chapter 18

Forensic Science and the Internet

By Andrew W. Donofrio

KEY TERMS

bookmark

broadband

browser

cookies

domain

download

e-mail

firewall

hacking

hypertext

Internet cache

Internet history

Internet protocol

Internet service provider (ISP)

mailing list

modem

newsgroups

router

search engine

uniform resource locator (URL)

VoIP (voiceover Internet protocol)

Wi-Fi

Learning Objectives

After studying this chapter you should be able to:

- Understand how the Internet is structured
- Know how to search for information on the Internet
- Describe informational retrieval sources, such as mailing lists and newsgroups, available through the Internet
- Learn how to retrieve information about forensic science on the Internet
- Relate various areas found on the computer where a user's Internet activities can be investigated
- Describe how e-mails, chat, and instant messages on the Internet can be traced and recovered
- List and describe three locations where investigators may pinpoint the origin of a hacker

Scott Peterson: A Case of Circumstantial Evidence

Scott Peterson was charged with the murder of his pregnant wife, Laci, and her unborn son, Conner. On the surface, this young couple lived a happy and content lifestyle in Modesto, California. The 30-year-old Peterson had married his college sweetheart, Laci, a 27-year-old substitute teacher. She was about one month away from delivering her first child. Scott Peterson had told investigators that he had last seen his wife on December 24, 2002, at 9:30 a.m. when he left home for a fishing trip off San Francisco Bay. The decomposed remains of Laci washed ashore in April 2003, not far from where Scott Peterson said he had gone fishing on the day she vanished. Peterson said she was dressed in a white top and black pants when he last saw her, but Laci's body was found with khaki pants. Her sister recalled that Laci was wearing khaki pants the night before her disappearance.

Peterson claimed that he had gone fishing for sturgeon or striped bass, but the police investigation revealed that he failed to bring the appropriate fishing rod and lines to catch such fish. Further revelations surfaced when it became known that Scott was having an affair with another woman. A search of Scott's warehouse led to the recovery of a black hair on a pair of pliers resting in Scott's boat. A mitochondrial DNA profile of the hair was consistent with Laci's DNA. Scott Peterson was charged with murder and convicted and currently awaits his fate on death row.

Visit WebExtra 18.1 to view the evidence that prosecutors presented to convict Scott Peterson.

Today, one cannot read a newspaper or turn on the television without seeing some reference to the Internet. The Internet, often referred to as the "information superhighway," is a medium for people to communicate with others and to access millions of pieces of information on computers located anywhere on the globe. No subject or profession remains untouched by the Internet, including forensic science. Every week many new pages of information are added to the Internet on the subject of forensic science, providing instant access to updated forensic science news and information. The Internet brings together forensic scientists from all parts of the world, linking them into one common electronic community.

The Internet was developed in 1969 by the U.S. Department of Defense with the purpose of providing a connection between computers in different locations. The project, called ARPANET, originated from a group of scientists and engineers funded by the Pentagon's Advanced Research Projects Agency (ARPA). Their idea was based on the premise that the network would still operate even if part of the connection failed. The first successful link was established between computers housed at UCLA and Stanford Research Institute. Shortly thereafter, USC–Santa Barbara and University of Utah computers were added to the system. In 1972, more than twenty sites were connected on the system when the first electronic mail (e-mail) message was sent. In the 1980s, this network of interconnected computers grew with the establishment of the National Science Foundation Network (NSFNet), which encompassed five supercomputing centers across the United States. At about the same time, regional networks were formed around the United States for the purpose of accessing NSFNet. By 1989, ARPANET had closed down and NSFNet, along with its regional networks, began to mushroom into a worldwide network known as the Internet.[1]

WHAT IS THE INTERNET?

The Internet can be defined as a "network of networks." A single network consists of two or more computers that are connected in some fashion to share information; the Internet connects thousands of these networks so information can be exchanged worldwide. Connections are sometimes made through a **modem**, a device that allows computers to exchange and transmit information through telephone lines. A modem passes digital information through a series of steps to convert it to analog signals that can be passed over a telephone line. The process is reversed when the modem converts analog signals coming in from the phone line. Modems transfer information at a rate of bits per second (bps). Obviously, a modem with high-speed capabilities will ensure a faster connection on the Internet. Currently, a modem that transmits at 56,000 bits (or 56 kilobits) per second is recommended for convenience and reasonable connection speed. This speed is roughly equivalent to transmitting 1,000 to 2,000 words per second. The trend, however, is to offer Internet users even higher-speed **broadband** connections to websites. Digital subscriber line (DSL) service is available from phone companies in many regions. DSL carries digital information on your regular telephone line without disturbing voice traffic. These lines can carry up to 1.1 megabits per second. Alternatively, one may opt to transmit over a TV cable line. Cable modems offer speeds comparable to DSL. Once your computer is hooked into a DSL or cable line you now have an additional option to link other computers in your home or office, through either network wire (typically Ethernet) or high-frequency radio waves via a wireless or **Wi-Fi** connection. A device called a **router** serves as a sort of splitter, designed to link computers and manage traffic between them. The router, whether wired or wireless, allows computers to share a connection to the Internet. The advantage of Wi-Fi technology is that it avoids messy wires. Once you have positioned a router in your home or office, another option awaits you—voice over Internet protocol (**VoIP**). The IP (Internet Protocol) portion of VoIP is the bloodline of the Internet—but more on that later.

A broadband Internet connection can send and receive the human voice in a manner indistinguishable from a traditional telephone line. If you're in the range of a router, your Wi-Fi phone (cost about $150) can operate like a traditional cell phone. Unlimited-calling plans are commercially available for $20–25 per month.

It is quite astonishing to think that there is no overriding network controlling the Internet. Rather, various larger, higher-level networks are connected through *network access points.* Many large **Internet service providers (ISPs)** (Verizon, AOL, Yahoo) connect to each other through these network access points. The ISP's customers can then connect to the network by connecting to the bank of modems or the cable/DSL connected routers present at the Internet service provider location and thus be connected to all the other networks. Because this places many individual computers on the network, an address system is needed so that all the data traveling on the network can get to its intended location.

On the Internet, the address is known as an **Internet protocol** (IP) address. This is derived from the protocol suite (transmission-control protocol/Internet protocol—TCP/IP) that defines how traffic is to be presented and transmitted over the Internet. TCP/IP is nothing more than a set of rules on how manufacturers and developers of both hardware and software must configure their products if they want to send traffic over the Internet. With all of the different computer manufacturers and software developers, some rules are necessary if computers are to successfully communicate on a global network. Just as any human language needs rules for people to communicate successfully, so does the language of computers. Computers that participate on the Internet, therefore, must be provided with an IP address from the Internet service provider to which they connect. IP addresses take the form ###.###.###.###, where, generally speaking, the ### can be any number from 0 to 255. A typical IP address might look like this: 66.94.234.13.

Not only do these IP addresses provide the means by which data can be routed to the appropriate location, but they also provide the means by which most Internet investigations are conducted (see Figure 18–1).

Once a computer is connected to the Internet, it becomes a node on this network of networks. **Domains** are human-readable names, such as www.nytimes.com, assigned to an IP address. Thus, www.nytimes.com is the registered name for the *New York Times*. A domain name usually consists of two or more labels separated by dots. The rightmost label is the *top-level domain.* Following are the most common abbreviations by which a top-level domain name is identified on the Internet:

.gov—government

.mil—military

.edu—educational institution

.com—commercial providers

.org—nonprofit organizations

To the left of the top-level domain is the subdomain; thus, nytimes is a subdomain of the .com domain. For the purpose of e-mail, the name of an individual at the *New York Times* may be added before the subdomain and the @ sign is used to separate them. An e-mail address may read as: Johndoe@nytimes.com.

At this point you may be wondering: If everything on the Internet uses an IP address to route data to the correct location, how can we use web addresses and e-mail addresses to access web-sites and send e-mail? The answer, although technically complex, is quite simple. Understanding the apparent limitations of the human mind to remember numbers, developers created the concept of the domain name system (DNS). *Domain name systems* are essentially large databases distributed over the Internet that relate domain names to their actual IP address. For instance, a person who wants to read the *New York Times* online only needs to know the web address. Even if the user is unsure of the actual address, the most logical place to start would obviously be www.newyorktimes.com or www.nytimes.com (both of which will work, by the way). In actuality, however, the address is 199.239.137.245. This can be verified by typing that IP address directly into your web browser where you would normally type the web address. Domain name systems makes it much easier for us to navigate the Web, but for investigative purposes it is important to realize that no names exist on the Internet; rather it's all about the IP address (see Figure 18–2).

WHERE TO GO ON THE INTERNET

The World Wide Web

The most popular area of the Internet is the World Wide Web. Also known as WWW, W3, or the web, it is a collection of documents, called *webpages,* that are stored in the computers connected to the Internet throughout the world. Web **browsers**, such as Netscape Navigator and Microsoft Internet Explorer, are programs that allow the user to explore information stored on the Web and to retrieve webpages the viewer wishes to read. Most browsers, such as the popular Netscape Navigator, perform within a toolbar interface. Various functions such as reload, back, forward, stop, open, and print appear on the toolbar so that with one click on an icon, the user can easily navigate the Internet. Web browsers permit the downloading and capture of documents, as well as printing of selected portions of websites. A browser also allows the user to explore the World

Wide Web and newsgroups.

Each webpage is stored in a specific website that has a unique web address that indicates where the document is actually located. The web address is called the **uniform resource locator (URL)**. The URL designates the site at which information is stored on the Internet. You can access a page by directly entering the URL into your browser. For example, the FBI has a website that can be accessed by typing in its URL: *www.fbi.gov*.

The URL for the FBI consists of the following components:

> *http://*—Hypertext Transfer Protocol is the programming language the browser uses to locate and read webpages.
>
> *www.*—Denotes the World Wide Web, the place on the Internet where the information is located.
>
> *fbi.*—Designates the subdomain or server; in this case, for the Federal Bureau of Investigation.
>
> *gov*—Designates the domain name.

Upon entering the FBI website, you are confronted with a multitude of services and information provided by the FBI, such as an overview of the bureau's operation, information regarding ongoing investigations, reports concerning crime statistics, and even a list of the ten most-wanted fugitives. The Internet has made browsing or exploring the Web easy through the existence of **hypertext**. Hypertext is not hard to find because it is highlighted with a different color within the webpage. When selected, hypertext enables the user to jump to another webpage related to the subject at hand. For example, if a user is interested in examining the FBI's Law Enforcement Bulletin webpage, one would merely look on the FBI's home page and search for the term *Law Enforcement Bulletin.* The user is immediately transferred to a hypertext page where, with a click of the mouse, one is taken to the website. This website has the URL *www.fbi.gov/publications/leb/leb.htm,* where *leb.htm* designates a document on the FBI website. Another FBI publication available online is *Forensic Science Communications.* This quarterly online forensic science journal presents technical articles, technical notes, case reports, and review articles in a paperless format.

The advantage of using hypertext is that the user can quickly switch back and forth between related webpages without having to retype the URL or start over at the beginning of the search. The existence of hypertext makes the Internet user-friendly and has given rise to the expressions *browsing* and *surfing the net.* Users can navigate from one website of the Internet to another, browsing at leisure through a succession of documents. Another quick way to reach a site is to designate it as a bookmark or favorite place. Most browsers allow the user to customize a list of favorite websites for easy access with one click (see Figure 18–3).

Hundreds of new sites are added every day, providing Internet users with a staggering amount of information. You may wish to explore popular website locations by using the list of the Top 100 Classic Websites compiled by *PC Magazine* (*http://www.pcmag.com*) (see Table 18–1). The *PC Magazine* Top 100 Classic Websites list includes sites related to business and finance, computing, news, entertainment, online shopping, and reference sources. This list opens a gateway to exploring the diversity of the World Wide Web. For example, you can easily visit Expedia.com, a popular site for those interested in finding information on traveling. Through this site you can book flights, hotel rooms, and cruises and plan most any type of vacation. The Amazon.com website provides one of the largest bookstores on the Internet. Here the user can search by keyword, author, subject, or title to locate or purchase books on any topic imaginable.

You can even ask Amazon.com to e-mail you when books related to your personal interests arrive in stock.

A favorite site of the author is the Switchboard website www.switchboard.com. This site is devoted to helping users locate long-lost friends or contact relatives who are spread across the country. A database of names, phone numbers, addresses, and e-mail addresses can easily be searched with a variety of options. Another interesting website that can be helpful to the user is the MapQuest website (*http://www.mapquest.com*). Here the user designates a location and the site generates a map, accompanied by directions explaining which roads to follow to best travel to that location. A fun site is BlueMountain.com (*www.bluemountain.com*), where the user can compose, personalize, and send greeting cards for all sorts of occasions.

Table 18–1 PC Magazine's Top 100 Classic Websites

AAPS: PC and Mobile

AvantGo

www.avantgo.com

Homestead

www.homestead.com

Mailblocks

www.mailblocks.com

MSN Hotmail

www.hotmail.com

Shutterfly

www.shutterfly.com

Vindigo

www.vindigo.com

WebEx

www.webex.com

Yahoo! Groups

groups.yahoo.com

Business and Finance

Bloomberg.com

www.bloomberg.com

Internal Revenue Service

www.irs.gov

MSN Money

www.moneycentral.msn.com

SmartMoney.com

www.smartmoney.com

The Motley Fool

www.fool.com

TheStreet.com

www.thestreet.com

U.S. Securities and Exchange Commission

www.sec.gov

Careers

Dice.com

www.dice.com

Monster.com

www.monster.com

Yahoo! HotJobs

hotjobs.yahoo.com

Computing

Annoyances.org

www.annoyances.org

Answers that Work

www.answersthatwork.com

Digital Photography Review

www.dpreview.com

EarthWeb

www.earthweb.com

eWeek

www.eweek.com

ExtremeTech

www.extremetech.com

Java Technology

www.java.sun.com

PalmGear.com

www.palmgear.com

PC Magazine

www.pcmag.com

Slashdot

slashdot.org

Technology Review

www.technologyreview.com

W3Schools

www.w3schools.com

Webopedia

www.webopedia.com

Current Events and News You CanUse

Electronic Privacy Information Center

www.epic.org

IEEE Virtual Museum

www.ieee-virtual-museum.org

NASA

www.nasa.gov

Project Vote Smart

www.vote-smart.org

World Health Organization

www.who.int

Lifestyle and Fun

Citysearch

www.citysearch.com

Discovery Kids

kids.discovery.com

Epicurious

www.epicurious.com

HowStuffWorks

www.howstuffworks.com

ifilm

www.ifilm.com

National Geographic Society

www.nationalgeographic.com

Nickelodeon Online

www.nick.com

Pogo

www.pogo.com

Smithsonian Institution

www.si.edu

Yahoo! Games

games.yahoo.com

News and Entertainment

AMG All Music Guide

www.allmusic.com

BBC News

www.bbc.co.uk

CNN

www.cnn.com

E! Online

www.eonline.com

ESPN.com

www.espn.com

Internet Archive

www.archive.org

Internet Movie Database (IMDb)

www.imdb.com

Slate

www.slate.com

The New York Times on the Web

www.nytimes.com

NPR

www.npr.org

The Onion

www.theonion.com

RollingStone.com

www.rollingstone.com

Salon.com

www.salon.com

ScienceDaily

www.sciencedaily.com

Television Without Pity

www.televisionwithoutpity.com

Wired News

www.wired.com

Security and the Net

Broadbandreports.com

dslreports.com

CERT Coordination Center

www.cert.org

GetNetWise

www.getnetwise.org

Gibson Research Corp.

www.grc.com

Internet Traffic Report

www.internetrafficreport.com

Netcraft

news.netcraft.com

SecurityFocus

www.securityfocus.com

TrendMicro

www.trendmicro.com

Search, Reference, and Portals

About.com

www.about.com

Centers for Disease Control and Prevention

www.cdc.gov

Dictionary.com

dictionary.reference.com

Encyclopaedia Britannica

www.britannica.com

FedStats

www.fedstats.gov

FirstGov.gov

www.firstgov.gov

Google

www.google.com

iVillage.com

www.ivillage.com

Librarians’ Internet Index

www.lii.org

The Library of Congress

www.loc.gov

MSN Encarta

encarta.msn.com

Nolo.com

www.nolo.com

WebMD

www.webmd.com

Yahoo!

www.yahoo.com

Shopping

Amazon.com

www.amazon.com

CarsDirect.com

www.carsdirect.com

ConsumerReview.com

www.consumerreview.com

eBay

www.ebay.com

Netflix

www.netflix.com

Overstock.com

www.overstock.com

PriceGrabber.com

www.pricegrabber.com

Shopping.com

www.shopping.com

Surprise.com

www.surprise.com

Techbargains.com

www.techbargains.com

Travel

Expedia

www.expedia.com

Fodors

www.fodors.com

Frommers.com

www.frommers.com

Lonely Planet

www.lonelyplanet.com

Orbitz

www.orbitz.com

Travelocity

www.travelocity.com

Search Engines

Sifting through the enormous amount of information on the World Wide Web can often resemble looking for a needle in a haystack. As the Internet grows, so does the need for automated search tools. Several directories and indexes known as **search engines** help users search the Internet for a particular topic. Typically, a user enters a keyword or phrase into a search engine to locate sites on the Internet that are relevant to a particular subject. The number of search engines continues to change with new technology, as faster, newer tools are adopted and slower, older ones are phased out. Interestingly, search engines have taken on a new look, becoming portals offering a wide variety of Internet services in addition to their traditional search functions. Some of the more popular search engines, along with their URLs, are listed here:

Search Engine	**URL**
Yahoo!	*www.yahoo.com*
Google	*www.google.com*
MSN	*www.msn.com*
Dogpile	*www.dogpile.com*
Lycos	*www.lycos.com*

The reader can also find a multitude of search engines at www. easysearcher.com/index.html. Most search engines contain tools called *spiders* or *crawlers* that search the Web seeking titles, subjects, and keywords to index the contents of individual webpages. Through search engines, the user can locate relevant webpages containing a particular piece of information on the Internet. When the user types in keywords and phrases related to the needed information, the search engines search their databases and list all the pages involving the keyword selection. Because each search engine has different capabilities, it is recommended that multiple search engines be used when researching a subject. For example, in compiling this edition, the author had the occasion to search on the term “Roger Severs” and found a variety of search engine matches:

Search Engine	**No. of Matches**
Google	43
Yahoo	31
Lycos	30
Dogpile	19

MSN	16

Automated search tools called *meta-engines* load a query into several of the Internet's leading search engines to compile a single list of results. One example of a meta-engine is MetaCrawler (*www.metacrawler.com*).

Electronic Mail and Mailing Lists

The service that is most commonly used in conjunction with the Internet is electronic mail. Also called **e-mail**, this communication system can transport messages across the world in a matter of seconds. In order to use e-mail, users must acquire an e-mail address, usually through an Internet service provider or a free e-mail server. Like regular postal mail, you will need an address to receive mail, and you will need to know a recipient's address to send messages. These messages are stored in an individualized mailbox that can be opened electronically at your convenience. Another interesting feature is that one can attach a file to an e-mail so that the recipient can **download** the attached file. The file can then be saved and stored on the recipient's computer. A file may consist of text, pictures, music, or video. For example, a text file can be viewed and modified by means of a word-processing program.

Also, having an e-mail account provides the opportunity to receive information through **mailing lists.** A mailing list is a discussion group for a selected topic in which related messages are sent directly to your mailbox through e-mail. For example, Forens-L is a mailing list dedicated to the discussion of forensic medicine and forensic science that provides a quick, useful way to exchange ideas or share information about forensics with people of similar interests around the world. To subscribe to the Forens-L mailing list, go to forensic.to/mailman/listinfo/forens-l_forensic.to for instructions.

Newsgroups

Another service much like mailing lists involves **newsgroups.** Like a mailing list, a newsgroup is devoted to a particular topic. Whereas a mailing list, however, is usually managed by a single site, a newsgroup networks many sites that are set up by local Internet service providers. The result is that a newsgroup joins together a significantly larger audience compared to a mailing list. A newsgroup is analogous to a bulletin board where articles (messages) are posted by subscribers. When you connect to a newsgroup, you have the ability to quickly scan through a list of article titles, selecting only those that interest you. To find a newsgroup of interest, you can explore Usenet, an index of the available newsgroups. The index can be located through search engines such as Google. Usenet can be searched through keywords in the same manner in which the World Wide Web is explored. For example, entering the keyword *law enforcement* into Usenet produces a list of articles from any of the newsgroups containing that keyword. The articles are hypertext so that with a click of the mouse, you can read the article instantaneously. A useful website (www.google.com/grphp?hl=en) lists newsgroups in hypertext so that the newsgroup can be accessed directly from the World Wide Web. Commercial services such as AOL allow the user to subscribe to a newsgroup and provide the program that will keep track of how many articles are available, which ones you have read, and which ones you have not read.

EXPLORING FORENSIC SCIENCE ON THE WORLD WIDE WEB

There are no limits to the amount or type of information that can be found on the Internet. The fields of law enforcement and forensic science have not been left behind by advancing computer technology. Extensive information about forensic science is available on the Internet. The types of webpages range from simple explanations of the different fields of forensics to intricate details of crime-scene reconstruction. You can also find information on which colleges offer programs for degrees in forensics or pages posted by law enforcement agencies that detail their activities, as well as possible employment opportunities. Table 18–2 lists a number of websites available in the forensic science field.

Reddy's Forensic Home Page (*www.forensicpage.com*) is a valuable starting point and a must for those with an interest in forensic science on the Internet. This site is a collection of forensic webpages listed under categories such as new links in forensics; general forensic information sources, associations, colleges, and societies; literature and journals; forensic laboratories; general webpages; forensic-related mailing lists and newsgroups; universities; conferences; and various forensic fields of expertise. Another website offering a multitude of information related to forensic science is Zeno's Forensic Webpage (*forensic.to/forensic.html*). Here you can find links to forensic education and expert consultation, as well as a wealth of information concerning specific fields of forensic science. A comprehensive and useful website for those interested in law enforcement is the Police Officer's Internet Directory (*www.officer.com*). This comprehensive collection of criminal justice resources is organized into easy-to-read subdirectories that relate to topics such as law enforcement agencies, police association and organization sites, criminal justice organizations, law research pages, and police mailing-list directories.

Table 18–2 Forensic Science Web Sites*

- *American Academy of Forensic Sciences*—professional society dedicated to the application of science to the law and committed to the promotion of education in the forensic sciences.
- *American Board of Forensic Entomology*—provides information about the field's history,

case studies, and professional status.

- *American Board of Forensic Toxicology Inc.*—working to establish and revise as necessary standards of qualification for those who practice forensic toxicology.
- *American College of Forensic Examiners (ACFE)*—dedicated to members of the forensic community. Also provides links to help people find expert witnesses in the forensic field.
- *American Society of Crime Laboratory Directors (ASCLD)*—nonprofit professional society devoted to the improvement of crime laboratory operations through sound management practices.
- *American Society of Questioned Document Examiners*—online references, technical notes on handwriting identification, and other areas of professional interest to the forensic document analyst.
- *ASCLD: Laboratory Accreditation Board*—offers the general public and users of laboratory services a means of identifying those laboratories which have demonstrated that they meet established standards.
- *Association for Crime Scene Reconstruction (ACSR)*—members are law enforcement investigators, forensic experts, and educators.
- *Association of Firearm and Tool Mark Examiners*—defines standards and ethics for individual workers in the field of firearms and toolmark examination.
- *California Association of Criminalists*—membership is offered to those who are presently employed as laboratory scientists and are professionally engaged in one or more fields directly related to the forensic sciences.

- *Canadian Society of Forensic Science*—includes journal abstracts, history, and meeting information.
- *European Association for Forensic Entomology*—created in May 2002 in Rosny sous Bois, France, to promote forensic entomology in Europe.
- *Forensic Science Service*
- *Forensic Science Society*—provides information on careers, conferences, and publications.
- *International Association for Identification*—professional association for those engaged in forensic identification, investigation, and scientific examination.
- *International Association of Bloodstain Pattern Analysis*—organization of forensic experts promoting education, establishing training standards, and encouraging research in the field of bloodstain pattern analysis.
- *International Association of Forensic Toxicologists*
- *International Organisation on Electronic Evidence (IOCE)*—formed to develop international principles for the procedures relating to digital evidence.
- *National Forensic Science Technology Center*—dedicated to supporting forensic science laboratories to achieve the highest possible quality of operations.
- *Natural Resources DNA Profiling and Forensic Centre*—undertakes research on natural populations of animals and plants for the purpose of providing information to managers charged with conserving biodiversity and ensuring the sustainable use of Canada’s biological resources.
- *Society of Forensic Toxicologists*

- *Southern Association of Forensic Scientists*—organization of professional forensic scientists.
- *Southern California Association of Fingerprint Officers*—an association for scientific investigation and identification.
- *U.S. Department of Justice: National Commission on the Future of DNA Evidence*—working to maximize the value of forensic DNA evidence in the criminal justice system.
- *Vidocq Society*—forensic experts who meet in the shadow of Independence Hall to investigate and solve unsolved homicides.

*Web sites are hypertexted at *http://dir.yahoo.com/Science/Forensics/Organizations/?o=a.*

WEBSITES YOU MAY WISH TO EXPLORE

The Internet contains hundreds of webpages for the reader who is interested in introductory information on forensic science and criminal investigation. The list on the pages that follow contains websites that serve such a purpose.

An Introduction to Forensic Firearm Identification. This website contains an extensive collection of information relating to the identification of firearms. An individual can explore in detail how to examine bullets, cartridge cases, and clothing for gunshot residues and suspect shooters' hands for primer residues. Information on the latest technology involving the automated firearms search system IBIS can also be found on this site.

Carpenter's Forensic Science Resources. This site provides a bibliography with hypertext references pertaining to different aspects of criminal investigations involving forensic evidence. For example, the user can find references about DNA, fingerprints, hairs, fibers, and questioned documents as they relate to crime scenes and assist investigations. This website is an excellent place to start a research project in forensic science.

Crime Scene Investigation. For those who are interested in learning the process of crime-scene investigation, this site provides detailed guidelines and information regarding crime-scene response and collection and preservation of evidence. For example, information concerning packaging and analysis of bloodstains, seminal fluids, hairs, fibers, paint, glass, firearms, documents, and fingerprints can be found through this website. This website explains the importance of inspecting the crime scene and the impact forensic evidence has on the investigation.

Crimes and Clues. Users interested in learning about the forensic aspects of fingerprinting will find this a useful and informative website. The site covers the history of fingerprints, as well as subjects pertaining to the development of latent fingerprints. The user will also find links to other websites covering a variety of subjects pertaining to crime-scene investigation, documentation of the crime scene, and expert testimony.

Interactive Investigator–Détective Interactif. This is an outstanding site. Visitors can obtain general information and an introduction to the main aspects of forensic science from a database on the subject. They can also explore actual evidence gathered from notorious crime scenes. Users will be able to employ deductive skills and forensic knowledge while playing an interactive game in which they must help Detective Wilson and Detective Marlow solve a gruesome murder.

The Chemical Detective. This site offers descriptions of relevant forensic science disciplines. Topics such as fingerprint, fire and arson, and DNA analysis are described in informative layperson's terms. Case histories describe the application of forensic evidence to criminal investigations. Emphasis is placed on securing and documenting the crime scene. The site directs the reader to other important forensic links.

Questioned Document Examination. This basic, informative webpage answers frequently asked questions concerning document examination, explains the application of typical document examinations, and details the basic facts and theory of handwriting and signatures. There are also links to noted document examination cases for the user to read and recognize the real-life application of forensic document examination.

FORENSIC ANALYSIS OF INTERNET DATA

It's important from the investigative standpoint to be familiar with the evidence left behind from a user's Internet activity. A forensic examination of a computer system reveals quite a bit of data about a user's Internet activity. The data described next would be accessed and examined using the forensic techniques outlined in Chapter 17.

Internet Cache

Evidence of web browsing typically exists in abundance on the user's computer. Most web browsers (Internet Explorer, Netscape, and Firefox) use a caching system to expedite web browsing and make it more efficient. This was particularly true in the days of dial-up Internet access. When a user accesses a website, such as the *New York Times* home page, the data is fed from that server (in this example the *New York Times*), via the Internet service provider, over whatever type of connection the user has, to his or her computer. If that computer is accessing the Internet via a dial-up connection, the transfer of the *New York Times* home page may take a while, because the data transfer rate and capabilities (bandwidth) of the telephone system is limited. Even with the high-speed access of a DSL line or cable connection, conservation of bandwidth is always a consideration. Taking that into account, web browsers store (cache) portions of the pages visited onto the local hard disk drive. This way, if the page is revisited, portions of it can be reconstructed more quickly from this saved data, rather than having to pull it yet again from the Internet and use precious bandwidth.

This **Internet cache** is a potential source of evidence for the computer investigator. Portions of, and in some cases, entire visited webpages can be reconstructed. Even if deleted, these cached files can often be recovered (see the section on deleted data in Chapter 17). Investigators must know how to search for this data within the particular web browser used by a suspect.

Internet Cookies

Cookies provide another area where potential evidence can be found. To appreciate the value of cookies you must first understand how they get onto the computer and their intended purpose. **Cookies** are placed on the local hard disk drive by websites the user has visited, if the user's web browser (such as Netscape Navigator or Internet Explorer) is set to allow this to happen. Netscape Navigator stores cookies in a *cookies.txt* file and Microsoft Internet Explorer places cookies in a dedicated directory. The website uses cookies to track certain information about its visitors. This information can be anything from history of visits and purchasing habits to passwords and personal information used to recognize the user for later visits. Consider a user who registers for an account at the Barnes and Noble bookstore website, and then returns to the same site from the same computer a few days later. The site will then display "Welcome, *Your User Name.*" This data is retrieved from the cookie file placed on the user's hard disk drive by the website during the initial visit and registration with the site.

It is helpful to think of cookies almost like a "Caller ID" for websites. The site recognizes and retrieves information about the visitor, as when a salesman recognizes the caller from a

Caller ID display and quickly pulls the client's file. Cookie files can be a valuable source of evidence. In Internet Explorer, they take the form of plain text files, which can typically be opened with a standard text viewer or word-processing program, revealing part of the data. The existence of the files themselves, regardless of the information contained within, can be of evidentiary value to show a history of web visits. A typical cookie may resemble the following: rsaferstein@forensicscience.txt. From this we can surmise that someone using the local computer login *rsaferstein* accessed the forensic science website. It is possible that the cookie was placed there by an annoying pop-up ad, but considered against other evidence in the computer data, the presence of this cookie may be of corroborative value.

Internet History

Most web browsers track the history of webpage visits for the computer user. This is probably done merely for a matter of convenience. Like the "recent calls" list on a cell phone, the **Internet history** provides an accounting of sites most recently visited, with some storing weeks' worth of visits. Users can go back and access sites they recently visited just by going through the browser's history. Most web browsers store this information in one particular file; Internet Explorer uses the *index.dat* file. On a Windows system, an *index.dat* file is created for each login user name on the computer. The history file can be located and read with most popular computer forensic software packages. It displays the uniform resource locator (URL) of each website, along with the date and time the site was accessed. An investigation involving Internet use almost always includes an examination of Internet history data.

In some respects, the term "*Internet* history" is wrong because it doesn't encompass all of its functions. Several browsers—Internet Explorer, for one—store other valuable evidence independent of Internet access. It is not uncommon to see files accessed over a network listed in the history. Similarly, files accessed on external media, such as floppy disks, CDs, or thumb drives, may also appear in the history. Regardless, the Internet history data is a valuable source of evidence worthy of examination (see Figure 18–4).

Bookmarks and Favorite Places

Another way users can access websites quickly is to store them in their **bookmarks** or “Favorite Places.” Like a preset radio station, web browsers allow users to bookmark websites for future visits. A lot can be learned from a user’s bookmarked sites. You might learn what online news a person is interested in or what type of hobbies he or she has. You may also see that person’s favorite child pornography or computer hacking sites bookmarked.

In Internet Explorer the favorite places are kept in a folder with link (shortcut) files to a particular URL. They can be organized in subfolders or grouped by type. The same is true for the Firefox web browser, except that Firefox bookmarks are stored in a document done in HyperText Markup Language (HTML), the same language interpreted by web browsers themselves.

FORENSIC INVESTIGATION OF INTERNET COMMUNICATIONS

Computer investigations often begin with or are centered on Internet communication. Whether it is a chat conversation among many people, an instant message conversation between two individuals, or the back-and-forth of an e-mail exchange, human communication has long been a source of evidentiary material. *Regardless of the type, investigators are typically interested in communication.*

Recall that in order to communicate on the Internet, a device needs to be assigned an Internet protocol (IP) address. The IP address is provided by the Internet service provider from which the device accesses the Internet. Thus the IP address may lead to the identity of a real person. If an IP address is the link to the identity of a real person, then it is quite obviously valuable for identifying someone on the Internet. To illustrate, let's assume that a user of the Internet, fictitiously named John Smith, connects to the Internet from his home by way of a Verizon DSL connection. Verizon in this case would be responsible for providing Smith with his IP address. Verizon was issued a bank of IP addresses to service its customers from a regulatory body designed to track the usage of IP addresses (obviously so no two were used at the same time). Smith, while connected to the Internet, decides to threaten an ex-girlfriend by sending her an e-mail telling her he is going to kill her. That e-mail must first pass through Smith's Internet service provider's computers—in this case Verizon—on its way to its destination—Smith's girlfriend. The e-mail would be stamped by the servers that it passes through, and this stamp would include the IP address given to Smith by Verizon for his session on the Internet. An investigator responsible for tracking that e-mail would locate the originating IP address stamped in the e-mail header. That IP address could be researched using one of many Internet sites (*www.samspade.org, www.arin.net*) to determine which Internet service provider was given this IP as part of a block to service its customers (see Figure 18–5). The investigator then files a subpoena with the Internet service provider (Verizon) asking which of its customers was using that IP address on that date and time.

IP addresses are located in different places for different methods of Internet communications. E-mail has the IP address in the header portion of the mail. This may not be readily apparent and may require a bit of configuration to reveal. Each e-mail client is different and needs to be evaluated on a case-by-case basis. For an instant message or chat session, the particular provider (the one providing the chat mechanism—AOL, Yahoo, and so on) would be contacted to provide the user’s IP address.

E-mail can be read by a number of *clients* or software programs. Two of the most popular ways to access, read, and store e-mail in today’s Internet environment, however, are Microsoft Outlook and through an Internet browser. Some people even use a combination of the two. If an e-mail account is linked through Microsoft Outlook, then the e-mail is stored in a compound file (a file with several layers). Typically, a compound file exists for received (inbox), sent, and deleted e-mail. Users can also create new categories (shown as folders in Outlook) and categorize saved e-mail there. Most computer forensic software applications can view (mount) these compound files so that the e-mail can be seen, including any file attachments. These files can also be imported into a clean copy (one not attached to an account) of Microsoft Outlook and the e-mail viewed there. Investigators must also be aware that in a computer network environment, the user’s outlook files may not reside on their workstation computer, but rather on a central mail or file server.

Most accounts offer the ability to access e-mail through a web-based interface as well. This way, users can access their e-mail remotely from other computers. For e-mail accessed through a web browser, the information presented earlier on Internet-based evidence applies. The web interface converts the e-mail into a document suitable for reading in a web browser. Consequently, web-based e-mail is often found in the Internet cache. This is particularly true of free Internet e-mail accounts such as Hotmail and Yahoo.

Much of the evidence from Internet communication is also derived from chat and instant message technology. This is particularly true in the world of child sexual exploitation over the

Internet. Various technologies provide chat and instant message services. Most chat and instant message conversations are not saved by the parties involved. Although most of the software does allow for conversation archiving, it is typically turned off by default. Therefore, conversations of this nature typically exist in the volatile memory space of random-access memory (RAM). Recall from Chapter 17 that RAM is termed volatile because it holds data only if it has power. Unplugging the computer will cause the data located in RAM to be lost. If, however, chat or instant message conversations are relevant as evidence and the computer was turned off, thus erasing the data in RAM, all might not be lost. Remember that there is an interaction between the computer system's RAM and the hard disk drive. RAM is a commodity and as such the computer's operating system makes an effort to conserve it as best as possible. This is done by swapping/paging that information back and forth into the swap space/paging file. Therefore remnants of chat conversations are often found in the swap space/paging file during a forensic examination of the hard disk drive. These remnants, however, are typically fragmented, disconnected, and incomplete. Therefore if the chat or instant message is still present on the screen (and thus probably still in RAM) the investigator needs a method by which to preserve and collect it.

A detailed discussion of capturing volatile data from RAM is beyond the scope of this chapter. Suffice it to say that many commercial forensic software packages can capture this data. Similarly, Linux-based tools can accomplish this as well. The examiner may even be able to export the data remotely to another device. Regardless of the method, the data needs to be acquired.

Furthermore, many programs such as America Online Instant Messenger, Yahoo Messenger, and mIRC (Internet Relay Chat) create files regarding the rooms or channels a user chatted in or the screen names with which a user sent instant messages. Each application needs to be researched and the computer forensic examination guided by an understanding of how it functions.

HACKING

Unauthorized computer intrusion, more commonly referred to as **hacking**, is the concern of every computer administrator. Hackers penetrate computer systems for a number of reasons. Sometimes the motive is corporate espionage; other times it is merely for bragging rights within the hacker community. Most commonly, though, a rogue or disgruntled employee with some knowledge of the computer network is looking to cause damage. Whatever the motivation, corporate America frequently turns to law enforcement to investigate and prosecute these cases.

Generally speaking, when investigating an unauthorized computer intrusion, investigators concentrate their efforts in three locations: *log files, volatile memory*, and *network traffic*. Logs and anomalons typically document the IP address of the computer that made the connection. Logs can be located in several locations on computer network. Most servers on the Internet track connections made to them through the use of logs. Additionally the router (the device responsible for directing data) may contain log files detailing connections. Similarly, devices known as **firewalls** may contain log files listing computers that were allowed access to the network or an individual system. Firewalls are devices (taking the form of either hardware or software) that permit only requested traffic to enter a computer system (or, more appropriately, a network). In other words, if a user didn’t send out a request for Internet traffic from a specific system, the firewall should blocks its entry. If the log files captured the IP address of the intruder, then revealing the user behind the IP is the same process as for e-mail. Investigating a computer intrusion, however, does get a bit more complicated than this.

Frequently, in cases of unlawful access to a computer network, the perpetrator attempts to cover the tracks of his or her IP address. In these instances, advanced investigative techniques might be necessary to discover the hacker's true identity. When an intrusion is in progress, the investigator may have to capture volatile data (data in RAM). The data in RAM at the time of an intrusion may provide valuable clues into the identity of the intruder, or at the very least his or her method of attack. As in the case of the instant message or chat conversation, the data in RAM needs to be acquired. Another standard tactic for investigating intrusion cases is to document all programs installed and running on a system, in order to discover malicious software installed by the perpetrator to facilitate entry. The investigator uses specialized software to document running processes, registry entries, open ports, and any installed files.

Additionally, the investigator may want to capture live network traffic as part of the evidence collection and investigation process. Traffic that travels the network does so in the form of data packets. In addition to data, these packets also contain source and destination IP addresses. If the attack requires two-way communication, as in the case of a hacker stealing data, then data needs to be transmitted back to the hacker's computer using the destination IP address. Once this is learned, the investigation can focus on that system. Moreover, the type of data that is being transmitted on the network may be a clue as to what type of attack is being launched, whether any important data is being stolen, or what types of malicious software, if any, are involved in the attack.

Chapter Summary

The Internet, often referred to as the "information superhighway," is a medium for people to communicate and to access millions of pieces of information from computers located anywhere on the globe. No subject or profession remains untouched by the Internet, including forensic science. The Internet brings together forensic scientists from all parts of the world, linking them into one common electronic community.

The Internet can be defined as a “network of networks.” A single network consists of two or more computers that are connected to share information; the Internet connects thousands of these networks so all of the information can be exchanged worldwide. Connections can be made through a modem, a device that allows computers to exchange and transmit information through telephone lines. Higher-speed broadband connections are available through cable lines or through DSL telephone lines. Computers can be linked or networked through wired or wireless (Wi-Fi) connections. Computers on the Internet have a unique numerical Internet protocol (IP) address and usually a name. Commercial Internet service providers connect computers to the Internet while offering the user an array of options. The most popular area of the Internet is the World Wide Web, a collection of pages stored in computers connected to the Internet throughout the world. Web browsers allow users to explore information stored on the web and to retrieve webpages they wish to read. Several directories and indexes on the Internet, known as search engines, help users locate information on a particular topic from the hundreds of thousands of websites on the Internet. A keyword or phrase entered into a search engine will locate websites that are relevant to that subject.

The service that is most commonly used in conjunction with the Internet is electronic mail (e-mail). This communication system can transport messages across the world in a matter of seconds. Extensive information relating to forensic science is available on the Internet. The types of webpages range from simple explanations of the different fields of forensics to intricate details of forensic science specialties.

Investigators seeking a history of an Internet user’s destinations can take advantage of the fact that computers store or cache portions of webpages visited, and websites often create cookies to track certain information about website visitors. An investigator tracking the origin of an e-mail will seek out the sender's IP address in the e-mail's header. Chat and instant messages can typically be located in a computer's random-access memory (RAM). Finding the origin of unauthorized computer intrusions (hacking) requires investigation of a computer's log file, RAM, and network traffic, among other things.

Review Questions

1. A(n) ____________ consists of two or more computers that are connected to share information.

2. The device that allows computers to exchange and transmit information through telephone lines is a(n) ____________.

3. The most popular area of the Internet from which information can be searched and retrieved is known as ____________.

4. The (URL, domain abbreviation) is a unique electronic address that indicates where a document is actually located.

5. True or False: The advantage of using hypertext is to be able to quickly switch back and forth between related webpages without having to retype the URL or to start over at the beginning of the search. ____________

6. Typically, a user enters a keyword or phrase into a(n) ____________ to locate sites on the Internet that are relevant to a particular subject.

7. (E-mail, Usenet) is a communication system that transports messages across the world in a

matter of seconds.

8. A device known as a(n) ___________ allows a network of computers to share a common connection to the Internet.

9. A(n) ___________ takes the form of a series of numbers to route data to an appropriate location on the Internet.

10. A user's hard disk drive ___________ portions of webpages that have been visited.

11. A(n) ___________ is placed on a hard disk drive by a website to track certain information about its visitors.

12. E-mails have the ___________ address of the sender in the header portion of the mail.

13. Chat and instant messages conducted over the Internet are typically stored in (RAM, ROM).

14. When investigating a hacking incident, investigators concentrate their efforts on three locations: ___________, ___________, and ___________.

15. Devices that permit only requested traffic to enter a computer system are known as (caches, firewalls).

Further References

Casad, J., *Sams Teach Yourself TCP/IP in 24 Hours,* 3rd ed. Indianapolis, Ind.: SAMS, 2004.

Chamakura, R. P., "Forensic Science and the Internet—Current Utilization and Future Potential," *Forensic Science Review* 9 (1997):97.

Leshin, C. B., *Internet Investigations in Criminal Justice.* Upper Saddle River, N.J.: Prentice Hall, 1997.

Prosise, C., K. Mandia, and B. Pepe, *Incident Response and Computer Forensics*, 2nd ed. New York: McGraw-Hill, 2003.

WebExtra 18.1

The Scott Peterson Case

www.prenhall.com/Saferstein

Modem

A device that connects a computer to another computer through a phone line.

Broadband

Describes any kind of Internet connection with a download speed of more than 56 kilobits per second.

Wi-Fi

Technology that uses high-frequency radio signals to transmit and receive data over the Internet. Allows for a wireless connection to the Internet.

Router

A device that manages traffic between computers belonging to a network, enabling them to share a connection to the Internet.

VoIP (Voice Over Internet Protocol)

Transmission of the human voice over the Internet, usually through a telephone.

Internet Service Provider (ISP)

A company that provides connections to the Internet.

Internet Protocol

The set of rules used to transmit packets of data over the Internet and route them to their destination.

Domain

A human readable name and abbreviation for a website—for example, com, org, or gov are common abbreviations.

Browser

A program that allows access to websites.

Uniform Resource Locator (URL)

A standard method by which Internet sites are addressed.

Hypertext

Links to other websites. The linked document is displayed by clicking on a highlighted word or icon.

Search Engine

A website devoted to searching for information on the Internet using keywords.

E-mail

Electronic mail.

Download

The transfer of a file through an Internet connection from a remote computer to a user's computer.

Mailing List

A list of people with a common interest who receive all the e-mails sent to the list.

Newsgroups

Large bulletin board systems that consist of several thousand specialized discussion groups.

Messages are posted to a bulletin board via e-mail for others to read.

WebExtra 18.2

An Introduction to Forensic Firearms Identification

www.prenhall.com/Saferstein

WebExtra 18.3

Carpenter's Forensic Science Resources

www.prenhall.com/Saferstein

WebExtra 18.4

Crime-Scene Investigation

www.prenhall.com/Saferstein

WebExtra 18.5

Crimes and Clues

www.prenhall.com/Saferstein

WebExtra 18.6

Interactive Investigator—Détective Interactif

www.prenhall.com/Saferstein

WebExtra 18.7

The Chemical Detective

www.prenhall.com/Saferstein

WebExtra 18.8

Questioned-Document Examination

www.prenhall.com/Saferstein

Internet Cache

Portions of visited webpages placed on the local hard disk drive to facilitate quicker retrieval once revisited.

Cookies

Files placed on a computer from a visited website; they are used to track visits and usage of that site.

Internet History

An accounting of websites visited. Different browsers store this information in different ways.

Bookmark

A feature that enables the user to designate favorite sites for fast and easy access.

WebExtra 18.9

Follow the trail of an e-mail as it travels through the Internet

www.prenhall.com/Saferstein

Hacking

Has various meanings, but is frequently used as a slang term for an unauthorized computer or network intrusion.

Firewall

Hardware or software designed to protect intrusions into an Internet network.

Figure 18–1 Two computers communicating by sending data to each other's IP address via the Internet. An IP address is assigned to each computer by their respective Internet service providers (ISPs).

Figure 18–2 A user wishing to visit the *New York Times* website types the user-friendly web address www.nytimes.com. Because all traffic on the Internet is routed by IP address, the web address needs to be resolved to the IP address. This is done by the user's ISP's domain name system (DNS).

Figure 18–3 Bookmarks or favorite places can be saved for quick access in most web browsers.

Figure 18–4 The Internet history displays more than just web browsing activity. Here we see Microsoft Word documents and a picture accessed on the current day.

Figure 18–5 Sites such as www.samspade.org can be used to track the origins of an IP address. *Courtesy Word to the Wise*

[1] An excellent history of the Internet is found in Katie Hefner and Matthew Lyon, *Where Wizards Stay Up Late: The Origins of the Internet* (New York: Simon & Schuster, 1996).