INVESTIGATING BY COMPUTER
SECOND EDITION

Association of Certified Fraud Examiners

GLOBAL HEADQUARTERS • THE GREGOR BUILDING
716 WEST AVE • AUSTIN, TX 78701-2727 • USA

VI. INVESTIGATING WITH DIGITAL FORENSICS

Digital information has become of paramount importance. Today more and more information is created, stored, and disseminated electronically, and digital data can be stored in large volumes and in a number of different locations. The increasing use of technology in all aspects of our personal and professional lives has created new opportunities for technology to be used to perpetrate almost every type of fraud. Fraud examiners must therefore be prepared to gather information from digital technology.

Fraud-related matters of all sorts deal with a multitude of data sources, which differ significantly depending on the case. The underlying event that prompts a fraud examiner's actions drives the prioritization of the types of data they analyze, what information they need, and the usefulness of that data concerning their responsibilities.

Although digital data is prevalent, its collection nevertheless poses challenges, such as:

- A lack of access to, awareness of, and control over data sources
- The evolving nature of technology
- The dependency on specific tools to acquire data
- The lack of any established standards and processes for collecting, storing, and preserving data
- The volatility of digital data
- Anti-forensics techniques (i.e., measures used to thwart digital forensic analysis, such as encryption and steganography)
- Increasing storage capacity of digital devices
- Fraud examiners must therefore be prepared to address the myriad issues related to examinations involving digital evidence.

This chapter covers the various aspects of conducting a digital forensics investigation.

Digital Forensics Versus Digital Investigations

While the terms *digital forensics* and *digital investigation* might appear synonymous, there are important distinction between the two terms. *Digital investigations* are investigations that involve relevant digital data processed or stored by *digital devices*—devices that process data in the form of numbers (digits). Digital devices can be used to communicate with others, create documents, access data online, enter data online, store information, and so on. An investigator leading an investigation into a crime that involves a digital device is not necessarily—and, in most cases, should not be—the forensic examiner.

Conversely, *digital forensics* refers to the process of recovering, investigating, and interpreting data found in digital devices for use in a court of law.

It is important to keep digital investigations and digital forensics separate. Combining the two, especially in cases in which the suspect is already named, can invite questions about the forensic examination's objectivity. Additionally, combining the two could subject investigators to unwelcome scrutiny regarding whether they suppressed exculpatory evidence that might have been found during an examination by a more objective investigator.

Digital Forensic Experts

Examiners do not need to be digital forensics experts to be involved in a forensic investigation. However, when conducting an examination involving computers, fraud examiners should determine whether they need a digital forensics expert. Generally, digital forensics should be performed by digital forensic experts specifically trained and experienced in investigative methods.

Digital forensic experts are individuals who specialize in identifying, recovering, collecting, preserving, processing, and producing digital data for use in investigations and litigation. These experts have a good basic understanding of computer science, operating systems, software, and data security issues, and they can uncover a large amount of data that relates to the use of a computer, including what is or has been stored on it, and the details about the computer's users.

Some organizations have their own in-house personnel whom they have trained and outfitted with the proper equipment and software tools to conduct the examination and analyze digital evidence, while others prefer the use of an outside examiner who can conduct a thorough examination, prepare a proper report, and deliver expert testimony if needed in legal proceedings.

Sometimes retrieving digital data is as easy as searching the target computer's hard drive, but other times retrieval requires a thorough knowledge of computers. Digital forensic experts might also be able to recover evidence that a non-expert cannot. For example, if the target of an investigation tried to delete electronic evidence, a forensic expert might be able to recover it. There are, in fact, a variety of ways to recover deleted data from a computer, and digital forensic experts are specially trained for such tasks. Deleted files are recoverable until they are overwritten because data is not erased from a computer's hard drive until it is overwritten. Digital forensics specialists usually cannot recover deleted files that have been overwritten.

A digital forensics expert can determine, among other things:

- Who used a computer
- Software recently used on a system
- Websites visited by a user
- Documents, letters, and images created, modified, or accessed by a user
- Who created a document and what software they used to create it
- What software, if any, has been used to "clean" the computer or wipe the hard drive

Additionally, digital forensics experts can recover, among other things, the following types of information from computer systems:

- Deleted files and other data that has not been overwritten (e.g., deleted documents, images, link or shortcut files, and email messages)
- Files deleted through computer-automated processes
- Temporary auto-save files
- Print-spool files
- The history of visited websites, even where the browser history and cache have been deleted
- Communications sent via email, chat, or instant messages
- Financial-based Internet transactions
- Documents, letters, and images created, modified, or accessed, even if the data was not saved on the computer in some situations
- Data that has been copied, corrupted, or moved
- The time and date information about files (e.g., when files were created, accessed, modified, installed, deleted, or downloaded)
- Data from a drive that has been defragmented or reformatted
- Hidden files

The increased sophistication of hardware and operating systems allows computer systems to store more information about how people use their computers, and therefore digital forensic examiners can uncover a large amount of data that relates to the use of a computer, what is or has been stored on it, and the details about the computer's user.

Moreover, digital forensic examiners have special tools and software designed to facilitate a thorough and legally sufficient analysis of items that contain digital evidence. It is important to allow a digital forensics expert to conduct a proper seizure and examination of digital evidence to ensure that the information can be used in a legal proceeding.

Failing to follow established forensic methods when searching and collecting electronic evidence can be disastrous, and therefore an inquiry that involves more than cursory analysis of electronic evidence

should involve a digital forensics expert. The more technical the nature of an inquiry, the more technical and specialized the forensics skills required.

Furthermore, whether technical specialists are needed to assist with an investigation depends on the complexity of the examination. The more technical the nature of the crime, the more technical and specialized the analytical skills required. Some computer investigations do not require extremely sophisticated technical skills. For example, in cases where the computer is the instrument used to commit the crime (e.g., scams involving payroll applications, cooking the books, fraudulent electronic funds transfers, or insider trading), the fraud examiner only needs to know how to search a database and analyze the results.

Additionally, the determination as to whether a specialist is needed depends on the fraud examiner's skill level. Therefore, a fraud examiner conducting an examination involving computers must have a basic understanding of their skill level and must be aware of their limitations in the field of computer technology because, in many instances, they need one or more technical specialists to assist with an investigation.

Within the digital forensics field, there are several different types of special experts:

- *Operating and file system experts*: These experts are proficient in certain operating systems (e.g., Windows, Linux, and OS X) and the various file systems they employ (e.g., Joliet, NTFS, FAT, VFS, Ext2), and they have the ability to convey operational characteristics and observe artifacts.
- *Data recovery experts*: These experts operate clean rooms designed to magnetically extract information from damaged media sources. Using special tools and equipment, these experts can disassemble a hard disk, separate the platters, and extract and reassemble the information for subsequent examination.
- *Forensic accounting experts*: Forensic accounting is the use of professional accounting skills in matters involving potential or actual civil or criminal litigation, and a forensic accountant can provide various services, including audits, accountant performance reviews, and examinations of financial documents for fraud, misconduct, or industry standard violations.
- *Recording and archival extraction experts*: These are experts in extracting information from tapes, digital media, or other system backups. Typically, backup solutions archive data in proprietary formats, making extraction very cumbersome.
- *Intrusion and malicious code experts*: These experts specialize in investigating computer network intrusions. Specialists can determine attack vectors, the tools employed, what occurred during access, and what, if anything, was taken.
- *Cloud-computing experts*: These experts specialize in investigating cloud networks.
- *Mobile-device experts*: These experts specialize in extracting information from mobile devices.

Given the diversity of computer-related fraud, no person can be an expert in all aspects of computer technology. However, generally, to conduct a successful examination involving digital technology, the digital forensics expert must know what to look for and where to look for it. This knowledge requires that the expert be familiar with a number of related topics, including:

- The various types of file systems used by different operating systems
- Where important evidentiary information can be stored in computer operating systems
- Where emails reside and how they can be recovered
- How to find information on the suspect's network, peripheral devices, and other forms of technology that can transmit or store data
- How to collect digital evidence in a forensically sound manner
- The various digital fraud detection tools
- The unique concerns related to electronic discovery

If a fraud examiner decides to hire a forensic expert, they must be diligent when hiring the expert. The best way to hire an expert is to follow a few simple guidelines:

- Ask colleagues for referrals.
- Make sure the expert is properly licensed (if required) and insured.
- Ask the experts if they have worked on the type of case that is at issue, and, if so, what they were retained for and the outcome.
- Set up a budget for the investigation.
- Make sure there is a comfortable relationship and good communication between the client and the expert.
- Listen to the expert because they might have new, unique ideas and suggestions that were not previously considered.

Digital forensics experts work for multiple clients on many different types of engagements, and therefore fraud examiners must perform a conflict of interest check to ensure that any prospective experts do not have any relationships that would create a conflict of interest.

Video

In the video titled "Chapter VI: Forensic Analytics," data analytics expert Vincent Walden, CFE, CPA, provides insight into forensic analytics and some of the tests involved in the field. [redacted] o view the video.)

Digital Evidence

Digital evidence is information stored or transmitted in binary form (i.e., ones and zeroes) that can be used to prove something. Again, the proliferation of digital technologies has created new opportunities for technology to be used in perpetrating almost every type of fraud. Consequently, fraud examiners gather some type of digital evidence in almost all fraud examinations.

Generally, when digital data is collected in an investigation, the facts at issue involve a computer, either as (1) a target of a criminal act, (2) an instrument of crime, or (3) a repository of evidence associated with the crime.

Computers themselves can be the targets of crime. Crimes committed against computers include computer and computer-component theft, system intrusions, software piracy, and software theft.

Computers can also be used to facilitate criminal conduct. When this occurs, the computer is known as the *tool* or *instrument* of crime.

Additionally, computers can be repositories of evidence associated with a crime. Generally, in an examination involving computers, the fraud examiner attempts to locate the storage of potential digital evidence, in one form or another, on the computer system. The computer system involved in the investigation is a potential repository of evidence whether the user intended to store an item or not. Therefore, in such situations, the fraud examiner is interested in incriminating evidence that the user intentionally or unintentionally stored on the computer system.

Volatility of Digital Evidence

Digital evidence is more volatile than tangible evidence because data can be altered or destroyed more easily than tangible information. Digital data is, by design, fragile and short-lived in nature. It is easily manipulated, substituted, modified, and deleted.

Generally, if a computer system is on, its files are changing. Operating systems and programs frequently alter, delete, and modify digital data, and this might happen automatically. Moreover, if a user interacts with a system, its files change. In fact, digital information can be altered by seemingly harmless actions, such as shutting down a running system, starting up a system, or looking through files on a running computer. In addition, if a computer system is connected to a network, its files will be changed.

The following is a list of some actions that can alter, delete, or modify data that is potentially relevant to an investigation:

- Using or interacting with a computer system

- Clicking on files or folders on a computer, which results in information being written to the system's hard drive and could potentially overwrite valuable evidence
- Turning systems on or off
- Visiting websites
- Using software applications
- Downloading or transferring files

Because digital evidence can be easily altered or destroyed, the integrity of digital evidence must be preserved. If data has been altered or destroyed, that is considered a violation of data integrity. Moreover, the alteration or destruction of digital evidence is irreversible. Thus, once the integrity of digital evidence has been violated, it usually cannot be restored.

The failure to preserve the integrity of digital evidence can result in several adverse consequences. First, the failure to preserve the integrity of digital evidence could result in the government's questioning of the integrity of any evidence collected in a fraud investigation.

Second, digital evidence that is destroyed when litigation is expected, or in progress, might give rise to claims of spoliation of evidence, which, if proven, could lead to monetary fines and sanctions, adverse inference jury instruction sanctions, or dismissal of claims or defenses. *Spoliation* is broadly defined as the act of intentionally or negligently destroying documents relevant to litigation.

Third, the failure to preserve the integrity of digital evidence could result in evidence being deemed inadmissible in a legal proceeding, or, if it is admitted, it might not be given much weight because evidence of questionable authenticity does not provide reliable proof.

Generally, evidence is not admitted (or it might not constitute reliable proof of what it is offered to prove) unless it is authenticated. Evidence is authenticated when the party offering the item produces some evidence (e.g., testimony from a person with firsthand knowledge) to show it is, in fact, what the party says it is and to show it is in the same condition from the moment it was seized until it is used in court. Thus, if it is determined that a piece of digital evidence cannot be authenticated, it might not be admitted even if it is plainly relevant, or it might be rejected because it is deemed unsuitable to prove the facts it purports to prove.

Although digital evidence differs from tangible evidence, the rules regarding the admissibility of digital evidence in court are really no different from the rules regarding the admissibility of any other type of evidence.

The rules of admissibility, however, differ in civil law and common law systems. Because civil law systems do not rely on juries like common law systems, there is a relative lack of restrictions on the admissibility of evidence in civil law systems. Thus, there are far fewer limitations on the admissibility of evidence in civil law trials than in common law ones.

Generally, in civil law systems, evidence is admitted if the presiding judge determines it is relevant, even if the authenticity of an item of evidence is in question. However, even though any relevant evidence is admissible, the court evaluates how much weight is to be given to an item of evidence, and courts consider authenticity when determining what weight is appropriate for evidence. Authentic evidence is reliable proof of what it purports to show, but evidence that lacks authenticity is not reliable proof of what it purports to show. Thus, even if evidence of questionable authenticity is admitted, it will not be given much weight because it is not reliable or helpful to a fact-finder.

In contrast, the common law system contains several rules that restrict admission of evidence. However, generally, to be admitted into evidence in common law systems, evidence must, in addition to being established as authentic, be:

- *Relevant to an issue that is in dispute in the case*: Relevant evidence is evidence that tends to make some fact in issue more or less likely than it would be without the evidence.
- *Material*: Material evidence is evidence that has important value to a case or that can be used to prove a point. Repetitive or additive evidence is nonmaterial evidence.

Therefore, if a fraud examiner collects digital evidence, they should be able to state unequivocally that the evidence was not changed in any way by their actions. This requires that they follow strict forensic methodologies to satisfy the stringent evidentiary standards necessary to ensure the integrity of the evidence beyond a reasonable doubt for presentation in court. That is, digital evidence must be properly preserved in a forensically sound manner so that it will be admissible.

Video

In the video titled "Chapter VI: Digital Evidence," fraud investigation expert Jean-François Legault, CFE, explains some of the concerns associated with digital evidence. Mr. Legault is the vice president, assistant director of high tech investigation at JPMorgan Chase, specializing in computer forensic investigations. ([redacted] view the video.)

Understanding File Systems

Introduction

To be effective, a digital forensics examiner must have a working knowledge of the various types of file systems used by different operating systems. This knowledge is necessary for several reasons. For example, if called to testify during trial, the examiner might be asked to explain how data is stored by a particular computer or operating system. This knowledge can also help identify potential sources of evidence that might otherwise be hidden from view.

A *file system* is "a method for storing and organizing computer files and the data they contain, to make it easy to find and access them."[5] It provides a method to access specific data, which is stored on a disk in files. A *file* is a named collection of related data that is used for organizing secondary memory (i.e., memory that the computer cannot directly access).

Most file systems make use of an underlying data-storage device that offers access to an array of fixed-size blocks, sometimes called *sectors*, which are most commonly a power of 2 in size (512 bytes or 1, 2, or 4 KB). The file system software is responsible for organizing these sectors into files and directories, and it is responsible for keeping track of which sectors belong to which file and which sectors are not being used.

There is a difference in most operating systems between the logical size of a file and the physical size of the file. The *logical size* of a file pertains to the number of bytes that it occupies and, in a directory listing, this is the number that is displayed for the file size. The *physical size* of a file, however, depends on how the operating system stores files. In most operating systems, the data in a disk is organized in fixed-size units—called *clusters* or *blocks*—that contain a certain number of disk sectors (usually 1–64), and the physical size is determined by the minimum number of whole clusters a file needs. For example, an 18 kb file, which takes up 4.5 clusters (one cluster equals 4 kb), needs 5 clusters for its physical size. And because 5 clusters are 20 kb, the file's physical size is 20 kb. The file's logical size—the actual size of the file—is 18 kb.

Moreover, files occupy a whole number of clusters, even if the logical size of the file is smaller than the cluster size. If the logical size of the file is smaller than the cluster size, the difference in space between the physical file space and the logical file space is called the *file slack* or *slack space*. In the above example, the file slack is 2 kb (20 − 18 = 2). The file slack might contain data from previously erased files, and

5. Lucio D. Jasio, Programming 16-bit PIC Microcontrollers in C: Learning to Fly the PIC 24 (Burlington, MA: Newnes, 2007).

such data could contain important evidence. Unless the fraud examiner has a special utility to identify and investigate the file slack area, potential evidence that might be located there could be missed.

A *volume*, however, is a collection of logical storage units that are represented by drive letters assigned by their operating systems. A single physical disk can contain several volumes.

Types of File Systems

There are many different kinds of file systems, and many operating systems support more than one type of file system.

Windows Operating Systems

Windows operating systems use two major file systems to organize data: the File Allocation Table (FAT) and the New Technology File System (NTFS).

FILE ALLOCATION TABLE

The File Allocation Table (FAT) is a file system designed to keep track of the allocation status of clusters on a hard drive. The initial version of FAT is now referred to as FAT12 and was designed as a file system for floppy diskettes. FAT16 was introduced in 1987 and allowed for larger volumes than FAT12—a maximum of 2 GBs. To overcome the size limitations of FAT16, Microsoft introduced FAT32 when it released Windows 95.

A FAT file system is composed of four different sectors: the reserved sector, the FAT region, the root directory region, and the data region.

First, the *reserved sector* is the boot sector that includes an area called the BIOS Parameter Block and usually contains the operating system's boot loader code.

Second, the *FAT region* typically contains (but might vary) two copies of the file allocation table for the sake of redundancy checking, although the extra copy is rarely used, even by disk-repair utilities. This table is the map to the data region, indicating which clusters are used by files and directories. The primary task of the file allocation table is to keep track of the allocation status of clusters, or logical groupings of sectors, on the disk drive. There are four different possible FAT entries: allocated (along with the address of the next cluster associated with the file), unallocated, end of file, and bad sector.

Third, the *root directory region* is a directory table that stores information about the files and directories located in the root directory. It is only used with FAT12 and FAT16. The root directory has a fixed maximum size that is pre-allocated at the creation of the volume. FAT32 stores the root directory in the data region along with files and other directories, allowing it to grow without restraint.

Finally, the data region is where files are actually stored. This is where the actual file and directory data is stored and takes up most of the partition.

NTFS

NTFS is the standard file system of Windows operating systems; it is based on Windows NT (2000, XP, Vista, etc.). "NTFS supersedes the FAT file system as the preferred file system for Windows operating systems. NTFS has several improvements over FAT and HPFS (High Performance File System) such as improved support for metadata and the use of advanced data structures to improve performance, reliability, and disk space utilization, plus additional extensions such as security access control lists (ACL) and file system journaling."[6]

MASTER FILE TABLE

In NTFS, all file data (e.g., file name, creation date, access permissions, and contents) are stored in the Master File Table (MFT) as metadata. The "MFT describes all the files and folders on the volume, including file names, time stamps, stream names, lists of cluster numbers where data streams reside, indexes, security identifiers, and file attributes."[7] The first 16 files referenced in the MFT are metafiles that define and organize the file system. These metafiles define files, back up critical file system data, buffer file system changes, manage free-space allocation, satisfy BIOS expectations, track bad allocation units, and store security and disk-space usage information.

Some specific metafiles of NTFS include $MFT, $MFTMirr, and $BITMAP. The $MFT is the first record in the MFT; it directly references the MFT. It describes all files on the volume (e.g., file names; timestamps; stream names; indexes; security identifiers; file attributes like "read only," "compressed," and "encrypted;" etc.) The second record is a partial copy of the MFT ($MFTMIRR) in case the MFT becomes corrupted. Another record is the bitmap file ($BITMAP), which should not to be confused with the image file format. The bitmap file is used to identify whether the cluster is in use (allocated to a file) or free (available for allocation). Every cluster on a volume contains an entry in the bitmap file where a bit is set to show whether the cluster is in use or free.

RECORDS

NTFS stores changes to MFT records. The Update Sequence Number Journal (USN Journal) is the feature that records changes to all files, streams, and directories on the volume. Each MFT entry contains a set of attributes pertaining to the file the entry references, and MFT records include attributes as well as their associated values. There are several attributes defined for NTFS, and such records exist for both files and folders. Folder records also include an index root attribute that points out the entries

6. AppleXSoft, "New Technology File System (NTFS)," www.applexsoft.com/glossary/ntfs.
7. Absolute Astronomy, "NTFS," www.absoluteastronomy.com/topics/NTFS.

for each file and subfolder. These records include the MAC times associated with each entry. MAC times are file metadata that record when certain most recent events happen to a file, such as its most recent access and edit.

FILE NAMES

File names—names assigned to a file—under NTFS are stored using Unicode characters and can be up to 255 characters. They can include embedded spaces, multiple periods, and special characters, all of which were not allowed in MS-DOS file names.

Mac Operating Systems

OS X—Apple's modern operating system—supports a variety of file systems, but the primary file system used by OS X is the HFS Plus file system. OS X also supports the UFS file system, and some versions of OS X can read and write to FAT file systems, which are common on Windows devices.

Locating Relevant Data

To conduct a successful examination, fraud examiners must know what to look for and where to look for it. However, this can be difficult because digital data can be stored in large volumes and in a number of different locations. For example, the fraud examiner should know where to look for:

- Information on any suspect computer systems
- Information on a suspect's workstation, including any peripherals or other portable media devices that contain data
- Information stored on any network from which the suspect's traffic flows
- Information stored in cloud storage services

This discussion examines where digital evidence might be located on computer systems, workstations and peripheral devices, networks, smartphones, and cloud environments.

Computer Systems

A wealth of information can be recovered and analyzed on seized computer systems. Common files containing evidence stored in computer systems include user-created files, user-protected files, and computer-created files.

User-Created Files

User-created files are digital files created under the user's direction. These files include text-based documents, spreadsheets, databases, emails, address books, presentation slides, audio and video files, image files, Internet bookmarks, and so on.

User-Protected Files

Often, users hide or protect files to prevent them from being found or accessed. There are a variety of techniques to hide files, but some of the most common methods include:

- Designating files as hidden
- Camouflaging files
- Using steganography
- Hiding or mislabelling hard copies of data
- Encrypting files
- Using alternate data streams

HIDDEN FILES

Most computer operating systems allow files to be designated as *hidden* files. When so designated, the files do not show up in file listings or searches. There are, however, ways to show hidden files, and fraud examiners should inspect hidden files to determine if the user intentionally hid them because they contain valuable information.

CAMOUFLAGED FILES

A user might camouflage certain files under an innocent name or different file extension to prevent others from discovering them. For example, a suspect might change a file name from "evidence.doc" to "install.exe" and place the file in a directory that stores program files. Therefore, fraud examiners should analyze a target's hard drive to determine whether any file types have been camouflaged. This is done by analyzing the file header—the first bits of data in a file—which contains data identifying the file format.

HIDING OR MISLABELING HARD COPIES OF DATA

Data can be stored on or off the computer. For instance, if, when searching for evidence in a suspect's workstation, investigators ignore the music CD cases in the suspect's desk, assuming that the cases do not contain anything of value, the investigators might miss the evidence potentially stored on a CD-R disc that is hidden in one of the music CD cases. Likewise, users might try to conceal data by storing the data on a commercially labeled disk. For example, in one case, a hacker stored all of his hacking programs on a disk that originally contained Microsoft Windows program installation files.

STEGANOGRAPHY

A user might also seek to protect files using steganography. *Steganography* is the process of hiding one piece of information within an apparently innocent file. For example, a user can use the least significant bits of a bitmap image to hide a message. By hiding the message in the least significant bits of an image,

there is almost no perceivable change in the bitmap image itself. Without directly comparing the altered image to the original, it is practically impossible to tell that the image was altered.

EXAMPLE

After a French defense contractor suspected that some of his critical designs were being compromised, the government conducted an investigation. The investigation revealed that an employee had obtained his position at the company to steal the contractor's trade secrets. To transmit the trade secrets out of the company, the employee embedded them into graphic images used on the company's public access website using steganography. Another party then copied the image files from the website and extracted the proprietary designs.

Digitized audio files are another type of file that can be used to hide messages. For example, by encoding a message in the least significant bits of a WAV file, a user can make the message almost impossible to detect.

There are a number of tools that investigators can use to detect steganography, and these tools utilize different methods to detect the use of steganography. Some common methods of detecting the use of steganography are:

- Visual detection by looking for visual anomalies in JPEG, BMP, GIF, and other image files
- Audible detection by looking for audible anomalies in WAV, MP3, MPEG, and other media files
- Statistical detection by determining whether the statistical properties of files deviate from the expected norm
- Structural detection by looking for structural oddities that suggest manipulation (e.g., size differences, date differences, time differences, or content modification)

ENCRYPTION

A user might also protect files by encrypting them. *Encryption* refers to procedures used to convert information using an algorithm (called a *cipher*) that makes the information unreadable to anyone without the encryption key. There are two primary ways of accessing encrypted information without a key. First, numerous utility programs can decrypt documents encrypted by many of the more common software applications. In addition, some encryption programs have a secret "key" for emergencies that may be used to decrypt encrypted data. There are also several companies that specialize in decryption.

Second, there is precedent for forcing a suspect (or employee) to divulge the decryption key. The computer is only a container and encryption is an additional lock prohibiting the investigator from reviewing the files for potential evidence. A fraud examiner with a legal right to examine encrypted data

might be able to force a suspect to divulge the decryption key because doing so is no different than ordering a suspect or employee to unlock a file cabinet to inspect its contents.

ALTERNATE DATA STREAMS

It is also possible to hide files in Windows systems by using alternate data streams. This method, however, is only possible in Windows New Technology File System (NTFS)—one of two basic types of file systems used to organize data in Windows systems. Alternate data streams are a feature in the NTFS file system that offer the ability to put data into existing files and folders without affecting their functionality, size, or display in many traditional file-browsing utilities like command line or Windows Explorer.

Although Windows Explorer does not provide a way to see what alternate data streams are in a file, alternate data streams are easy to make, and an individual with basic knowledge can use them to hide files without having to use any third-party tools.

Fortunately, there are forensic tools that can search NTFS partitions for alternative data streams.

Computer-Created Files

Evidence might also be found in computer-created files, which are files generated by a computer's operating system. This type of information is important because it can identify that a certain activity has taken place, and in most cases the user is not aware that this information is being written. Some common examples of computer-generated data available for examination are discussed below.

METADATA

Metadata is data about data, and these file tidbits contain a tremendous amount of information.

There are two sources of metadata in Windows operating systems: file-system metadata and file metadata.

FILE-SYSTEM METADATA

Every file contained on a computer has a record entry in the master file table (MFT) that includes information about attributes such as name, size, and relevant dates (e.g., the date created, accessed, written, or modified).

FILE METADATA

File metadata (or embedded metadata) is stored within a file and provides information about its host. This metadata provides information about the files in which it is stored. For example, in a Word

document, the file metadata might show who authored the document; who received, opened, copied, edited, moved, or printed the document; when the document was altered; the number of revisions made to the document; who last saved the document; and the document's last print date.

Similarly, when an individual takes a photograph with a digital camera, the camera stores Exchangeable Image File Format (EXIF) metadata into the image file. Some of the metadata that might be stored along with digital photos include:

- The date and time the photo was taken
- Details about the camera's settings (e.g., ISO speed, aperture, etc.)
- The camera's make and model
- The location coordinates (latitude and longitude) where the photo was taken

However, users can minimize or delete metadata. For instance, minimizing metadata is easy for Microsoft Office applications like Excel, PowerPoint, and Word. To delete metadata from files in such applications, a user simply clicks *File* → *Check for Issues* → *Inspect Document*. Then the user inspects the document for metadata and deletes any they want to remove.

WINDOWS REGISTRY

The Windows registry is a central database that stores settings and configurations for the operating system and most applications installed on the system. The system and its users, applications, and hardware make use of the registry to store configuration details and for reference while operating the system. The information stored in the registry includes settings for hardware, software, installed programs, and user preferences.

Registry files contain vast amounts of information, and they can show:

- User names and passwords for email, websites, and programs
- Internet sites visited, along with the dates and times of visits
- Search terms used on Google and other search engines
- Recent file activity
- List of software installed on system
- Whether the system requires users to log in
- Last date the user logged into the system
- Last failed log-in attempt
- Drives that the user mounted to the computer (i.e., drives that the user made accessible through the computer's file system)
- List of removable storage devices

Often, when a fraudster uses software, they create a footprint in the registry that can provide valuable insight into relevant activity that occurred in the system.

The Windows registry is a hierarchical database consisting of five sections called *hives*. These hives comprise keys, subkeys, and value entries.

Registry keys are similar to folders because each key can contain subkeys, which can contain further subkeys, and so forth. Keys and subkeys are referenced with a syntax similar to Windows' path names—using backslashes to indicate levels of hierarchy—and they work in basically the same way as the folders in the Windows Explorer.

Registry values are name/data pairs stored within keys. Values are referenced separately from keys. Although value names can contain backslashes, backslashes make the values difficult to distinguish from their key paths.

The five hives are provided below:

- HKEY_CLASSES_ROOT (HKCR): This hive stores information about registered applications, such as file associations and OLE Object Class IDs, tying them to the applications used to handle these items. On Windows 2000 and above, HKCR is a compilation of HKCU\Software\Classes and HKLM\Software\Classes. If a given value exists in both of the subkeys above, the one in HKCU\Software\Classes is used.
- HKEY_CURRENT_USER (HKCU): This hive stores settings that are specific to the currently logged-in user. The HKCU key is a link to the subkey of HKEY_USERS that corresponds to the user; the same information is reflected in both locations. On Windows NT-based systems, the users' settings are stored in their own files called NTUSER.DAT and USRCLASS.DAT inside their own Documents and Settings subfolder (or their own Users subfolder in Windows Vista). Settings in this hive follow users with a roaming profile from machine to machine.
- HKEY_LOCAL_MACHINE (HKLM): This hive stores settings that are general to all users on the computer. On NT-based versions of Windows, HKLM contains four subkeys: SAM, SECURITY, SOFTWARE, and SYSTEM; they are found within their respective files in the %SystemRoot%\System32\config folder. A fifth subkey, HARDWARE, is volatile, is created dynamically, and, as such, is not stored in a file. Information about system hardware drivers and services are located under the SYSTEM subkey, while the SOFTWARE subkey contains software and Windows settings.
- HKEY_CURRENT_CONFIG (HKCC): This hive contains information gathered at runtime. Information stored in this key is not permanently stored on disk, but rather regenerated at the boot time.
- HKEY_USERS: This hive contains user configuration for the system's user accounts.

A user can view the registry structure by launching the Registry Editor utility.

Windows Paging File

Operating systems have a limited supply of Random Access Memory (RAM) and, if one runs out of physical RAM, the user can move some of the data outside physical memory and store it temporarily in a special file on the hard disk. This file, which is called a *paging file* or a *swap file*, is used to simulate RAM. The paging file exists in a file called pagefile.sys and is located on the root of the Windows installation drive. Because the paging file stores information that is supposed to be stored in RAM, it might be possible to recover data never intended to be written to the hard drive by analyzing the pagefile.sys file.

Windows Hibernation File

When a system goes into hibernation mode, it writes the contents of RAM to the hibernation file before powering off. This is done so that the system can be restored to the state it was in when hibernation was implemented. The hibernation file, which exists in a file called hiberfil.sys, reflects the content of the RAM when the system went into hibernation, and it may be analyzed to recover information never meant to be written to the hard drive.

EVENT LOGS

Every operating system generates an *event log*—files that record events or transactions on a computer. Event logs provide a history of activity on a system, and they are a critical source of information when investigating cybersecurity incidents.

In fact, a log entry is created for each event or transaction that takes place on any computer; consequently, there are numerous types of event logs. Some common logs include system logs, application logs, and security logs.

SYSTEM LOGS

System logs record events executed on an operating system, including miscellaneous events and those generated during system start-up, like hardware and controller failures. Other common types of system events include starting up and shutting down, configuration updates, and system crashes.

APPLICATION LOGS

Application logs record events regarding access to application data, such as data files being opened or closed; reading, editing, deleting, or printing application files; or modifying records in an application file.

SECURITY LOGS

Security logs track security-related events like valid and invalid log-in attempts and times, password changes, and changes to access rights. Potential sources of security logs include:

- Server and workstation operating system logs
- Security tool logs (e.g., anti-virus and intrusion detection and prevention system)
- Outbound proxy logs

Internet Activity

Internet browsers create temporary files that store information about websites that a user has visited. These files can show websites that were recently visited and usually include time and date information relevant to the visit; they can also show images previously viewed online. This information allows the fraud examiner to recover websites and images previously viewed by the system's users.

Depending on the browser used, this information might be contained in a single file, a single folder containing numerous files, or multiple folders containing a number of files.

Typically, Internet browsers have three logging facilities that can be used to reconstruct a suspect's browsing activities: browser history, temporary Internet files, and cookies.

BROWSER HISTORY

All Web browsers store Internet browsing history, which refers to the list of Web pages (and associated data) a user has visited for a certain period.

TEMPORARY INTERNET FILES (CACHE)

Many browsers store temporary Internet files (or cache) that are downloaded and cached from the Internet when a user visits a website. Browsers use these files to store website data for Web pages or URL addresses visited by a user. Browsers store many different types of temporary Internet files when a user visits a site, such as entire Web pages, images, JavaScript, style sheets, video files, cookies and more.

Deleted temporary Internet files can be recovered from unallocated space and can be a valuable source of evidence.

COOKIES

Cookies are parcels of text sent by a server to a Web browser and then sent back by the client each time they access that server. HTTP cookies are used for authenticating, session tracking, and maintaining specific information about users, such as site preferences or the contents of electronic shopping carts.

Temporary Files

Applications create temporary files like those created by Internet browsers. For example, certain communication software like instant messaging (IM) and chat software keep a history of the user's conversations in a proprietary and sometimes encrypted format.