectual property theft, computer misuse (recall the discussion in Chapter 6 about the Indian IT Act – Tables 6.6, 6.7 and 6.8), corporate policy violation, mobile device (PDA, cell phone) data acquisition and analysis, malicious software/application, system intrusion and compromise, encrypted, deleted and hidden files recovery, pornography, confidential information leakage, etc.

Computer forensics is still a relatively new discipline in the domain of computer security. It is a rapidly growing discipline and a fast growing profession as well as business. The focus of computer forensics is to find out digital evidence – such evidence required to establish whether or not a fraud or a crime has been conducted. There is a difference between computer security and computer forensics. Although "computer forensics" is often associated with "computer security," the two are different.

Computer forensics is primarily concerned with the systematic "identification," "acquisition," "preservation" and "analysis" of digital evidence, typically after an unauthorized access to computer or unauthorized use of computer has taken place; while the main focus of "computer security" is the prevention of unauthorized access to computer systems as well as maintaining "confidentiality," "integrity" and "availability" of computer systems.

Information security aspects are explained in detail in Ref. #14, Books, Further Reading. Thus, the goal of computer forensics is to perform a structured investigation on a digital system. For those who are reading this chapter directly before visiting Chapter 1, computer crime is any criminal offense, activity or issue that involves computers.

There are two categories of computer crime: one is the criminal activity that involves using a computer to commit a crime, and the other is a criminal activity that has a computer as a target.

Information security experts consider "cyberlaw compliance" as one of the many aspects of "techno-legal information security." They advise organizations to formulate an appropriate plan of action to comply with cyberlaws as a part of the IS practice. This association of cyberlaw into the information security domain has gained additional importance due to some amendments that have been made to ITA 2000. Typical types of data requested for a digital forensics examination by the law enforcement agencies include: investigation into electronic mail (E-Mail) usage, website history, cell phone usage, cellular and Voice over Internet Protocol (VoIP) phone usage, file activity history, file creation or deletion, chat history, account login/logout records and more. Therefore, it becomes necessary to address the legal issues involved in cyberforensics. This is addressed in Section 7.16.2. Tables 7.9, 7.10 and 7.11 provide the list of forensic tools available in the market. URLs are also provided in Refs. #15, #16, #17 and #18, Additional Useful Web References, Further Reading with information about various laws/statutes pertaining to cybercrime. This information would be particularly useful for cyberlaw students.

Forensics means a "characteristic of evidence" that satisfies its suitability for admission as fact and its ability to persuade based upon proof (or high statistical confidence level).

In precise terms, "forensics science" is the application of science to law and it is ultimately defined by use in court. Forensics science is the application of physical sciences to law in search for truth in civil, criminal and social behavioral matters to the end that injustice shall not be done to any member of society. An alternative definition for digital forensics science is:

> *the use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation and presentation of digital evidence derived from digital sources for the purpose of facilitating or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorized actions shown to be disruptive to planned operations.*

The goal of digital forensics is to determine the "evidential value" of crime scene and related evidence.

The roles and contributions of the digital forensics/computer forensics experts are almost parallel to those involved as forensics scientists in other crimes, namely, analysis of evidence, provision of expert testimony, furnishing training in the proper recognition, and collection and preservation of the evidence. Now, let us understand the term "digital forensics science."

7.3 Digital Forensics Science

Digital forensics is the application of analyses techniques to the reliable and unbiased collection, analysis, interpretation and presentation of digital evidence. There is a number of slightly varying definitions. The term *computer forensics*, however, is generally considered to be related to the use of analytical and investigative techniques to identify, collect, examine and preserve evidence/information which is *magnetically stored or encoded*. The objective of "cyberforensics" is to provide digital evidence of a specific or general activity. Following are two more definitions worth considering:

1. **Computer forensics:** It is the *lawful and ethical seizure, acquisition, analysis, reporting and safeguarding of data and metadata derived from digital devices which may contain information* that is notable and perhaps of evidentiary value to the trier of fact in managerial, administrative, civil and criminal investigations. In other words, it is the collection of techniques and tools used to find evidence in a computer.
2. **Digital forensics:** It is the use of *scientifically derived and proven methods* toward the *preservation, collection, validation, identification, analysis, interpretation, documentation and presentation of digital evidence* derived from digital sources for the purpose of facilitation or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorized actions shown to be disruptive to planned operations.

Box 7.1 COFEE Time!

Computer Online Forensics Evidence Extractor (COFEE) is a USB thumb-drive gadget onto which Microsoft have loaded 150+ "commands" that can, among other things, decrypt passwords, display Internet activity and uncover all data stored on the computer. This interestingly named tool was developed by Anthony Fung, a former Hong Kong police officer now working as a senior investigator on Microsoft's Internet Safety Enforcement Team. Microsoft's intent behind creation of COFEE has been exclusively for use by law enforcement agencies. It is believed that while creating the design of this tool, Microsoft perhaps did not take into consideration the feedback about problems faced by law enforcement agencies worldwide. During their ongoing fights against a variety of cybercrimes, law enforcement agencies around the world face some common challenges.

Box 7.1 COFEE . . . (*Continued*)

Law enforcement professional need to capture critical evidence on a computer at the *scene of an investigation* before the evidence is powered down and removed for forensics analysis. Digital evidences, when "live" (e.g., active system processes and network data), must be handled with care because the "live" evidence is volatile. There is always the risk of "volatile evidence" getting lost when the computer is turned off. Therefore, the challenge is how does an officer on the scene effectively do this if he/she is not a trained computer forensics expert?

COFEE helps the law enforcement agencies even when there are no on-the-scene computer forensics capabilities. It enables them to collect live "volatile evidence" more easily, reliably and cost-effectively. Even a law enforcement officer with minimal computer experience, once he/she is taken through the tool tutorial, can use a preconfigured COFEE device. The officer can take advantage of the same common digital forensics tools as used by experts to gather volatile evidence that can prove critical for the investigation. The officer can undertake investigation tasks by simply inserting a USB device into the computer.

On-the-scene, agents can run more than 150 commands on a live computer system. COFEE tool also provides reports in simple format that are easy for later interpretation. These reports can be used by experts and can also be used as supportive evidence for subsequent investigation and prosecution. The COFEE tool and its underlying framework can be tailored to effectively meet the needs of a particular investigation, that is, it can be fully customized. On the lighter side, one wonders if there will also be Total Evidence Analyzer (TEA) soon.

More information on the COFEE[1] tool can be obtained by visiting the links provided in References. We have provided a link in Ref. #2, Video Clips, Further Reading, where a computer forensics expert explains how digital evidence is seized as part of forensics investigation.

It is difficult to provide a precise definition of "digital evidence" because the evidence is recovered from devices that are not traditionally considered to be computers. Some researchers prefer to expand the definition by including the "collection" and "examination" of all forms of digital data, including the data found in cell phones, PDAs, iPods and other electronic devices. In general, the role of digital forensics is to:

1. Uncover and document evidence and leads.
2. Corroborate evidence discovered in other ways (E-Discovery – see Box 7.3).
3. Assist in showing a pattern of events (data mining has an application here).
4. Connect attack and victim computers (Locard's Exchange Principle – see Box 7.5).
5. Reveal an end-to-end path of events leading to a compromise attempt, successful or not.
6. Extract data that may be hidden, deleted or otherwise not directly available.

The typical scenarios involved are:

1. Employee Internet abuse – more about this is mentioned in Chapter 9;
2. data leak/data breach – unauthorized disclosure of corporate information and data (accidental and intentional);
3. industrial espionage (corporate "spying" activities);
4. damage assessment (following an incident);
5. criminal fraud and deception cases;
6. criminal cases (many criminals simply store information on computers, intentionally or unwittingly) and countless others;
7. copyright violation – more about this is mentioned in Chapter 10 (in CD).

Figure 7.1 shows the kind of data you "see" using forensics tools.

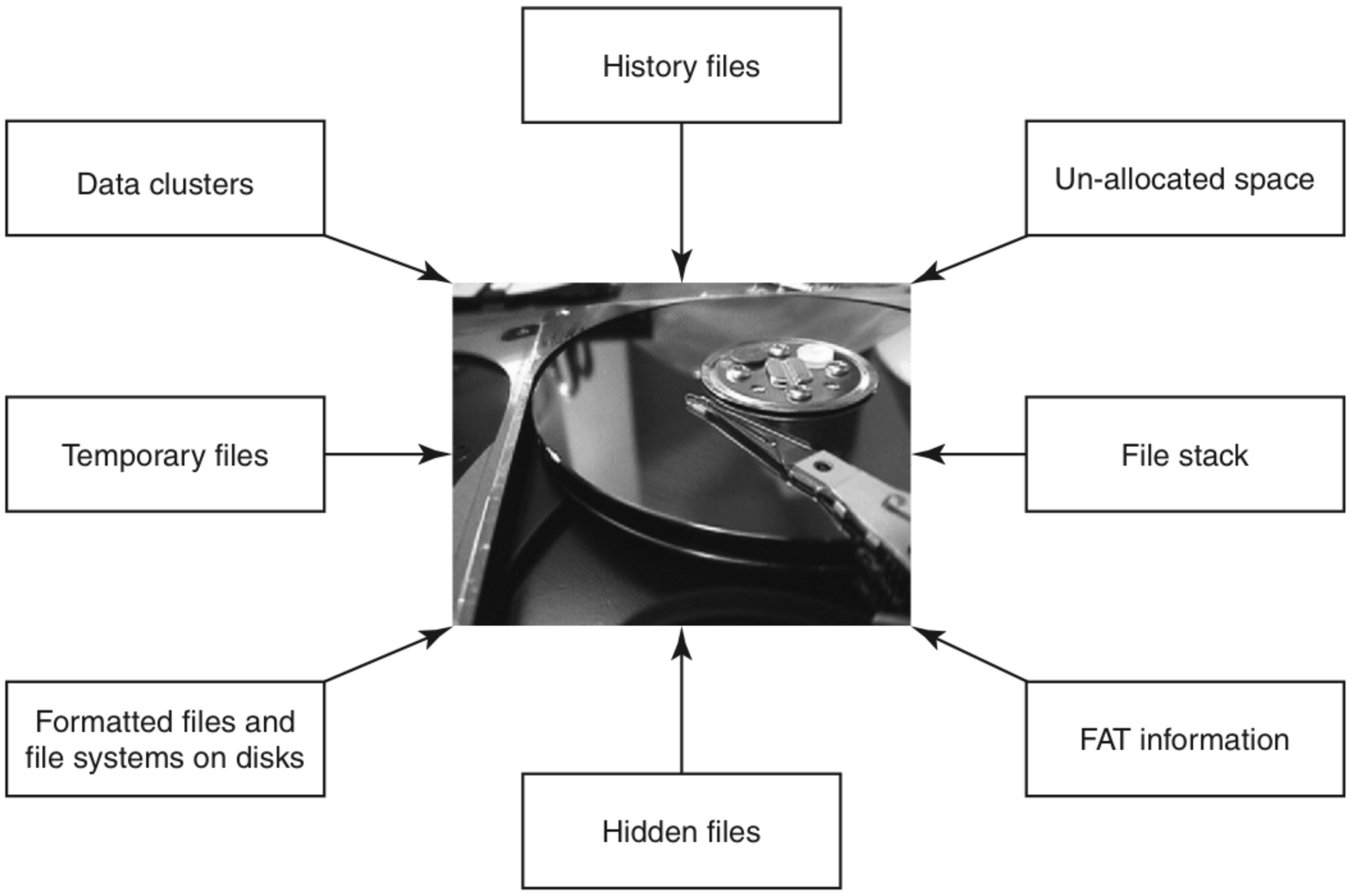


Figure 7.1 | Data seen using forensics tools. FAT means file allocation table.

Using digital forensics techniques, one can:

1. Corroborate and clarify evidence otherwise discovered.
2. Generate investigative leads for follow-up and verification in other ways.
3. Provide help to verify an intrusion hypothesis.
4. Eliminate incorrect assumptions.

Box 7.2 Differences between Forensics Policy and Security Policy

Often people get confused between "forensics policy" and "security policy." They think that the meaning of two terms is the same; however, this is not true. *Security policy* is a statement that clearly specifies the allowed and disallowed elements with regard to security. It partitions the system states into "secure" and "unauthorized" security policy that helps implement mechanisms to enforce system security policy. On the other hand, *forensics policy* is a statement that clearly states which assets are forensically important. It also specifies data needed for investigation into breach of those assets.

Forensics policy partitions space of all possible breaches or criminal activity into sets of events that are forensically noteworthy and those that are not. It allows for mechanisms or design decisions to enforce the policy. Here is another way to understand the *difference between security policy and forensics policy* – violation of security policy leads to insecure information systems/application with vulnerabilities arising due to consequences of break-in or insider misuse. On the other hand, violation of forensics policy means lack of evidence which results in the loss of ability of an organization to prove guilty the people who are involved in cybercrime incidence.

The "goals" defined by forensics policies are different than those defined by security policies. Goals of forensics policies deal with assets, data and possible storage issues. They capture digital evidence; therefore, forensics integrity of data is preserved. They capture enough data to ensure

Box 7.2 Differences between . . . (*Continued*)

that prosecution is possible. Forensics policy goals specify events that must be handled and data that must be preserved. Events not included in the policy will not need associated data. Here is an *example of forensics policy*:

1. Goal is to capture data from network intrusions for possible prosecution.
2. This (forensics) policy states that all events identified as intrusions will have their associated data captured and preserved.
3. Enforcement mechanisms: routine preservation of IDS, firewall, router and web server logs for some configurable length of time.

Here is another example of organizational level operating rules derived from an organization's forensics policy:

1. All access to Oracle DB must be monitored.
2. Access logs and administration logs to Oracle DB will be preserved for no less than 1 year.
3. Access and activity to web server is monitored.
4. Apache web server logs will be preserved for 1 year/6 months.
5. Firewall and Snort logs will be preserved for 1 year.
6. Router logs will be preserved for 6 months.
7. Network will be tested every 6 months for congestion situation by overloading it until it begins to drop traffic.
8. Network capacity will be increased before traffic hits the level where packets will be dropped.

7.4 The Need for Computer Forensics

The convergence of Information and Communications Technology (ICT) advances and the pervasive use of computers worldwide together have brought about many advantages to mankind. At the same time, this tremendously high technical capacity of modern computers/computing devices provides avenues for misuse as well as opportunities for committing crime. This has lead to new risks for computer users and also increased opportunities for social harm. The users, businesses and organizations worldwide have to live with a constant threat from hackers who use a variety of techniques and tools to break into computer systems, steal information, change data and cause havoc. The topic of "threats to information systems" is thoroughly discussed in Ref. #12, Books, Further Reading. The widespread use of computer forensics is the result of two factors: the increasing dependence of law enforcement on digital evidence and the ubiquity of computers that followed from the microcomputer revolution.

Box 7.3 Digital Forensics Investigations and E-Discovery

Digital evidence plays a crucial role in the threat management life cycle, from incident response to high-stakes corporate litigation. Forensics discoveries provide the ability to search and analyze various pieces of potential evidence of electronic nature. Evidence can involve computer hard drives, portable storage, floppy diskettes, portable music players and PDAs, just to name a few.

All forms of evidence are verified and duplicated prior to investigation to ensure the *integrity of the evidence* for litigation purposes if needed. Managers who are responsible for litigation tend to take help from forensics professionals to solve a growing range of evidentiary and investigative challenges.

Key evidence often resides on more than a user hard drive or file server, requiring the capture and analysis of evidence from enterprise productivity servers, network logs or proprietary databases.

Box 7.3 Digital Forensics . . . (*Continued*)

Many threats arise from illegal Internet activities that extend beyond the firewall and require new investigative and forensics approaches. Users are becoming more sophisticated and so are their efforts to circumvent security policies or encrypt, delete or destroy digital evidence. Forensics professionals need supporting solution for the acquisition, management and analysis of digital evidence. Such computer forensics services include the following:

1. Data culling and targeting;
2. discovery/subpoena process;
3. production of evidence;
4. expert affidavit support;
5. criminal/civil testimony;
6. cell phone forensics;
7. PDA forensics.

Specific client requests for forensics evidence extracting solution support include:

1. Index of files on hard drive;
2. index of recovered files;
3. MS Office/user generated document extraction;
4. unique E-Mail address extraction;
5. Internet activity/history;
6. storage of forensics image for 1 year (additional charges then apply);
7. keywords search;
8. chain of custody (see Section 7.8, Figs. 7.10 and 7.11, Boxes 7.4 and 7.12);
9. mail indexing;
10. deleted file/folder recovery;
11. office document recovery;
12. metadata indexing;
13. conversion to PDF;
14. log extraction;
15. instant messaging history recovery;
16. password recovery;
17. format for forensics extracts (DVD, CD, HDD, other);
18. network acquisitions.

Such types of computer evidences are important because quite often the evidence becomes the deciding factor in a criminal, civil or employee dismissal action. Investigations involving *trade secrets*, commercial disputes, and misdemeanor and felony crimes can be won or lost solely with the introduction of recovered E-Mail and other electronic documentation. If someone makes an attempt to delete, erase or otherwise hide critical evidence, you need the competent data recovery capabilities of forensics discoveries. Evidence that may not be known to attorneys may exist and often can be found during the forensics process. Also, timelines of computer usage is of help in crafting deposition questions and in targeting witnesses for interview.

Computer users typically "delete" incriminating and/or sensitive computer files (e.g., using tools such as "Deep Freeze," a software tool that is actually meant to protect your computer) but the information may still exist in slack space on the computer's hard drive that is hidden (do see the list of links provided at the end of this box). This computer data may linger for months or even years. However, it can be recovered and documented using computer forensics methods and techniques. Unfortunately, there are many examples of computer usage in violation of company policy. Sexual harassment, embezzlement, theft of trade secrets, abuse of the Internet and unauthorized outside employment on company time are just a few examples of violation that warrant a forensics examination of a computer. Even in investigations where hard drives are reformatted in an attempt to hide evidence, forensics discoveries can still potentially recover critical information (do see the list of links provided at the end of this box). Forensic discoveries can also aid in recovering passwords for critical files that have been maliciously set or changed.

Box 7.3 Digital Forensics . . . (*Continued*)

There are further challenges; for example, many times, computers are reissued when employees leave. Computer that is used continuously may destroy the incriminating evidence that can be used against a former disgruntled employee. Also, constant use of the computer may raise questions as to who created the incriminating evidence and when. To prevent these problems and to preserve potentially valuable information, it is recommended that a strict chain of custody should be followed and the subject computer should be shutdown, that is, the computer on which digital evidence is believed to be residing.

Useful links – The following link has the video where a digital forensics expert explains about E-Discovery and other aspects showing usefulness of digital forensics:
http://www.youtube.com/watch?v=y_BLtefQv40 (27 February 2010).

The following links (accessed on 28 March 2010) provide information about various tools including Deep Freeze:

1. http://software.informer.com/getfree-deep-format-recover/ (Deep Format Recover Tools);
2. http://www.astahost.com/info.php/Deep-Freeze-Partition_t2571.html (Deep Freeze-related Blog);
3. http://www.hochstadt.com/protecting-your-computer-using-deep-freeze (an article here explains how you can protect your computer using "Deep Freeze");
4. http://technodata.blogspot.com/2006/11/how-to-format-hard-disk-by-disk.html (this article explains how to format the hard disk);
5. http://www.softlist.net/search/deep-freeze-2000-xp/ (Deep Freeze 200 XP Free Downloads);
6. http://www.pctechguide.com/forums/ubbthreads.php/topics/4391/Hard%20Disk%20re-format (technical blog);
7. http://www.bluescreengone.com/comparison.htm (Table showing Comparison of various Data Recovery and Data Formatting Tools);
8. http://www.soft82.com/free/remote-yahoo-password-stealer/ (Yahoo Password Recovery Tools and many similar utilities);
9. http://www.soft82.com/free/free-youtube-downloader-mp3/ (link to many free downloads of useful utilities).

The media, on which clues related to cybercrime reside, would vary from case to case. There are many challenges for the forensics investigator because storage devices are getting miniaturized due to advances in electronic technology; for example, external storage devices such as mini hard disks (pen drives) are available in amazing shapes (Fig. 7.2).

Looking for digital forensics evidence (DFE) is like looking for a needle in the haystack. Here is a way to illustrate why there is always the need for forensics software on suspect media – the capacity of a typical regular hard disk is 500 GB (gigabytes). In an A4 size page, there are approximately 4,160 bytes (52 lines × 80 Characters = 4,160 bytes assuming 1 byte per character). This is equivalent to 4 KB (kilobytes). An A4 size of paper sheet has thickness of 0.004 inches. Data of 4 MB (megabyte; 1,000 times of 4 KB) when printed on A4 size of paper would be 4 inches thick. Data of 4 GB if printed on A4 sheet would be 4,000 inches, that is, 1,000 times of 4 MB. This would turn out to be 4 inches thick. The printout of 500 GB would be 500,000 inches! It would be virtually impossible to "retrieve" relevant forensics data from this heap!! There comes the help from forensics software – it helps sieve relevant data from the irrelevant mass (vital few from trivial many as the proverb goes).

The term "chain of custody" is important (see Box 7.4).

Chain of custody means the chronological documentation trail, etc. that indicates the seizure, custody, control, transfer, analysis and disposition of evidence, physical or electronic.

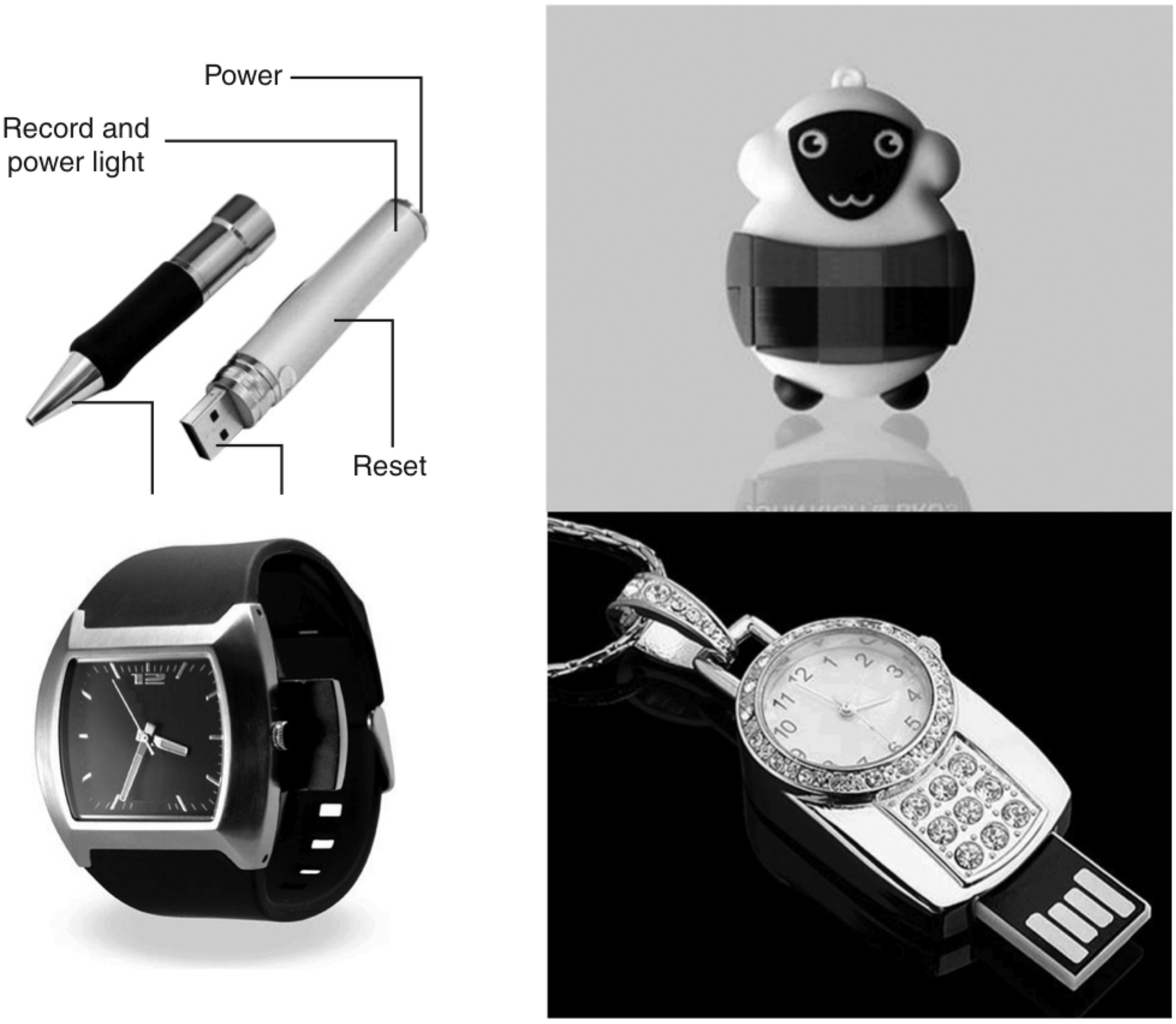


Figure 7.2 Hidden and miniaturized storage media.
Sources: http://www.ghdigital.com; http://www.technology-guide.co.uk; http://designyoutrust.com; http://gadgethobby.com

Box 7.4 Chain of Custody Example

The basic idea behind ensuring "chain of custody" is to ensure that the "evidence" is NOT tampered with. The recovery of a "crime weapon" at the murder scene would be an example of "chain of custody." This is explained below.

Case Study

Officer Amar collects the knife and places it into a container, then gives it to forensics technician Balan. Forensics technician Balan takes the knife to the laboratory and collects fingerprints and other evidence from the knife. He then gives the knife and all evidence gathered from the knife to evidence clerk Charu. Charu then stores the evidence until it is needed, documenting everyone who has accessed the original evidence (the knife and original copies of the lifted fingerprints).

The chain of custody requires that from the moment the evidence is collected, every transfer of evidence from one person to another person should be documented as it helps to prove that nobody else could have accessed that evidence. It is advisable to keep the number of evidence transfers as low as possible. In the courtroom, if the defendant challenges the chain of custody of the evidence, it can be proven that the knife in the evidence room is the same knife as found at the crime scene. However, if due to some discrepancies it cannot be proven who had the knife at a particular point in time, then the chain of custody is broken and the defendant can ask to have the resulting evidence declared inadmissible.

In a broader perspective "evidence" includes everything that is used to determine or demonstrate the truth of an assertion. Evidence can be used in court to convict people who are believed to have committed crimes; therefore, evidence must be handled in a scrupulously careful manner to avoid later allegations of tampering or misconduct that can compromise the case of the prosecution toward acquittal or to overturning a guilty verdict upon appeal.

The purpose behind recording the chain of custody is to establish that the alleged evidence is, indeed, related to the alleged crime, that is, the purpose is to establish the integrity of the evidence. In the context of conventional crimes, establishing "chain of custody" is especially important when the evidence consists of fungible goods.

"Fungibility" means the extent to which the components of an operation or product can be interchanged with similar components without decreasing the value of the operation or product.

For a person to be considered as "identifiable person," he/she must always have the physical custody of a piece of evidence. Practically speaking, this means that a police officer or detective will take charge of a piece of evidence, document its collection and hand it over to an evidence clerk for storage in a secure place. All such transactions as well as every succeeding transaction between evidence collection and its appearance in court need to be completely documented chronologically to withstand legal challenges to the authenticity of the evidence. Documentation must include conditions under which the evidence is collected, the identity of all those who handled the evidence, duration of evidence custody, security conditions while handling or storing the evidence and the manner in which evidence is transferred to subsequent custodians each time such a transfer occurs (along with the signatures of persons involved at each step).

Chain of custody is also used in most evidence situations to maintain the integrity of the evidence by providing documentation of the control, transfer and analysis of evidence.

Chain of custody is particularly important in situations where sampling can identify the existence of contamination and can be used to identify the responsible party. In Section 7.8, the relevance of chain of custody is explained in the context of computer/digital forensics.

7.5 Cyberforensics and Digital Evidence

Cyberforensics can be divided into two domains:

1. Computer forensics;
2. network forensics.

Many security threats are possible through computer networks (to know more on this, readers can refer to Ref. 11, Books, Further Reading). Therefore, "network forensics" [26] assumes importance in the context of cybercrime.

Network forensics is the study of network traffic to search for truth in civil, criminal and administrative matters to protect users and resources from exploitation, invasion of privacy and any other crime fostered by the continual expansion of network connectivity.

As compared to the "physical" evidence, "digital evidence" is different in nature because it has some unique characteristics. First of all, digital evidence is much easier to change/manipulate! Second, "perfect" digital copies can be made without harming original. At the same time the integrity of digital evidence can be proven. Another subtle aspect (of digital evidence) is that it is usually in the form of the "image" – this means that it is convenient and possible to create a defensible "clone" of storage device. Different information (clues) can be found at different levels of abstraction. Understanding the uniqueness of digital evidence is important for appreciating the phases involved in a digital forensics investigation and maintaining the "chain of custody" (refer to Section 7.8, Figs. 7.10 and 7.11, Boxes 7.4 and 7.12).

There are many forms of cybercrimes: sexual harassment cases – memos, letters, E-Mails; obscene chats or embezzlement cases – spreadsheets, memos, letters, E-Mails, online banking information; corporate espionage by way of memos, letters, E-Mails and chats; and frauds through memos, letters, spreadsheets and E-Mails. In case of computer crimes/cybercrimes, computer forensics helps. Computer forensics experts know the techniques to retrieve the data from files listed in standard directory search, hidden files, deleted files, deleted E-Mail and passwords, login IDs, encrypted files, hidden partitions, etc. Typically, the evidences reside on computer systems, user created files, user protected files, computer created files and on computer networks. Computer systems have the following:

1. Logical file system that consists of
   - File system: It includes files, volumes, directories and folders, *file allocation tables* (FAT) as in the older version of Windows Operating System, clusters, partitions, sectors.
   - Random access memory.
   - Physical storage media: It has magnetic force microscopy that can be used to recover data from overwritten area.
     (a) Slack space: It is a space allocated to the file but is not actually used due to internal fragmentation and
     (b) unallocated space.
2. User created files: It consists of address books, audio/video files, calendars, database files, spreadsheets, E-Mails, Internet bookmarks, documents and text files.
3. Computer created files: It consists of backups, cookies, configuration files, history files, log files, swap files, system files, temporary files, etc.
4. Computer networks: It consists of the Application Layer, the Transportation Layer, the Network Layer, the Datalink Layer.

Readers who are not savvy with these terms and concepts can refer to Ref. #11, Books, Further Reading. The Open System Interconnection (OSI) Layer Model (Application Layers, Transportation Layer, Datalink Layer, etc.) is also explained in Ref. #11, Books, Further Reading.

Box 7.5 The Father of Forensics Science – the Sherlock Holmes of France

The year 1877–1966 was the era of Dr. Edmond Locard. He is considered as the pioneer in forensics science and was popularly known as the Sherlock Holmes of France. He formulated the basic principle of forensics science: "Every contact leaves a trace." This came to be known as *Locard's exchange principle*. Following is one of his most famous quotes:

> *Wherever he steps, wherever he touches, whatever he leaves, even without consciousness, will serve as a silent witness against him. Not only his fingerprints or his footprints, but his hair, the fibers from his clothes, the glass he breaks, the tool mark he leaves, the paint he scratches, the blood or*

Box 7.5 The Father of . . . (*Continued*)

semen he deposits or collects. All of these and more bear mute witness against him. This is evidence that does not forget. It is not confused by the excitement of the moment. It is not absent because human witnesses are. It is factual evidence. Physical evidence cannot be wrong, it cannot perjure itself, it cannot be wholly absent. Only human failure to find it, study and understand it, can diminish its value. In other words, Whenever two human beings come into contact, something from one is exchanged to the other, that is, dust, skin cells, hair, etc."

For a short video clip that demonstrates how "Locard Principle" works, one can visit the link: http://science.howstuffworks.com/locards-exchange-principle.htm/printable (11 September 2009).

Locard studied medicine and law at Lyon, eventually becoming the assistant of Alexandre Lacassagne, a criminologist and professor. He held this post until 1910, when he began the foundation of his criminal laboratory. He produced a monumental seven-volume work, Traité de Criminalistique, and in 1918, developed 12 matching points for fingerprint identification. He continued with his research until his death in 1966.

7.5.1 The Rules of Evidence

This is a very important discussion, especially, for those who are students of legal courses. It was mentioned in Chapter 6 (Section 6.4) that the Indian IT Act amended the Indian Evidence Act. According to the "Indian Evidence Act 1872," "Evidence" means and includes:

1. All statements which the court permits or requires to be made before it by witnesses, in relation to matters of fact under inquiry, are called *oral evidence*.
2. All documents that are produced for the inspection of the court are called *documentary evidence*.

Legal community believes that "electronic evidence" is a new breed of evidence. They also, at times, have an apprehension that the law of evidence as per Indian Evidence Act of 1872 may not hold good for electronic evidence. Some lawyers express doubts and apprehensions about the process of leading electronic evidence in the courts. However, this is not true; the traditional principles of leading evidence, along with certain newly added provisions in the Indian Evidence Act 1972 through the Information Technology Act (ITA) 2000, constitute the body of law applicable to electronic evidence. The challenges, however, need to be understood from the "rules of evidence" perspective.

Paper evidence, the process is clear and intuitively obvious. Digital evidence by its very nature is invisible to the eye. Therefore, the evidence must be developed using tools other than the human eye.

It is only logical that the process used in the case of digital evidence mimic the process that is used for paper evidence. As each step requires the use of tools or knowledge, the process must be documented, reliable and repeatable. The process itself must be understandable to the members of the court. Acquisition of digital evidence is both a legal and technical problem. In fact, these two aspects are irrevocably related. The law specifies what can be seized, under what conditions, from whom and from where. It requires to determine what particular piece of digital evidence is required for examination, that is, is it a particular file or a word processing document or an executable program, etc. It may also require examination to determine where a particular piece of evidence is physically located. Is the file on a local hard drive or is it on a server

located in another legal jurisdiction? In short, it may be necessary to show a technical basis for obtaining the legal authority to search. Likewise, it may require technical skills to actually accomplish the search. The product of this phase is usually raw media, devoid of meaning or usefulness.

There are number of contexts involved in actually identifying a piece of digital evidence:

1. **Physical context:** It must be definable in its physical form, that is, it should reside on a specific piece of media.
2. **Logical context:** It must be identifiable as to its logical position, that is, where does it reside relative to the file system.
3. **Legal context:** We must place the evidence in the correct context to read its meaning. This may require looking at the evidence as machine language, for example, American Standard Code for Information Interchange (ASCII).

The path taken by digital evidence can be conceptually depicted as shown in Fig. 7.3.

Digital evidence originates from a number of sources such as seized computer hard drives and backup media, real-time E-Mail messages, chat room logs, Internet service provider records, webpages, digital network traffic, local and virtual databases, digital directories, wireless devices, memory cards, digital cameras, etc. Digital forensics examiners must consider the trustworthiness of this digital data. Many vendors provide technology solutions to extract this digital data from these devices and networks. Once the extraction of the digital evidence has been accomplished, protecting the digital integrity becomes paramount concern for investigators, prosecutors and those accused.

Similarly for the evidence in regular crimes, it is important to "isolate" the potential evidence. Some important tips are – do not turn ON the computer or review media, restrict physical and remote access, unplug computer power, network and phone line, and document times, people and steps taken. A point to note is that the need to unplug the computer power depends on the crime situation and the type of analysis required. For example, for live analysis of the digital evidence, it would not be advisable to unplug the power. Therefore, it is best to involve qualified specialists early in the process. Following are some guidelines for the (digital) evidence collection phase:

1. Adhere to your site's security policy and engage the appropriate incident handling and law enforcement personnel.
2. Capture a picture of the system as accurately as possible.

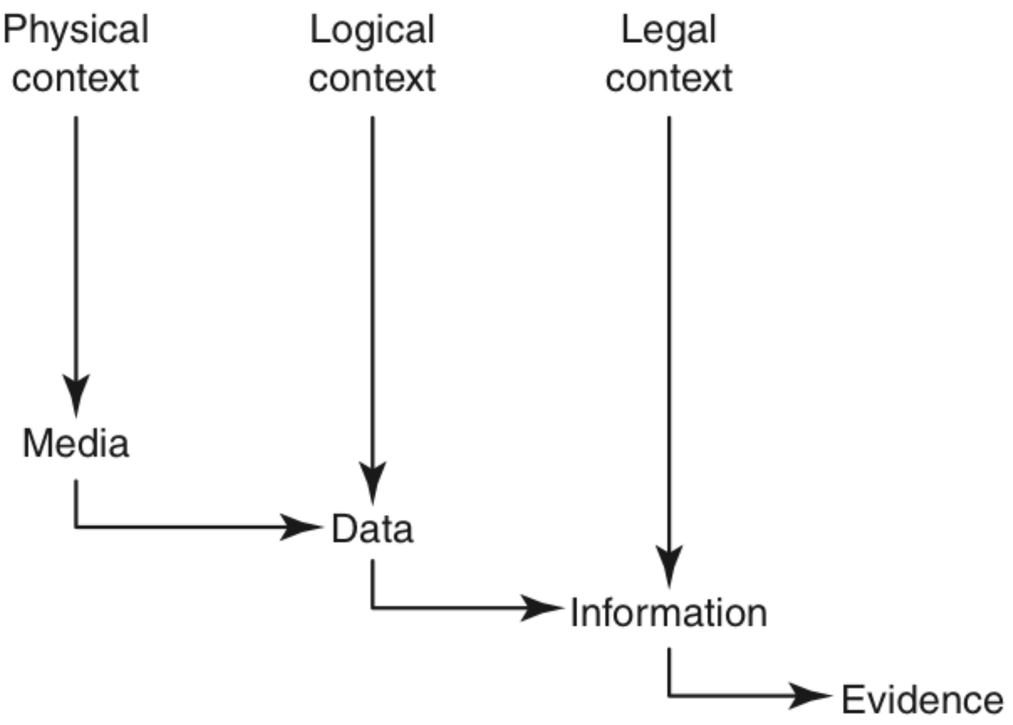


Figure 7.3 | Path of the digital evidence.

3. Keep detailed notes with dates and times. If possible, generate an automatic transcript (e.g., on Unix systems the "script" program can be used; however, the output file it generates should not be given to media as that is a part of the evidence). Notes and printouts should be signed and dated.
4. Note the difference between the system clock and Coordinated Universal Time (UTC). For each timestamp provided, indicate whether UTC or local time is used (since 1972 over 40 countries throughout the world have adopted UTC as their official time source).
5. Be prepared to testify (perhaps years later) outlining all actions you took and at what times. Detailed notes will be vital.
6. Minimize changes to the data as you are collecting it. This is not limited to content changes; avoid updating file or directory access times.
7. Remove external avenues for change.
8. When confronted with a choice between collection and analysis you should do collection first and analysis later.
9. Needless to say, your procedures should be implementable. As with any aspect of an incident response policy, procedures should be tested to ensure feasibility, particularly, in a crisis. If possible, procedures should be automated for reasons of speed and accuracy. Being methodical always helps.
10. For each device, a systematic approach should be adopted to follow the guidelines laid down in your collection procedure. Speed will often be critical; therefore, where there are a number of devices requiring examination, it may be appropriate to spread the work among your team to collect the evidence in parallel. However, on a single given system collection should be done step by step.
11. Proceed from the volatile to the less volatile; order of volatility is as follows:
    - Registers, cache (most volatile, i.e., contents lost as soon as the power is turned OFF);
    - routing table, Address Resolution Protocol (ARP) cache, process table, kernel statistics, memory;
    - temporary file systems;
    - disk;
    - remote logging and monitoring data that is relevant to the system in question;
    - physical configuration and network topology;
    - archival media (least volatile, i.e., holds data even after power is turned OFF).
12. You should make a bit-level copy of the system's media. If you wish to do forensics analysis you should make a bit-level copy of your evidence copy for that purpose, as your analysis will almost certainly alter file access times. *Try to avoid doing forensics on the evidence copy.*

Address Resolution Protocol (ARP) is a very important part of IP networking. ARP is used to connect OSI Layer 3 (Network) to OSI Layer 2 (Datalink). For most of us this means that ARP is used to link our IP addressing to our Ethernet addressing (MAC Addressing). For you to communicate with any device on your network, you must have the Ethernet MAC address for that device. If the device is not on your LAN, you go through your default gateway (your router). In this case, your router will be the destination MAC address that your PC will communicate with. There are two types of ARP entries: static and dynamic. Most of the time, you will use dynamic ARP entries. What this means is that the ARP entry (the Ethernet MAC to IP address link) is kept on a device for some period of time, as long as it is being used. The opposite of a dynamic ARP entry is static ARP entry. With a static ARP entry, you are manually entering the link between the Ethernet MAC address and the IP address. Because of management headaches and the lack of significant negatives to using dynamic ARP entries, dynamic ARP entries are used most of the time.

7.6 Forensics Analysis of E-Mail

In Chapter 2 (Section 2.3.1), it was mentioned how criminals can use fake mails for various cybercrime offenses. There are tools available that help create fake mails. *Forensics analysis of E-Mails* is an important aspect of cyberforensics analysis – it helps establish the authenticity of an E-Mail when suspected. This aspect is explained in this section – we start with understanding E-Mail components and then the E-Mail header structure is explained. E-Mails are now the most common means of communication worldwide and are often the subject of forensics analysis if this happens to constitute "digital evidence." Owing to the rising pressures from regulatory agencies and also due to possible litigations in global businesses, organizations are obligated to electronically store information to support discovery and disclosure requests. In this section, we want to discuss how E-Mail messages/IDs can help in forensic analysis of cybercrimes.

An E-Mail system is the hardware and software that controls the flow of E-Mail. The two most important components of an E-Mail system are the E-Mail server and the E-Mail gateway. *E-Mail servers* are computers that forward, collect, store and deliver E-Mail to their clients and E-Mail gateways are the connections between E-Mail servers. *Mail server software* is a network server software that controls the flow of E-Mail and the mail client software helps each user read, compose, send and delete messages. An E-Mail consists of two parts, the header and the body. *Message headers* are the important part for investigating E-Mail messages and hence it will be discussed in detail in this section. The "header" of an E-Mail is very important from forensics point of view – a full header view of an E-Mail provides the entire path of E-Mail's journey from its origin to its destination. The header view includes the originating Internet Protocol (IP) address and other useful information (see Table 7.1). There is usually a link provided on the E-Mail from its origin to its destination.

Box 7.6 Electronic Messages and the Indian Evidence Act

Section 88 of the *Indian Evidence Act* is about *Presumption as to telegraphic messages*. It states the following:

> *Presumption as to telegraphic messages. The Court may presume that a message, forwarded from a telegraph office to the person to whom such message purports to be addressed, corresponds with a message delivered for transmission at the office from which the message purports to be sent; but the Court shall not make any presumption as to the person by whom such message was delivered for transmission.*

As per Section 66A(C) of Indian IT Act *any electronic mail or electronic mail message for the purpose of causing annoyance or inconvenience or to deceive or to mislead the addressee or recipient about the origin of such messages (Inserted vide ITAA 2008) shall be punishable with imprisonment for a term which may extend to 2 or 3 years and with fine* – refer to URL http://cybercrime.planetindia.net/ch11_2008.htm

In terms of the amended Indian ITA 2000, that is, the ITA 2008 (notified on 5 February 2008), it is interesting to know how the court thinks when it comes to *evidences based on E-Mails*. As mentioned at the beginning, according to the Section 88A of the *Indian Evidence Act*, the court may presume that an electronic message forwarded by an originator through an E-Mail server to the addresses, to whom the message purports to be addressed, corresponds with the message as fed into the person's computer for transmission. However, the court shall not make any "presumption" as to the person by whom such message was sent. Furthermore, it is interesting to note the word "may presume" and the expressions "shall presume" and "conclusive proof."

1. The word "may presume" denotes that the court may either regard such fact as proved until it is disapproved, or may call for proof of it (Section 4 of the Indian Evidence Act 1872).
2. The word "shall presume" denotes that whenever it is directed by this Act that the court shall presume a fact, it shall regard such fact as proved, unless and until it is disapproved.

Box 7.6 Electronic Messages . . . (*Continued*)

3. The word "conclusive proof" denotes that when one fact is declared by this Act to be conclusive proof of another, the court shall, on proof of the one fact, regard the other as proved, and *shall not allow evidence to be given for the purpose of disapproving it*.

Source: Refer to Section 4 of the Indian Evidence Act. It can be downloaded from http://chddistrictcourts.gov.in/THE%20INDIAN%20EVIDENCE%20ACT.pdf (22 November 2008).

Table 7.1 | E-Mail header example

1. Return-Path: <secret@hotmail.com>
2. Received: from mailhub-1.net.treas.gov ([10.7.14.10]) by nccmail.usss.treas.gov for <avenit@usss.treas.gov>; Fri, 18 Feb 2000 11:46:07-0500
3. Received: from mx-relay.treas.gov ([199.196.144.6]) bytias4.net.treas.gov via smtpd (for mailhub.net.treas.gov [10.7.8.10]) with SMTP; 18 Feb 2000 16:55:44
4. Received: from hotmail.com (f7.law4.hotmail.com [216.33.149.7]) by mx-relay2.treas.gov for <avenit@usss.treas.gov>; Fri, 18 Feb 2000 11:55:44 – 0500 (EST)
5. Message-ID: <20000218165543.56965.qmail@hotmail.com>
6. Received: from 199.196.144.42 by www.hotmail.com with HTTP; Fri, 18 Feb 2000 08:55:43
7. X-Originating-IP: [199.196.144.42]
8. From: "Secret" <secret@hotmail.com>
9. To: avenit@usss.treas.gov
10. CC: smith@aol.com

Header information varies with E-Mail service provider, E-Mail applications and system configuration. As we know, the header part carries information that is needed for E-Mail routing, subject line and time stamps whereas the body contains the actual message/data of an E-Mail. The header and the body are separated by a blank line. The header contains several mandatory and optional fields, trace information and heading fields. The E-Mail header is a sequence of fields (it may not be in a particular order), each consisting of a field name and a field value. An example of a heading field would be: To: xyz@abccom.in. Headers on E-Mail can easily be "spoofed" by *spammers* and other irresponsible network users. E-Mails, when used in cyberforensics investigation, must get uniquely identified, if, for example, an E-Mail is suspected to be one of the evidence sources.

As mentioned earlier, the body of an E-Mail is separated from the header and it might also contain attachments in the form of MIME or SMIME (also known as S/MIME – secure/multipurpose Internet mail extensions). It is a protocol that provides digital signatures and encryption of Internet MIME messages. It is an encoding protocol (readers can visit the link at http://email.about.com/cs/standards/a/mime.htm to understand how MIME works for E-Mails).

We have mentioned previously that an E-Mail has two parts and header is one of those two parts. However, there is a header protocol: when an E-Mail message is sent, the user typically controls only the recipient line(s), that is, *To*, *Cc* and *Bcc*, if mentioned, and the *Subject* and *Date*. The rest of the header information is added by mail software while it is processed. Along the E-Mails route, a server can add or delete lines (anonymous remailer). Table 7.1 shows example of a mail header; for ease of understanding, each element of the header has been numbered. The discussion that follows is with reference to those numbers. *Header Protocol Analysis* is important for investing evidence that may come in form of an E-Mail.

In Table 7.1, elements 2, 3 and 4 show the route taken by the message from sending to delivery. Every computer that receives this message adds a "*Received:* field" with its complete address and time stamp; this

helps in tracking delivery problems. Element 5 of the mail header is the Message-ID, a unique identifier for this specific message. The Message-ID is logged and it can be traced through computers that are on the message route if there is a need to track the mail. Element 6 of the E-Mail header shows where the E-Mail was first received from with the IP address of the sender. It also shows the date and time when the message was sent. In this regard, it is important to understand the difference between simple mail transfer protocol (SMTP) and Hypertext Transfer Protocol (HTTP). HTTP is used to transfer displayable webpages and related files whereas SMTP is used to transfer E-Mail. Thus, SMTP is a protocol for sending E-Mail messages between servers whereas HTTP is a set of rules used to browse through Internet commonly used with web browsers such as Internet Explorer, Firefox). When you request, E-Mail logs you and you should ensure that you get them from the right server.

Next, consider element 7 of the sample mail header shown in Table 7.1 – it shows the originating IP address of the sender, but without the date and time the IP address will not allow you to identify the specific user. This may or may not be present in headers. If the IP address is a "Static" Address, you *will* be able to identify the specific user (most IP addresses are "dynamically" assigned). Element 8 indicates the name and E-Mail address of the message originator, that is, the "sender." Generally, this is the domain name we want to trace. Element 9 shows the name and E-Mail address of the primary recipient; the address may be for a mailing list (sales_dep@company.com) or systemwide *alias* (avenit@usss.treas.gov) or a personal username. The next element, element 10, of the sample mail header lists the names and E-Mail addresses of the "courtesy copy" recipients of the message. Some E-Mails may have "Bcc:" recipients as well; these "blind carbon copy" recipients get copies of the message, however their names and addresses are not visible in the headers.

Once we get the IP address our task is to find the Internet service provider details. The following links are worth visiting:

1. www.all-nettools.com (among other tools; a link to E-Mail tools is available here);
2. www.ip2location.com (there is a utility here that helps you know where your Internet visitors are coming from, that is, which country, which state, which city, which Internet service provider, which domain name, which connection type, which ZIP code, etc. It helps you trace an IP address to Country, Region, City, Latitude, Longitude, ZIP Code, Time Zone, Connection Speed, Internet service provider, Domain Name, IDD Country Code, Area Code, Weather Station Code and Name).
3. www.domaintools.com (there are DNS tool and many other tools available here);
4. www.dnsstuff.com (domain/E-Mail-related tools are accessible from this site);
5. http://www.hackingspirits.com/cyb_forensics/fsic_articles/trace_emails.html is a good link to know tracing the origin of an E-Mail, that is, locating countries from an IP address.

The Internet service provider plays an important role in E-Mail forensics. The Internet service provider provides Internet access to businesses, organizations, schools, colleges and individuals. Examples of Internet service provider are VSNL (Videsh Sanchar Nigam Limited), Sify, Hathway, Rolta, MTNL/BSNL, Reliance, etc. The details available from the Internet service provider are name, address and contact number of the subscriber of the Internet facility, type of IP address, any other relevant information with regard to IP address at a particular given date and time, usage details, etc.

Box 7.7 Points to Remember when you Use E-Mail as an Evidence

1. Ensure the use of E-Mail is subject to agreed procedures, which are supported and enforced by management at a high level. Acceptable Use Policies ought to prescribe good usage and identify bad usage.

Box 7.7 Points to Remember . . . (*Continued*)

2. Train users of E-Mail about acceptable use of E-Mail, and about their rights and the obligations expected of them.
3. Implement access control mechanisms to computer systems – so that its use can be attributed to a person, a terminal, a date and a time.
4. Ensure computer systems are kept safe and secure so that the systems and the data within are protected from unauthorized access and accidental or deliberate loss and damage.
5. Retention and deletion of E-Mail should be organization-defined and not user-defined. Individual users should not have any discretion as to the categories of E-Mails that should be retained or deleted.
6. Implement a solution that archives and stores E-Mails centrally. The archive should support all the main file formats and also retain metadata.
7. The archive should classify E-Mails entering the archive at the point of entry. The archive should prevent the entry of duplicates.
8. Make sure that the archiving platform facilitates the exporting of evidence as files as a part of the E-Discovery process.
9. Implement an archiving solution that allows full search and retrieval. Metadata should be searchable as should content.
10. Enable logging of all events acting on the archive. The logs should be retained as part of the archive, for auditing and verification purposes.
11. Provide contingency for continuity of both archiving and discovery in the event of an outage.
12. Ensure the archiving platform supports the marking-up of files so that privileged materials can be withheld and/or redacted during E-Discovery.

E-Mail headers are organized bottom-up. This means that the E-Mail was handed from the machines at the bottom of the E-Mail header to the ones at the top of it. These machines are referred to as Message Transfer Agents (MTAs) and each of them adds a "received" section to the E-Mail header, sometimes referred to as "received header." This is similar to postmarks used in conventional postal systems. The order of the "received" sections is like a stack of pancakes, with the one receiving the E-Mail last at the top of the stack. Refer to Fig. 7.4 – note that there were three received sections (elements 2, 3 and 4 in Table 7.1). This means that three MTAs were involved in the delivery of the E-Mail message with the one at the bottom being the one receiving the original message from the sender.

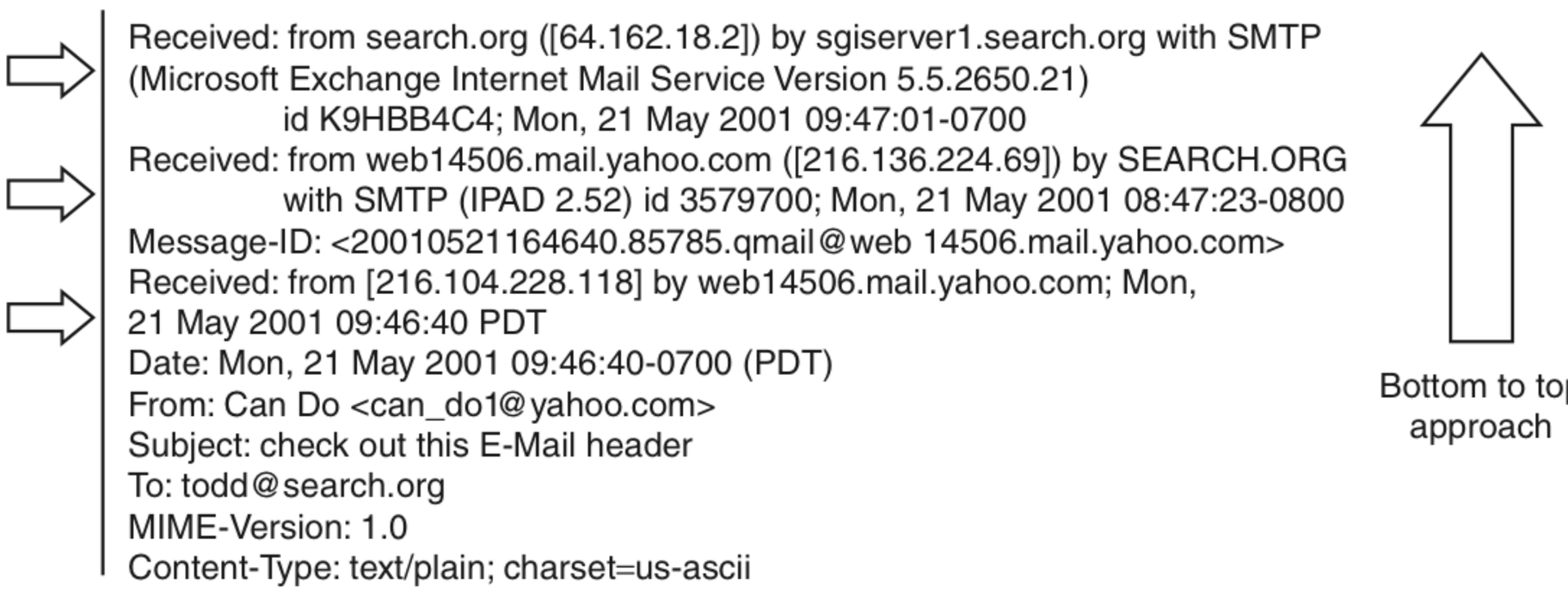


Figure 7.4 Bottom-up approach to tracing E-Mail source.
Source: "Tracing-eMail-Headers.pdf," Marwan Al-Zarouni, School of Computer and Information Science, Edith Cowan University, Perth, Western Australia. To see a visual example of tracing an E-Mail, readers can visit: http://www.youtube.com/watch?v=hSvswzSy3oA&feature=related (15 March 2010). It shows a video clip with demonstrating an E-Mail.

E-Mail tracing is done by examining the header information contained in E-Mail messages to determine their source.

While tracing E-Mails, the "header information" is included along with E-Mails either at the beginning or the end of E-Mail messages. A typical E-Mail header looks like as shown in Table 7.2.

To determine the source of the E-Mail, investigators must first examine the received section at the bottom of the header and work their way up in a bottom to top approach (see Fig. 7.4).

It is also important that during E-Mail investigation cases the logs of all servers in the received chain are examined as soon as possible. The time stamp is very important in E-Mail investigation cases because HTTP and SMTP logs get archived frequently, especially by large Internet service providers. When a log is archived, a considerable amount of time and effort is involved to retrieve and decompress the log files needed to trace E-Mails. Fake E-Mail creation tools are rampantly used by cybercriminals. Therefore, it is possible that some E-Mails have fake headers with fake "from" E-Mail addresses to fool investigators; however, extreme caution and careful scrutiny should be practiced in investigating every part of the E-Mail header (recall it in Chapter 2 in which it is explained that there are tools available that help create fake mails).

Typically, the sender's E-Mail address can be found after the "From" section of the header. However, that is not the only place it can be found. It can also be found under other sections depending on the E-Mail client uses. These sections include the following (this is not the exhaustive list; it is just an example to give you some idea):

1. . X-originating E-Mail;
2. . X-sender;
3. . return-path.

At times, E-Mail addresses can suggest the method used to generate the E-Mail and the server that the E-Mail originated from (i.e., hotmail, outlook, corporate server, Internet service provider, etc.). However, E-Mail addresses should be viewed with caution by investigators as they can be easily faked. Note that some headers begin with an "X-," this means that they are X-headers. You can use X-headers to sort and filter

Table 7.2 | Typical E-Mail header

1. Received: from search.org ([64.162.18.2]) by sgiserver1.search.org with SMTP (Microsoft Exchange Internet Mail Service
   Version 5.5.2650.21)
   id K9HBB4C4; Mon, 21 May 2001 09:47:01-0700
2. Received: from web14506.mail.yahoo.com ([216.136.224.69]) by SEARCH.ORG
   with SMTP (IPAD 2.52) id 3579700; Mon, 21 May 2001 08:47:23-0800
   Message-ID: <20010521164640.85785.qmail@web14506.mail.yahoo.com>
3. Received: from [216.104.228.118] by web14506.mail.yahoo.com; Mon, 21 May 2001 09:46:40 PDT
   Date: Mon, 21 May 2001 09:46:40 -0700 (PDT)
   From: Can Do <can_do1@yahoo.com>
   Subject: check out this E-Mail header
   To: todd@search.org
   MIME-Version: 1.0
   Content-Type: text/plain; charset = us-ascii

E-Mails sent by SourceForge. Depending on the context of the message, custom E-Mail headers (X-headers) are added to the E-Mail. The customer headers can be used by E-Mail agents and clients to filter and sort E-Mail. For example, both Outlook and Thunderbird support filtering on custom E-Mail headers. Thus, X-headers are inserted by E-Mail client programs or applications that use E-Mail to pass information to E-Mail handling programs for processing. They may be introduced by large vendors and picked up for use by others. In this way an X-header can be considered as a de facto standard. An example of this is the "X-Mailer" header which many E-Mail clients use to define the E-Mail client application and version used.

Next, let us understand how fake E-Mail addresses can be detected. The sender's E-Mail address can be easily faked and can be hard to detect. If the server mentioned in the bottom "received" section does not match the server of the E-Mail address, this suggests that the E-Mail address is a fake one. An example is shown in Table 7.3.

Note that in Table 7.3 the E-Mail address in the "From" field has "hotmail.com" as the domain for the E-Mail whereas in the received section of the header there is no hotmail server mentioned at all. This is clearly a forged (fake) E-Mail and it is very likely to have a fake "From" address. Also note that the time on the received section is Central European Standard Time (CEST), and hotmail.com servers are not in Europe.

Now let us consider Sendmail. Sendmail is a general purpose Internet work E-Mail routing facility that supports many kinds of mail-transfer and delivery methods, including SMTP used for E-Mail transport over the Internet. Sendmail is a descendant of the delivermail program that was written by Eric Allman. Sendmail is a well-known project of the free and open-source software (OSS) and Unix communities, and has spread both as free software and proprietary software. It is a very widely used MTA. MTAs send mail from one machine to another. Sendmail is not a client program, which you use to read your E-Mail but rather a behind-the-scenes program that actually moves your E-Mail over networks or the Internet to where you want it to go. If there is ever a situation to reconfigure Sendmail, you will also need to have the sendmail.cf package installed. In case you need documentation on Sendmail, you need to install the sendmail-doc package.

To uniquely identify each E-Mail, all MTAs use some sort of unique identifier. This identifier is referred to as "Message-ID." *Message-ID field* is inserted into a header either by mail user agent (MUA) or the first MTA. Even though the Message-ID is optional as per RF2822, it recommends using it. Sendmail, for example, is one MTA that handles E-Mail delivery and relaying process. Sendmail uses Message-ID for tracing E-Mails and for logging process IDs. It recommends including Message-ID in E-Mails and also recommends setting relevant macros in its configuration file to implement compulsory checking of Message-IDs.

A point to note is that unlike Spoofing other fields in the header, Spoofing Message-ID needs special knowledge. Sendmail-related FAQs are available at http://www.sendmail.org/faq/section2.

Deep analysis on Message-IDs may reveal some sort of information that will open a window to trace the source of an E-Mail. Also Message-ID will help to find a particular E-Mail log entry within a log file of E-Mail server.

Table 7.3 | Header of a fake mail (an example only)

Received: from infvic.it (adsl-98-201.38-151.net24.it [151.38.201.98]) by mail-relay2.bpvi.it (Postfix) with ESMTP id 2887550074 for <redazione@infvic.it>; Mon, 19 Apr 2004 10:41:54 +0200 (CEST) From: sfiorillo@hotmail.com

Only technical envy spammers can spoof the Message-ID cleverly. So deep analysis on Message-IDs may reveal some sort of information that will open a window to trace the source of an E-Mail. Also the Message-ID will help to find a particular E-Mail log entry within a log file of E-Mail server. There are some commonalities between conventional mails and E-Mails; for example, like conventional mail service, when E-Mail is routed from source to destination all intermediate relay servers (SMTP) insert their stamp at the beginning of the header. This stamping procedure helps to trace the E-Mail if such a demand arises. The stamp consists of three fields, namely. "From," "SMTP-ID" and "For."

The text in Table 7.4 shows an E-Mail header that passed through several MTAs, that is, E-Mail header with several identifiers. Each MTA inserted a unique ID in the header of E-Mail. There are several identifiers in the header field of an E-Mail that may help to trace the source of the E-Mail but the context for Fig. 7.4 is limited to Sendmail Message-ID only. Analyzing intermediate SMTP-IDs is beyond the scope of this book. However, in this section, we have briefly discussed intermediate SMTP-IDs.

In the context of E-Mail, "messages" are viewed as having an "envelope" and "contents." The envelope contains information required to accomplish transmission and delivery. The contents contain the object to be delivered to the recipient. The delivery has to be at a valid E-Mail address and this is where the RFC2822 comes into picture.

7.6.1 RFC2822

RFC2822 is the Internet Message Format. According to the Internet specification RFC2822, there are several formats of valid E-Mail addresses, like joshi@host.net, john@[10.0.3.19], "Joshi Ganesh"@host.net or "Joshi Ganesh"@[10.0.3.19]. Many E-Mail address validators on the Web fail to recognize some of those valid E-Mail addresses. Some examples of invalid E-Mail addresses are as follows:

1. joshi@box@host.net: Two at signs (@) are not allowed;
2. .joshi@host.net: Leading dot (.) is not allowed;
3. joshi@-host.net: Leading dash (–) is not allowed in on domain name;
4. joshi@host.web: Web is not a valid top level domain;
5. joshi@[10.0.3.1999]: Invalid IP address.

The RFC2822 standard applies only to the Internet Message Format and some of the semantics of message contents. It contains no specification of the information in the envelope. RFC2822 states that each E-Mail

Table 7.4 | E-Mail header with several identifiers

1. Received: from search.org ([64.162.18.2]) by sgiserver1.search.org with SMTP (Microsoft Exchange Internet Mail Service Version 5.5.2650.21)
   id K9HBB4C4; Mon, 21 May 2001 09:47:01-0700
2. Received: from web14506.mail.yahoo.com ([216.136.224.69]) by SEARCH.ORG with SMTP (IPAD 2.52) id 3579700; Mon, 21 May 2001 08:47:23-0800
   Message-ID: <20010521164640.85785.qmail@web14506.mail.yahoo.com>
3. Received: from [216.104.228.118] by web14506.mail.yahoo.com; Mon, 21 May 2001 09:46:40 PDT
   Date: Mon, 21 May 2001 09:46:40-0700 (PDT)
   From: <can_do1@yahoo.com>
   Subject: check out this E-Mail header
   To: todd@search.org

must have a "globally unique identifier." This must be included into the header of an E-Mail. The RFC2822 also defines the syntax of Message-ID. It should be like a legitimate E-Mail address and it must be included within a pair of angle brackets. According to RFC2822, Message-ID can appear in three header fields: "Message-ID header," "in-reply-to header" and "references header." But Message-ID of the present E-Mail must be included against the "Message-ID" header. Remember that there are SPAM problems and to that, there is no simple solution. E-Mail headers cannot be trusted; not all E-Mail can be traced or authenticated. Only a legitimate mail typically can be traced. However, for SPAM and virus-generated E-Mail it is difficult to know if the headers are absolutely trustworthy.

Readers may refer to Ref. #2, Video Clips, Further Reading in which it is explained how to trace an E-Mail. To conclude this section, we say that tracing an E-Mail is an important forensics activity in instances where an E-Mail is believed to hold a queue for a cybercrime. Understanding the E-Mail header structure is important while tracing an E-Mail and we have discussed that so far. Readers, who are interested in learning about tracking E-Mails, can try out the tutorials available at the following links accessed on 6 December 2009:

1. http://www.visualware.com/resources/tutorials/email.html (both download and live demonstrations are available at this link);
2. http://www.visualware.com/resources/tutorials/emailX.html (a tutorial on E-Mail tracking tutorial is available here).

7.7 Digital Forensics Life Cycle

As per FBI's (Federal Bureau of Investigation) view, digital evidence is present in nearly every crime scene. That is why law enforcement must know how to recognize, seize, transport and store original digital evidence to preserve it for forensics examination. Figure 7.5 shows the process model for understanding a *seizure and handling of forensics evidence* legal framework. The cardinal rules to remember are that evidence:

1. is admissible;
2. is authentic;
3. is complete;
4. is reliable;
5. is understandable and believable.

Let us now understand what is involved in the digital forensics process.

7.7.1 The Digital Forensics Process

The digital forensics process needs to be understood in the legal context starting from preparation of the evidence to testifying. Digital forensics evidence consists of exhibits, each consisting of a sequence of bits, presented by witnesses in a legal matter to help jurors establish the facts of the case and support or refute legal theories of the case. The exhibits should be introduced and presented and/or challenged by properly qualified people using a properly applied methodology that addresses the legal theories at issue. The tie between technical issues associated with the digital forensics evidence and the legal theories is the job of "expert witnesses."

As part of the court procedure, the exhibits are introduced as evidence by either side. *Testimony* is presented to establish the process to identify, collect, preserve, transport, store, analyze, interpret, attribute, and/or reconstruct the information contained in the exhibits and to establish, to the standard of proof required by the matter at hand, that the evidence reflects a sequence of events that is asserted to have produced it. The party must show not only the evidence to be admitted but must also establish that the evidence is

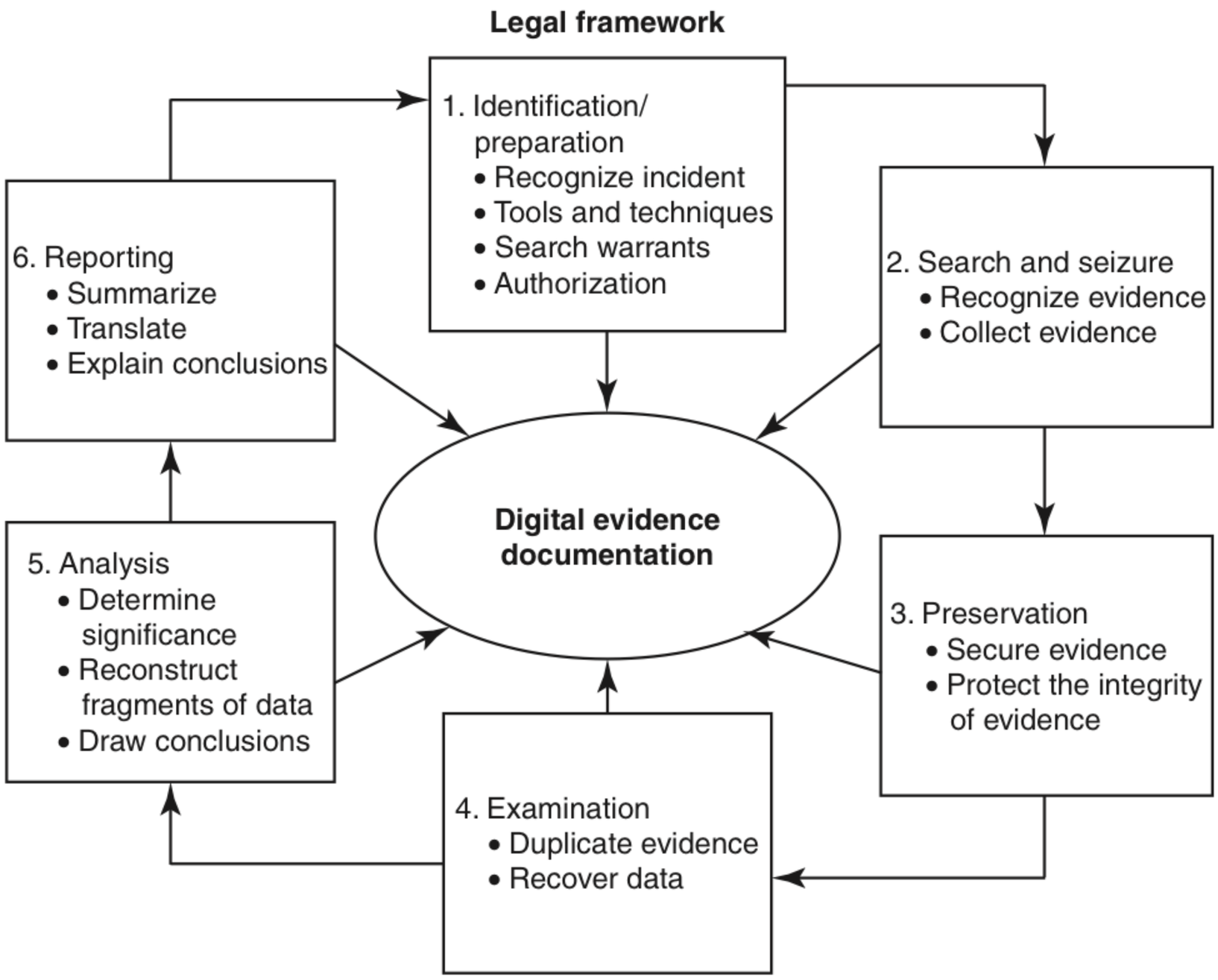


Figure 7.5 | Process model for understanding a seizure and handling of forensics evidence legal framework.

relevant, authentic and that the evidence presented is not the result of hearsay, original writing or the legal equivalent thereof, and more probative than prejudicial.

Usually the assumption is that adequate facts can be established for the introduction of an evidence exhibit. Under this assumption, people involved in the "chain of custody" need to testify a number of aspects relating to the evidence – the testimony would typically include the processes used for creating, handling and introducing the evidence, the method used for collecting the exhibit (i.e. the evidence artifacts) as well as the manner in which the exhibit is brought to court. These people also get involved to testify about the event sequences that may have produced the evidence exhibit. Digital forensics evidence is usually latent, that is, "hidden" in that it can only be seen by the trier of fact at the desired level of detail through the use of tools. In order for tools to be properly applied to a legal standard, it is required that the people who use these tools properly apply their scientific knowledge, skill, experience, training and/or education. They should also use a methodology that is reliable within defined standards to show the history, pedigree and reliability of the tools, proper testing and calibration of those tools, and their application to functions they perform within the limitations of their reliable application.

Non-experts can make statement about evidence to the extent that they can clarify non-scientific issues by stating what they observed. Digital forensics evidence can be challenged by establishing that, by intent or accident, content, context, meaning, process, relationships, ordering, timing, location, corroboration and/or consistency are made or missed by the other side, and that this produced false positives or false negatives in the results presented by the other side. The trier of fact then must determine how the evidence is applied to the matter at hand so as to weigh it against and in conjunction with all of the other evidence and to render judgments about the legal matters that the evidence applies to.

Box 7.8 Forensics Experts – What do they Do?

The role of forensics experts has become a very special one in digital forensics and there are many reasons for it. Handling of digital evidence requires special expertise that comes from training and experience. A lot of protocols come into picture depending on the nature of the evidence; for example, the, complexity, volume and delicate nature of relevant electronic evidence. Depending on such nature of the evidence, even expensive hardware and software tools will be required along with the investigator's experience to achieve optimal results. In most cases, it is best to address this through partnership with a third party expert forensics firm.

"Peeking around the data" on your own may destroy relevant date and time stamps and other metadata, and more importantly, it may expose you to sanctions for spoliation. Using overly generic discovery requests ("please provide all electronic data … .") can produce excessively broad or burdensome requests. The court may reject such request as it may be too expensive and time consuming to fulfill these requests. If digital forensics evidence is properly managed, then a computer forensics expert will be able to focus on the relevant electronic discovery targets, and will be able to lower the eventual cost of litigation and increase the probability of a favorable outcome.

A forensics expert team brings the following additional benefits:

1. **Technology expertise:** This is perhaps the biggest advantage of partnership with a computer forensics expert. As an example of the technological complexity, consider the proliferation of operating systems in the last decade: mainframe operating systems, Windows 95/98, UNIX, Linux, Windows NT, Windows Server, Macintosh, Windows 2000, Windows XP and Novell Netware. Specific forensics tools must be used with each of these file systems, along with training and experience to interpret search results. Although some evidence may be found easily, other evidence may have been deleted, altered, hidden or encrypted. Forensics experts routinely deal with such complexities and nuances.
2. **Forensics methodology:** A comprehensive forensics methodology, repeatable and defensible, has become a key attribute in choosing a forensics expert firm. Proper use of a repeatable process prevents making the same mistake twice, ensures proper chain of custody, leverages successful techniques from prior cases, supports clear and concise testimony, and generally guarantees efficient forensics case management.
3. **Experience and efficiency:** The tools and methods of computer forensics examination are still in their infancy. Experts know how to quickly navigate through the variety of esoteric tools and procedures. Experts also have the experience to cull thousands of files based on patterns and keywords. Therefore, working with experts will efficiently produce relevant results for counsel.

The "chain of custody" concept, too, is a very important one in digital forensics (see Section 7.8). We have provided links in Ref. #2, Video Clips, Further Reading about evidence seizure as part of forensics investigation.

Once the forensics experts know the landscape of the computers and other artifacts involved, they formulate a cost proposal governing all needed activities in the forensics search and analysis. This is combined with a proposed timeline of activities, lists of anticipated deliverables and a plan for production and turnover of evidence. In addition to this, forensics experts also submit a preliminary risk analysis for the forensics service being proposed. This will detail any technical and political obstacles that were envisaged. For forensics findings of any type to be used as admissible evidence in court, the data acquisition, also known as "imaging," of the subject computers must be flawless and defensible in substance and technique. Forensics examiners are trained to follow a carefully developed set of protocols for acquisition of electronic evidence designed to ensure authenticity and diligent chain of custody.

7.7.2 The Phases in Computer Forensics/Digital Forensics

The investigator must be properly trained to perform the specific kind of investigation that is at hand. Tools that are used to generate reports for court should be validated. There are many tools to be used in the process.

One should determine the proper tool to be used based on the case. Broadly speaking, the forensics life cycle involves the following phases:

1. Preparation and identification;
2. collection and recording;
3. storing and transporting;
4. examination/investigation;
5. analysis, interpretation and attribution;
6. reporting;
7. testifying.

To mention very briefly, the process involves the following activities:

1. **Prepare:** Case briefings (see Box 7.9), engagement terms, interrogatories, spoliation prevention, disclosure and discovery planning, discovery requests.
2. **Record:** Drive imaging, indexing, profiling, search plans, cost estimates, risk analysis.
3. **Investigate:** Triage images, data recovery, keyword searches, hidden data review, communicate, iterate.
4. **Report:** Oral vs. written, relevant document production, search statistic reports, chain of custody reporting, case log reporting.
5. **Testify:** Testimony preparation, presentation preparation, testimony.

Let us take a brief look at each of the activites mentioned. Table 7.5 shows phase-wise outcome from the phases mentioned above.

Preparing for the Evidence and Identifying the Evidence

In order to be processed and applied, evidence must first be identified as evidence. It can happen that there is an enormous amount of potential evidence available for a legal matter, and it is also possible that the vast majority of the potential evidence may never get identified. Consider that every sequence of events within a single computer might cause interactions with files and the file systems in which they reside, other processes and the programs they are executing and the files they produce and manage, and log files and audit trails of various sorts. In a networked environment, this extends to all networked devices, potentially all over the world. Evidence of an activity that caused digital forensics evidence to come into being might be contained in a time stamp associated with a different program in a different computer on the other side of the world that was offset from its usual pattern of behavior by a few microseconds. If the evidence cannot be identified

Box 7.9 Case Briefings

In case briefings, consider the following:

1. Ensure that you know both your client's and the adverse party's position, and have seen all relevant paperwork.
2. Try not to project a bias in the case description; the intent should be to consider the case objectively, and provide you with the good news and the bad news (bad news early can be good news).
3. Be upfront in discussing any limitations or restrictions on the forensics investigation, including budgetary constraints, time deadlines, cooperation levels to be expected from the adverse party, required travel, onsite or after-hours forensics imaging requirements, etc.

as relevant evidence, it may never be collected or processed at all, and it may not even continue to exist in digital form by the time it is discovered to have relevance.

Collecting and Recording Digital Evidence

Digital evidence can be collected from many sources. Obvious sources include computers, cell phones, digital cameras, hard drives, CD-ROM, USB memory devices and so on. Non-obvious sources include settings of digital thermometers, black boxes inside automobiles, RFID tags and webpages (which must be preserved as they are subject to change). Special care must be taken when handling computer evidence: most digital information is easily changed, and once changed it is usually impossible to detect that a change has taken place (or to revert the data back to its original state) unless other measures have been taken. For this reason, it is common practice to calculate a cryptographic hash of an evidence file and to record that hash elsewhere, usually in an investigator's notebook, so that one can establish at a later point in time that the evidence has not been modified as the hash was calculated. Figures 7.6 and 7.7 show the media that typically holds digital evidence.

Figure 7.6 Media that can hold digital evidences.
Sources: http://www.homeofficebuddy.com; http://oldcomputers.net; http://www.homecomputertalk.com; http://www.cyberindian.net; http://www.srs-electronicmall.com; http://transcriptdivas.co.uk and http://images.google.co.in; http://www.mobileshop.com; http://images.google.co.in; http://www.slipperybrick.com; http://images.google.co.in; http://www.letsgodigital.org; http://www.computerrepairmaintenance.com; http://www.indigoshop.co.uk; http://www.adorama.com, http://sp.sony-europe.com/media/4/1914, http://www.video99.co.uk/dat.jpg

Figure 7.7 | Some more media that can hold digital evidences.

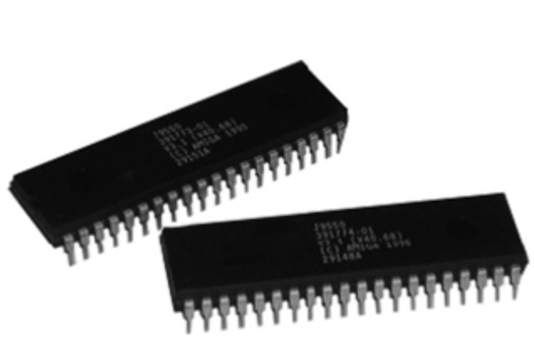
Collecting volatile data requires special technical skills. If the machine is still active, any intelligence that can be gained by examining the applications currently open is recorded. If the machine is suspected of being used for illegal communications, such as terrorist traffic, not all of this information may be stored on the hard drive. If information stored solely in random access memory (RAM) is not recovered before powering down, it may be lost. This results in the need to collect volatile data from the computer at the onset of the response.

Embedded flash memory falls under the family of solid state non-volatile memory; it is used in thumb drives (USB stick), cell phones, game console, secure digital cards (SD cards) and multimedia cards (MMC). This technology differs from the normal hard disk by not containing any moving parts such as arms and cylinders. In addition, the physical size of the embedded memory chips makes it a good candidate to be used in every device that interacts with our daily life. The benefits of embedded memory continue to increase life expectancy of the memory chip due to a reduction of mechanical failure even if it had been used in high vibrated or trembling environment. Figure 7.8 shows the various types of "embedded memories" inside a computer (ROM, PROM, EPROM, EEPROM).

Storing and Transporting Digital Evidence

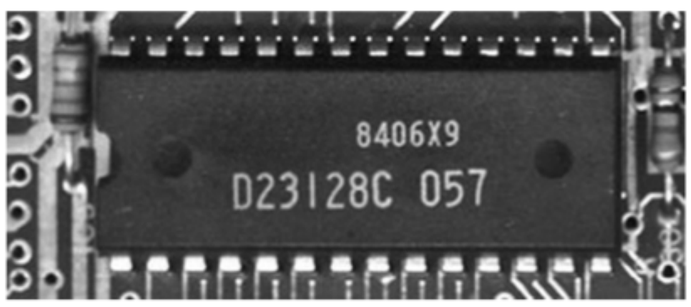
The following are specific practices that have been adopted in the handling of digital evidence:

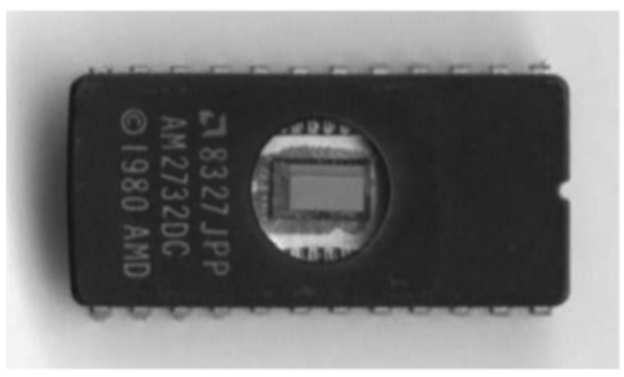
1. Image computer media using a write-blocking tool to ensure that no data is added to the suspect device;
2. establish and maintain the chain of custody (refer to Section 7.8);



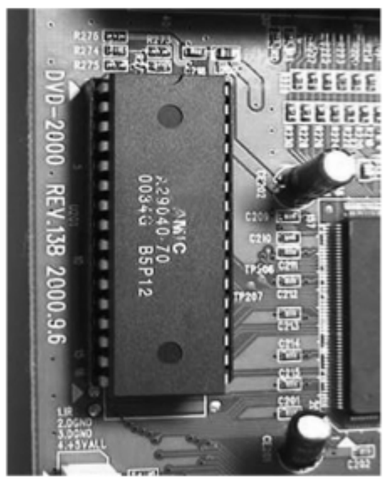
(a)

(b)





(c)



(d)

Figure 7.8 Embedded memories inside computer. (a) Read-only memory (ROM) chips; (b) erasable programmable read-only memory (EPROM) chip; (c) programmable read-only memory (PROM) chips; (d) electrify erasable programmable read-only memory (EEPROM) chips. *Sources:* http://amigakit.leamancomputing.com; http://www.old-computers.com; http://upload.wikimedia.org and http://www.electrongate.com; http://wiki.laptop.org and http://www.dv-rec.de

3. document everything that has been done;
4. only use tools and methods that have been tested and evaluated to validate their accuracy and reliability.

Some of the most valuable information obtained in the course of a forensics examination will come from the computer user. An interview with the user can yield valuable information about the system configuration, applications, encryption keys and methodology. Forensics analysis is much easier when analysts have the user's passphrases to access encrypted files, containers and network servers.

In storage, digital media must be properly maintained for the period of time required for the purposes of trial. Depending on the particular media (see Figs. 7.6 and 7.7), this may involve any number of requirements ranging from temperature and humidity controls to the need to supply additional power or to re-read media. Storage must be adequately secure to assure proper "chain of custody" (refer to Section 7.8), and typically, for evidence areas containing large volumes of evidence, paperwork associated with all actions related to the evidence must be kept to assure that evidence does not go anywhere without being properly traced. Many things can go wrong in storage, including decay over time; environmental changes resulting in the presence or absence of a necessary condition for preservation; direct environmental assault on the media; fires, floods and other external events reaching the evidence; loss of power to batteries and other media-preserving mechanisms; and decay over time from other natural and artificial sources.

Sometimes evidence must be transported from place to place. For example, when collected from a crime scene, the evidence must somehow be moved to a secure location or it may not be properly preserved through a trial. Digital forensics evidence can generally be transported by making exact duplicates, at the level of bits, of the original content. This includes the movement of content over networks, assuming adequate precautions are taken to assure its purity during that transportation.

Evidence is often copied and sent electronically, on compact disks or on other media, from place to place. Original copies are normally kept in a secure location to act as the original evidence that is introduced into the legal proceedings. If there is any question about the bits contained in the evidence, it can be settled by returning to the original evidence. Facsimile evidence, printouts and other similar depictions of digital forensics evidence may also be transported, but they are not a good substitute for the original digital forensics evidence in most cases, among other reasons, because they make it far harder, if not impossible, to properly analyze what the original bits were. For example, many different bit sequences may produce the output depictions, and identical bit sequences may produce different output depictions.

Adequate care must be taken in transportation to prevent spoliation as well. For example, in a hot car, digital media tends to lose bits. Increasingly, evidence is transported electronically from place to place, and even the simplest errors can cause the data arriving to be incorrect or improperly authenticated for legal purposes. Care must also be taken to preserve chain of custody and assure that a witness can testify accurately about what took place, using and retaining contemporary notes, and taking proper precautions to assure that evidence is not spoilt and is properly treated along the way.

Examining/Investigating Digital Evidence

In an investigation in which the owner of the digital evidence has not given consent to have his or her media examined (as in some criminal cases) special care must be taken to ensure that the forensics specialist has the legal authority to seize, copy and examine the data. Sometimes authority stems from a search warrant.

As a general rule, one should not examine digital information unless one has the legal authority to do so. Amateur forensics examiners should keep this in mind before starting any unauthorized investigation.

Now let us understand the difference between live and dead analysis. After that we explain about "imaging of the media." Traditionally, computer forensics investigations were performed on data at rest, for example, the content of hard drives. This can be thought of as a "dead analysis." Investigators were told to shutdown computer systems when they were impounded for fear that digital time bombs might cause data to be erased. In recent years, there has been increasingly an emphasis on performing analysis on live systems. One reason is that many current attacks against computer systems leave no trace on the computer's hard drive; the attacker only exploits information in the computer's memory. Another reason is the growing use of cryptographic storage: it may be that the only copy of the keys to decrypt the storage is in the computer's memory; turning OFF the computer will cause that information to be lost.

For the purpose of digital evidence examination, "imaging of electronic media" (on which the evidence is believed to be residing) becomes necessary.

The process of creating an exact duplicate of the original evidentiary media is often called "Imaging." Computer forensics software packages make this possible by converting an entire hard drive into a single searchable file – this file is called an "image." Using a stand-alone hard drive duplicator or software imaging tools such as DCFLdd, IXimager or Guymager, the entire hard drive is completely duplicated. This is usually done at the sector level, making a bit-stream copy of every part of the user-accessible areas of the hard drive which can physically store data, rather than duplicating the file system. The original drive is then moved to secure storage to prevent tampering. During imaging, a write protection device or application is normally used to ensure that no information is introduced onto the evidentiary media during the forensics process. The imaging process is verified by using the SHA-1 message digest algorithm (with a program such as sha1-sum) or other still viable algorithms such as MD5. At critical points throughout the analysis, the media is verified again, known as "hashing," to ensure that the evidence is still in its original state. In corporate environments seeking civil or internal charges, such steps are generally overlooked due to the time required to perform them. They are essential for evidence that is to be presented in a courtroom, however.

Analysis, Interpretation and Attribution

Analysis, interpretation and attribution of evidence are the most difficult aspects encountered by most forensics analysts. In the digital forensics arena, there are usually only a finite number of possible event sequences that could have produced evidence; however, the actual number of possible sequences may be almost unfathomably large. In essence, almost any execution of an instruction by the computing environment containing or generating the evidence may have an impact on the evidence. Basically, all digital evidence must be analyzed to determine the type of information that is stored upon it. For this purpose, specialty tools are used that can display information in a format useful to investigators. Such forensics tools include but are not limited to the following list. Readers can refer to links in References to know more about this toolkit. (Also refer to Appendix I in CD.)

1. Access Data's FTK[2];
2. guidance Software's EnCase[3];
3. Dr. Golden Richard III's file carving tool Scalpel[4]; "file carving" is the process of recovering files from an investigative target, potentially without knowledge of the file system structure;
4. Brian Carrier's Sleuth Kit[5]: The Sleuth Kit (TSK) is a library and collection of Unix- and Windows-based tools and utilities to allow for the forensics analysis of computer systems.

Typical forensics analysis includes a manual review of material on the media – an example of OS-specific investigation is reviewing the Windows registry. Through this registry inspection, the investigators objective is to look for suspect information, discovering and cracking passwords, performing keyword searches for topics related to the crime, and extracting E-Mail and images for review. Numerous other tools are used in digital forensics investigations to analyze specific portions of information. See Box 7.10 regarding file carving technique. In Chapter 11 (Section 11.6.2), we have provided case studies based on TSK and EnCase.

Box 7.10 File Carving – a Powerful Technique for Digital Forensics

File carving is the process of recovering files from an investigative target, potentially without knowledge of the file system structures. The process is based on information about the format of the file types of interest, as well as on assumptions about how files are typically laid out on block devices. If the file system metadata is used at all, it is typically used only for establishing cluster sizes and avoiding carving of undeleted files (which can be extracted without file carving).

Box 7.10 File Carving . . . (*Continued*)

File carving is an important technique for digital forensics investigation and for simple data recovery. By using a database of headers and footers (essentially, strings of bytes at predictable offsets) for specific file types, file carvers can retrieve files from raw disk images, regardless of the type of file system on the disk image. Perhaps more importantly, file carving is possible even if the file system metadata has been destroyed. File carving is a particularly powerful technique because files can be retrieved from raw disk images, regardless of the type of file system. File retrieval is possible even if the files system metadata has been completely destroyed. For example, a file deposited on a FAT partition can often be recovered even if the partition is reformatted as NTFS, then ext2, then FAT again, even if bad block checks (which are generally read-only operations) are applied. Although a file system's metadata can be quite fragile, file data is much more resilient. One limitation of the current generation of automatic file carvers is that a file's data must be contiguous to be carved properly.

With some manual intervention or additional work, even non-contiguous data can be carved. Luckily, modern file systems, such as ext2/3 (for Linux) and NTFS (for Windows), are actually quite kind to file carvers. This is because they strive to perform disk allocation which minimizes file fragmentation to reduce seek time and improve file system performance. Even under legacy file systems such as FATx, which are prone to fragmentation, the data of many files of modest size is likely to be unfragmented. This is because file fragmentation, if present, is on cluster boundaries and cluster sizes under FATx tend to be rather large.

File carvers ignore the file system and carve the images directly from data blocks. In cases of fragmented files, the carver returns an imperfect photo, but this image might be sufficient to identify the subject (see Fig. 7.9).

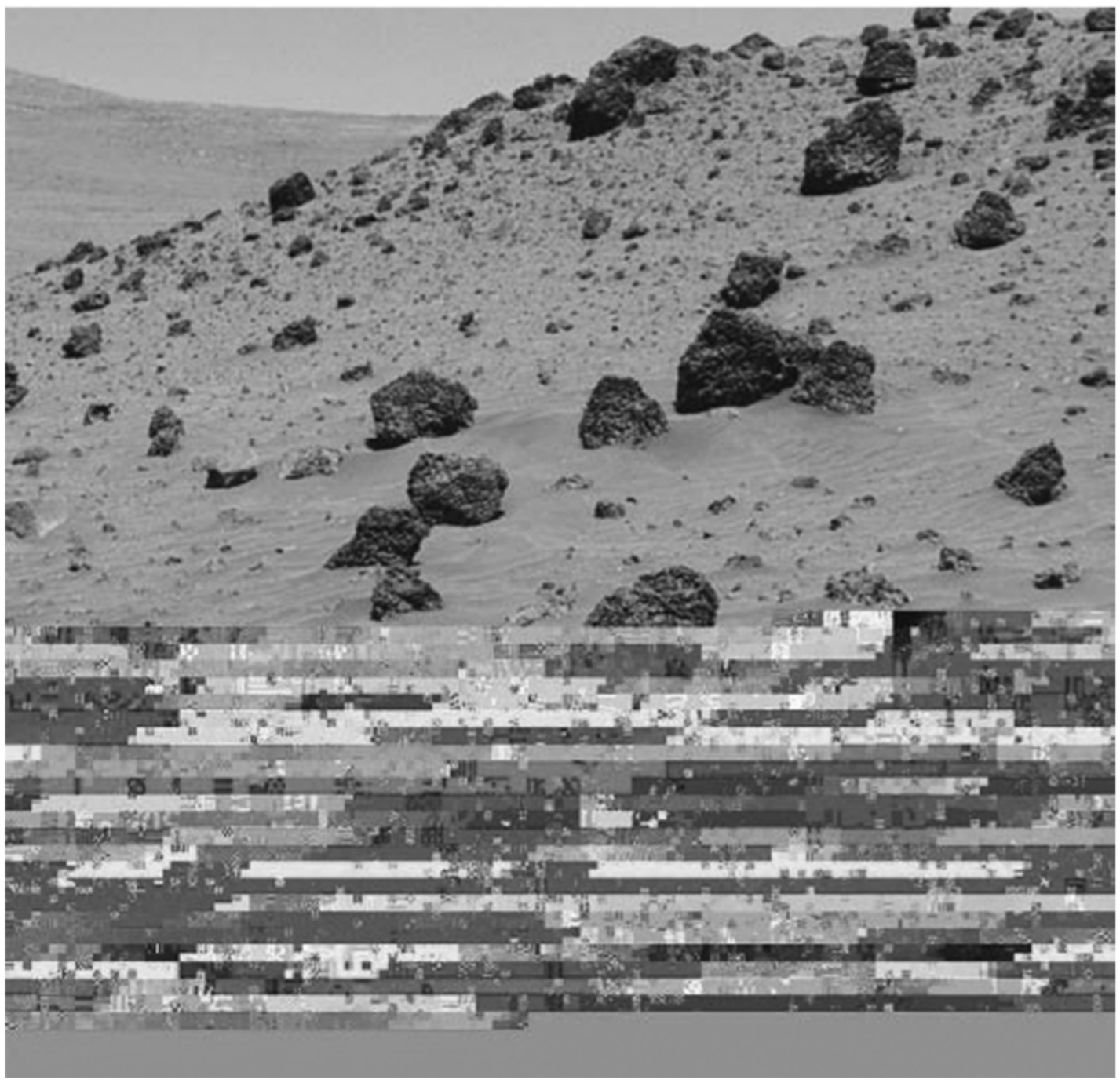


Figure 7.9 An image constructed from a fragmented file.
Source: Linux Magazine available at the link: http://www.linux-magazine.com/w3/issue/93/Foremost_Web.pdf (24 December 2009).

As it is not feasible to reconstruct every possible sequence to investigate all the sequences that may have produced the actual evidence in any particular case, forensics analysts focus only on large sets of sequences of events. They tend to characterize evidence aspects in those terms. For example, if the evidence includes a log file that appears to be associated with a file transfer, the name of the file transfer program included in the log file will typically be associated with common behavior of that program and will be used as a basis for the analysis (readers will understand this when they refer to cyberforensics investigation case – Digital Forensics Case Illustration 2: Analysis of Seized Floppy – the Drug Peddler Case in Section 11.6.2 of Chapter 11). The user identity indicated in the log file may be associated with a human or group, and this creates an initial attribution that can then be used as a basis for further efforts to attribute to the standard of proof required. Note, however, that the presence of this record in an audit trail does not mean that the program was ever run at all or that the thing the record indicates ever took place or that the user identified caused the events of interest. There are many possible sequences of events that could result in the presence of such a record. For example, without limiting the totality of possible event sequences, the record could have been placed there maliciously; it could be a record produced by another program that looks similar to the program being considered; it could have been a record produced by the program even though the file transfer failed; the record could have been produced by a Trojan Horse acting for the user or the record could be there because of a failure in a disk write that produced a crosslink between disk blocks associated with different sorts of records.

Analysis, interpretation and attribution of digital forensics evidence can be reconciled with non-digital evidence and digital forensics evidence can be externally stipulated or they can be demonstrated facts. For example, suppose the digital forensics evidence appears to show that person X was present at the local console of a computer in Los Angeles, California, two hours after passing through customs and immigration in London, UK. Suppose further that the network logs from distant systems show that the transfer indeed took place; even then, it is not a reasonable interpretation to assert that the individual was in Los Angeles. Another explanation is possible, whether there are two distinct individuals involved rather than a single individual, a remote control mechanism, alteration of multiple logs in multiple systems, alteration of customs and immigration logs, altered time clocks or any of a long list of other possibilities. Although in some venues, the "do not confuse me with the facts" approach may apply, in a legal setting, digital forensics evidence should reconcile with external reality.

Several open-source tools are available to conduct an analysis of open ports, mapped drives (including through an active VPN connection) and open or mounted encrypted files (containers) on the live computer system. Utilizing open-source tools and commercially available products, it is possible to obtain an image of these mapped drives and the open encrypted containers in an unencrypted format. Open-source forensics tools[6] for PCs include Knoppix and Helix by US e-fense Inc. These are Unix-based tools used in Linux environment. Commercial imaging tools include Access Data's Forensics Toolkit and Guidance Software's EnCase application.

The above-mentioned open-source tools mentioned can also scan RAM and Registry information to show recently accessed Web-based E-Mail sites and the login/password combination used. Additionally, these tools can also yield login/password for recently accessed local E-Mail applications including MS Outlook. In the event that partitions with Encrypted File System (EFS – file system driver that provides filesystem-level encryption in Microsoft Windows operating systems) are suspected to exist, the encryption keys to access the data can also be gathered during the collection process. With Microsoft's most recent addition, Vista and Vista's use of BitLocker and the Trusted Platform Module (TPM), it has become necessary in some instances to image the logical hard drive volumes before the computer is shutdown.

RAM can be analyzed for prior content after power loss. Although as production methods become cleaner, the impurities used to indicate a particular cell's charge prior to power loss are becoming less common. However, data held statically in an area of RAM for long periods of time are more likely to be detectable using these methods. The likelihood of such recovery increases as the originally applied voltages,

operating temperatures and duration of data storage increases. Holding unpowered RAM below –60°C will help preserve the residual data by an order of magnitude, thus improving the chances of successful recovery. However, it is impractical to do this during a field examination.

Now let us understand *types of digital analysis*. It is important, because, a digital investigation may encounter many formats of digital data and, therefore, there exist several types of analysis. The different analysis types are based on interpretation, or abstraction, layers, which are generally part of the data's design. For example, consider the data on a hard disk, which has been designed with several interpretation layers. The lowest layer may contain partitions or other containers that are used for volume management. Inside each partition is data that has been organized into a file system or database. The data in a file system is interpreted to create files that contain data in an application-specific format. Each of these layers has its own analysis techniques and requirements. Examples of common digital analysis types include:

Box 7.11 The RAID Levels

Explanation of RAID is important in forensics context. RAID data acquisitions are performed as part of computer forensics. RAID stands for *R*edundant *A*rray of *I*ndependent (or inexpensive) *D*isks. It is a category of disk drives that employs multiple drives in combination for fault tolerance and performance. Although use of RAID disk drives is frequent on servers, the use is not generally necessary for personal computers. With RAID, you can store the same data redundantly, that is, in multiple paces in a balanced way to improve overall performance. In late 1980s and early 1990s, computer information servers had to sustain a dramatic increase in capacity expectation in terms of amount of data served and stored on them. Storage technologies had become too expensive to place a large number of high-capacity hard drives in the servers. The response to this situation came through concept of RAID; subsequently RAID became very popular. Note that "data striping" means spreading out blocks of each file across multiple disk drives.

RAID was a system developed as a solution to link together a large number of low-cost hard drives with a view to form a single large capacity storage device that provided superior performance, storage capacity and reliability as compared to older storage solutions. Since then RAID became widely used and is deployed as an enterprise storage method in server markets. However, in the last 5 years it has become much more common in end-user systems.

Attractiveness of RAID comes from the fact that the array of disks distributes data across multiple disks; however, computer user and operating system sees the array as one single disk. The array of disks (RAID) can be set up to serve multiple purposes and offers many advantages such as redundancy, increased performance and lower costs.

Those who are technically savvy may know that there are number of different RAID levels as follows:

1. **Level 0:** This is nothing but a striped disk array without fault tolerance. It provides data striping (spreading out blocks of each file across multiple disk drives) but no redundancy. This results in an improved performance; however, it does not deliver fault tolerance. All data in the array is lost if one drive fails.
2. **Level 1:** This is mirroring and duplexing to provide disk mirroring. Level 1 provides double the rate of read transaction for single disks, but provides the same write transaction rate as single disks.
3. **Level 2:** This is error-correcting coding; however, it is not a typical implementation. This level is rarely used. It stripes data at the bit level rather than the block level.
4. **Level 3:** This is bit-interleaved parity. Level 3 provides byte-level striping with a dedicated parity disk. It is rarely used; probably because it cannot service simultaneous multiple requests.
5. **Level 4:** This is dedicated parity drive. Its use is common for implementation of RAID. Level 4 offers block-level striping (like Level 0) with a parity disk. If a data disk fails, the parity data is used to create a replacement disk. There is a disadvantage to Level 4 in that the parity disk can create write bottlenecks.

Box 7.11 The RAID . . . (*Continued*)

6. **Level 5:** This is block interleaved distributed parity. The idea here is to provide data striping at the byte level and also to stripe error correction information. Level 5 results in excellent performance and good fault tolerance. It is most popular among RAID implementation methods.
7. **Level 6:** This is independent data disks with double parity. This level provides block-level striping with parity data distributed across all disks.
8. **Level 0+1:** This is nothing but a mirror of stripes. It is not one of the original RAID levels. With this level used, two RAID 0 stripes are created and one RAID 1 mirror is created over them. The use of this level is typically seen for both replicating and sharing data among disks.
9. **Level 10:** This is stripe of mirrors. However, it is not considered to be an original RAID level. With this level, multiple RAID 1 mirrors are created, and a RAID 0 stripe is created over these.
10. **Level 7:** This is a trademark of STC (Storage Computer Corporation). It adds caching to Levels 3 or 4.
11. **RAID S:** This is also known a Parity RAID. It is an EMC Corporation's proprietary striped parity RAID system used in its Symmetrix storage systems.

For desktop computer systems, there are typically three forms of RAID used: RAID 0, RAID 1 and RAID 5.

1. **Media analysis:** It is analysis of the data from a storage device. This analysis does not consider any partitions or other operating system (OS)-specific data structures. If the storage device uses a fixed size unit, such as a sector, then it can be used in this analysis.
2. **Media management analysis:** It is analysis of the management system used to organize media. This typically involves partitions and may include volume management or redundant array of independent (or inexpensive) disks (RAID, see Box 7.11) systems that merge data from multiple storage devices into a single virtual storage device.
3. **File system analysis:** It is the analysis of the file system data inside a partition or disk. This typically involves processing the data to extract the contents of a file or to recover the contents of a deleted file.
4. **Application analysis:** It is the analysis of the data inside a file. Files are created by users and applications. The format of the contents is application-specific.
5. **Network analysis:** It is the analysis of data on a communications network. Network packets can be examined using the OSI Model to interpret the raw data into an application-level stream. Application analysis is a large category of analysis techniques because there are many application types. Some of the most common ones are as follows:
   - *OS analysis:* An OS is an application, although it is a special application because it is the first one that is run when a computer starts. This analysis examines the configuration files and output data of the OS to determine what events may have occurred.
   - *Executable analysis:* Executables are digital objects that can cause events to occur and they are frequently examined during intrusion investigations because the investigator needs to determine what events the executable could cause.
6. **Image analysis:** It was mentioned that the "image" is a single searchable file. Digital images are the target of many digital investigations because some are contraband. This type of analysis looks for information about where the picture was taken and who or what is in the picture. Image analysis also includes examining images for evidence of steganography (steganography in the context of forensics is discussed in Section 7.12).
7. **Video analysis:** Digital video is used in security cameras and in personal video cameras and webcams. Investigations of online predators can sometimes involve digital video from webcams. This type of analysis examines the video for the identification of objects in the video and the location where it was shot.

Reporting

Once the analysis is complete, a report is generated. The report may be in a written form or an oral testimony or it may be a combination of the two. Finally, evidence, analysis, interpretation and attribution must ultimately be presented in the form of expert reports, depositions and testimony. After extracting and analyzing the evidence collected, the results may need to be presented before a wide variety of audience including law enforcement officials, technical experts, legal experts, corporate management, etc. Depending on the nature of the incident or crime, it may become mandatory to present the findings in a court of law. It could be a police investigation or a presentation to appropriate corporate management or it could be an internal company investigation. As a result of the findings in this phase, it should be possible to confirm or discard the allegations with regard to particular crime or suspected incident. The presentation of evidence and its analysis, interpretation and attribution have many challenges.

Presentation of the report is more of an art than a science, but there is a substantial amount of scientific literature on methods of presentation and their impact on those who observe those presentations. Aspects ranging from the order of presentation of information to the use of graphics and demonstrations, all present significant challenges and are poorly defined. In general, reporting is a complex and tricky process and beyond the scope of discussion here. The following are the broad-level elements of the report:

1. Identity of the reporting agency;
2. case identifier or submission number;
3. case investigator;
4. identity of the submitter;
5. date of receipt;
6. date of report;
7. descriptive list of items submitted for examination, including serial number, make and model;
8. identity and signature of the examiner;
9. brief description of steps taken during examination, such as string searches, graphics image searches and recovering erased files;
10. results/conclusions.

In Chapter 11, we present illustrative examples of a digital forensics investigation report (refer to Section 11.6.3).

Testifying

This phase involves presentation and cross-examination of expert witnesses. Depending on the country and legal frameworks in which a cybercrime case is registered, certain standards may apply with regard to the issues of expert witnesses. Digital forensics evidence is normally introduced by expert witnesses except in cases where non-experts can bring clarity to non-scientific issues by stating what they observed or did. For example, a non-expert who works at a company may introduce the data he/she extracted from a company database and discuss how the database works and how it is normally used from a non-technical standpoint. To the extent that the witness is the custodian of the system or its content, he/she can testify to matters related to that custodial role as well.

Only expert witnesses can address issues based on scientific, technical or other specialized knowledge. A witness qualified as an expert by knowledge, skill, experience, training or education may testify in the form of an opinion or otherwise if (a) the testimony is based on sufficient facts or data, (b) the testimony is the product of reliable principles and methods, and (c) the witness has applied the principles and methods reliably to the facts of the case. If facts are reasonably relied upon by experts in forming opinions or inferences, the facts need not be admissible for the opinion or inference to be admitted; however, the expert may in any event be required to disclose the underlying facts or data on cross-examination.

Experts typically have very specialized knowledge about specific things of import to the matter at hand and anyone put up as an expert who does not have the requisite specialized knowledge can be seriously challenged by competent experts and counsel on the other side. Experts who are shown to be inadequate to the task are sometimes chastised in the formal decisions made by the courts, and such witnesses are often unable to work in the field for a period of many years thereafter because counsel for the opposition will bring this out at trial.

Now that we have explained the phases involved in the digital forensics investigation process, to conclude this section, we have summarized in Table 7.5 the outcomes from those phases. After that we have explained about precautions to be taken while collecting electronic evidence.

7.7.3 Precautions to be Taken when Collecting Electronic Evidence

So far we have established how important the digital/computer evidence is for cyberforensics. Therefore, collection of the evidence must happen with due care. Special measures should be taken while conducting a forensics investigation if it is desired for the results to be used in a court of law. One of the most important

Table 7.5 | Digital forensics – phase-wise outputs

Phase	*Activities/Processes*	*Outputs*
Evidence Preparation and Identification	• Monitoring authorization and management support, and obtain authorization to do the investigation. • Ensuring that operations and infrastructure are able to support an investigation. • Providing a mechanism for the incident to be detected and confirmed. • Creating an awareness so that the investigation is needed (identify the need for an investigation). • Planning for getting the information needed from both inside and outside the investigating organization. • Identifying the strategy, policies and previous investigations. • Informing the subject of an investigation or other concerned parties that the investigation is taking place.	Plan Authorization Warrant Notification Confirmation
Collection and Recording, Preserving and Transportation	• Determine what a particular piece of digital evidence is, and identifying possible sources of data. • Determine where the evidence is physically located. • Translating the media into data. • Ensuring integrity and authenticity of the digital evidence, for example, write protection, hashes, etc. • Packaging, transporting and storing the digital evidence. • Preventing people from using the digital device or allowing other electromagnetic devices to be used within an affected radius. • Recording the physical scene. • Duplicating digital evidence using standardized and accepted procedures. • Ensuring the validity and integrity of evidence for later use.	Crime type Potential Evidence Sources Media Devices Event

(*Continued*)

Table 7.5 | (*Continued*)

Phase	*Activities/Processes*	*Outputs*
Examination/ Investigation and Analysis, Interpretation and Attribution	• Determining how the data is produced, when and by whom. • Determine and validating the techniques to find and interpret significant data. • Extracting hidden data, discovering the hidden data and matching the pattern. • Recognizing obvious pieces of digital evidence and assessing the skill level of suspect. • Transform the data into a more manageable size and form for analysis. • Confirming or refuting allegations of suspicious activity. • Identifying and locating potential evidence, possibly within unconventional locations. • Constructing detailed documentation for analysis and drawing conclusions based on evidence found. • Determining significant based on evidence found. • Testing and rejecting theories based on the digital evidence. • Organizing the analysis results from the collected physical and digital evidence. • Eliminating duplication of analysis. • Build a timeline. • Constructing a hypothesis of what occurred, and comparing the extracted data with the target. • Documenting the findings and all steps taken.	Log files, file Events log Data Information
Presentation and reporting	• Preparing and presenting the information resulting from the analysis phase. • Determine the issues relevance of the information, its reliability and who can testify to it. • Interpreting the statistical from analysis phase. • Clarifying the evidence and documenting the findings. • Summarizing and providing explanation of conclusions. • Presenting the physical and digital evidence to a court or corporate management. • Attempting to confirm each piece of evidence and each event in the chain either along with each other, or independent of one evidence and/or other events. • Proving the validity of the hypothesis and defend it against criticism and challenge. • Communicating relevant findings to a variety of audiences (management, technical personnel, law enforcement).	Evidence, Report
Disseminating the case	• Ensuring physical and digital property is returned to proper owner. • Determining how and what criminal evidence must be removed. • Reviewing the investigation to identify areas of improvement. • Disseminating the information from the investigation. • Closing out the investigation and preserving knowledge gained.	Evidence Explanation New policies and investigation Procedures Evidence disposed Investigation closed

measures is to ensure that the evidence has been accurately collected and that there is a clear chain of custody right from the scene of the crime to the investigator and ultimately to the court (the "chain of custody" concept is explained in Section 7.8 – see Box 7.4 and Box 7.12).

In order to comply with the need to maintain the integrity of digital evidence, certain rules must be complied with. In general, the following principles are applicable:

1. **Principle 1:** No action taken by law enforcement agencies or their agents should change data held on a computer or storage media, which may subsequently be relied upon in court.
2. **Principle 2:** In exceptional circumstances, where a person finds it necessary to access original data held on a computer or on storage media that person must be competent to do so and be able to give evidence explaining the relevance and the implications of his/her actions.
3. **Principle 3:** An audit trail or other record of all processes applied to computer-based electronic evidence should be created and preserved. An independent third party should be able to examine those processes and achieve the same result.
4. **Principle 4:** The person in-charge of the investigation (the case officer) has overall responsibility for ensuring that the law and these principles are adhered to.

7.8 Chain of Custody Concept

Chain of custody is the central concept in cyberforensics/digital forensics investigation.

Recall the discussion we had in Section 7.4 and Box 7.4. A chain of custody is the process of validating how many kinds of evidences have been gathered, tracked and protected on the way to a court of law. It is essential to get in the habit of protecting all evidences equally so that they will hold up in court. Forensic investigation professionals know that if you do not have a chain of custody, the evidence is worthless. They learn to deal with everything as if it would go to litigation.

The purpose of the chain of custody is that the proponent of a piece of evidence must demonstrate that it is what it purports to be.

In other words, there is a reliable information to suggest that the party offering the evidence can demonstrate that the piece of evidence is actually, in fact, what the party claims it to be and can further demonstrate its origin and the handling of the evidence because it was acquired.

The chain of custody is a chronological written record of those individuals who have had custody of the evidence from its initial acquisition until its final disposition.

A chain of custody begins when an item of relevant evidence is collected, and the chain is matintained until the evidence is disposed off (Figs. 7.10 and 7.11). The chain of custody assumes continuous accountability. This accountability is important because, if not properly maintained, an item (of evidence) may be inadmissible in court.

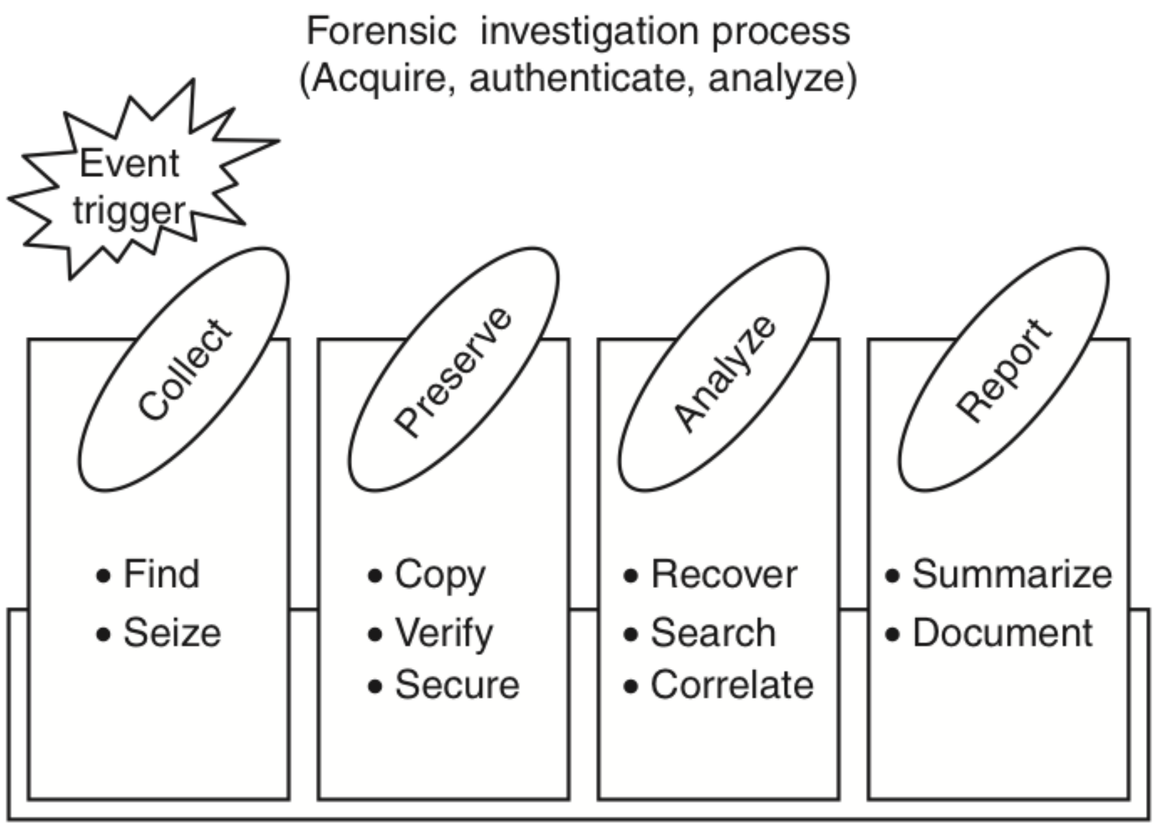


Figure 7.10 | Maintaining chain of custody – 1.

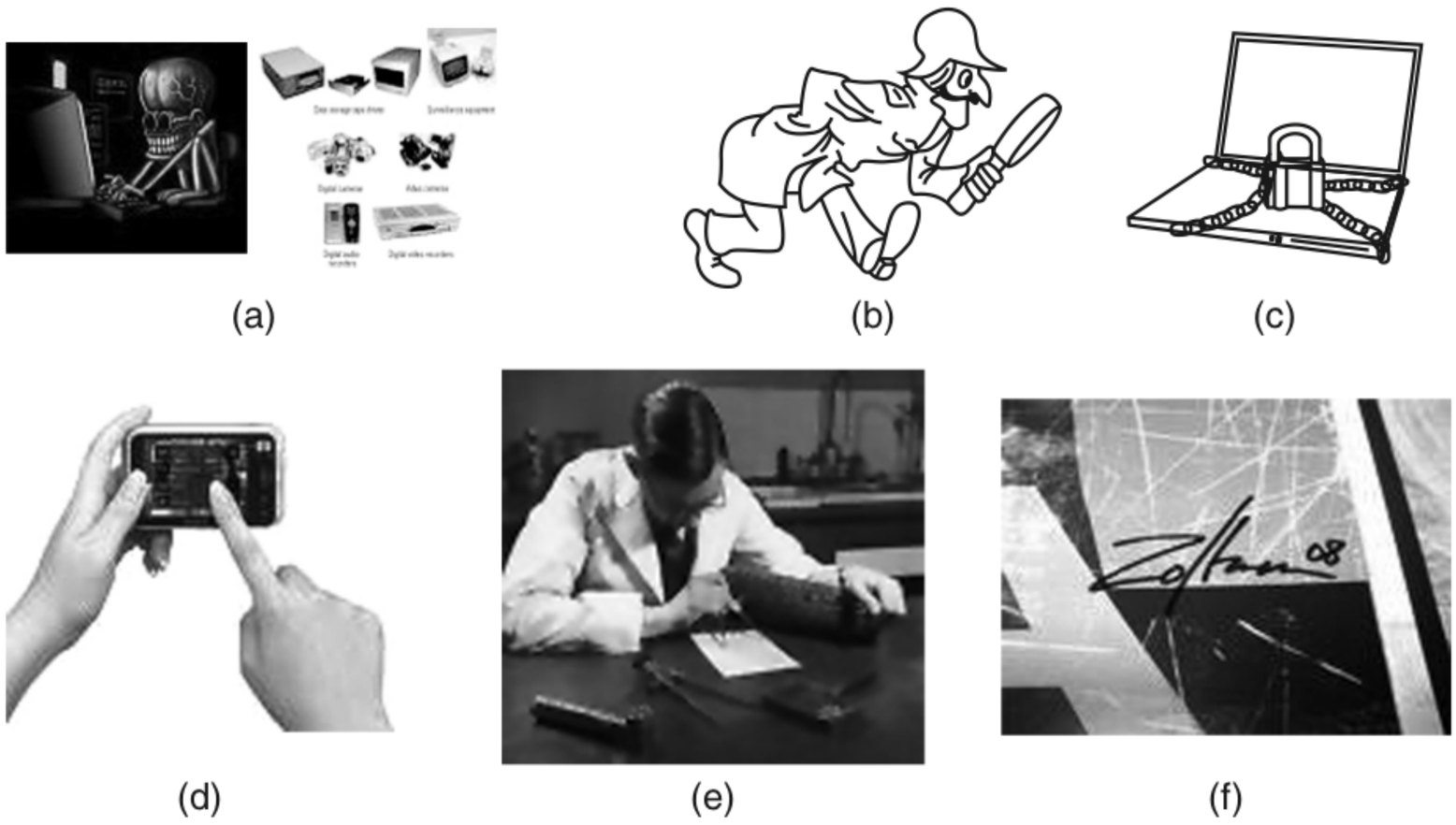


Figure 7.11 | Maintaining chain of custody – 2. (a) Source of evidence – where did it come from? (b) Who found it? (c) Where was it stored/locked up? (d) Who touched it/tampered with it? (e) What did they do to it? What did they do with it? (f) Human signature is always required.

Box 7.12 The Chain of Evidence Concept

A trial can be lost quickly if opposing counsel can show that the evidence chain of custody was violated. Before the world became computerized, proving chain of custody was easier. Attorneys and law enforcement filled out a chain of custody form that showed who had handled the evidence and the dates and times. With so much evidence coming from computers, especially in civil cases, many law firms need chain of custody software that is foolproof.

Box 7.12 The Chain . . . (*Continued*)

To avoid the risk of mishandling evidence, proper chain of custody procedure should always be strictly observed. This includes a thorough documentation of sources of data, the use of "write-blocking" devices to ensure no data changes take place inadvertently; initial forensics screening of disk drives for relevant data and making bit-for-bit copies of hard drives (images). The chain of custody also ensures that digital fingerprints (file hashes) match up at all stages of investigation, and that documentation (including photography and serial number inventory) of all evidence artifacts, as well as maintaining case logs for all evidence-related activities, is performed.

When preparing for trial, you need reliable chain of custody software supported by experts that know the latest rules. From the time we receive the media to the time we ship the results back, chain of custody is upheld. In a forensics data recovery, the senior individuals on forensics staff are the only ones who have access to the media. This allows for an efficient recovery and limits amount of individuals who have contact with the evidence.

Collecting information regarding the environment and use of the computer of machine under investigation, in an attempt to answer questions such as the following prior to the arrival of the forensics investigator, should be of utmost importance (refer to Figs. 7.10 and 7.11):

1. Who had access to the machine?
2. What level of authorization did all of those individuals having access to the machine have?
3. What was the machine used for?
4. What external devices did the machine connect to or interact with?
5. Which and how many servers did the machine "touch"?
6. Where and how will you store and safeguard the machine and the evidence after seizure?
7. Will you or an external third party be responsible for the storage and safeguarding of the seized machine and associated evidence?

At the very least, the evidence or property custody document should include the following information:

1. Name or initials of the individual collecting the evidence;
2. each person or entity subsequently having custody of it;
3. dates on which the evidence items were collected or transferred;
4. department (or Agency or Unit or Team) name and case number;
5. a brief description of the item seized.

7.9 Network Forensics

Recall the mention of network forensics in Section 7.5. We have already discussed that open networks can be the source of many network-based cyberattacks. In fact, a recent survey done[7] by the Cop-Tech forum (a joint initiative of cybercrime cell of police and city IT firms) in a leading IT business city revealed that 50% of the Wi-Fi Internet connection in the city continued to be unprotected! A situation like this leads to the point that network forensics professionals need to understand how wireless networks work and the fundamentals of related tehcnology. The topic of wireless network forensics is too vast and this section is aimed at providing an overview only here – from security perspective, they are the most risky ones. To know more about security of wireless networks, refer to Ref. #18, Books, Further Reading. In 1997, Marcus Ranum coined the term "wireless forensics."

Wireless forensics is a discipline included within the computer forensics science, and specifically, within the network forensics field. The goal of wireless forensics is to provide the methodology and tools required to collect and analyze (wireless) network traffic that can be presented as valid digital evidence in a court of law.

The evidence collected can correspond to plain data or, with the broad usage of VoIP technologies, especially over wireless, can include voice conversations. The wireless forensics process involves capturing all data moving over Wi-Fi network and analyzing network events to uncover network anomalies, discover the source of security attacks and investigate breaches on computers and wireless networks to determine whether they are or have been used for illegal or unauthorized activities. When performing wireless forensics, the security analyst must follow the same general principles that apply to computer forensics: identify, preserve and analyze the evidence to impartially report the findings and conclusions. There are many technical challenges for Wi-Fi traffic acquisition. That discussion is beyond the scope of this chapter. We have provided some useful resources on the topic of wireless network forensics in References Section.[8]

7.10 Approaching a Computer Forensics Investigation

From the discussion so far, we can appreciate that computer forensics investigation is a detailed science. The main phases are: *secure the subject system* (from tampering during the operation); *take a copy of hard drive* (if applicable); *identification and recovery of files* (including those deleted); *access/copy hidden, protected and temporary files*; *study "special" areas on the drive* (e.g., residue from previously deleted files); *investigate data/settings from installed applications/programs*; *assess the system as a whole, including its structure*; *consider general factors relating to the users activity*; *create detailed report*. Throughout the investigation, it is important to stress that a full audit log of your activities should be maintained. In Chapter 12 (in CD), there is guidance on the topics of building career in cybersecurity.

Cyberforensics experts go by their experience which shows that computer criminals always leave tracks (Locard's Exchange Principle – Box 7.5); it is just a matter of finding these tracks. However, this is not always easy. Computer technology is continuously evolving. Computers and other communication systems are becoming very complex. People businesses and organizations are connected through all kinds of networks. At the same time, computer crime techniques are becoming more sophisticated and better coordinated (refer to Chapter 4 – Tools and Methods Used in Cybercrime). If evidence collection is done correctly, it is much more useful in apprehending the attacker and stands a much greater chance of being admissible in the event of a prosecution.

Now, let us understand how a forensics investigation is typically approached and the broad phases involved in the investigation. The phases involved are as follows:

1. Secure the subject system (from tampering or unauthorized changes during the investigation);
2. take a copy of hard drive/disk (if applicable and appropriate);
3. identify and recover all files (including deleted files);
4. access/view/copy hidden, protected and temp files;
5. study "special" areas on the drive (e.g., the residue from previously deleted files);
6. investigate the settings and any data from applications and programs used on the system;
7. consider the system as a whole from various perspectives, including its structure and overall contents;
8. consider general factors relating to the user's computer and other activity and habits in the context of the investigation;
9. create detailed and considered report, containing an assessment of the data and information collected.

Certain things should be avoided during the forensics investigation depending on the nature of the computer system being investigated. For example, one should avoid changing date/time stamps (of files for example) or changing data itself. The same applies to the overwriting of unallocated space (which can happen on reboot for example). "Study it but Do NOT Change" is a useful catch phrase!

While there are some things that should be avoided, there are also other things that cannot be/should not be avoided before taking up a forensics investigation. The engagement contract and non-disclosure agreement (NDA) are some of those crucial not-to-forget things. This is because, customers of computer forensics laboratory must agree to be bound by terms and conditions of service set forth for any services offered by a computer forensics laboratory. In the context of a typical NDA, "customer" means the person, firm or company ordering products or services; "default" means any breach by either party of its obligations or any act, omission, negligence or statement by either party, its employees, agents or subcontractors arising out of or in connection with a contract and in respect of which either party may be legally liable; "the company" means the computer forensics laboratory; "engagement" means any job or jobs assigned to the computer forensics laboratory by the customer.

7.10.1 Typical Elements Addressed in a Forensics Investigation Engagement Contract

Typically, the following important elements are addressed before while drawing up a forensics investigation engagement contract:

1. **Authorization:** The customer will be asked to authorize the computer forensics laboratory or its agents to conduct an evaluation of the data/media/equipment onsite or offsite to determine the nature and scope of the engagement and to enable the company to provide an estimate of the cost of forensics investigation and/or the turnaround. Furthermore, the customer will be asked to agree on facilitating the engagement by providing all authorizations, security or legal clearances as required prior or throughout the course of the forensics investigation engagement.

   The customer will be required to authorize the computer forensics laboratory, its employees, independent contractors and agents to securely receive and transport the media/equipment/data to, from and between their premises required to deliver the services contracted by the customer.

   The customer will need to represent, warrant and affirm that he/she, or it is the owner or the authorized representative of the owner of the property or the equipment and all of the information and data stored on said property or equipment. By entering the NDA, the customer is supposed to declare that the representations are true and correct. The customer needs to agree to indemnify concerned computer forensics laboratory for any claims against the company related to any jobs assigned to the computer forensics laboratory whose services are engaged for the forensics investigation.
2. **Confidentiality:** The concerned computer forensics is supposed to use any information contained in the data, media and/or equipment provided to the company by the customer only for the purpose of fulfilling the engagement, and is expected to hold such customer information in the strictest confidence. Any confidential information disclosed by the customer under the agreement remains the owner's sole property, and computer forensics laboratory shall employ reasonable measures to prevent the unauthorized use of customer information. Such measures shall not be less than those measures employed by computer forensics laboratory in protecting its own confidential information. The involved computer forensics laboratory cannot disclose confidential information except to its employees, consultants or subcontractors as needed for the sole purpose of performing the engagement. Such information is not to be disclosed to any other party except as required by law. Computer forensics laboratory is to employ appropriate technical and organizational measures to safeguard any customer information, including personal data, and will act only on the instruction of the customer with regard to such information.
3. **Payment:** Customer agrees to pay the computer forensics laboratory all sums authorized from time to time by customer, which will typically include (a) charges for computer forensics laboratory

services; (b) reasonable travel and per diem expenses for onsite work; (c) shipping and insurance and actual expenses, if any, for parts; (d) media and/or off-the-shelf software used in the forensics service engagement. Unless otherwise agreed to in advance by computer forensics laboratory, all such sums are due and payable in advance by company check, bank wire transfer or credit card.

4. **Consent and acknowledgment:** Any consent required of either party becomes effective only if provided in a commercially reasonable manner; this includes but is not limited to, verbal authorization if followed by written confirmation, electronic or otherwise, by the computer forensics laboratory at the earliest possible opportunity. Customer needs to acknowledge that the equipment/data/media may be damaged prior to computer forensics laboratory receipt. Customer also needs to acknowledge that the efforts of the engaged computer forensics laboratory to complete the forensics investigation engagement may result in the destruction of or damage to the equipment/data/media. The computer forensics laboratory will not, however, assume responsibility for additional damage that may occur to the customer's equipment/data/media during computer forensics laboratory efforts to complete the engagement.
5. **Limitation of liability:** The concerned computer forensics laboratory will not consider itself to be liable for any claims regarding the physical functioning of the equipment/media or the condition or existence of data stored on the media supplied before, during or after services. In no event will the forensics laboratory be liable for any loss of data or loss of revenue or profits, goodwill or anticipated savings or any consequential loss whether sustained before, during or after services even if computer forensics laboratory has been advised of the possibility of damages or loss to persons or property.

The customer must be made aware of the inherent risks arising out of possible damage to media or equipment during the course of forensics investigation. Such risks include but are not limited to risks arising from possible destruction or damage to the media or equipment and/or data stored and inability to recover data, or inaccurate or incomplete forensics data recovery, including those that may result from the negligence of computer forensics laboratory involved in the investigation. The customer will be expected to agree that he/she will not hold responsible any of the involved computer forensics laboratory for any direct or indirect damage or loss of equipment or media or data loss. In the case of any damage or loss to the original media or equipment, the liability of the forensics laboratory shall be limited to providing the customer with similar media or equipment of comparable price or capacity.

The maximum aggregate liability of computer forensics laboratory to the customer whether in contract, tort or otherwise for any direct loss or damage including to tangible property suffered by the customer as a result of any default of computer forensics laboratory shall be limited in aggregate to the lesser of the stated sum or an amount equal to the sums paid by the customer under the contract during the preceding number of days stated in the contract (typically 30 days).

Any advice or recommendations given to the customer by the forensics laboratory or its employees or agents as to storage, application and use or preference of the equipments which is not confirmed in writing by the computer forensics laboratory is followed or acted upon entirely at the customer's own risk. Accordingly computer forensics laboratory shall not be liable for any such advice or recommendation which is not so confirmed. Although the computer forensics laboratory is to make every effort to preserve the integrity of any data or equipment related to the engagement, the customer has to agree not to hold the forensics laboratory responsible for any accidental damages to the data or equipment in its possession including but not limited to surface scratches, deformations and cracks.

1. **Customer's representation:** Customer needs to warrant the forensics laboratory that he/she is the owner of, and/or has the right to be in possession of, all equipment/data/media furnished to the laboratory and that collection, possession, processing and transfer of such equipment/data/media are in compliance with data protection laws to which customer is subject to.

2. **Legal aspects/the law side:** Both the parties need to agree that the agreement shall be governed by prevailing law in every particular way including formation and interpretation and shall be deemed to have been made in the country where the contract is signed.
3. **Data protection:** The computer forensics laboratory (engaged in the investigation) will hold the information that the customer has given verbally, electronically or in any submitted form for the purpose of the forensics investigation to be carried out as per contracted services from the forensics laboratory. Customer may apply for a copy of the information that the laboratory hold about customer and customer has the right to have any inaccuracies corrected.
4. **Waiver/breach of contract:** The waiver by either party of a breach or default of any of the provisions on this agreement by either party shall not be construed as a waiver of any succeeding breach of the same or other provisions, nor shall any delay or omission on the part of either party to exercise or avail itself of any right, power or privilege that it has, or may have hereunder operates as a waiver of any breach or default by either party.

7.10.2 Solving a Computer Forensics Case

A real-life example, showing how a case is solved using forensics, is available in Section 11.6.2 (Digital Forensics Case Illustration 2: Analysis of Seized Floppy – the Drug Peddler Case). As for this chapter, we summarized this section by presenting the steps involved in solving a computer forensics case. These are just some broad illustrative steps and they may vary depending on the specific case in hand.

1. Prepare for the forensics examination.
2. Talk to key people to find out what you are looking for and what the circumstances surrounding the case are.
3. If you are convinced that the case has a sound foundation, start assembling your tools to collect the data in question. Identify the target media.
4. Collect the data from the target media. You will be creating an exact duplicate image of the device in question. To do this, you will need to use an imaging software application like the commercial *EnCase* or the open-source *Sleuth Kit/Autopsy*.[9]
5. To extract the contents of the computer in question, connect the computer you are investigating to a portable hard drive or other storage media and then boot the computer under investigation according to the directions for the software you are using. It is imperative that you follow the directions precisely because this is where the "chain of custody" starts (refer to Section 7.8, Figs. 7.10 and 7.11, and Boxes 7.4 and 7.12). Make sure that you use a write-blocking tool when imaging the media under investigation. This makes sure that nothing is added to the device when you are creating your image.
6. When collecting evidence, be sure to check E-Mail records as well. Quite often, these messages yield a great deal of information (see Section 7.6 for E-Mail forensics).
7. Examine the collected evidence on the image you have created. Document anything that you find and where you found it. There are tools available to help look into open files, encrypted files, and mapped drives and to even analyze network communications. You can look into both commercial products and open-source products.[6]
8. Analyze the evidence you have collected by manually looking into the storage media and, if the target system has a Windows OS, check the registry. Be sure to look into Internet searches as well as E-Mail and pictures that are stored on the target computer. Many times, criminals will hide incriminating information in pictures and E-Mails through a process called *steganography* (see Section 7.12).
9. Report your findings back to your client. Be sure to provide a clear, concise report; this report may end up as evidence in a court case.

7.11 Setting up a Computer Forensics Laboratory: Understanding the Requirements

There are four broad types of requirements, namely, the physical space, the hardware equipment, the software tools and the forensics procedures to be followed to aid those involved in the cybercrime investigation. Figures 7.12 and 7.13 show how a typical laboratory looks.

First of all there is a physical facility in which the laboratory is set up. This is meant to be the home base for secure storage of evidentiary materials, for the analysis of captured data, for the operation of cloned systems, for the production of final evidence reports and for the physical premises where the forensics professional will perform most of their duties and work. Therefore, it should be designed as a secure storage facility that can also house an office, an operational laboratory and a production facility all rolled into one. The lab home should also have a separate interview facility or a small office/cabin where interviews and/or collaborative investigative procedures can be carried out without disturbing any ongoing technical or forensics work. This is because an investigating officer or attorney with an in-depth knowledge of the case may have queries that can be answered more effectively in collaboration with the forensics investigator. The forensics professionals generally perform specific analysis and/or search actions to find the answer to questions posed by the investigating officer or attorney.

Physical floor space requirement for a forensics laboratory varies depending on the size of the group that will occupy it. The bottom line is that the laboratory space should be in a secure location or contain appropriate measures that will stop unauthorized access to the premises. It should have an adjacent and secure walk-in lock-up vault that can keep intruders from gaining access to its contents as well as to protect the contents from fire/heat, smoke, water and electromagnetic emanations and should generally not be near radio equipment. Figure 7.14 shows cyberforensics equipments and Fig. 7.15 shows different types of connectors that are used with forensics tools.

Figure 7.12 | Cyberforensics laboratory – 1.

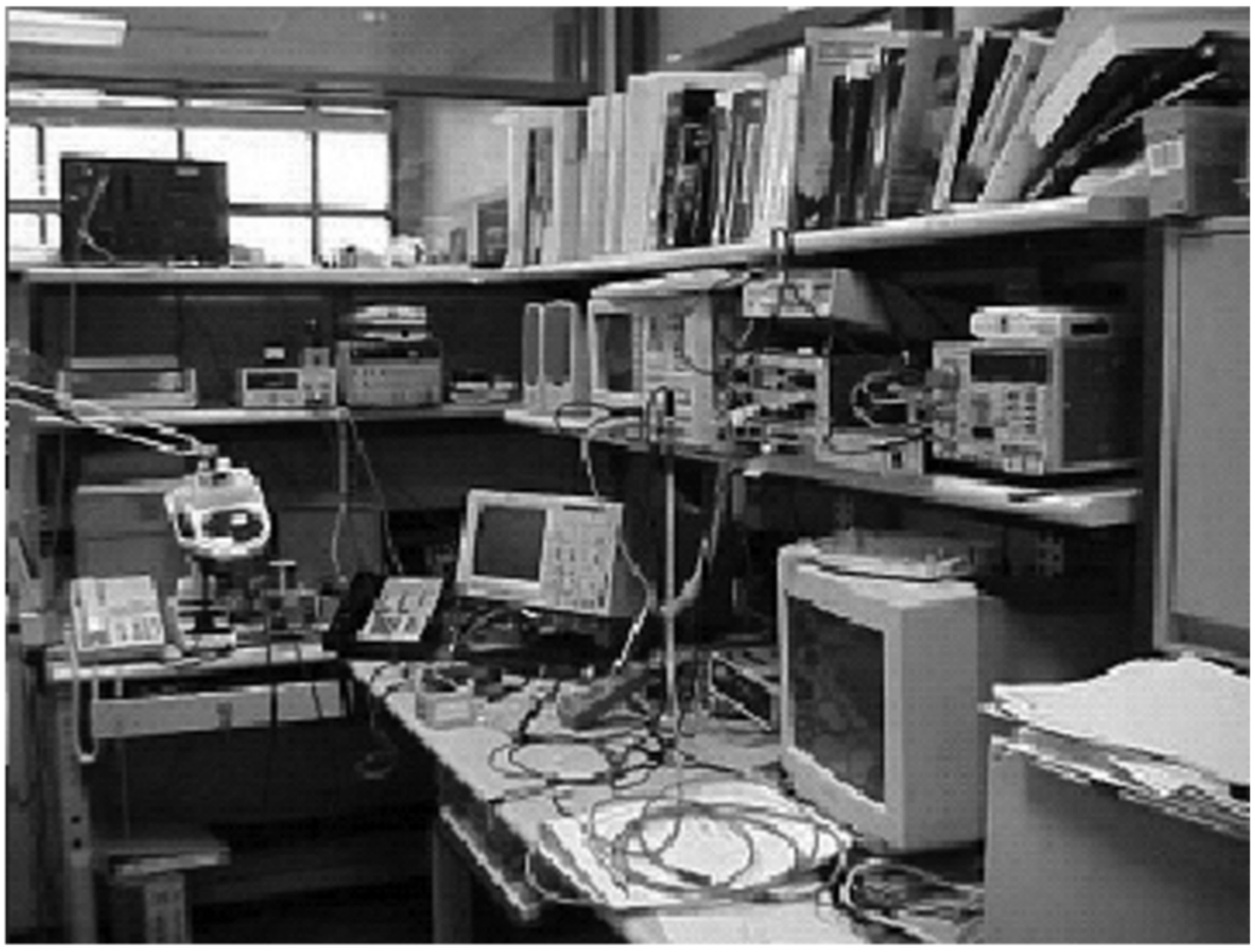


Figure 7.13 | Cyberforensics laboratory – 2.

Figure 7.14 | Cyberforensics equipments.

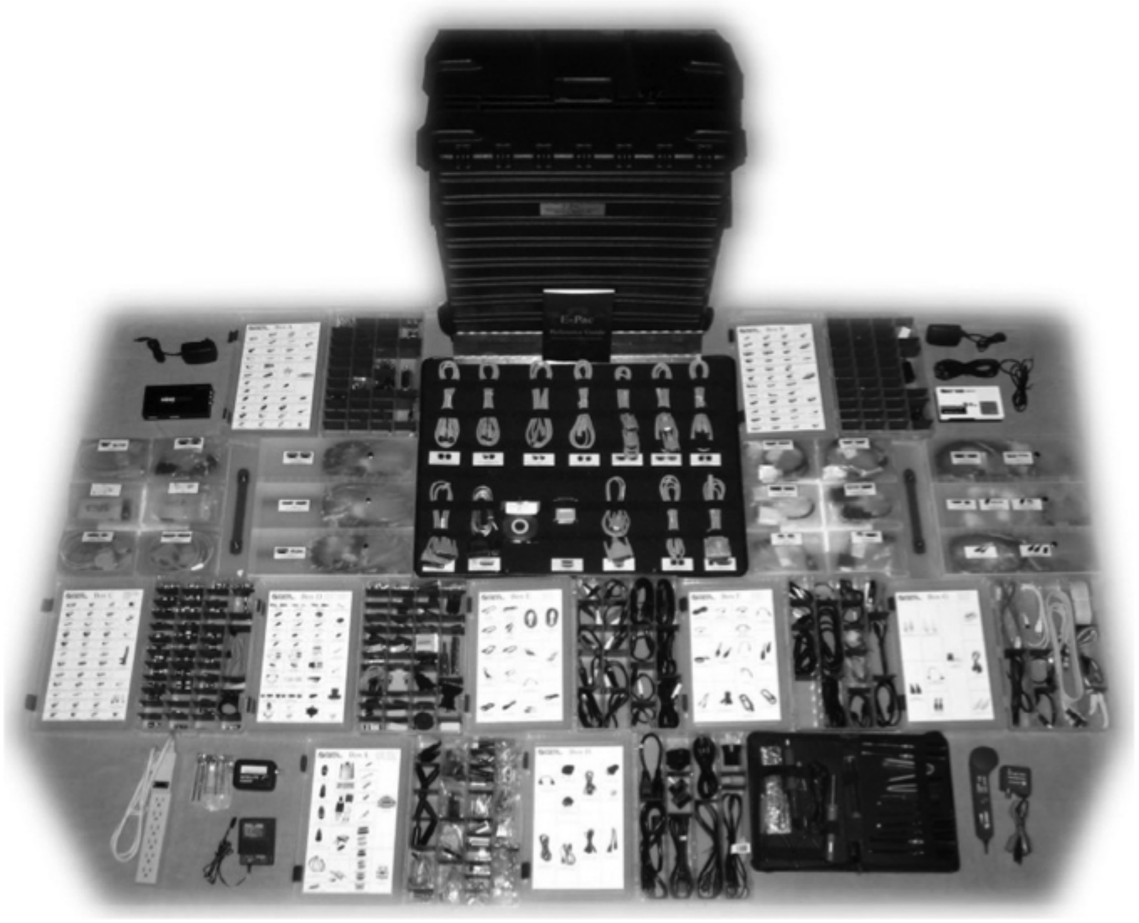


Figure 7.15 | Connectors used with cyberforensics tools.

The laboratory stores valuable items from investigation perspective – seized equipment, as well as official certified evidentiary copies of seized data, will be stored in the fire-proof vault. With the appropriate enforced sign-out/in procedures, it will serve to maintain the chain of evidence (refer to Section 7.8, Figs. 7.10 and 7.11, and Boxes 7.4 and 7.12. Therefore, access to the vault and its contents should be logged as well as locked when not under use, and should be monitored at all times. There must also be adequate lockable storage space for various specialized equipments that will, over the course of investigations, be acquired and used for other investigation. This space must also accommodate consumables like CDs, DVDs, removable hard drives of various capacities, paper, toner cartridges, etc.

Apart from the physical space requirement, another key requirement for a computer forensics laboratory is the hardware items. The laboratory requires a number of computers, including a network server with a large storage capacity (preferably configured for the standard removable hard drives).

The server in the computer forensics laboratory is typically used to manage, document and administer cases, store various software tools and manage one-off specialist hardware. The hardware that must be managed includes, for example, devices like Rimage DVD Publishing Systems[10] CD production units, CopyPro floppy disk readers, printers, etc. Figure 7.16 shows a disk duplicator equipment to give an idea about such hardware equipment required in a computer forensics laboratory.

The *evidentiary copy of seized data* is usually written to CD or DVD and, because of the large capacity of current hard drives, this can be a time-consuming process (see the paper quoted in References).[11] Some disk duplication devices (the Rimage, and other units like it), make it possible to create, number and label the media unattended, producing as many as 50 CD/DVDs without intervention. Capturing the contents of floppy disks is even more time consuming, and devices like the CopyPro also can acquire as many as 50 floppy disks without intervention. The capabilities of these types of devices may vary from model to model; the two mentioned (Rimage and CopyPro) are merely examples with specific capacities. There should

Figure 7.16 | Disk duplication equipment in a forensics laboratory.

also be separate Internet connection(s) but NEVER connected to the forensics server. The Internet will be useful for finding and sharing forensics information and techniques and for communicating with other forensics professionals. Staying abreast of developments in this field is a vital part of staying current and updated in the forensics arena. The Internet provides one source to help accomplish this need.

The forensics laboratory also needs a number of workstations for connecting them to the internal network. The number of workstations required in the laboratory will depend on how many forensics expert staff members are employed to work in the laboratory. The workstations will enable them to work on individual cases simultaneously and have access to the shared devices and resources. Portable acquisition computers, that is, the *portable forensics kits* (Fig. 7.17) are also required as many times the forensics staff will need to work on the crime incident site. Ideally, each portable kit should be configured identically with the standard forensics suite of tools and removable hard drives (the same standard hard drives as mentioned earlier) of various capacities.

Each kit should have a robust carrying case that can accommodate extra hard drives. An array of associated connection plugs, converters and a hard drive write blockers (such as "FastBloc" for example) are available. An EnCase Accessory [12] comes in two types – field edition and laboratory edition (Fig. 7.18). The field versions of the forensics kits will be used for onsite acquisition and/or seizure. It is usually preferable for acquisition to be undertaken in the controlled conditions of the laboratory; however, there are circumstances where this is not practical and an evidentiary acquisition must be undertaken onsite (e.g., when dealing with an Internet service provider).

These kits must also have an assortment of forms, such as labels, tags, pens, tape, evidence bags, an electronic camera, a GPSS, etc., all of which are vital to the process of seizure and acquisition. There will

Figure 7.17 | Portable forensics kits.

The lab edition

The field edition

Figure 7.18 | FastBloc – the Field Kit and the Lab Kit. (a) The lab edition and (b) the field edition. *Source:* Guidance Software's Product Catalogue/Data Sheet titled "The Next Generation of write blockers" can be accessed at the following link: http://www.forensics.ie/images/products/guidance_fastbloc_datasheet.pdf (8 January 2010).

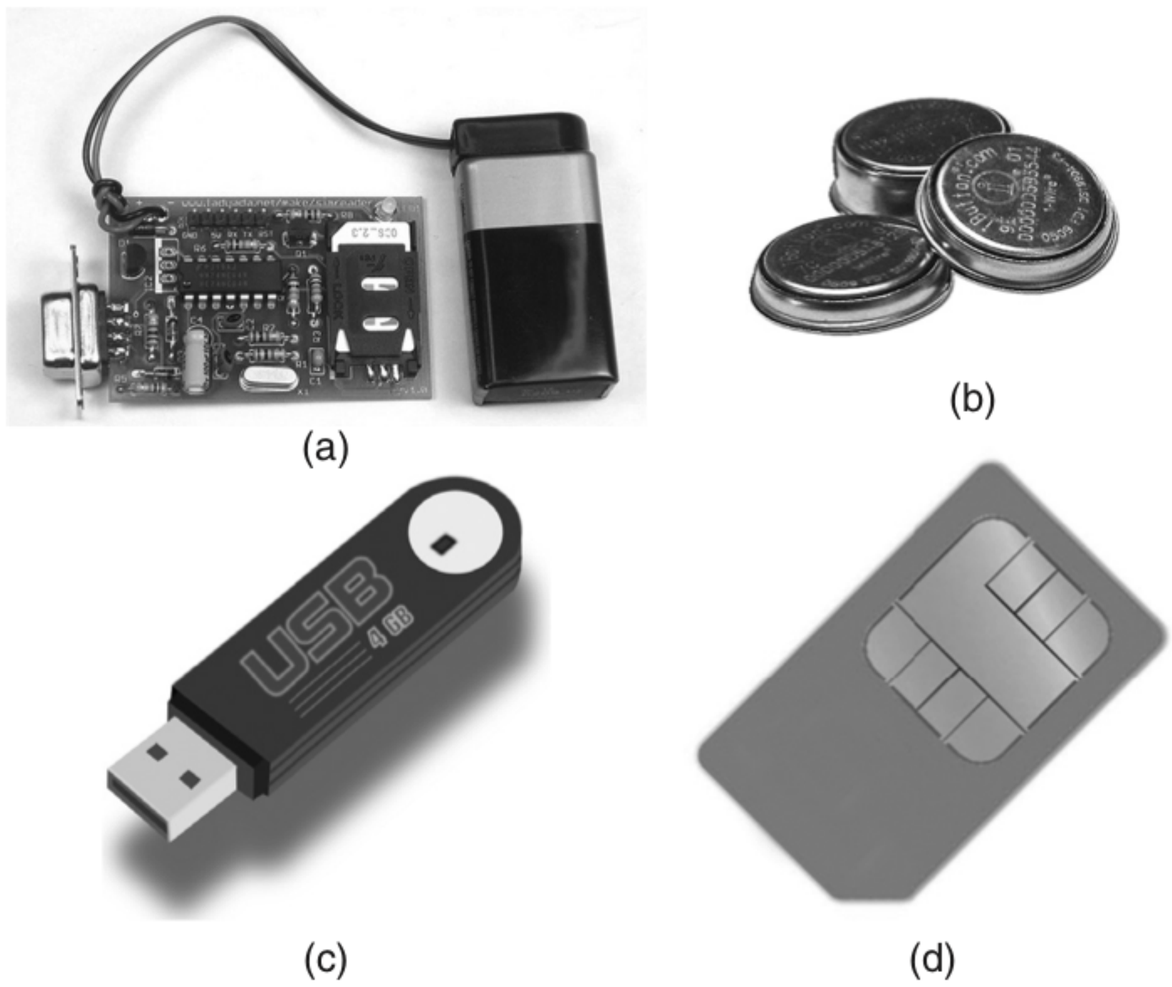


Figure 7.19 | (a) SIM card reader, (b) iButtons, (c) flash memory, (d) SIM card.

be an ongoing need to obtain devices, media, cables, converters and specialized media readers of various types, both for experimental purposes and for the acquisition of evidence from media other than hard drives or floppies (e.g., SIMs, flash memory of various description, iButtons, etc. – see Fig. 7.19). The hardware and physical premises constitute the largest outlay of funds. This, however, is an ongoing process and funds must be allocated regularly for the purchase of new hardware as it finds its way into the public arena.

On the software side, there are several requirements for setting up a forensics laboratory. The standard forensics software package, such as EnCase, WebCase, Forensics Tool Kit, Password Recovery Tool Kit, etc. are expensive products.

Some forensics kits require physical dongles to work and that these must be managed. A dongle is a physical security device (generally connected on the parallel port of the computer) that allows software to be used only when the device is present. In most cases, the capabilities of the software tool outweigh the nuisance and inconvenience of the required dongle. These products tend to be upgraded annually and, in each case, funds must be allocated for the upgrades. However, the software tools that are used comprise a far wider range than just those cited above. Many are freeware and many are not. No single tool performs the entire job of forensics acquisition, analysis and reporting; therefore, we need to use the right tool for the right task. Therefore, the forensics software tool library is extensive and it continues to grow. Having the right tool may make the difference between capturing relevant evidence and not being able to do so – as a result a case may be won or lost. In addition, the standard operational software will be required too; typically, this includes LAN software, OS, administrative software, graphics software, etc. These tools need to be upgraded occasionally; therefore, funds must be allocated for this ongoing process too. The continued cost of software acquisition and upgrades is usually smaller than that of hardware; however, it still constitutes a significant

portion of any forensics laboratory budget and must not be overlooked. The physical price of operating a forensics laboratory is not insignificant.

Orientation/mindset for procedure-based working is very important for a person working in a cyberforensics/computer forensics laboratory. One has to be very meticulous and should have a mindset for attention to details. This is important to be successful in digital forensics work like any other forensics work. Methods and procedures are an important part of operating a successful forensics laboratory.

The main issues that are attacked when evidence is presented in a court of law are credentials and methodology. In some countries, the court may prefer the forensics evidence from government appointed and/or neutral party laboratories rather than the evidence from private agencies where opportunities for manipulation/exploitation are perceived.

Close attention must be paid to strictly following and documenting the methodology adopted by the laboratory for the acquisition, analysis and reporting processes. Moreover, it is equally important to have a formal procedure in handling documents and control of evidence to be able to document the "chain of evidence." These are the main aspects that are unique to a forensics laboratory. There are other procedures and policies that should be in place and enforced, but they are the standards like Internet usage, E-Mail rules, back-up methods, etc. (See Appendix C in CD.)

7.12 Computer Forensics and Steganography

Steganography is the art of information hiding. The threat raised by steganography is very real. Its use is not easy to detect or intercept, as the information does not need to be broadcast across the Internet. The hidden message can reside unsuspectingly on a website, for example, and can be viewed from around the world. The technology is undoubtedly being used for other immoral purposes.

"Steganalysis" is of increasing importance to cybersecurity. Hiding messages in image data, called steganography, is used by criminals and by noncriminals as well to send information over the Internet. The term "steganography" originates from the Greek term for "covered writing."

Steganography was primarily used for "secret communications" to conceal the very existence of the message. Steganography is the art of "hiding information" in seemingly innocuous carriers in an effort to conceal the existence of the embedded information. Interestingly, steganography is not only the art of information hiding, but also the art and science of hiding the fact that communication is even taking place. Although "cryptography" is considered the predecessor of "steganography," steganography differs from cryptography. Cryptography is the art of secret writing, which is intended to make a message unreadable by a third party but does not hide the very existence of the secret communication. While steganography is separate and distinct from cryptography, there are many analogies between the two and, in fact, steganography is categorized as a form of cryptography since *hidden* communication certainly is a form of *secret* writing. In a way, "'steganography" and "cryptology" are guided by the same motive that is, to make a message useless for those who want to read it. When steganographic techniques are used, the *message cannot be seen* because it

is hidden and with cryptology techniques applied the *message cannot be deciphered* although it is not hidden. There is one distinct difference, cryptography is dependent on hiding the meaning of the message, whereas steganography is dependent on hiding the presence of a message altogether. Steganography accomplishes its objective through exploiting the Internet technology. The sheer size of the Internet and its vast amounts of data are what accomplishes this fete, and for this reason, it can be a very effective method of securing data transfer, which most of the time is used by the cybercriminals.

Steganography, by its very nature, poses a threat to forensics analysts, as they now must consider a much broader scope of information for analysis and investigation. Steganography is, therefore, considered as one of the "antiforensics methods." "Computer antiforensics" are methods of removal and subversion of evidence with the objective to mitigate results of computer forensics. There are other methods too that act as computer antiforensics, namely, encryption, self-splitting files plus encryption, database rootkits, BIOS rootkits, bypassing integrity checkers, etc. Let us understand "rootkits."

Box 7.13 Steganography, Cryptography and Digital Watermarks

It has been a long-standing desire of human beings to keep sensitive communications secret. Guillermito classification of steganographic software[13] mentions that the following is possible:

1. Adding data at the end of the carrier file;
2. inserting data in some junk or comment filed in the header of the file structure;
3. embedding data in the carrier byte stream, in a linear, sequential and fixed way;
4. embedding data in the carrier byte stream, in a pseudorandom way depending on a password;
5. embedding data in the carrier byte stream, in a pseudorandom way depending on a password, and changing other bits of the carrier file to compensate for the modifications induced by the hidden data to avoid modifying statistical properties of the carrier file.

As mentioned before, "Steganography" is the art of covered or hidden writing. The purpose of steganography is "covert communication" to hide a message from a third party. However, this is different from "cryptography," that is the art of "secret writing," which is intended to render a message unreadable by a third party. However, cryptography does not hide the existence of the secret communication. Although steganography is separate and distinct from cryptography, there are many similarities between the two; therefore, some people categorize steganography as a form of cryptography, as hidden communication is a form of secret writing.

Now let us understand the difference between encryption and steganography. Contents of information kept private and confidential using "encryption" so that only those who have the proper keys (privacy key and public key) can extract the secret contents. On the other hand, the sole purpose of using steganography is to hide the fact of secret message (possibly containing the evidence) if it exists. Therefore, the military calls this "covert communication" and the path for this kind of communication is called "covert channel."[14] Some good reference links on "covert channels" are provided in Ref. #10, Additional Useful Web References and in Refs. #14, #15 and #20, Articles and Research Papers, Further Reading.

Note yet another difference: steganography hides the covert message but not the fact that two parties are in communication. Steganography techniques generally involve placing a hidden message in some transport medium called the "carrier."

"Digital Watermarking"[15] is conceptually similar to steganography; however, its technical goals are different. Generally, only a small amount of repetitive information is inserted into the carrier and it is not necessary to hide the watermarking information. It is useful for the watermark to be able to be removed while maintaining the integrity of the carrier. Refer to Appendix N in CD.

Cryptography is discussed in Ref. #16, Books, Further Reading.

7.12.1 Rootkits

The term rootkit is used to describe the mechanisms and techniques whereby malware including viruses, Spyware and Trojans attempt to hide their presence from Spyware blockers, antivirus and system management utilities. Rootkits can be classified as – persistent rootkits, memory-based rootkits, user-mode rootkits and kernel-mode rootkits. These classifications are made depending on whether the malware survives reboot and whether it executes in user mode or kernel mode. Basically, a "rootkit" is a set of tools used after cracking a computer operating system that hides logins, processes, password, etc., which would carefully hide any trace that those commands normally display. Rootkits are installed after an attacker has exploited a system vulnerability and gained root access. Rootkits by themselves do not give an attacker root access; they only work after a system compromise. Rootkits consist of tools that generally have three functions: (a) maintain root access to the system, (b) hide the presence of the attacker and (c) attack (or accelerate attacks) against other systems. From attacker's perspective, rootkits serve following important functions:

1. The first, and primary, function of a rootkit is to maintain access to the compromised system. This access can happen via any communication channel from an easily detectable telnet shell to a secure shell to covert communication channels overlaid over commonly used protocols. An attacker who cannot maintain access to a host cannot exert his or her control over it.
2. The second main function of a rootkit is to hide, or otherwise obfuscate, the presence of the attacker. The ability to hide the presence of an attacker is what makes the rootkit such a powerful tool and is critical to the success of an attacker in maintaining root access to a system. This is achieved by removing evidence of the compromise and taking measures to misrepresent the system state to curious or confused system administrators. Removing evidence can be achieved through cleaning various log files and temporarily disabling any monitoring demons. How a rootkit chooses to hide its presence and the presence of the attacker is the qualifying characteristic of the rootkit. For example, an attacker could replace commonly used system executables, re-route system calls and install a loadable kernel module in a single rootkit installation. Attackers typically choose more than one method of hiding their presence (like for "defense in depth" think "offense in depth"!).
3. The third function of a rootkit is to perform actions that meet the attacker's objectives, mainly by attacking or aiding in attacking other systems. This usually means compromising host security (e.g., by using keyloggers), gathering packet traces on the local network, performing vulnerability scans or even launching automated attacks from the compromised host.

Now let us understand "binary rootkits."

Binary rootkits take administrative utilities and modify them to hide specific connections, processes and activities of specific users. These utilities would also include tools to provide root access to a particular user or when supplied with a particular argument. For example, an attacker could modify the "w" binary to hide his/her user account while logged on, the "ps" command to hide any processes he/she is running, and the "su" command to always allow root access whenever a specific password is supplied. These changes are not limited to the binary files only but source files can also be directly modified by attackers. If the source code is not examined for these inconsistencies, a rebuilt binary that is assumed to be "clean" can be compromised. When the binary tools are deployed, they are often placed inside of a hidden directory until the administrative programs can be fully compromised. An aspect of social engineering is used when creating these directories. Some of the common locations include confusing or unsuspecting directory names, such as /dev/.hdd or /dev/.lib. Others include commonly overlooked directory names such as /etc/… or /etc/".. " (dot-dot-space). There are, of course, defenses available for rootkits. Let us take a brief look at them.

Binary rootkits can be defeated through the use of file integrity scanners. Most file integrity scanners work by computing checksums, cryptographic checksums or even digital signatures. The file signatures must be created when the system is in an uncorrupted state and are useless if not prepared before a rootkit has been installed. The cryptographic checksum itself is not immune to attack, and care should be taken that it ise not recomputed with the corrupted media and overwritten (by writing them to immutable media and storing them offline, for example). However, successful use of this technique also means that any legitimate patches or updates must be followed up by recomputing the checksum (a technique that may fail in practice for frequently updated systems). Binary rootkits can also be detected by system integrity tools. However, certain commonly changed or temporary directories may be ignored by the system integrity scans, so care should be taken by the system administrator to inspect these directories for unusual activities. Further discussion of each of these methods is beyond the scope of this chapter. Readers interested in greater details of these topics should refer to Ref. #4, Additional Useful Web References, Further Reading and Ref. #4, Video Clips, Further Reading.

7.12.2 Information Hiding

Let us now have an overview of some characteristics of information hiding and then we discuss about steganalysis methods for determining the existence of and potential locations of hidden information. Today we are in the "digital age" where the messages can be hidden in images, sound files, text and other digital objects. These messages are invisible to a casual observer. The use of "steganography" on public networks, such as the Internet, is unknown due to its stealthy nature. Unless it is being actively looked for, one would not know that it is there. For example, pop-up ads, photos on sales and purchase portals such as eBay, and

Box 7.14 Hair Splitting Experience for Forensics Investigation Experts!

"Self-splitting files" is a method developed by the Intruders Tiger Team Security. The technique consists of the following:

1. A framework that interprets input files (binary or texts) and places them in sectors marked as bad blocks (but in the truth, they are not).
2. An input file (binary or text) that is divided in several asymmetric parts that are encrypted and placed outside of order in bad blocks;
3. An input file (binary or text) that is wiped after having been processed.
4. An input file (binary or text) that can be in a HTTP(s) or FTP server, and is processed without touching the disk (all in memory).
5. A library that allows the easy interaction of framework with sniffers, etc.
6. A tool that returns a sequence of blocks and a pseudorandom key that must be known to read, remount or execute the file.

Another technique used is the Wipe utility; it makes things difficult for the cyberforensics experts and investigators. Wipe is a name given to the method of safe deletion. In the current file systems, the files are not totally extinguished. When we delete a file (with "rm," "del," etc.), the field "link count" is set to zero and the field "deleted time" to the hour that the file is excluded. Therefore, the files can be easily recovered via software methods used in computer forensics. A method to make the recovery of files difficult is the use of the Wipe that does nothing more than open the file and overwrite it several times with pseudorandom (or predefined) content and later unlinks them from "inode" and "directory entries."

Examples of utility are necrofile e klismafile (the Defiler's Toolkit). Even after Wipe's method, files can be recovered through a "ferromagnetic" phenomenon that is called "hysteresis loop." However, this type of computer forensics method requires equipment with highest cost, and the recover process being each day more and more complex because of the increase of hard disk density and the number of overwrites made by the Wipe.

other recreational sites, all have the potential of containing hidden messages. An average organization is not expected to take on the vast responsibility of searching large websites and newsgroup areas for potential steganographic images. Most organizations can and do, however, monitor network traffic that is entering and exiting the local area network.

The first question is why one would want to "hide" information. In the first category, there are those who are trying to protect their intellectual property rights. There is a high availability of information via the Internet and that makes it increasingly difficult to protect intellectual property and enforce copyright laws. The use of "digital watermarks" [15] provides a way to insert a copyright notice into a document or image. The watermark is often a small image or text that is repeated frequently throughout the document or image. A similar technique is to embed a digital fingerprint or serial number. Fingerprint presents a certain advantage in that it can be used to trace the copy back to the original and thus it is a powerful tool for prosecuting copyright violators. Therefore, this is about the first category of people who would be interested in "hiding" information.

In the second category are people who are interested in hiding information to convey information in a covert manner and avoid observation by unintended recipients. In this case, the hidden message is more significant than the "carrier" object that is used to transport it. Steganography is often compared to cryptography in its ability to restrict unauthorized access to information. Cryptography is used to encrypt or scramble the data in such a fashion that only the intended recipient can decrypt it. At the outset of this discussion, we explained the difference between "cryptography" and "steganography." There are three common approaches of hiding information in digital images[16] (a) least significant bit insertion, (b) masking and filtering and (c) algorithms and transformations.

Box 7.15 Hide and Seek in the World of Information Communication

An interesting question is "how do we hide information in the electronic age?" Computers use binary, a combination of zeros and ones to represent text and graphics. The ASCII (American National Standard Code for Information Interchange) is the de facto standard for representing text and certain control characters. ASCII uses one parity bit and seven data bits to represent each character in the English language. For example, an uppercase "A" is represented by 1000001. A digital image is composed of picture elements or "pixels." Each pixel contains information as to the intensity of the three primary colors: red, green and blue. This information can be stored in a single byte (8 bits) or in three bytes (24 bits). For example, in an 8-bit image white is represented by the binary value of 11111111 and black by 00000000.

Current information hiding techniques rely on the use of a cover object (image, document, sound file, etc.) − sometimes known as a carrier. The secret message is then broken down to its individual bits by a steganographic tool (stego-tool) and embedded in the cover object. Many tools will utilize a password or pass-phrase that is necessary to extract the hidden message and is referred to as a stego-key. The result of this process is known as the stego-object.

One would wonder where can information be hidden. Almost anywhere on the Internet! The standard protocol suite used on the Internet is the Transmission Control Protocol/Internet Protocol (TCP/IP). The headers used to transfer data between computers allow the use of flags and certain reserved fields. With the appropriate tool, information can be inserted into these fields. The advantage of this technique is that headers are rarely read by humans and thus makes an ideal place to hide data. The disadvantage of this method is that firewalls can be configured to filter out packets that contain inappropriate data in the reserved fields, thus defeating the steganographic transmission.

Another popular technique for hiding information is to include extra spaces in documents. These spaces may contain hidden characters. Again this is a simple technique for hiding information and consequently is easy to detect and defeat. By opening such a document in a word processor the unusual spacing becomes readily apparent. Reformatting the document can remove the hidden message. The use of audio files can provide a good carrier for hidden messages. By their very nature, sound files tend to be large in size and thus do not attract attention. In particular, the MP3Stego tool can be used to hide information and maintain nearly CD quality sound.

Criminals do a number of activities with their data – they delete data, destroy data, hide data or may also encrypt data – all with the purpose that the traces of their criminal deeds do not fall in the hands of the investigator. As for data destruction, "data sanitization" is an important term to understand. By definition, 'sanitization method' is the specific way of using a data destruction program to overwrite the data on a hard drive or other storage device. Technically speaking, there are other methods as well for destroying data. These methods are not based on software overwriting and are also referred to as "data sanitization methods." Often a question is asked as to which method for data sanitization is the best one. An often quoted method in this regard is the *DoD Data Sanitization Method* – for example the DoD 5220.22-M data sanitization method is usually implemented as a '3-pass method':

1. In Pass 1: "0" is written and the write is verified.
2. In Pass 2: "1" is written and the write is verified.
3. In Pass 3: a random character is written and the write is verified.

There seems almost a universal agreement that overwriting an entire hard drive once with a single character prevents recovery of data from a hard drive even if any software based file recovery method is used. It is believed that most data sanitization methods are over-kill. They claim that a single overwriting of data is adequate to prevent data recovery even if advanced; hardware-based methods are used for extracting information from hard drives. There, however, does not seem to be a universal agreement about this.

We conclude this discussion on the note that there are concerns that terrorists are using steganographic techniques for hiding their messages, terror-related events inside images, text that is displayed on the Internet sites. As the research in steganographic method advances, there can be hopes of abetting the secretive communication methods used by the terrorists in propagating their illicit messages on the Internet about their destructive work. To conclude steganography is one of the threats that needs to be considered and understood for possible future occurrences.

7.13 Relevance of the OSI 7 Layer Model to Computer Forensics

The OSI 7 Layer Model is useful from computer forensics perspective because it addresses the network protocols and network communication processes. The basic familiarity with the OSI 7 Layer Model is assumed for the discussion in this section. To know more about OSI Model and network protocols, refer to Ref. #15, Books, Further Reading. The OSI Model depicted in Fig. 7.20 shows Internet Protocols involved at each of the seven layers. Forensic analyst needs to very well understand how the TCP/IP works.[17] To effectively perform forensics network analysis, forensics professionals must have a strong understanding of underlying network processes and protocols. Explaining these concepts is beyond the scope of this chapter. Interested readers may refer to the relevant chapter mentioned in Ref. # 15, Books, Further Reading.

In Chapter 1, it was mentioned that "hacking" is one of the cybercrimes – it involves unauthorized access to a computer system. Effective penetration testing requires complete understanding of the methods and motivations of a typical hacker. The steps taken by attackers who hack networks are shown in Fig. 7.21; each is briefly described.

7.13.1 Step 1: Foot Printing

Foot printing includes a combination of tools and techniques used to create a full profile of the organization's security posture. These include its domain names, IP addresses and network blocks. Some of the tools used

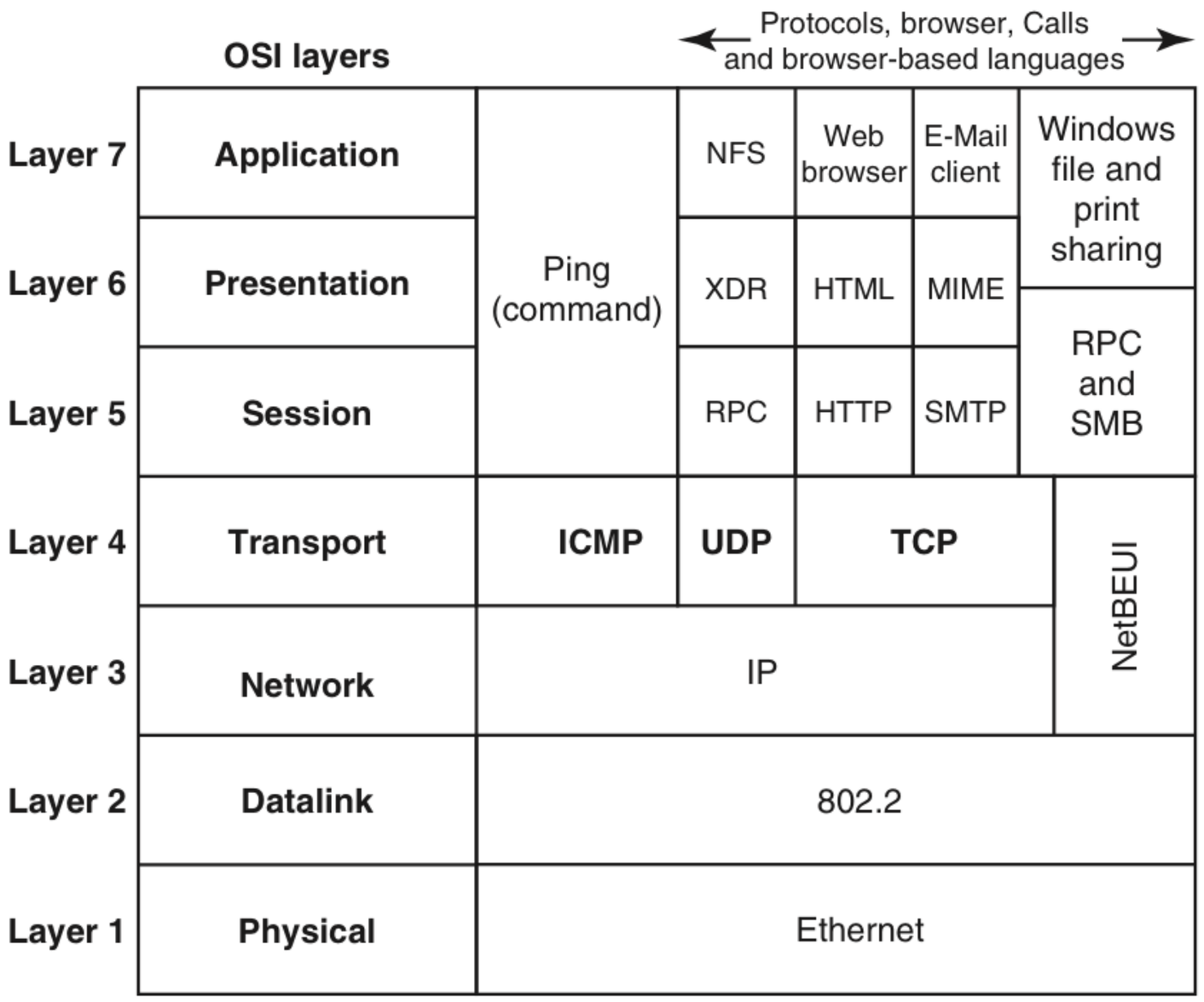


Figure 7.20 The OSI 7 Layer Model with Internet Protocols. Abbreviations: UDP, User Datagram Protocol; IP, Internet Protocol; ICMP, Internet Control Message Protocol; TCP/IP, Transmission Control Protocol/Internet Protocol; RPC, remote procedure calls; HTTP, Hypertext Transfer Protocol; SMTP, Simple Mail Transfer Protocol; XDR, eXternal data representation; HTML, hypertext markup language; MIME, multipurpose Internet mail extensions; SMB, server message block; NFS, network file system.

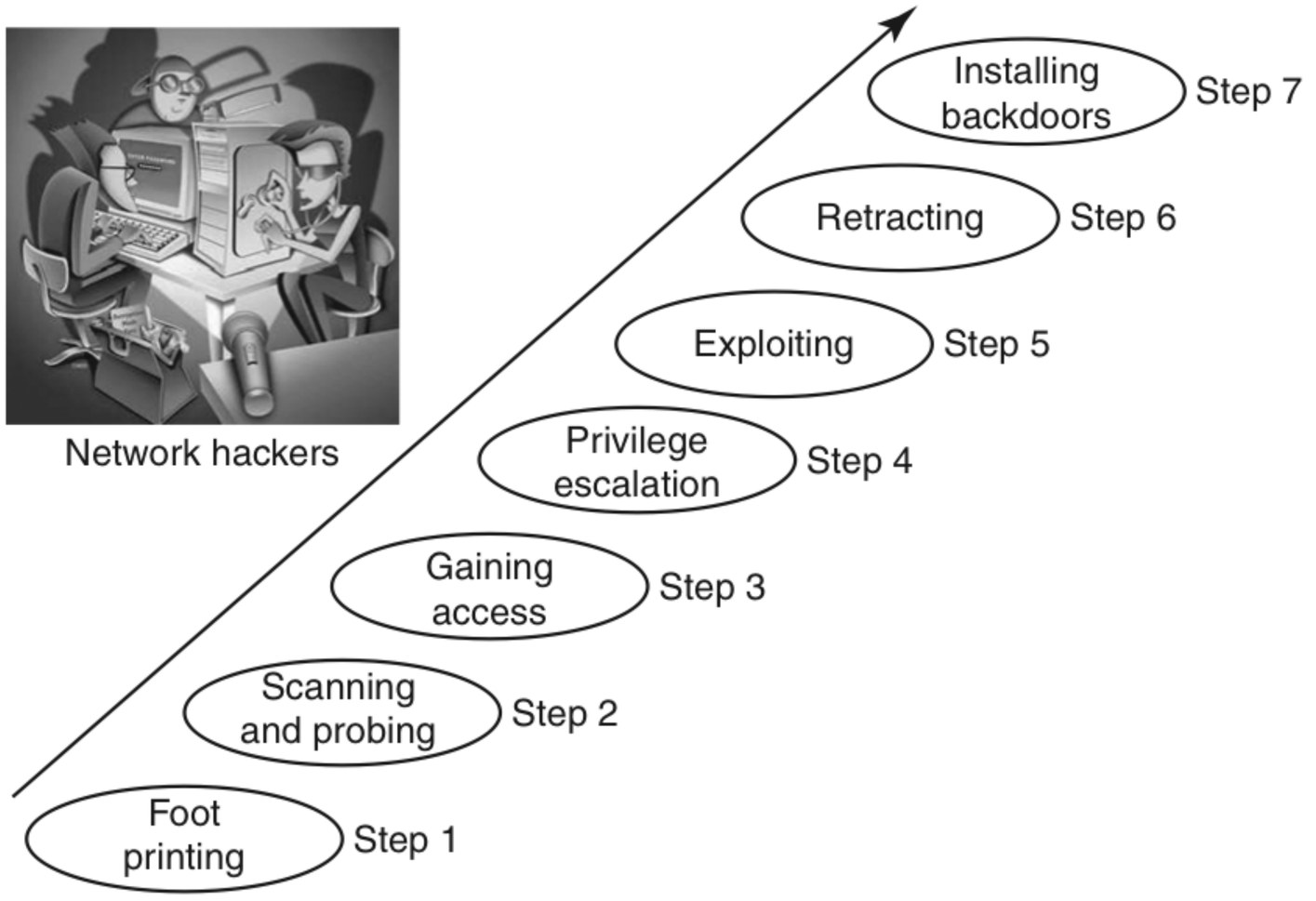


Figure 7.21 Network hacking steps.
Note: This is being illustrated how hacking takes place.

for foot printing are SamSpade, nslookup, traceroute and neotrace (refer to Tables 2.1 and 2.2 in Chapter 2 for a list of tools used during the "active attack" phase). Once the IP address and domain names are known, a hacker will typically perform a series of scans or probes to gather more information about individual machines for the purpose of gaining unauthorized access to the system at a later date. These scans may include ping sweeps, TCP/UDP scans and OS identification. All of these actions can be performed with a single tool called Nmap. Nmap is a free security scanner written by Fyodor.

The tool called "Metasploit" (mentioned in Table 4.1, Chapter 4) was developed as an automated tool to provide useful information to people who perform penetration testing. To know more on penetration testing, refer to Ref. #19, Books, Further Reading. In addition to providing this service, it has also become an automated tool to gain access to insecure computer systems. Metasploit groups exploit both OS and application. Performing a successful exploit has now become as simple as finding out the target's OS or application, entering the target's IP address and pressing a button.

7.13.2 Step 2: Scanning and Probing

The hacker will typically send a ping echo request packet to a series of target IP addresses. As a result of this exploratory move by the hacker, the machines assigned to one of these IP address will send out echo response thereby confirming that there is a live machine associated with that address. Similarly, a TCP scan sends a TCP synchronization request to a series of ports and to the machines that provide the associated service to respond. Finally, using tools such as Nmap, the hackers can determine device type and OS details by interpreting the responses. System scanning and probing can provide insights about the easiest path into the targeted system to a hacker. Gaining access takes advantage of specific security weaknesses in the system to allow access via an individual machine.

Uneducated hackers were known as "script kiddies" (refer to Chapter 10) because they could gain access to target machines via scripts that were published on hacker websites. These hackers are increasingly known as "click kiddies" because the process has become as easy as clicking a button. Figure 7.22 shows the network hacker categories.

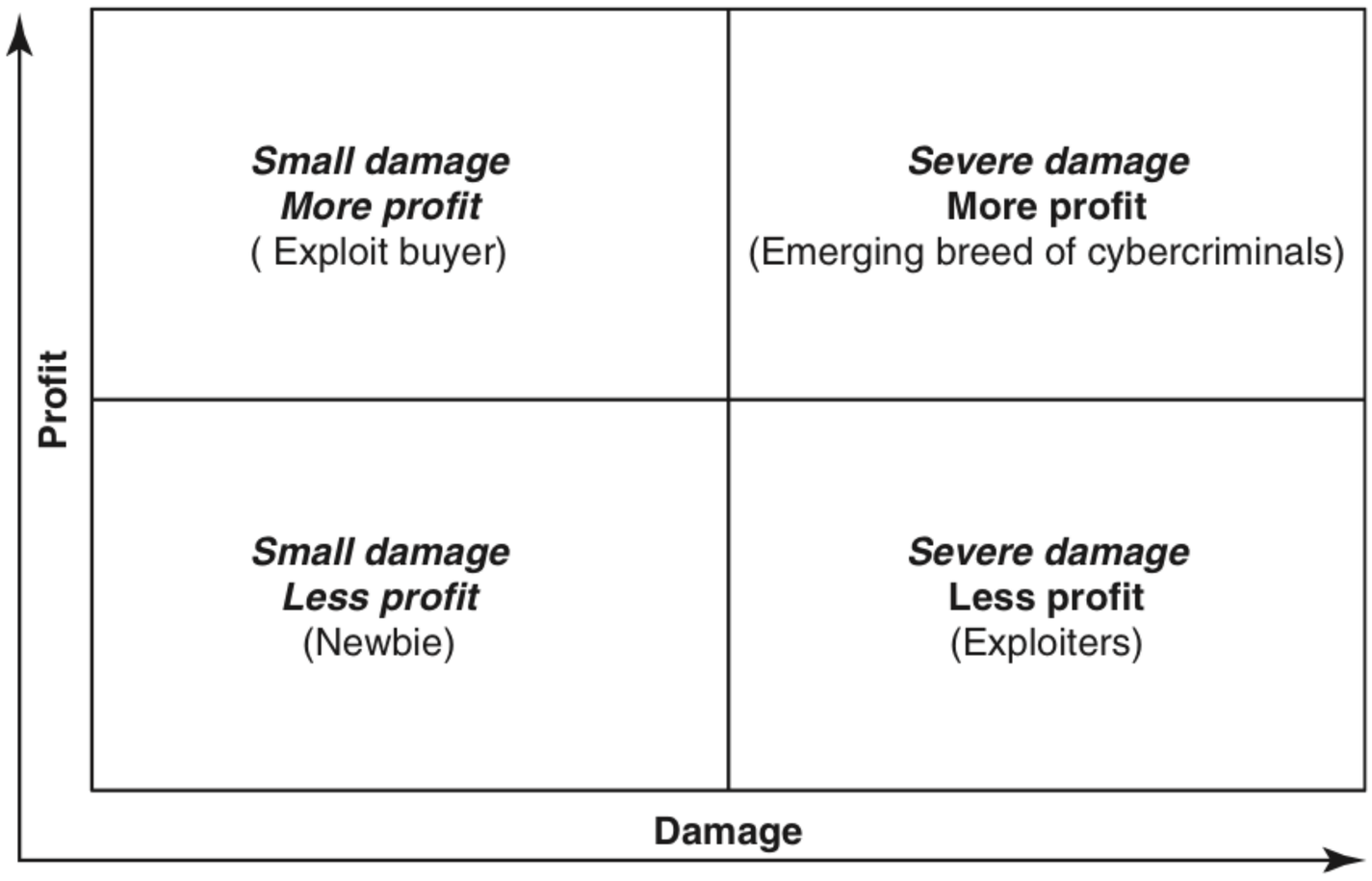


Figure 7.22 | Hacker categories (profit and damage).

7.13.3 Step 3: Gaining Access

The hacker's ultimate goal is to gain access to your system so that he/she can perform some malicious action, such as stealing credit card information, downloading confidential files or manipulating critical data. As each device and OS in your network has a unique security posture, the information provided during system scanning and probing can give the hacker insight as the easiest path into your system. Gaining access takes advantage of specific security weaknesses in the system to allow access via an individual machine.

7.13.4 Step 4: Privilege

When a hacker gains access to the system, he will only have the privileges granted to the user or account that is running the process that has been exploited. Gaining access to root or administrator will allow the hacker more access or greater power throughout the network. All hackers, therefore, would like to gain root or administrator privileges on the network. An exploited application that is running under a root user will give the hacker immediate root access. However, if the application that is exploited is not running under a root, the hacker must perform additional actions to earn it. This usually entails trying to crack the passwords.

7.13.5 Step 5: Exploit

Gaining root access gives the hacker full control on the network. Every hacker seems to have his/her own reasons for hacking. Some hackers do it for fun or a challenge, some do it for financial gain and others do it to "get even" (see Fig. 7.21). Exploiting the system, therefore, will take many forms. Hackers who do it for fun or a challenge will generally change a webpage or leave a "calling card" to let his peers know that he/she was successful. Hackers fall into this category. More often, hackers try to break into systems for financial rewards. This will generally help them to download valuable information that can later be sold to other parties. Sometimes, there can be a disgruntled employee who may gain access to sabotage an important project.

7.13.6 Step 6: Retracting

There are many reasons (as mentioned in Section 7.13.5) that drive cybercriminals to hacking. Whatever the motive, hackers do not want to be caught and sent to jail. Therefore, the next step in the hacker methodology is covering tracks. The hacker will usually take the time and effort to modify system logs to hide his/her actions and try to mislead forensics investigators that a crime has been committed. Refer to Table D.II.13 in Appendix D (in CD).

7.13.7 Step 7: Installing Backdoors

Finally, most hackers will try creating provisions for entry into the network/hacked system for later use. This, they will do by installing a backdoor (see Chapter 4) to allow them access in the future. A backdoor is a security hole deliberately left in place to allow access from an uncommon/unobvious path. These can usually be easily detected by skilled security professionals. In fact, almost all of the actions performed by a hacker can be detected by a forensics investigator. They just need to know where to look.

Professional working in the domains data networking and security would understand the terms "Layer 3" or "Layer 2" or "Application Layer." This terminology stems directly from the ISO Model and how it is applied to practical solutions. The model concepts are conventionally used to design and troubleshoot networks, and the 7 Layer Model is the standard for designing the network protocols. Careful study of the OSI 7 Layer Model can show us support for concepts that we have learnt from more conventional forms

of information security theory. Understanding and applying the model to information security scenarios can also help us assess and address information security threats in a network environment, allowing us to organize efforts to make security assessments and perform forensics analysis of compromised systems and threats presented in theory and found operating in the wild.

Cyberforensics experts have the technical skills that are useful for reading "obfuscated IP addresses." Those who send Spam, unsolicited commercial junk mail, usually try to keep their true identities secret – otherwise, they would take the brunt from the disgruntled Internet citizens who wish to retaliate. In addition to using a bogus return E-Mail address, they often include obfuscated URLs. Instead of having a human-readable name, or the dotted-decimal format such as 135.17.243.191, a URL may appear in 10-digit integer format (base 256), so it appears like this: http://2280853951. More details like this to illustrate the technical skills in cyberforensics investigation are addressed in Chapter 11.

7.14 Forensics and Social Networking Sites: The Security/Privacy Threats

Social networking is one of the most popular activities across the globe in today's digitally connected networked world. Orkut, Facebook, MySpace, Bebo, "Bigadda" (an Indian social networking site), etc. are some familiar names of social networking sites. They are not the only ones, however. There are a surprisingly large number of social networking sites.[18] Technically speaking, a "social networking site" is defined as Web-based services that allow individuals to:

1. Construct a public or semi-public profile within a bounded system;
2. articulate a list of other users with whom they share a connection;
3. view and traverse their list of connections and those made by others within the system; the nature and nomenclature of these connections may vary from site to site.

Using a social networking site brings like-minded people together for chat, conversation, exchanging ideas and even meeting in real life. Social networking is a popular activity in the Internet days because, it enables people to reach out to their old/long lost friends and classmates, relatives, etc. who may have migrated to other countries. Social networking sites even help connect like-minded people, people with the same professions or collaboration and discussion of ideas. Social networking, thus, makes people part of a worldwide community and so the sites are getting popular. Social networking has thus, become an extension of "sitting around the camp fire" and discussing life events. It is, then, no wonder that the usage of social network sites has increased rapidly in recent years.

Kids, teenagers are the ones who are known to be making the maximum use of social networking sites. There are professional networking sites too; for example, the most famous "LinkedIn." LinkedIn is a social networking website oriented toward professional networking with over 10 million users spanning 150 industries and more than 400 economic regions. Users are able to and encouraged to create a profile including such information as resumes, job offers and past employers and then to build up a list of connections. The user profile and list of connections can help to gain introduction to someone the user wishes to know through a mutual contact, to find jobs and business opportunities, search for potential candidates (if the user is an employer), etc. While this may be so, there is a lot of discussion these days about the security threats through careless use of social networking sites. Almost everyday, there is some news or other about how people (especially the young generation) are victimized as a result of indiscrete use of social networking sites.

Current litigations, regarding social networking sites, have raised a number concerns: (a) whether the content social networking site violate people's intellectual property rights, (b) whether social networking sites

infringe the privacy of their own users or (c) whether fraudulent or other illegal activity, such as the sale of "knock-off" luxury goods or the promotion of prostitution, occurs with actual or collaborative knowledge shared in the social networking media community. Although these concerns may be perceived by some people as mere exaggerations or over-reaction to the rise of social media networking, the fact remains that adequately preserving the content at issue is critical. Content preservation can be challenging given the dynamic, short-lived and often multi-format nature of social media. In order to properly collect and authenticate social networking content, there is indeed a need for using tools and for executing emerging forensics methodologies. Like in other forensics circumstances, here too, it becomes essential to maintain robust chain of custody documentation to ensure that this highly relevant evidence is authenticated for admissibility. There is generally no control over the content posted on social media networking sites. Even if forensics data has been preserved, high level of forensics skill is required to analyze and quantify the preserved data to answer questions such as:

1. Who posted the offending content?
2. Is there a 'real live' person to whom the offending content can be attributed even when evidence exists?
3. Can we identify the time frame associated with the posting of the offending content?
4. How much of the offending content exists across the entire social networking platform?
5. Is there other evidence that corroborates or supports interpretation of the relevant content?
6. How accurate is the reported physical location?

Such forensics analysis of social media websites/social media networking sites is the need of the hour given criminal activities abound.

Criminal activities can arise from the use of social network sites.

For example, a mother was convicted of computer fraud for her involvement in creating a phony account on MySpace to trick a teenager, who later committed suicide. [19] There seems to be no respect for "privacy" as if it has become a thing of the past. For example, although LinkedIn is generally well perceived by the media, there seem to be two main privacy issues: (a) LinkedIn does not allow you to remove your profile and (b) it shows member profiles. The good news is that the first privacy concern is now addressed in LinkedIn's privacy policy. The second privacy concern is, surprisingly, perceived as one of the good things about LinkedIn.

All is not well even with "better of kind" professional networking sites like LinkedIn. For example, according to some people LinkedIn first builds the trust of the user, then hooks the user into the system by using easy-to-use simplified forms and finally gets the user to invest through the use of nagging tools.

Facebook (www.facebook.com) is one of the biggest online social networking websites, hosting approximately 25 million users. The website was originally created in February 2004 for college and university students, but since September 2006, it has become available to anyone with an E-Mail address. Facebook has many enticing features that allow users to join many networks (including university networks, geographic networks, vocation networks, etc.), to join groups of common interest and to upload pictures. Figures 7.23 and 7.24 show some interesting data about privacy concerns with regard to social networking sites and the age distribution of social networking site users.

The success of a social networking site depends on the number of users it attracts. In the attempt of wooing the users to their sites and thereby generating revenues, the designers of social networking sites also incline toward making available some material on these sites which may not always be decent. Generally, the

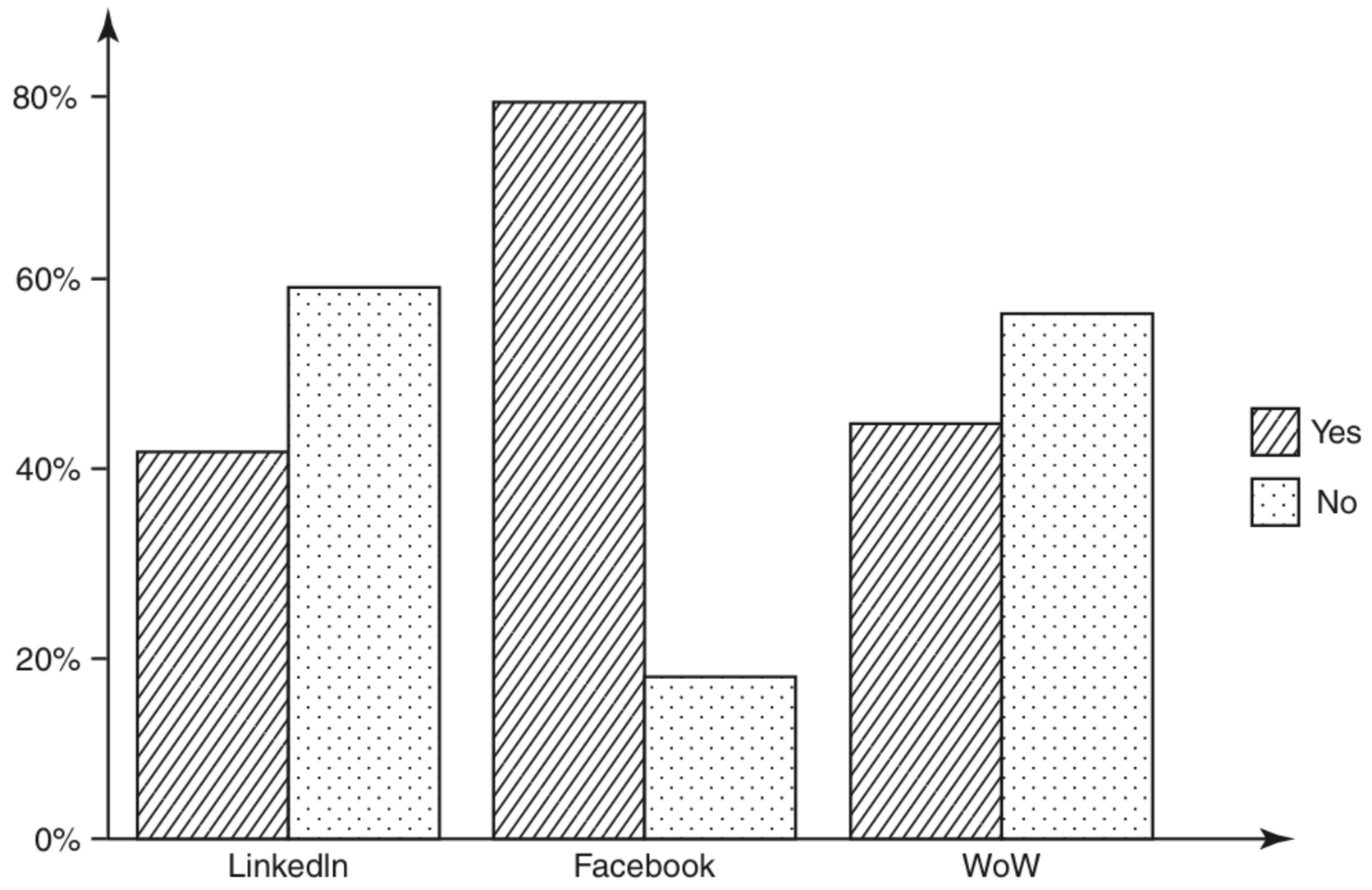


Figure 7.23 User concerns about privacy on social networking sites (LinkedIn, Facebook, WoW).
Source: The Paper by Helen Drislane and Kelly Heffner, 14 May 2007, is available at http://www.eecs.harvard.edu/cs199r/fp/HelenKelly.pdf (25 January 2010).

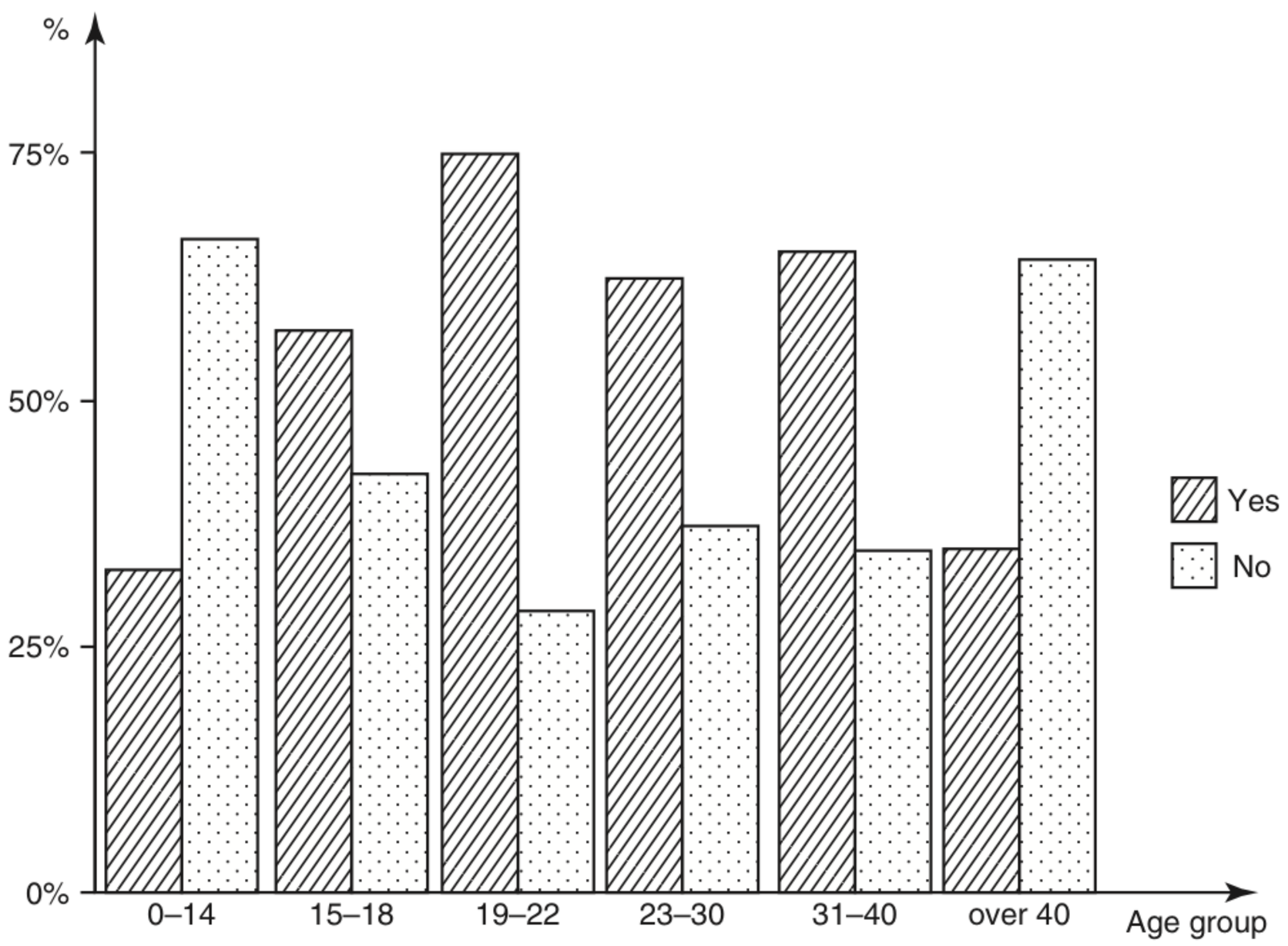


Figure 7.24 Privacy concerns about social networking sites vary with age.
Note: This is with regard to openness about sharing personal information.
Source: The paper by Helen Drislane and Kelly Heffner (14 May 2007) is available at http://www.eecs.harvard.edu/cs199r/fp/HelenKelly.pdf (25 January 2010).

aim of a social networking site is to design the site in such a way that encourages behavior to increase both the number of users and their connections. Unfortunately, like any fast growing technology, security has not been a high priority in the development of social network sites. Owing to this, that is, lack of security thinking in the design of social networking sites, significant risks have resulted. At times, these risks seem to overweigh the benefits obtained from the social networking websites.

There are privacy concerns that arise from the use of social networking sites. For example, Facebook's Beacon service tracks activities from all users in third-party partner sites, including people who never signed up with Facebook or who have deactivated their accounts. Beacon captures data details on what users do on the external partner sites and sends it back to Facebook server, along with users' IP addresses, the addresses of webpages the user visits, etc. This is an example of vulnerability in Facebook. Table 7.6 lists the security features on the top 10 social networking websites of 2009 and Table 7.7 shows security features for the top ten social security networks rated in the year 2010.

There are a number of security issues associated with social networking sites.

Security issues that are associated with social networking sites are listed below:

1. Corporate espionage.
2. Cross-site scripting.
3. Viruses and worms.
4. Social networking site aggregators.
5. Spear Phishing and social networking specific Phishing.
6. Infiltration of networks leading to data leakage.
7. ID theft (Phishing and Identity Theft is discussed in Chapter 5).

Table 7.6 | Top 10 social networking sites (year 2009) – security features

Social Networking Site Name	*Supports HTML?*	*Does the Site Track Visitors?*	*Does the Site have Customizable Privacy Settings?*
Facebook	No	No	Yes
Orkut	No	Yes	Yes
MySpace	Yes	No	Yes
Bebo	No	No	Yes
Friendster	No	Yes	Yes
Hi5	Yes	Yes	Yes
Netlog	No	No	Yes
PerfSpot	Yes	Yes	Yes
Yahoo!360	Yes	No	Yes
Zorpia	No	No	Yes

Note: The listing in Table 7.6 is not in the order of 2009 rating for the sites. It is in the order of "familiarity" of the site name assuming general public familiarity with those names of the sites. In year 2009 "Bebo" had rating 1 while in Year 2010 "Facebook" has got Number 1 rating. In year 2009 "MySpace" had No. 5 rating while in year 2010, it was rated No. 2. "Friendster" enjoyed No. 3 rating in year 2009 whereas its rating in year 2010 has gone down to No. 4. (Source of these changes rating is the link mentioned at the bottom of Table 7.7).

Table 7.7 | Top 10 social networking sites (year 2010) – security features

Site Name	*Privacy Settings Available?*	*User Blocking Available?*	*Spam Reporting Available?*	*Abuse Reporting Available?*	*Saftety Tips Available?*
Facebook	Yes	Yes	Yes	Yes	Yes
MySpace	Yes	Yes	Yes	Yes	Yes
Bebo	Yes	Yes	Yes	Yes	Yes
Friendster	Yes	Yes	Yes	Yes	No
Hi5	Yes	Yes	Yes	Yes	Yes
Orkut	Yes	Yes	Yes	Yes	Yes
PerfSpot	Yes	Yes	No	Yes	Yes
Yahoo!360	Yes	Yes	No	Yes	Yes
Zorpia	Yes	Yes	No	Yes	No
Netlog	Yes	Yes	Yes	Yes	No

Note: There is full features comparison available in addition to the security features mentioned in the table above. The site overall rating for the social networking sites is available at: http://social-networking-websites-review.toptenreviews.com/ (24 January 2010). (Sites are listed in the order of year 2010 rating).

8. Bullying.
9. Digital dossier aggregation vulnerabilities, secondary data collection vulnerabilities, face recognition vulnerabilities.
10. Content-based image retrieval (CBIR).
11. Difficulty of complete account deletion.
12. Spam.
13. Stalking.

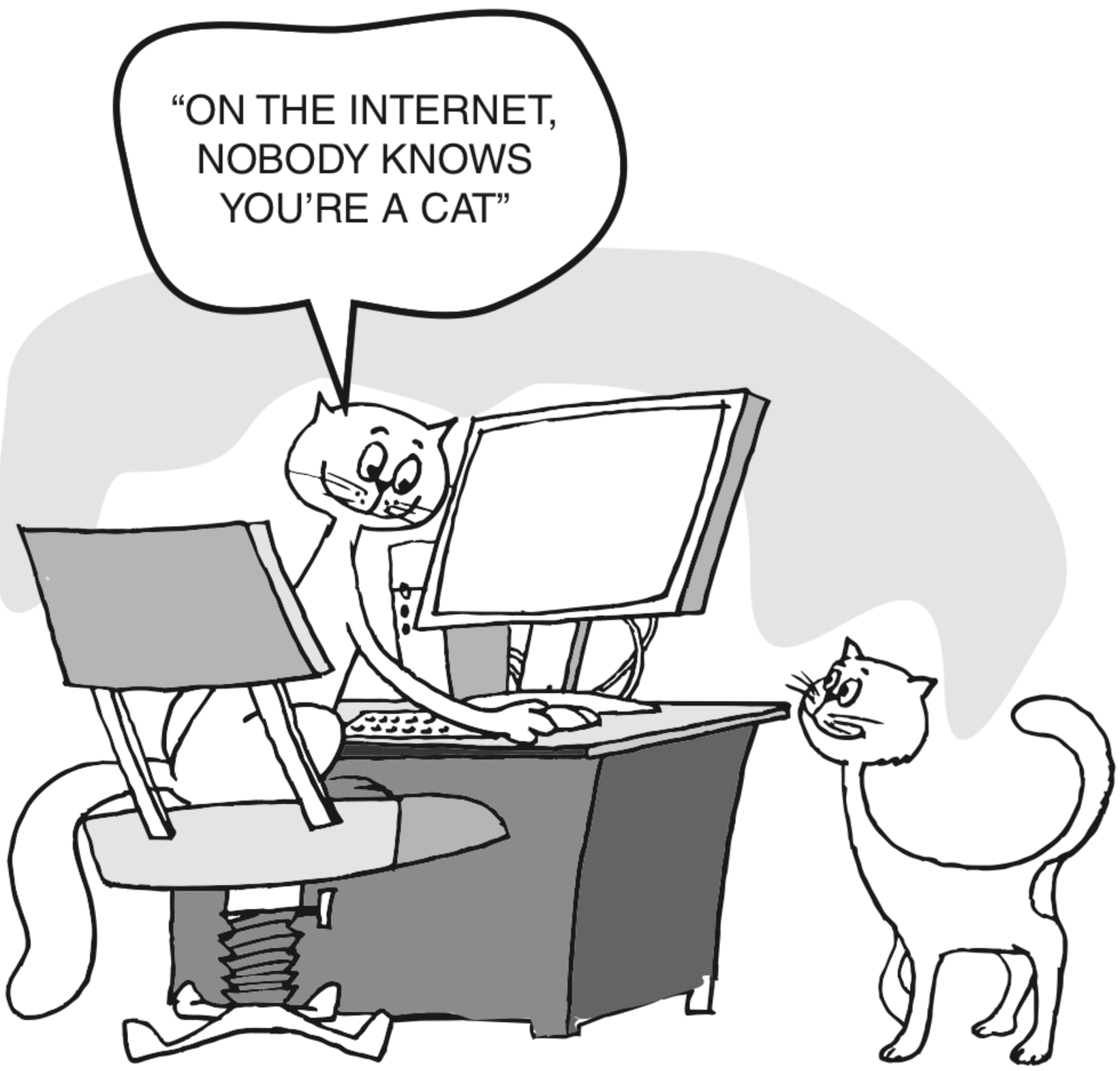
MySpace does not allow Java Script code to be used in their site. However, it does allow HTML code. MySpace and Hi5 allow their users to have the ability to add HTML code that can link to scripts and objects designed to retrieve other visiting user's information. The users of PerfSpot cannot add HTML code to their pages, but there are restrictions on the contents of this code. Yahoo!360 lacks detailed security controls as to who can view the user's profile. Their option is either everyone can view their blog or just friends. In other social networks such as Facebook and MySpace, specific security features are available to provide users with more control over their personal information. Friendster, Hi5, Orkut and PerfSpot users have the benefit of monitoring their profile visitors. This option increases awareness, protection and eases user tracking. Orkut is the only social network that forces the user's viewing history to be revealed if he/she decides to track other users who had viewed his/her profile.

It is possible to retrieve information about the users who visit your profile on social networking sites. Some of these sites provide built-in applications that show the usernames of the people who visit your profile. Other sites store log transcripts, which capture chat session information such as username and date. The function of user data retrieval can be accomplished within a website with user incorporation of either Java Script or Hypertext Preprocessor (PHP) code. Some social networking sites have restrictions in place to make Java Script and/or PHP code inactive when users try to incorporate into their site. There are different methods for capturing the non-personal identifiable information of users who communicate with each other in the virtual world. We have established that user data retrieval can be achieved with use of Web-based scripts.

User data can be retrieved and user data retrieval methods take place in the online environments of websites, IM chat sessions (virtual meetings) and E-Mails. From the user data retrieval methods used, the most important non-personal-identifiable user information that can be retrieved is the IP address. An IP address

Table 7.8 | Retrieving sender's IP address from E-Mail received

Gmail	*Hotmail*	*Yahoo mail*
1. Access your inbox. 2. Select the message you would like to trace for its IP. 3. Click on the upside down triangle located on the right, next Reply. You will see options such as "Reply to all," "Forward," "Filter Messages like This,"etc. 4. Select "Show Original."	1. Make sure you are in classic Mode. 2. Right click on the message. 3. Select "View Message Source."	1. Select the message. 2. Right click on the message. 3. Click on "View Full Headers."



"Annonymity breeds crime"
"Stranger is Danger"

Figure 7.25 | On the Internet, it does not matter "who" you are as long as you have "ID"!!

can be used for tracking back to a user's location or the user's Intenet service provider location. After retrieving the IP address, there are many links available for retrieving the geographical location of the user. It is possible to write a utility to track the visitors to a social networking site. The scripts for developing such utilities can work on any social networking website which allows you to embed HTML code in your homepage. The social networking sites and sales transaction websites that support visitor tracking program utilities are AOL Instant Messenger, eBay, Friendster, Hi5, MySpace and Yahoo!360. The sites that do not support visitor tracking utility programs are Bebo, DeviantArt.com, Facebook, Netlog, Orkut, Perfspot and Zorpia.

Currently, creating a visitor log for Facebook using the Facebook programming API is against the Facebook Developer terms and conditions. Users who attempt to create an application which tracks profile hits without visitor knowledge will be banned from the site if the application is found. Recall the discussion in Section 7.6 – it is possible to retrieve the sender's IP address from an E-Mail you received at your Gmail, Hotmail and Yahoo mail accounts. The steps are presented in Table 7.8.

As a responsible Netizen, you can take a few basic precautions; for example, knowing more about the person communicating with you online can protect you. Owing to the increased amount of crimes being committed over the internet, knowing the true identity and location of others can add another layer of protection on the lighter side, see Fig. 7.25.

Details of tools used for digital forensics with each type of OS (Window, Unix, MAC, etc.) can be studied by referring to technical books on cyber/computer/digital forensics. Some such books are mentioned in Refs. #1–10, Books, Further Reading. There are also links provided in Video Clips, Further Reading for video demonstrations. To conclude this section, we say that providing social networking site users with tools which will help protect them is ideal. Such tools are developed for installation on a user's computer to provide them the ability to retrieve other online user information via chat and social network websites. These tools will also benefit law enforcement agents when crimes are committed. Such tools typically help retrieve the IP address, OS (used on the machine from where the browsing is done) and the browser types associated with the used session for visiting a social networking site. Retrieval of this information occurs upon the virtual contact from that other person, be it by them simply browsing our personal page or by other person contacting via virtual meeting, for example chatting. Now let us understand computer forensics from compliance perspective.

7.15 Computer Forensics from Compliance Perspective

With the rampant use of the Internet, there is so much at stake; corporate data is not safe anymore given that almost all information assets lie on the corporate networks. We are in the era of Net-centric digital economy.

Criminals can gather small pieces about you, about your confidential data to generate what is known as "digital persona," that is, they keep track about your Internet activities, what resides on your corporate networks, etc.

Recall the discussion in Chapter 2 about how cybercriminals plan their attacks. There are cybercrimes and therefore there are investigations. Investigations require "evidences." This takes us to the legal territory where there are a number of legal compliance requirements. Information security compliance requires the precise enforcement of policies and controls. Investigations, in which computer forensics techniques are utilizied, become an essential part of this enforcement. Let us revisit our thinking on the key information security laws and regulations that mandate computer forensics for compliance. It is appropriate to address this here in this section because we want to understand how the need for mandatory legal compliance can affect cyberforensics.

7.15.1 The Regulatory Perspective for Forensics at the International Level

Internationally, there are a few laws and regulations that indicate the need for digital investigations: *Sarbanes Oxley* (the SOX), *California SB 1386* (see Box 7.16), *Gramm Leach Bliley Act* (the GLBA) and *Health Insurance Portability and Accountability Act* (HIPAA) of 1996.

These laws/regulations specify investigation and response to security breaches or policy violations. Computer forensics makes it easier to meet these requirements.

IT businesses are "global" with customers and IT service supplier organizations operate all over the world. We focus on the aforementioned "Big-4 laws" (SOX, California SB 1386, the GLBA and HIPAA) because they have the broadest implications for commercial organizations. They impact companies that are publicly traded, store financial or medical information or do business in California. To illustrate the implication, let us say that there is a BPO center in India that deals with support to medical entities in California; as such the HIPAA may become relevant under regulatory compliance. Such business scenarios can affect most medium to large businesses as international companies that do business in the US. Such requirements can impact global businesses.

Let us understand the role of computer forensics in achieving compliance to the Big-4 laws mentioned. After recourse to these *Big-4 laws*, we shall examine computer forensics expertise situation in India with regard to the Indian ITA 2008. We will learn that through the requirement for "adequate security practices," which include "security incident handling," these laws/legislations become relevant in the context of forensics with cybercrimes.

Box 7.16 California Senate Bill 1386 (Another Angle to Cyberforensics)

The Bill is important for those involved in doing business with an entity in the US. California's first-of-its-kind information security legislation (SB 1386) went into effect on 1 July 2003. This legislation requires entities or individuals who do business in California to notify California residents whenever their unencrypted personal information is reasonably believed to have been compromised. According to California Senate Bill 1386, personal information includes "an individual's first name or first initial and last name in combination with one or more of the following": a social security number, driver's license number or California Identification Card number, account number and/or credit or debit card information including numbers and passwords, PINs and access codes. The bill also limits coverage to personal data that is "unencrypted." The statute provides a broad definition for "security breach" as an "unauthorized acquisition of computerized data that compromises the security, confidentiality or integrity of personal information maintained by the person or business." The statute does not define the term "unauthorized" or specify what evidence of a breach is necessary to trigger notification obligations. The statute also does not resolve on the question whether companies have an affirmative duty to actively monitor and detect security breaches.

According to the Bill, it is mandatory for an organization to follow certain disclosure obligations following the discovery of a security breach that may have compromised customer data. "Notice must be given to any resident of California whose personal information is or is reasonably believed to have been acquired by an unauthorized person." Notice must be given in "most expedient time possible" and "without unreasonable delay" subject to certain provisions that define what reasonable is for your organization. Organizations that hold personal data on California residents, it becomes essential for IT departments to review the security of consumers' personal information.

1. **The Sarbanes Oxley Act (SOX):** The Act was enacted to fight corporate fraud. Massive financial fraud at Enron, WorldCom, Global Crossing and Arthur Anderson led to the passing of this legislation in 2002. The Securities and Exchange Commission (SEC) is responsible for enforcement of SOX and all publicly traded companies must report yearly on the effectiveness of their financial controls. Corporate Governance has become a critical operational focus of organizations to ensure that they have the proper controls and audit processes in place to prevent and detect fraud.

   The legislation has serious consequences for non-compliance. This includes civil and criminal penalties. Section 302 of SOX specifies that CEOs and CFOs are directly responsible for the accuracy of their company's financial reports. Much of the focus on SOX has been regarding Section 404. Section 404 requires management to specify their responsibility for financial controls and report on the adequacy and shortcoming of the controls. Many companies offer products and services to help companies achieve Section 404 compliance. SOX has other provisions that have not received the same attention from technology and service providers.

   Many companies recognize the need for computer forensics as part of normal business operations and controls and it therefore indirectly supports Section 404 compliance. For Section 301, case law has established that computer forensics is required to properly investigate fraud. In addition, computer forensics is widely accepted as the only precise and reliable method to determine if digital records have been deleted and/or altered; therefore, computer forensics is needed to maintain compliance with Section 802 (this section of SOX impacts electronic evidence). Computer forensics has proven itself as a tool in fighting against wrongful termination litigation, HR investigations, *theft of intellectual property* and *E-Discovery management;* all of these issues enhance the accuracy of financial reporting, thus supporting Section 404. Sections 301 and 802 compliance will require the use of computer forensics as established by case law and by best practices. In today's world, it becomes essential for organizations to have computer forensics capability anywhere and anytime in their organizations to ensure compliance with Sarbanes Oxley.
2. **California SB 1386 (refer to Box 7.16):** This Bill requires organizations doing business in California to report security breeches that result in the unauthorized disclosure of a resident's private or financial information. The objective for this legislation is to thwart identity theft and consumer fraud. Given that many organizations (IT companies as well as non-IT companies) worldwide are engaged in business with the US, this law affects most domestic and international companies. Disclosure is required if an exposure to individual's "personal information" is involved (see Box 7.16). Notification, however, is not required if the information disclosed was encrypted.

   The law allows for civil actions to be brought against non-complying businesses or they may be enjoined by the court. It is crucial for any business to conduct a thorough investigation to determine if it "reasonably" believes that information has been compromised or not. The legislation does not provide a clear definintion for "reasonable investigation of a security breach." However, security organizations and government agencies have documented current incident response processes.

   The National Institute of Standards and Technology (NIST) provides clear guidance for government and commercial organizations to investigate security incidents.[20] NIST publication on the *Computer Security Incident Handling Guide*[20] specifically outlines incident investigation and the role of computer forensics to properly acquire and analyze the incident. NIST also clearly identifies "unauthorized access" as a type of security breach that their process addresses.
3. **Gramm-Leach Bliley Act (GLBA):** This Financial Modernization Act of 1999 (known as the GLBA) has a broad spectrum of qualifications, requirements and regulating parties. Eight agencies and states are charged with managing and enforcing the regulations. The GLBA applies to financial organizations or any organization that collects or transfers private financial information for the purpose of doing business or providing a service to its customers.

There are two aspects of GLBA: (a) the *Financial Privacy Rule* and (b) the *Safeguards Rule*. The Financial Privacy Rule addresses the collection and dissemination of customers' information whereas the Safeguards Rule governs the processes and controls in an organization to protect customers' financial data. The Federal Trade Commission enforces the "Safeguards Rule" in GLBA. In addition to the public embarrassment of non-compliance, organizations may be fined thousands of dollars a day while they are non-compliant.

The Safeguards Rule of GLB calls for financial institutions to:

1. Ensure the security and confidentiality of customer information;
2. protect against any anticipated threats or hazards to the security or integrity of such information;
3. protect against unauthorized access to or use of such information that could result in substantial harm or inconvenience to any customer.

The GLBA is relevant to forensics because computer forensics is an integral part of investigating and auditing all of GLB Safeguards Rule elements mentioned above. For response to incidents, GLBA guidelines require – "*Response programs that specify actions to be taken when the bank suspects or detects that unauthorized individuals have gained access to customer information systems, including appropriate reports to regulatory and law enforcement agencies.*"

Technical guidelines that support GLBA call for extensive Intrusion Detection System (IDS) response by utilizing computer forensics investigations. To know more on IDS, refer to Ref. #17, Books, Further Reading. From the guidelines for security controls and the guidelines for incident response, we believe that GLBA compliance requires the utilization of computer forensics both proactively and for incident response to ensure the privacy of client information and to exhibit due diligence in GLB compliant efforts.

4. **HIPAA (Health Insurance Portability and Accountability Act of 1996):** HIPAA has the primary goal for healthcare providers to improve the privacy and security of their clients' medical information. In the US, healthcare providers (hospitals, medical insurance agents, medical professionals and many allied entities involved in delivery of healthcare services) records clearing houses and health plans must comply with HIPAA. Trading partner organizations that handle medical records electronically would fall under HIPAA rules.

   Finalized HIPAA rules include "information security" which encompasses incident response. HIPAA definition of security incident is "… *the attempted or successful unauthorized access, use, disclosure, modification or destruction of information or interference with system operations in an information system.*" HIPAA specifies that there should be thorough analysis and reporting of security incidents – with that comes the need for the forensics investigation. Organizations must, therefore, consider their incident response policies carefully. Generally, such policies are part of organization's overall security policies; much about this is addressed in Chapter 9. Security policy and procedures aspects are addressed in detail in Ref. #13, Books, Further Reading.

   Computer forensics software is often specified to be part of any reasonable incident response policy to clearly understand the scope of the incident. Determining, with forensics precision, what information has been compromised, when the compromise took place, what systems were affected, and if malware or backdoors (see Chapter 4) that are invisible to non-forensics tools are still present, are examples of the types of investigations that are essential to having an effective incident response program. Even beyond security incidents, computer forensics plays a natural role in supporting overall information security by providing the investigation of any anomalies that could indicate policy or use violations that could jeopardize HIPAA privacy rules.

Having understood why the Big-4 laws/legislations are important from the forensics compliance perspective, let us understand the changing face of computer forensics in terms of "traditional forensics" (see Fig. 7.26) vs. "remote forensics." Basically, computer forensics requires accessing system data in a least-intrusive manner. Traditionally, this has meant removing the hard drive from the suspect computer and connecting it to a forensics workstation using a write blocker[21] (Fig. 7.18).

Compliance requirements in the forensics field require that the (digital) evidence must be "forensically sound." According to the literature,[22] a forensically sound copy of a hard drive is as follows:

> *… created by a method that does not, in any way, alter any data on the drive being duplicated. A forensically sound duplicate must contain a copy of every bit, byte and sector of the source drive, including unallocated empty space and slack space; precisely as such data appears on the source drive relative to the other data on the drive. Finally, a forensically-sound duplicate will not contain any data… other than which was copied from the source drive.*

Furthermore, it can be said that the manner used to obtain the evidence must be documented and should be justified to the extent applicable.

As learned in this chapter, one of the key requirements of a sound forensics examination of digital evidence is that the original evidence must not be modified, that is, the examination or capture of digital data from the hard disks of a seized computer must be performed so that the disk contents are not changed (recall the discussion in Sections 7.5 and 7.7). This is because, during forensics acquisition and analysis, it is possible to write to the evidence drive accidentally.

For example, in the US, there are "Federal Rules of Evidence"; to be admissible in a US court, evidence must be both relevant and reliable. The reliability of scientific evidence, such as the output from a digital forensics tool, is ascertained by the judge (as opposed to a jury) in a pre-trial "Daubert Hearing."[23] The Daubert Test is an expansion of the court's prior approach to the admissibility of scientific evidence. To avoid the immediate dismissal of the evidence from court, the investigator should take care not to compromise the evidence.

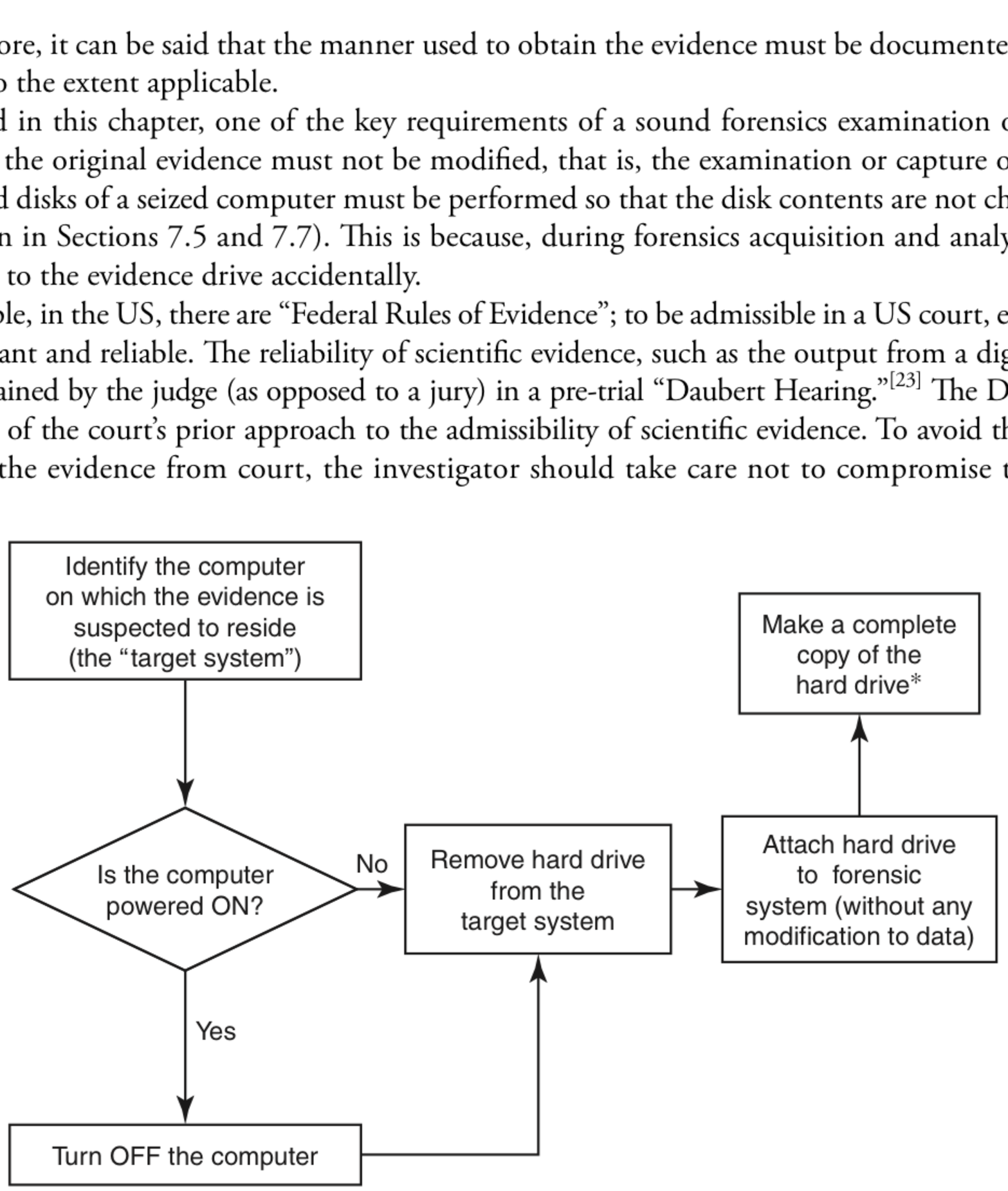


Figure 7.26 Traditional approach to forensics analysis. *denotes tools/devices mentioned in Fig. 7.14.

The easiest way to ensure this is to use a *write blocker*. The investigator follows a set of procedures that are designed to prevent the execution of any program that would likely modify the disk contents. With the write blocker tool, the forensics investigator can examine the data on disk without alteration. There are several drawbacks with this methodology. First, it is costly and slow to physically remove the disk from a server, and the removal of the disk means downtime and user dissatisfaction. Second, although the data is not altered, important volatile data is lost because the system must be powered off to remove the disk. Although the investigator may utilize system tools and freeware to access this data before the system is shutdown, the data gathered is suspect as it may be compromised by malware running in the server. In other words, the "chain of custody" (the concept was explained in Sections 7.4 and 7.8, and Box 7.4) can be suspected/questioned when the digital evidence is admitted in the court.

Owing to these issues, a new generation of computer forensics tools[22] has been developed with the capability to examine a live system through your corporate network. By placing a small read-only agent on the server, disk data and volatile system data can be accessed and imaged via the network. There are literature references[24] (reports/articles, etc.) indicating that this methodology called "remote forensics," though new, has been successfully utilized in some courts. With the new generation of network-enabled computer forensics, the investigations can be done quickly and at a lower overall cost than traditional computer forensics.

7.15.2 Computer Forensics Compliance Requirements: Implications for Evidential Aspects

Due to the current cyberspace scenario, there is an increasing need for evidence in organizations. There are rising cybersecurity threats as well as compliance needs (as part of security incidence response handling) for presenting the digital evidences/computer evidences. Good digital evidence is thus becoming a business enabler. In the face of the security threats, the incidents that follow and the need for forensics investigations, organizations are turning to logs to provide a continuous trail of everything that happens with their IT systems and, more importantly, with their data. Refer to Table D.II.13 in Appendix D.

Today there are many tools available that generate logs of different types from different sources; such logs allow organizations to generate a picture of the IT activity. Let us consider an example; if a disgruntled employee harbors the ill intention to steal data accesses, a database containing confidential information, there would likely be a log of that activity that can be reviewed to determine the who, what and when of the access to such a database. Such "logs of IT activity" provide the pieces of information and organizations can use those pieces of information to follow the paths of all of their users, bad intentioned or not. Managing these logs can benefit an organization in a number of ways. These logs offer situational awareness and thereby help organizations identify new threats as well as allow their effective investigation. Review of routine logs and detailed analysis of stored logs are beneficial for identifying security incidents, policy violations, fraudulent activity and operational problems shortly after they have occurred, as well as for providing information useful for resolving such problems. Refer to Table D.II.13 in Appendix D.

Considering that log management provices inherent benefits of log management, it is not surprising that log data collection and analysis are generally considered to be a security industry "best practice." However, there are a number of regulations that mandate the collection, storage, maintenance and review of logs, turning log management from a "should do" to a "must do." Some of these regulations draw on National Institute of Standards and Technology Computer Security Special Publications (NIST SP) toward detailed log requirements.[25]

7.15.3 Computer Forensics Expertise Status in India

There is a rise in cybercrimes – recall the overview provided in Chapter 1. In Chapter 6 (Section 6.5) some of the challenges with regard to cybercrime scenario in India were described. With that as a background for this section, it can be inferred that in India, computer forensics is a much needed expertise. At the present, there seems to be a shortage of these skills. India seems to suffer from a two-fold problem: lack of availability of cyberforensics expertise as well as lack of awareness about cyberforensics/digital forensics/computer forensics. Involvement of cyberforensics in the day-to-day activities of individuals as well as corporations is going to increase due to the rising rate of cybercrimes in India (recall the Indian cybercrime data presented in Chapter 1 (Tables 1.1–1.4). ICT has been strongly rooted in India. India has a large youth population with computer usage going high. Due to this it is no surprise that cybercrimes rates are on the rise. This brings the need for increased use of cyberforensics for resolving both civil and criminal issues in India – for example, computer forensics can be used to investigate tax evasion of the accused ones.

To sum up this section, we note that compliance efforts require computer forensics investigation capabilities to investigate privacy and security incidents. The reach of computer forensics must be enterprise-wide and ideally, the response time should be immediate in order to demonstrate that the organization are utilizing best practices in managing and controlling their information security compliance. Organizations need to have a combination of in-house capability supplemented with external expert services. The fact remains that India does not have cyberforensics capabilities and the required manpower in tune with growth of ICT-related crimes and contraventions. Data breaches and cybercrimes in India cannot be reduced till we make strong cyberlaws. Cyberlaws of India need to be supported by sound cybersecurity and effective cyberforensics. In the absence of legal enablement of ICT systems in India, there is not much to expect. We need a good team of techno-legal experts who not only help in the drafting of good laws but also its amendments and enforcement.

7.16 Challenges in Computer Forensics

Although computer forensics has well-developed techniques, investigation of cybercrime is by no means easy. A microcomputer may have 200 GB or more storage capacity. There are more than 5.2 billion messages expected to be sent and received in the US alone per day. There are more than 3 billion indexed webpages worldwide. There are more than 550 billion documents online. Terabytes of data are stored on tape or hard drives. Therefore, looking for forensics evidence among these is like looking for the proverbial needle in the haystack! Over and above this, there is another challenge; most of existing tools and methods allow anyone to alter any attribute associated with digital data. The form of digital data to be analyzed is usually transformed in some way and always processed before scrutiny. Encryption is a major antiforensics technique and key word search can be defeated by renaming file names.

Cybercrime investigators are faced by the challenge of how to collect the specific, probative and case-related information from very large groups of files. They need to use approaches such as link analysis and visualization.[27] They need to use enabling techniques for lead discovery from very large groups of files; typical techniques used to look for "patterns" are the techniques called text mining and data mining[27] along with the techniques of intelligent information retrieval; all this is part of "data mining" which is a vast subject not within the scope of this chapter. Readers interested in learning about data mining may refer to Ref. #24, Books, Further Reading . Computer forensics must also adapt quickly to new products and innovations with valid and reliable examination and analysis techniques.

On the network forensics side, there are many challenges. The networks may span multiple time zones and multiple jurisdictions and this makes it necessary to use absolutely trusted timestamps (to ensure the authentication and integrity of timestamps for each piece of network evidence) and ensuring that all jurisdictions collaborate. Moreover, the network data will be available in both offline and real-time modes, the latter requiring the ability to capture and analyze data on the fly. The data could involve many different protocols and the amount of data could potentially be very large due to the increasing size of network bandwidth. A protocol could also involve multiple layers of signal (e.g., voice-over IP or VoIP, HTTP tunneling – for the discussion on tunneling protocols readers can refer to Ref. #15, Books, Further Reading). The current set of computer forensics tools will not be able to handle the real-time and data size/volume. Techniques are required for rapidly tracing a computer criminal's network activities (e.g., IP address) and for mapping a network's topology. There need to be a paradigm shift for network forensics techniques to analyze the rate and size of captured data.

The increasing volume of potential data to examine can create problem for law enforcement. Seizing all the computers at a search site and examining them at the deepest levels are the most significant factors contributing to the examination backlog and yet they are difficult to achieve. In order to alleviate this problem, new data intake and data reduction strategies must be implemented. It is always a good idea to adapt data acquisition strategies to the goals involved in each specific case. These strategies must also be pragmatic with regard to data volume and time constraints. Technological obsolescence must be recognized – yesterday's computer is not the equivalent of today's computer and is not even remotely similar to tomorrow's computer. Failure to recognize this will inevitably result in lost investigative leads and ineffective prosecutions.

Many forensics matters do not go to trial, especially in the business arena where a convincing set of data often suffices to induce an out-of-court settlement, or where investigative techniques are applied on a "need-to-know" basis, such as to determine whether internal or external corporate espionage or malicious activity has occurred. Experts may assist in preparing legal briefs, and they can be requested to provide sworn testimony and opinions in city, state and federal hearings conducted by legislative bodies and their commissions or task forces. They frequently work hand-in-hand with computer security teams to assist in the development of procedural, policy and control techniques to help prevent (or assist in mitigating) losses. Durations of forensics investigations vary; they may take a few hours in complex case analysis or a few days for simple cases, or the investigations can persist over the course of years for complex cases. Although some experts are engaged for the full range of investigative and testimonial tasks, those who are valued for their highly persuasive verbal skills and who can react well to on-the-spot challenges, may only review and present evidence prepared by other forensics computing specialists. Certain digital information, beyond the contents of the data itself, may be pertinent to case development. Such information can include the time and date stamps of files, folder structure hierarchies and message transmission tags. Real-time data collection efforts are more complex because they may need to address legalities and privileges involved in surveillance, and must avoid inadvertent damage claims (such as may occur when a server is made inaccessible for a period of time). Things to be wary of include alterations to the digital media that could occur when the electronic device is turned ON or OFF, and inadvertent activation of Trojan Horse or time-bomb malware that was left behind to corrupt data and confound forensics efforts. One caveat is that "you should only find what is actually there" however, ensuring this can involve the development and implementation of collection, blocking, prevention and tracking techniques. This is wheress evidence collection kits comprising of software and hardware tools, can be applied and these kits become useful.

7.16.1 Technical Challenges: Understanding the Raw Data and its Structure

In this section, our objective is to understand some of the technical aspects encountered in digital forensics analysis; only a few examples will be provided. Treating all the technical challenges and their greater details

is not within the scope of this chapter as digital forensics is a very large domain. The phases involved in digital forensics investigation (Section 7.7.2) should be kept in mind while going through the discussion here. Familiarity with the basic computer science fundamentals will help to appreciate the discussion in the following paragraphs. We explain only the technical aspects of data representation that a digital forensics investigator must understand.

There are two aspects of the *technical challenges faced in digital forensics investigation* – one is the "complexity" problem and the other is the "quantity" problem involved in a digital forensics investigation. A digital forensics investigator often faces the "complexity problem" because acquired data is typically at the lowest and most raw format. Non-technical people may find it too difficult to understand such format. For resolving the complexity problem, tools are useful; they translate data through one or more "layers of abstraction" until it can be understood. For example, to view the contents of a directory from a file system image, tools process the file system structures so that the appropriate values are displayed. The data that represents the files in a directory exist in formats that are too low level to identify without the assistance of tools.

The directory is a layer of abstraction in the file system. Examples of non-file system layers of abstraction include:

1. ASCII;
2. HTML Files;
3. Windows Registry;
4. Network Packets;
5. Source Code.

Digital forensics is also challenged by the "quantity problem" – it involves the hugeness of digital forensics to analyze. It is inefficient to analyze every single piece of it. Data reduction techniques need to be used to solve this. Data reduction is done by grouping data into one larger event or by removing known data. Examples of abstraction layers are data reduction techniques; for example:

1. Identifying known network packets using IDS signatures;
2. identifying unknown entries during log processing;
3. identifying known files using hash databases;
4. sorting files by their type.

Digital forensics analysis tools aim at accurately presenting all data at an appropriate layer of abstraction and format, so that the tools can be effectively used by an investigator to identify evidence. The required layer of abstraction is dependent on investigator's skill level as well as the investigation requirements. For example, in some cases viewing the raw contents of a disk block is appropriate whereas other cases will require the disk block to be processed as a file system structure. Tools must exist to provide both options. Let us understand the abstraction layer properties with regard to digital forensics. Large amounts of data are analyzed in a more manageable format using the abstraction layers. Abstraction of data layers is a core feature in the design of modern digital systems. This is because, all data, regardless of application, is represented on a disk or network in a generic format, bits that are set to one or zero. For using this generic storage format for custom applications, the data bits are translated by the applications to a structure that meets its needs. The custom format is a layer of abstraction.

ASCII is one basic example of abstraction. Every letter of the English alphabet is assigned to a number between 32 and 127 (for detailed information about ASCII scheme refer to the link http://en.wikipedia.org/wiki/ASCII). When a text file is saved, the letters are translated to their numerical representation and the value is saved on the media as bits. When the file is viewed in the raw, it shows a series of ones and zeros. When the ASCII layer of abstraction is applied, the numerical values get mapped to their corresponding

characters and the file is displayed as a series of letters, numbers and symbols. A text editor is an example of a tool operating at this layer of abstraction.

Next, let us consider fat file system example; "FAT" is a file allocation table. FAT file system is one of the most basic file systems that is still used in many computers. It is broken up into three main areas. The first area is the *Boot Sector* that contains the addresses and sizes of structures in this specific file system. The next two areas are the FAT and the *Data Area*. The locations of which are identified in the Boot Sector. The *Data Area i*s divided into consecutive sectors called *clusters.* Clusters store the contents of a file or directory. Each cluster has an entry in the FAT that specifies if the cluster is unallocated or which cluster is the next in the file that has allocated it. Files are described by a *directory entry* structure. The directory entry structures are stored in the clusters allocated to the parent directory. The structure contains the file name, time, size and starting cluster. The remaining clusters in the file, if any, are identified using the FAT.

The FAT file system has seven layers of abstraction. The first layer uses just the partition image as input, assuming that the acquisition was done of the raw partition using a tool such as the UNIX "dd" tool. This layer uses the defined Boot Sector structure and extracts the size and location values. Examples of extracted values include:

1. Starting location of FAT;
2. size of each FAT;
3. number of FATs;
4. number of sectors per cluster;
5. location of Root Directory

The abstraction layers of the FAT file system are as follows:

1. Layer 0: Raw file system image;
2. Layer 1: File system image and values from Boot Sector and FAT Entry Size;
3. Layer 2: FAT Area and Data Area;
4. Layer 3: Starting Cluster, FAT Entries;
5. Layer 4: Clusters, Raw Cluster Content and Content Type;
6. Layer 5: Formatted Cluster Content;
7. Layer 6: List of Clusters.

Thus, we see that the digital forensics investigator has to have highly technical skills of understanding the data structure and its representation inside the computer systems where the digital evidence is suspected to reside.

7.16.2 The Legal Challenges in Computer Forensics and Data Privacy Issues

Evidence, to be admissible in court, must be relevant, material and competent, and its probative value must outweigh any prejudicial effect.

Although digital evidence is not unique with regard to relevancy and materiality, there is still a challenge involved. Digital evidence can be easily duplicated and modified; often it can be without even leaving any traces; it can present special problems related to competency. What is more to even reach the point where specific competency questions are answered, digital evidence needs to satisfy the legal admissibility

requirements. Modern computers have enormous data storage facilities. Gigabyte disk drives are common and a single computer may contain several such drives. Seizing and freezing of digital evidence can no longer be accomplished just by burning a single CD-ROM. Failure to freeze the evidence prior to opening the files, coupled with the fact that merely opening the files changes them, can and has invalidated critical evidence. There is also the problem of locating the relevant evidence within massive amounts of data. Examining such volumes of information to find relevant evidence is a daunting task. Owing to this, people often tend to think that there are complicated technical aspects in digital forensics. However, the reality is far from this. Although there are many technical aspects of digital forensics (recall the discussion in the preceding section), they involve understanding the raw data stored inside computer systems, retrieving data from existing or deleted files, interpreting their meaning and putting them within the context of the investigation. Actually, the real challenges involve artificial limitations imposed by constitutional, statutory and procedural issues – we often loose sight of the goal of retrieving evidence!

There are many types of personnel involved in digital forensics/computer forensics: (a) technicians, (b) policy makers and (c) professionals.

Technicians who carry out the technical aspects of gathering evidence. They have sufficient technical skills to gather information from digital devices, understand software and hardware as well as networks. The technical skills include – familiarity with computer hardware – thoroughly knowing the inside of the computer, understanding how hard drives work and their settings, understanding motherboards, understanding how the computer power supply units work and knowing about computer power connections, knowing how computer memory chips work (refer to Fig. 7.8). On the software side, the skills involve thorough understanding of various types of computer OS (e.g., Microsoft OS products, Linux, Unix. etc.), forensics products (diagnostic utilities and forensics diagnostic software and hardware equipment that are available in the market (see Tables 7.9–7.11). In addition, professional forensics training is a must to enter this domain. They explain many technical aspects such as – difference between "clone" and "image" of drive, how do you make a "forensically sound" duplicate of a drive, how can you prove that the duplicate drive is forensically sound, steps to preserve the forensics evidence, etc. For technicians, forensics analysis of E-Mails is also important – it is explained in Section 7.6.

There are also *Policy Makers*; they establish forensics policies that reflect broad considerations – their main focus is on the big picture, but they must be familiar with computing and forensics also. The other entity involved in a digital forensics investigation is the *Professionals* – the link between policy and execution – who must have extensive technical skills as well as good understanding of the legal procedures. Skills for digital forensics professionals are the following:

1. Identify relevant electronic evidence associated with violations of specific laws;
2. identify and articulate probable cause necessary to obtain a search warrant and recognize the limits of warrants;
3. locate and recover relevant electronic evidence from computer systems using tools;
4. recognize and maintain a chain of custody;
5. follow a documented forensics investigation process.

To know more on technical details, legal professionals and those who are practicing in the cybercrime area may refer to Ref. #16, Additional Useful References, Further Reading.

Box 7.17 Drama in Court! Impact of Cyberforensics on Legal Practitioners

It would be no exaggeration to say that the public is primarily educated about forensics science by Hollywood films and television shows. There is no dearth of investigative news shows, documentaries, docudramas, Hollywood films and crime dramas that show us the horrifying and scary details of computer crimes. For example, on the popular YouTube media, there are hundreds of video clips available on these crimes. The media often focuses on law enforcement personnel who use "forensics" techniques to solve crimes. In a way, this is an update to the Sherlock Holmes crime investigation novels, as he used logic and scientific technique to single out the real suspect, often from a plethora of viable candidates. Probably the most popular recent movie series to focus on the use of forensics techniques are those based on the character of criminal, Hannibal Lector. Each of the three movies (Silence of the Lambs, Red Dragon and Hannibal) featured an FBI forensics profiler as one of its main characters.

In a scenario such as described above, an interesting question to explore is: Given the widespread popularity of forensics crime portrayals, do prosecutors and defense lawyers sense a change in jury expectations? Given the media's current emphasis on the importance of forensics science to resolve criminal investigations, a logical question to ask is if prospective jurors now have a higher expectation on the presentation of physical evidence by forensics experts. When criminal cases rely on testimonial and circumstantial evidence, do jury members feel that something is amiss? Is there a possibility that jurists might acquit when forensics evidence is not presented during a trial, but substantial circumstantial and testimonial evidence exists? If these changes are being perceived by attorneys as genuine, have lawyers reacted to changes in jurors by adjusting trial preparation and the presentation of evidence at trial? Additionally, are attorneys questioning jurors' viewing habits to either strike or attempt to retain avid viewers of forensics crime dramas?

Detection and recovery is the heart of computer forensics. This is the aspect which matters in the legal presentation of a cybercrime case in the court. The goal of detection and recovery is to recognize the digital objects that may contain information about the incident and document them. "Acquisition" is to copy and preserve the state of data that could be evidence. By "forensic acquisition of media" we mean the process of making a bit-for-bit copy, or image file, of a piece of media, where image files are frequently used in civil or criminal court proceeding. Therefore, completeness and accuracy of acquisition process is required. In addition, the source of evidence must remain and not get altered by attackers or by normal processes innocently.

Technical persons involved in digital forensics/computer forensics need simple technical skills such as understanding the various kinds of file systems (e.g., the FAT), system software, data organization and specific OS – Windows, Mac and Linux/Unix being the main operating systems in the market today. For the hand-held digital mobile devices there are operating systems such as the Symbian, the Palm OS, file systems and evidence recovery, etc. The legal professionals need to understand the working of court system, the legislations, laws (for cybercrime), and the investigative process and the evidential value of the electronic artifacts recovered/seized as potential evidences to be presented in the court while putting up the case.

Data of digital nature can be very easily deleted or altered; for example, by turning ON the computer or by simply opening/viewing a file or by password protecting files or even by saving data to another platform. Information may be available in areas that necessiate the use of special tools and techniques to identify and review. If data is not properly recovered and analyzed, it may not be admissible and/or credible in a court of law. Therefore, forensics investigators need to be careful in the matter of capturing the "perceived" evidence. They should understand that before they seize a computer or other electronic hardware they must consider whether they require a search warrant. They should be aware that if they wish to access stored electronic communications, they will need to comply with the Privacy Act/Privacy Law if applicable in their country. In order to conduct real-time electronic surveillance, they will need to obtain a wiretap order from a judge.

If access to digital evidence does not come through from an confiscation agency, court orders may become necessary to obtain the data as well as use of the extraction tools to determine whether protocols had been appropriately applied. Conversely, a prosecution or defense team may wish to suppress evidence from discovery, if they believe it could be damaging to the case. This is where the time-consuming aspects of the forensics examination may occur. In general, it is difficult to perform a comprehensive decomposition and logging of all materials (such as the contents of every sector of a terabyte hard drive, or thousands of hours of digital video from a surveillance camera), so a "scratch-and-sniff" approach might be used to yield promising information. Even though cost-effective, tactical decisions to proceed with only a partial investigation may be regretted in hindsight if a post-mortem comprehensive analysis shows that an alternative outcome might have prevailed.

Box 7.18 Beware – Forensics Acts and Laws!

It is said that forensics is a territory of dilemmas! The foremost dilemma with the study of electronic law is that it is very complicated to confine its study within simple parameters; Internet and electronic commerce do not define a distinct area of law as with contract and tort law. Electronic law crosses many legal disciplines, each of which can be studied individually. Cybercriminals abound – there will always be those in the world who wish to gain some benefit without actually paying for it. As a result, the electronic law will also cross-over certain aspects of criminal law.

In Chapter 1, many forms of cybercrimes are mentioned. Whether it is racial or sexual harassment, stalking, bullying at work or neighbors from hell, harassment is a form of discrimination that is generally prohibited by legislation. Harassment in the workplace is something employers must not tolerate, and includes any form of unwelcome, unsolicited, or unreciprocated behavior that a reasonable person would consider offensive, humiliating or intimidating. Chapters 1 and 2 mention about cyberstalking. Cyberstalking is the distribution of malicious communication through E-Mail and the Internet. Although based on new technology, it is in principle precisely the same as many other form of malicious communication and can be dealt with through the usual civil and criminal law methods. The distribution of offensive E-Mail through the Internet and such communication will also constitute an offense under a variety of statues (such as the Malicious Communication Act in the UK).

Chapter 1 mentioned Children's Online Privacy Protection Act (COPPA). Pornography is a big business on the Internet and has even been seen by some as its foundation. In the US, pornography is protected as speech under the First Amendment of the Constitution. Obscenity, on the other hand, is not protected. Obscenity may be legally possessed in an individual's private home, but generally its distribution is illegal.

In much of the common law world, law enforcement needs to obtain a legal authorization in order to search and seize evidence. Generally, this power is granted through a request for a search warrant that states the grounds for the application, including the law that has been broken. In the US, the requirements demand that the application describe the specific premises need to be searched, as well as the items being sought. In the physical world, there is a real limit on the length of time during which a search can be conducted, This rule does not impose much of a limit on electronic searches. As investigators make a copy of the digital evidence (such as a hard drive), they are able to continue to search those files for "string," which are beyond the scope of the original warrant, and do so at their leisure. According to the Fourth Amendment rule, an investigator executing a warrant is able to look in any place listed on the warrant where evidence might conceivably be concealed. Text of the Fourth Amendment states the following:

> "*The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized*."

As per traditional practices, an investigator was precluded from looking into any location beyond the area of the evidence he/she wishes to seize. Electronic evidence, however, may be stored anywhere. The result is that in an investigator can electronically look anywhere in search of digital evidence. For more details on the Fourth Amendment in the cyberspace, refer to the link mentioned

Box 7.18 Beware – Forensics . . . (*Continued*)

at the end of this box. Chapter 6 mentions about "evidence." The Indian IT Act impacts the Indian Penal Code 1860, the Indian Evidence Act 1872, The Bankers' Books Evidence Act 1891, The Reserve Bank of India Act 1934 (see Chapter 6) to make them in tune with the provisions of the IT Act. Electronic evidence in law is the legal recognition and evidential value in litigation of evidence in digital format. An Anton Piller order is a civil court order providing for the right to search premises and seize evidence without prior warning. In the US, the Business Software Alliance has used those orders as a remedy when they are attempting to stop illegal software use (termed software piracy) and copyright infringement to achieve the recovery of property.

Refer to the following URL mentioned for a complete presentation on the Fourth Amendment: http://law.uoregon.edu/faculty/shoar/docs/cc10/2010-fourthamendment.ppt (7 September 2010). Also see the link in Ref. #6, Video Clips, Further Reading.

7.17 Special Tools and Techniques

We present in this section ready reckoner of forensics tools as well as some special techniques such as "data mining." Forensics tools have also been mentioned in various sections of this chapter so far – for example "file carving" is mentioned in Box 7.10. Most of the "file carvers" tools operate by first looking for file headers and/or footers and then by "carving out" the blocks between these two boundaries. Outlook files with extension "jpeg" and MS Word files are fragmented and, therefore, appear corrupted or missing to a user using traditional data carving. "Disk duplication" equipments were mentioned in Section 7.11. Embedded memories inside the computer were mentioned in Section 7.7.2. Recall that Computer Online Forensics Evidence Extractor (COFEE) was mentioned in Box 7.1. Helix is another well-known tool for digital forensics investigation. Helix is a tool specially tailored for incident response, system investigation and analysis, data recovery and security auditing. It is attuned to experienced users and system administrators working in small-to-medium, mixed environments, where threats of data loss and security breach are high. Helix has two modes – pure Linux bootable live CD and the Windows mode, where it can be used *in vivo* on top of a desktop running Windows OS. Helix is available for download by E-Mail registration. For further technical details on Helix, visit http://www.dedoimedo.com/computers/helix.html (10 February 2010).

The list of carving tools is presented in Table 7.9. The list of reviewed tools is presented in Table 7.11. Cyberforensics expert Peter Stephenson, when asked, about how important tools are in digital forensics, responded by saying "… *Essentially, incident management is a forensics problem. That challenge demands a serious toolkit of computer forensic, network-enabled forensic, network forensics and analytical tools.*" In the words of another well-known forensics analyst Steve Hailey, "*If the tools being used are the mechanism to find evidence on a computing device, and several different tools can replicate the process, then it doesn't matter what tools were used.*"

Most tools have the same underlying principles:

1. Creating forensics quality or sector-by-sector images of media;
2. locating deleted/old partitions;
3. ascertaining date/time stamp information;
4. obtaining data from slack space;
5. recovering or "undeleting" files and directories, "carving" or recovering data based on file headers/file footers;
6. performing keyword searches;
7. recovering Internet history information.

7.17.1 Digital Forensics Tools Ready Reckenor

The list of "carving" tools presented in Table 7.9 is divided in three main categories (a) data recovery (b) partition recovery and (c) carving. The associated websites are also mentioned for more information on these tools. Readers may like to re-visit Box 7.10.

Table 7.9 | List of carving tools

Sr. No.	*Name of the Tool*	*Brief Description*
		Data Recovery Tools
1.	Norton Disk Edit	The master boot record is required to boot your computer. Having a current backup of your master boot record is an excellent way to ensure that, in the event of a virus or hardware failure, you will be able to recover your system in the shortest amount of time possible.
2.	HD Doctor Suite	It is a set of professional tools used to fix firmware problem. To know more on this, visit http://www.salvationdata.com/data-recovery-equipment/hd-doctor.htm
3.	SalvationDATA	This tool claims to have a program that can read the "bad blocks" of Maxtor drives with proprietary commands. To know more on this, visit http://www.salvationdata.com/
4.	BringBack	The tool offers easy to use, inexpensive and highly successful data recovery for Windows and Linux (ext2) operating systems and digital images stored on memory cards, etc. To know more on this, visit http://www.toolsthatwork.com/bringback.htm
5.	RAID Reconstructor	Recall the RAID levels explained in Box 7.11. Runtime Software's RAID Reconstructor will reconstruct RAID Level 0 (Striping) and RAID Level 5 drives. To know more on this, visit http://www.runtime.org/raid.htm
6.	e-ROL	e-Rol allows you to recover through the Internet files erased by mistake. Recover your files online for free. To know more on this, visit http://www.e-rol.com/en/
7.	Recuva	Recuva is a freeware Windows tool that recovers accidentally deleted files. To know more on this, visit http://www.piriform.com/recuva
8.	Restoration	Restoration is a freeware Windows software that will allow you to recover deleted files. To know more on this, visit http://www.snapfiles.com/get/restoration.html
9.	Undelete Plus	Undelete Plus is a free deleted file recovery tool that works for all versions of Windows (95-Vista), FAT12/16/32, NTFS and NTFS5 filesystems and can perform recovery on various solid state devices. To know more on this, visit http://www.undelete-plus.com/
10.	R-Studio	R-Studio is a data recovery software suite that can recover files from FAT(12-32), NTFS, NTFS 5, HFS/HFS+, FFS, UFS/UFS2 (*BSD, Solaris), Ext2/Ext3 (Linux) and so on. To know more on this, visit http://www.data-recovery-software.net/
11.	Stellar Phoenix	Data recovery software services and tools to recover lost data from hard drive. Visit http://www.stellarinfo.com/
12.	DeepSpar Disk Imager	It is a dedicated disk imaging device built to handle disk-level problems and to recover bad sectors on a hard drive. To know more on this, visit http://www.deepspar.com/
13.	Adroit Photo Recovery	This is a photo recovery tool that uses validated carving and is able to recover fragmented photos. Adroit Photo Recovery is able to recover high definition RAW images from Canon, Nikon, etc. To know more on this, visit http://photo-recovery.info/

(*Continued*)

Table 7.9 | (*Continued*)

Sr. No.	*Name of the Tool*	*Brief Description*
		Partition Recovery Tools
1.	Partition Table Doctor	It helps recover deleted or lost partitions (FAT16/FAT32/NTFS/NTFS5/EXT2/EXT3/SWAP). To know more on this, visit http://www.ptdd.com/index.htm
2.	NTFS Recovery	DiskInternals NTFS Recovery is a fully automatic utility that recovers data from damaged or formatted disks. To know more on this, visit http://www.diskinternals.com/ntfs-recovery/
3.	gpart	Gpart is a tool which tries to guess the primary partition table of a PC-type hard disk in case the primary partition table in sector 0 is damaged, incorrect or deleted. To know more on this, visit http://www.stud.uni-hannover.de/user/76201/gpart/ (resticted permission)
4.	TestDisk	This is an OSS tool (open-source software) and is licensed under the GNU Public License (GPL). To know more on this, visit http://www.cgsecurity.org/wiki/TestDisk
5.	Partition Recover Software	Partition Recovery software for NTFS and FAT system that examines lost windows partition of damaged and corrupted hard drive. To know more on this, visit http://www.stellarinfo.com/partition-recovery.htm
		File Carving Tools
1.	DataLifter® - File Extractor Pro	Data carving runs on multiple threads to make use of modern processors. Visit http://www.datalifter.com/products.htm
2.	Simple Carver Suite	This is a collection of unique tools designed for a number of purposes including data recovery, forensics computing and E-Discovery. The suite was originally designed for data recovery and has since expanded to include unique file decoding, file identification and file classification. Visit http://www.simplecarver.com/
3.	Foremost	Foremost is a console program to recover files based on their headers, footers and internal data structures. Visit http://foremost.sourceforge.net/
4.	Scalpel	Scalpel is a fast file carver that reads a database of header and footer definitions and extracts matching files from a set of image files or raw device files. Scalpel is filesystem-independent and will carve files from FATx, NTFS, ext2/3, or raw partitions. Visit http://www.digitalforensicssolutions.com/Scalpel/
5.	CarvFs	A virtual file system (fuse) implementation that can provide carving tools with the possibility to do recursive multi-tool zero-storage carving (also called in-place carving). Patches and scripts for scalpel and foremost are provided. Works on raw and EnCase images. For more information, visit http://www.forensicswiki.org/wiki/CarvFs
6.	LibCarvPath	A shared library that allows carving tools to use zero-storage carving on carvfs virtual files. For more information, visit http://www.forensicswiki.org/wiki/LibCarvPath
7.	PhotoRec	This is file data recovery software designed to recover lost files including video, documents and archives from hard disks and CDRom and lost pictures (thus, its "Photo Recovery" name) from digital camera memory. Visit http://www.cgsecurity.org/wiki/PhotoRec

(*Continued*)

Table 7.9 | (*Continued*)

Sr. No.	*Name of the Tool*	*Brief Description*
8.	PhotoRescue	Datarescue PhotoRescue Advanced is picture and photo data recovery solution made by the creators of IDA Pro. PhotoRescue will undelete, unerase and recover pictures and files lost on corrupted, erased or damaged compact flash (CF) cards, SD Cards, Memory Sticks, SmartMedia and XD cards. For more information, visit http://www.datarescue.com/photorescue/
9.	Revlt	RevIt (Revive It) is an experimental carving tool, initially developed for the DFRWS 2006 carving challenge. It uses "file structure-based carving." Note that RevIt currently is a work in progress. For more information, visit https://www.uitwisselplatform.nl/projects/revit
10.	Magic Rescue	Magic Rescue is a file carving tool that uses "magic bytes" in a file contents to recover data. For more information, visit http://www.student.dtu.dk/~s042078/magicrescue/
11.	FTK	FTK2 includes some file carvers. *Note:* see Table 7.11
12.	SmartCarving	SmartCarving is a file carving technique to recover fragmented files. SmartCarving utilizes a combination of structure-based validation along with validation of each file's unique content.
13.	GuidedCarving	This is a technique to recover fragmented files introduced in *Adroit Photo Forensics*. GuidedCarving allows a user to attempt to recover a fragmented file that failed to fully recover using SmartCarving.
14.	Adroit Photo Forensics	Adroit Photo Forensics supports data carving of popular image formats. Also supports fragmented carving using *SmartCarving* and *GuidedCarving.*

Note: The term "Data Recovery" is frequently used to mean forensics recovery, but the term really should be used for recovering data from damaged media.

The information in Table 7.10 is based on Dr. Peter Stephenson's in-depth review of forensics tools conducted in year 2006. The survey by Dr. Stephenson is considered to be one of the best known among similar other surveys. Links for other survey are mentioned after the table.

Prices indicated are "as at point in time" that is the exact prices should be found out at the time of purchase. Versions indicated are as per testing done in 2006 – for more information, refer to website www.scmagazine.com

Links to other surveys/reviews about forensics tools are provided below:

1. *Best Forensics Tools – 2007 Edition* can be accessed at the following link: http://www.dragoslungu.com/2007/04/17/best-forensics-tools-2007-edition/ (4 April 2010).
2. *EnCase Forensic 6 – Review* is available at the following link: http://whereismydata.wordpress.com/2008/08/31/encase-forensic-6-review/ (3 April 2010).
3. The following link will take you to *Computer Forensics Tool Testing (CFTT) Survey:* http://blogs.sans.org/computer-forensics/2010/03/04/computer-forensics-tool-testing-cftt-survey/ (1 April 2010). Table 7.11 provides information about the tools compared in Table 7.10. Reference websites are also provided.

Table 7.10 | Forensics tools features comparison at a glance

Product Name	*Coroner's Toolkit*	*Encase Forensics*	*Forensics Toolkit*	*i2 Analyst's Notebook*	*LogLogic LX2000*	*Mandiant First Response*	*NetWitness*	*ProDiscover Incident Response*	*Sleuth Kit/Autopsy Browser*
Supplier	Open source	Guidance software	Access data	i2Inc.	LogLogic	Mandiant	Man Tech Intl.	Technology Partners	Open source
UNIX/Linux platform	Yes	No	No	No	Yes	No	Yes	No	Yes
Windows Platform	No	Yes	Yes	Yes	No	Yes	No	Yes	No
Analyzes W – Windows U – UNIX/Linux	U	W, U	W, U	—	—	W	—	W, U	W, U
Remote Capture	No	No	No	—	Yes	Yes	—	Yes	No
GUI	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Requires Remote Agent	No	No	No	No	No	Yes	No	No	No
Preforensics Audit	No	Yes	Yes	No	No	Yes	No	Yes	No

Table 7.11 | Top tools in digital forensics

Name of the Tool	*Brief Information*
The Coroner's Toolkit (TCT) Version 1.16 Visit www.porcupine.org	• OSS tool (open-source software) • No cost – OSS • Not a GUI-based product • Used with UNIX platform – it is a collection of command line tools • Written by Dan Farmer and Wietse Venema • Deep UNIX knowledge is the prerequisite to use TCT because it is a UNIX-only tool • Documentation is not detailed • Support to this tool is not much available – like most OSS tools, users are expected to fend for themselves
EnCase Forensics Version 5.0 Visit www.guidancesoftware.com	• GUI-based • Expensive • Adequate documentation • Simple to operate and use • Can acquire many different media • Plenty of web support • Costs around $3000 Targeted at large organizations
Forensics Toolkit (FTK) Version 1.61 Visit www.accessdata.com	• Comprehensive with many features but not simple to use • Overwhelming program interface • Good documentation • Comes with a USB-pluggable hardware device • Costs around $1100
i2 Analyst's Noted Version 6.0.55 Visit www.i2inc.com	• Very different type of forensics analysis tool • Can import metadata from EnCase • Data can be imported from spreadsheet using CSV file • Has the ability to analyze complex crimes • Help system is very good • People-based support is expensive (online support and phone support is also available • Costs around $3700
LogLogic's LX 2000 Visit www.loglogic.com	• Highly expensive – costs around $50000 • Excellent tool for log analysis – analyzes logs in real-time mode • Mature and useful product • High functionality but complex to set up • Excellent interface but high learning curve (many hidden features to be understood) • Adequate documentation
Mandiant First Respolnse Version 1.1 Visit www.mandiant.com	• No cost – freeware forensics audit tool with strong audit features • Not easy to use – awkward interface • Useful features once understood • Gathers forensics information in an organized and simple-to-read fashion • Limited support and limited documentation • Deploys agents across network computers to gather a snapshot before evidence is gathered

(*Continued*)

Table 7.11 | (*Continued*)

Name of the Tool	*Brief Information*
NetWitness Version 6.0 Visit www.netwitness.com	• Network traffic security analyzer and works as a security intelligence tool. Helps automate IDS analysis process • Good user interface • Set up is easy due to installation wizard • Good user interface • Low scalability • Inadequate documentation • Web-based support and E-Mail-based support for registered users • Costs $30,000 • Not suitable for large enterprises
ProDiscover Incident Response Version 4.55 Visit www.techpathways.com	• Complete IT forensics tool can access computer over the network • Fairly easy to use • Supports remote analysis of running processes, open files, open port and services running on open ports • Not for first time users – experience in forensics required • Well laid-out documentation • Full disk imaging capability, ability to find hidden data, file metadata information, hash keeping as well as across network data gathering • Costs $8000
Sleuth Kit and Autopsy Browser Visit www.sleuthkit.org	• No cost – it is a freeware • Good documentation • Good support • Straightforward to use for those familiar with UNIX environment but difficult to use by those not familiar with UNIX • Can use Non-UNIX file systems too • The browser can run on any HTML environment • Adequate documentation • Support is much better as compared to many other OSS – E-Mail-based support is available and many active user forums are available

7.17.2 Special Technique: Data Mining used in Cyberforensics

Data mining is a very vast topic; full treatment of this topic is beyond scope of the chapter. In this section, we are only going to explain how data mining techniques are applied in cyberforensics. Chapter 1 presents the various categories of cybercrimes. A criminal act can encompass a wide range of activities, from civil infractions such as illegal parking to internationally organized mass murder such as the 9/11 attacks in New York, US and 26/11 attacks in Mumbai, India. Law enforcement agencies across the world compile crime statistics. Depending on the type of cybercrimes, the impact and the impacted parties can vary. Some examples of impact and impacted parties are national security and government, financial impacts and individuals, brand image and organizations. More on this is presented in Chapter 9.

Traditional data mining techniques such as association analysis, classification and prediction, cluster analysis, and outlier analysis identify patterns in structured data. To know more on these analyses, visit http://www.thearling.com/glossary.htm. Newer techniques identify patterns from both structured and unstructured data. Literature [28] shows that as with other forms of data mining, crime data mining raises privacy concerns. Automated data mining techniques are being researched for both local law enforcement and national security applications. A brief explanation of some of the data mining techniques is as follows:

1. **Entity extraction:** This technique is used to identify particular patterns from data such as text, images or audio materials. It has been used to automatically identify persons, addresses, vehicles and personal characteristics from police narrative reports.[29] In computer forensics, the extraction of software information such as the data structure, program flow, organization and quantity of comments, and use of variable names, can facilitate further investigation by grouping similar programs written by hackers and tracing their behavior. Entity extraction technique provides basic information for crime analysis, but its performance depends greatly on the availability of extensive amounts of clean input data.
2. **Clustering techniques:** This involves grouping data items into classes with similar characteristics to maximize or minimize intraclass similarity. For example, in order to identify suspects who conduct crimes using similar methods or to distinguish among groups that belong to different gangs. These techniques are not featured with a set of predefined classes for assigning items. The statistics-based *concept space algorithm* is used to automatically associate different objects such as persons, organizations and vehicles in crime records.[30] The Financial Crimes Enforcement Network AI System (AI is the branch of computer science, known as "Artificial Intelligence")[31] uses link analysis techniques (to identify transactions' patterns) to exploit Bank Secrecy Act data to support the detection and analysis of money laundering and other financial crimes. The technique of "clustering" crime incidents can automate a major part of crime analysis but is limited by the high computational intensity typically required.
3. **Association rule mining:** This technique discovers frequently occurring item sets in a database and presents the patterns as rules. This technique has been applied to detect network intrusion and to derive association rules from users' interaction history. Investigators can also apply this technique to network intruders' profiles to help detect potential future network attacks. [32]

Automated techniques to analyze different types of crimes need a unifying framework describing how to apply them. In particular, there is a need for understanding the relationship between analysis capability and crime type characteristics. This understanding can help investigators more effectively to use those techniques to identify trends and patterns, address problem areas and even predict crimes. After having completed a brief overview of data mining techniques in cyberforensics, we now discuss forensics auditing.

7.18 Forensics Auditing

"Forensics auditing" is also known as "forensics accounting."

Forensic auditing includes the steps needed to detect and deter fraud. Forensics auditors make use of the latest technology to examine financial documents and investigate white-collar crimes like such as frauds, identity theft, funds embezzlement, securities fraud, insider trading, etc. Forensics accounting is a specialized form of accounting; it uses accounting, auditing and investigative techniques. Forensics accounting professionals are assigned specialty tasks, such as analyzing and tracking evidence of economic transactions. In some cases, they are asked to present this evidence to a court of law. Forensics auditors are responsible for detecting fraud, identifying individuals involved, collecting evidence, presenting the evidence in criminal proceedings, etc. From career perspective, forensics auditors can work in both large and small organizations like insurance companies, banks, courts, government agencies and law firms. Cybersecurity Careers are addressed in Chapter 12.

There is almost always some legal/evidential angle in forensics auditing. Consider as an example of forensics auditing, the investigation of a fraud or presumptive fraud with the objective of gathering evidence that could be presented in a court of law. There is an increasing use of auditing skills to prevent fraud by identifying and rectifying situations that could lead to frauds being perpetrated (i.e., risks). It might be useful, therefore, to categorize forensics auditing as being either "reactive" or "proactive."

"Insider trading" needs some explanation. According to the American Heritage Dictionary, "insider trading" is the illegal buying or selling of securities on the basis of information that is unavailable to the public. It involves trading of a corporation's stock or other securities (e.g., bonds or stock options) by individuals with potential access to non-public information about the company.

"Insider trading"[33] refers to two separate financial transactions – one being perfectly legal and the other being subject to massive civil fines and possible prison time. There is a legal form of insider trading – it involves the sale of securities or stocks by officers of a company or stockholders who own more than 10% of the company. In many countries, trading by corporate insiders, for example, officers, key employees, directors and large shareholders, etc., may be considered legal if this trading is done in a way that does not take advantage of non-public information. However, the term is frequently used to refer to a practice in which an insider or a related party trades based on non-public information obtained during the performance of the insider's duties at the corporation, or otherwise in breach of a fiduciary or other relationship of trust and confidence or where the non-public information was misappropriated from the company.

Recall the discussion about regulatory perspective for forensics (Section 7.15.1). Government departments/agencies can possibly use the techniques of forensics auditing to assess compliance with regulations governing payments of grants/subsidies. Compliance auditors could also use these techniques while auditing such governmental programs. We have mentioned about "steganography" in Section 7.12. Antiforensics tools can hide data with cryptography or steganography. "Antiforensics" is addressed in the next section. Although steganography has not yet come either under the direct scope of IT Audits or under cyberforensics investigation, because it is not yet considered a direct threat by auditors and cyberforensics investigators, it is the one that needs to be considered and understood for possible future occurrences.

Box 7.19 Auditing vis-à-vis Cyberforensics Investigation

For many people, both the terms auditing and cyberforensics investigation are the same but actually they are not. Typically, an "audit" involves examination of information and operations for accuracy, legality and propriety. Internal audits are meant to report risks and to make recommendations to promote sound-operating practices. Audit examinations typically involve documents, records, reports, internal control systems, accounting procedures and actual operations. On the other hand, "cyberforensics investigation" is the process of extracting information and data from computer storage media and guaranteeing its accuracy and reliability (typically in an evidential context).

Box 7.19 Auditing vis-à-vis . . . (*Continued*)

Objective of an "audit" is to determine whether all transactions are properly recorded in the accounts, and appropriately reflected in the organization's statement and reports. The objective of a "cyberforensics investigation," on the other hand, is to identify "digital evidence" using scientifically derived and proven methods that can be used to facilitate or to help reconstruct events in an investigation. The secondary objective is to identify the responsible person and seriousness of the misconduct.

Auditing and cyberforensics investigation vary in their "scope" too. Scope of an audit typically includes – audit objectives, risk assessment, nature and extend of controls testing (compliance testing or substantive testing) and the extent of auditing procedures to be performed, reliance on previous audits. The scope of cyberforensics investigation involves scientific methods to identify, collect, analyze, validate, interpret, preserve, document and present electronic evidence derived from digital sources.

Results of audit are generally communicated via written report to management. On the other hand, report of a cyberforensics investigation is submitted to the investigator, prosecutor, law enforcement, organizational management and others. The impacts differ too. An audit is conducted in a non-confrontational manner, with generally helpful cooperation by the auditee whereas a cyberforensics investigation may be adversarial. Each investigation is independent and unique in itself.

Recall the discussion in Section 7.5.1 The Rules of Evidence – "secure, auditable digital date/time stamps" will have the following attributes:

1. **Accuracy:** The time presented is from an authoritative source and is accurate precision required by the transaction, whether day/hour/millisecond.
2. **Authentication:** The source of time is authenticated to a suitable timing laboratory so that a third party can verify the precision and accuracy.
3. **Integrity:** The time should be secured and not subjected to corruption during "handling." If it is corrupted, either inadvertently or deliberately, the corruption will be apparent to a third party.
4. **Non-repudiation:** An event or document should be bound to its time so that association between event or document and the time cannot be later denied.
5. **Accountability:** The process of acquiring time, adding authentication and binding it to the subject event should be accountable, so that a third party can be assured that due process was applied and that no corruption transpired.

"Digital Evidence Collection" is a game of patience. The potential for fraud, unintended errors can be eliminated by adding secure and auditable time to digital evidence. The use of secure date/time stamps can not only improve the integrity digital evidence, but also provide higher assurance required for digital chain of evidence. Using secure and auditable time helps to ensure that any important electronic time stamp that cannot be corrupted has an evidentiary trail of authenticity. The secure issuance of timestamps for digital evidence has some critical components associated with them:

1. First, it must be remembered that digital data needs binding of time. Such binding of time with digital data must occur within a trusted computing environment to assure the "efficacy of the time stamping process."
2. Next, it is important to consider the accuracy of the clock used as the source for time stamping. The clock used should be appropriate for the application. For example, the accuracy of a timestamp

indicating access to a secure facility through the use of a card access or biometric device of 30 seconds may be reasonable. However, the time stamp on an electronic stock transaction or money transfer may warrant a finer resolution.

3. When a local trusted clock is used as the source for time stamping, its calibration and audit must be routine, continuous and traceable. Furthermore, to make the audits reliable, the audit of such clocks must be performed by a trusted, disinterested third party.
4. Finally, the issuer must verify the validation of the resulting timestamps. The verification/validation records ought to be made available to any party that has the need to evaluate the accuracy, validity, trustworthiness or traceability of a timestamp.

7.19 Antiforensics

"Antiforensics" is the application of scientific method to digital media to invalidate factual information for judicial review. There are four categories of antiforensics: (a) *Data destruction;* (b) *data hiding*; (c) *data encryption* and (d) *data contraception*. Antiforensics is a combination of people, process and tools. There are several counter-forensics commercial software tools available in the market. They are designed to eliminate specific records and files but leave system otherwise functional, that is, overwrite deleted data to thwart recovery and cope with system files, like the Registry. Counter-forensics tools are increasingly reported as important factors in legal action. Organizations must seek to understand the mindset, skill set and capabilities of those employing antiforensics techniques. By way of a situational context, this is like trying to understand from cyberattackers' perspective so that threats posed to the information systems can be understood. Cybercriminals exploit the fact that forensics takes time. Recall "Locard Exchange Principle" (Box 7.5) – conventional wisdom tell us that an attacker will attempt to leave as little evidence as possible. From another angle, however, there are significant advantages to an attacker for creating extraneous evidence. The first is the time factor mentioned earlier; forensics investigation takes time and time is money! Multisystem compromises against enterprise networks resulting in non-linear increase in the amount of effort that goes in accurate analysis of suspect system. In situations where an extraordinarily large number of computers are under suspicion, businesses can rarely perform a full forensics analysis of all the computers.

Recall the discussion in Section 7.16.2. In that context, here is another point on forensics from "privacy" perspective: Modern OS and applications that run on the OS generate a high amount of data about users' activities. Current trends in computer use raise concerns about recovering "privacy-sensitive" data from computer systems, especially given the mobile computing trends and "remote working" trends. The line between "home" and "office" is getting thinner as work takes place round the clock in a "work anywhere" and "any time" mode. Employees use company computers (desktops, laptops, etc.) for personal E-Mails, banking, shopping, listening to music, watching video, etc. (although some organizations have strict rules and guidelines on allowed usage and security restrictions on sites to be visited using company computers). Refer to Chapter 9 and Appendix C.

When companies provide employees with laptops to work from home,[34] other family members may also use these computers. In such a scenario, company computers often may contain "sensitive personal information" (along with business confidential information) which individuals want to keep private. The laptops may also contain records that companies would like to protect and examine. Users are getting increasingly aware of their "privacy exposure" from these records and the "digital artifacts" that linger even after files are "deleted" on the computer they use. As a result of this, a range of "counter-forensic" privacy tools have emerged. These tools are nothing but software designed to irretrievably eliminate records of

computer system usage and other "sensitive data." Following are some of the lists of well-known tools with "counter-forensics features":

1. Windows Washer;
2. Windows and Internet Cleaner;
3. CyberScrub Pro;
4. Evidence Eliminator;
5. Acronis Privacy Expert;
6. SecureClean.

A table[35] showing feature comparison of 1, 2, 3 and 4 is available in a link in References. However, those are not the only antiforensics tools available in the market place today. More are described in the next paragraph.

Metasploit antiforensics investigation arsenal includes tools such as (a) *timestomp*,[36] (b) *Slacker*, (c) *transmogrify* and (d) *Sam Juicer*. Let us first understand how "timestomp" acts as an antiforensics tool. Timestomp uses Windows system calls *NtQueryInformationFile()* and *NtSetInformationFile()*. However it does not use the call *SetFileTime()*. Timestomp features include display and set MACE attributes to mess up with EnCase and MS Anti-Spyware. Timestomp leverages a series of Win32 system calls to modify the Last Modified (M), Last Accessed (A), Creation Date (C) and Entry Modified (E), together referred to as "MACE." During a forensics analysis, the examiner will use these values to attempt to piece together a timeline of event. If an attacker is able to undetectably modify these entries then the examiner can no longer rely on timestamps to create a timeline for the crime committed. By performing a series of standard Win32 function calls, an attacker can place data at the end of a cluster. A subsequent series of function calls allows an attacker to retrieve the stored data. Timestomp allows the techie criminal to hide information on a system that may not be immediately distinguished from other random slack space information. Little can be done to examine hidden slack space information that has been properly obfuscated or encrypted. This explains how "timestomp" may permit an attacker to subvert file time stamps to corrupt a forensics analysis; however, it can also be used to validate various forensics tools for reliability. Thus, it works like a double-edged sword!

"Transmogrify" is a simple search and replace engine that allows for the file signatures to be changed between various types. An attacker might change a JPG file to show up as a Windows Executable. Most popular forensics tools only perform the most basic pattern matching and file extension examination to identify file type. Therefore, an investigator relying on these tools is most likely to misidentify the file type and allow it to go as unexamined! The only way to tell if a file is a JPG would be to open it, and the only way to determine if a file is an EXE would be to execute it. This is because a JPG hidden as an EXE would never run and an EXE hidden as a JPG would never display!

Timestomp and *transmogrify*, mentioned in the previous paragraphs, are tools that focus on changing, hiding or planting misleading evidence. "Sam Juicer" is designed to help advanced attackers to prevent evidence from ever being created. As a result, no evidence will ever come to disk and, therefore, a postmortem forensics analysis of the disk will not reveal any clues as to how the compromise occurred. The analysis will not even throw light about the extent of control the attacker had on the computer system. There is no easy solution (as at the time of writing this) available to prevent Sam Juicer from running once the machine has been compromised.

A well-known tool, which is used for data hiding, is called *Slacker*[37] – it is part of the Metasploit framework mentioned earlier in the section. Slacker can hide data within the slack space of FAT or New Technology File System (NTFS) file system. Slacker breaks a file into pieces and places each piece of that file into the

slack space of other files, thereby hiding it from the forensics examination software. At the beginning of this section, we mentioned that there are four categories of antiforensics – data hiding is one of those categories. Data hiding technique involves the use of bad sectors. When performing this technique, the user changes a particular sector from good to bad and then data is placed onto that particular cluster. It is a common belief that forensics examination tools will see these clusters as bad and continue without any examination of their contents.

Forensics tools "Sleuth Kit" and "EnCase" are mentioned in Section 7.10 and Sections 7.7.2 and 7.10.2, and Tables 7.10 and 7.11 describe their features. EnCase relies on the Windows API to perform timestamp translation. However, a glitch in the Windows API results in a blank value being displayed when the time stamp values are set below a certain threshold. Thus, we need to appreciate that computer forensics tools are neither panacea nor magic; after all, they are only complex software tools that like all software may be subject to certain attacks. Although these tools play such a critical role in our legal system, it is important that they be as accurate, reliable and secure against tampering as possible. Vulnerabilities would not only question the admissibility of forensics images, but could also create a risk that if undetected tampering occurs, courts may come to wrong decisions in cases that affect lives and property. Antiforensics is more than a technology. This approach to criminal hacking can be summed as follows: *Make it hard for them to find you and impossible for them to prove they found you*. If an attacker succeeds in making a cyberforensics investigation extremely costly, then he/she can actually create a business case against in-depth forensics analysis!

Summary

The field of digital forensics/computer forensics has grown rapidly in the 21st century, most notably due to the increased trend in mobile devices found at technical, non-technical and violent crime scenes and the rise in mobile workforce in the global economy. Mobile computing/remote working, etc. are some of the emerging patterns of work. In this chapter, fundamentals of cyberforensics and its associated aspects were presented. Cyberforensics has become an important domain given the kind of world in which we live now and the way businesses now operate. Cybersecurity has become a mission-critical component in modern times. When security threats are not analyzed effectively, the result can be unpredictable catastrophes. The emergence of information forensics comes from the incidence of criminal, illegal and inappropriate behaviors. We are living in the knowledge age where information and knowledge are the most sought after commodities. Criminals, competitors and even employees exploit loopholes in current security architectures and control structures; use antiforensics techniques and tools to hide their traces; and apply forensics tools and techniques to obtain the required information to commit cybercrimes. Steganography is a dynamic tool with a long history and the capability to adapt to new levels of technology. As steganographic tools reach the stage of advanced technological features, the steganalyst and the tools they use must also advance. In reference to steganography used by cybercriminals, we also explained about rootkits, which is a set of software tools inserted by an intruder into a computer in order to allow that intruder to enter the computer again at a later date and use it for malicious purposes without being detected. These purposes include (a) collecting data about computers (including other computers on a network) and their users (such as passwords and financial information), (b) causing such computers to malfunction and (c) creating or relaying Spam. Like any tool, steganography (and steganalysis) is neither inherently good nor evil; it is the manner in which it is used which will determine whether it is a benefit or a detriment to our society.

Security has become a major concern on social networks. It is very important that we find the right solutions to tackle the different security problems on the social networking sites today. The chain of evidence and accuracy of digital evidence is very important in cyberforensics investigation. Experienced human investigators can often analyze crime trends precisely, but as the incidence and complexity of crime increases, human errors occur, analysis time increases and criminals have more time to destroy evidence and escape arrest. By increasing efficiency and reducing errors, crime data mining techniques can facilitate police and enable investigators to allocate their time to other valuable tasks. Attackers' objective is to make forensics investigation difficult. They aim at foiling the investigations. Antiforensics behaviors, tools and technologies are, therefore, an area of concern when they do not get caught, as discussed in the chapter.

Review Questions

1. Is there a difference between computer security and computer forensics? Explain.
2. Can a cybercrime investigation be done without involving a forensics expert? Explain with reasons.
3. Explain how the "chain of custody" concept applies in computer/digital forensics.
4. Explain the importance of strong documentation in cyberforensics profession.
5. Is there a difference between "digital forensics" vis-à-vis "computer forensics"? Explain.
6. Explain the role of digital forensics. What do you think is the reaction of traditional legal communities about role of "digital evidence" in crime? Prepare a debate note by considering your own view as well as by talking to the legal community professionals and/or the professors in the institute where you are studying.
7. Explain the importance of "chain of custody" concept. Provide illustrations to support your answer.
8. Do you think the *Indian Evidence Act* is adequate to handle digital evidence? Explain your answer with supporting illustrations.
9. Explain some of the best practices in handling digital evidence. Explain what "rules of evidence" are.
10. Explain how an E-Mail can be traced for forensics purpose. Outline the various key steps involved.
11. What are the various phases and activities involved in the life cycle of a forensics investigation process? Support your answer through various relevant examples.
12. What are the different types of digital analysis that can be performed on the captured forensics evidence?
13. What are the typical elements of a digital forensics investigation report?
14. What would be the nature of evidence collected for network forensics?
15. What role does an "expert witness" play in a cyberforensic/digital forensics case?
16. What precautions should be taken while collecting electronic evidence? What are the things to be avoided during a cyberforensic/digital forensics investigation? Support your answers with examples. What are the things that *cannot* be avoided?
17. Explain why the NDA (non-disclosure agreement) is important in a forensics investigation. What do you think are the risks that may arise if an NDA is not signed before commencing the investigation?
18. Highlight the key steps to be performed in solving a computer forensics case.
19. Explain what is required in setting up a computer forensics laboratory. What tools are required on hardware and software side?
20. What steps do the network hackers execute, as explained in this chapter?
21. What is a "social networking" site? What are the security threats that can emanate from social networking sites?

22. What are "rootkits." Why are they dangerous? How do rootkits help cyberattackers?

23. What are the major international regulations (the "Big 4 Laws") as mentioned in this chapter that impact forensics?

24. Explain the "complexity" and "quantity" problems faced in digital forensics investigation.

25. Explain the "data privacy" challenge in cyberforensics. Support your point with suitable illustrative examples.

26. Do you think that using "counter-forensics privacy tools" is a good idea? Why? Explain with examples.

27. Provide an overview of how "data mining" techniques can be applied in cyberforensics.

28. Highlight some of the key differences between an "audit" and a "cyberforensics investigation."

29. What, do you think, has led to antiforensics behaviors and tools? Elaborate your answer with suitable examples. Explain how the criminals exploit the situations.

References

[1] Following are the links for COFEE:

- http://en.wikipedia.org/wiki/Computer_Online_Forensic_Evidence_Extractor (6 November 2009).
- http://www.groundreport.com/Media_and_Tech/Microsoft-Makes-the-COFEE/2860183 (6 November 2009).
- http://www.postchronicle.com/cgi-bin/artman/exec/view.cgi?archive=68&num=144908 (6 November 2009).
- http://www.wired.com/threatlevel/2008/04/microsoft-gives (6 November 2009).
- http://www.technovelgy.com/ct/Science-Fiction-News.asp?NewsNum=1616 (6 November 2009).
- http://www.ghacks.net/2008/04/29/computer-online-forensic-evidence-extractor/ (6 November 2009).

[2] Following are some useful links on *Access Data's FTK tool kit*:

- Access Data's Home Page: http://www.accessdata.com/ (21 December 2009).
- Access Data's products for various types of forensics investigations: http://www.accessdata.com/Products.html (21 December 2009).
- The forensics features of access data's toolkit: http://www.accessdata.com/forensictoolkit.html (21 December 2009).
- This is about FTK 2.0.2 to 2.1 Upgrade Instructions: http://ftk21.accessdata.com/Upgrade_from_2-02_to_2-1.pdf (21 December 2009).

[3] Useful links on *Guidance Software's EnCase* can be visited at:

- http://www.digitalintelligence.com/software/guidancesoftware/encase/ (21 December 2009). Following link is about news accolade on this tool:
- http://investors.guidancesoftware.com/releasedetail.cfm?ReleaseID=416497 (22 December 2009).
  The SC Maganize has announced this tool to be one of the best; a report on that can be seen at:
- http://www.scmagazineus.com/guidance-software-encase-forensic-v-6/review/159/ (18 December 2009).
  To learn about Guidance Software's EnCase Portable, visit a video demo clip at:
- http://vimeo.com/5702414 (10 December 2009).
  To know more about Guidance Software's EnCase® Portable having won Cygnus Law Enforcement Group Award in Forensics Category, visit:
- http://finance.yahoo.com/news/Guidance-Softwares-EnCase-bw-4161315896.html?x=0 (25 December 2009).

[4] Following are some useful links on "*File Carving*" that explains what is "file carving"

- http://www.fim.uni-linz.ac.at/Lva/IT_Recht_Computerforensik/File_carving.pdf (16 December 2009).
  A good presentation on the *Advances and Challenges in File Carving* can be read at: