Data Analytics

Lesson 11

KEY CONCEPTS

■ Data ■ Data Recovery Tools ■ Evidence ■ Internet ■ Data Protection ■ Swap Files ■ Temporary Files ■ Cache Files

Learning Objectives

To understand:

- The procedure and ethical norms associated with data recovery in the overall perspective of Data Analytics
- Function of computer forensics tools for the different file systems
- Different nuances of identification, preservation, and analysis of evidence related to Data Analytics

Lesson Outline

- Data Recovery Tools
- Data Recovery Procedures and Ethics
- Gathering Evidence- Precautions, Preserving and safely handling original media for its admissibility
- Document a Chain of Custody and its importance
- Complete time line analysis of computer files based on file creation, file modification and file access
- Recover Internet Usage Data
- Data Protection and Privacy
- Recover Swap Files/Temporary Files/Cache Files
- Introduction to Encase Forensic Edition, Forensic Toolkit
- Use computer forensics software tools to cross validate findings in computer evidence-related cases
- Lesson Round-Up
- Test Yourself
- List of Further Readings

INTRODUCTION TO DATA ANALYTICS

Data Analytics is the science of analyzing raw datasets in order to derive a conclusion regarding the information they hold. It enables us to discover patterns in the raw data and draw valuable information from them. Data analytics processes and techniques may use applications incorporating machine learning algorithms, simulation, and automated systems. The systems and algorithms work on unstructured data for human use. These findings are interpreted and used to help organizations understand their clients better, analyze their promotional campaigns, customize content, create content strategies, and develop products. Data analytics help organizations to maximize market efficiency and improve their earnings.

Process of Data Analytics

Below are the common steps involved in the data analytics method:

Step 1: Determine the criteria for grouping the data

Data can be divided by a range of different criteria such as age, population, income, or sex. The values of the data can be numerical or categorical data.

Step 2: Collecting the data

Data can be collected through several sources, including online sources, computers, personnel, and sources from the community.

Step 3: Organizing the data

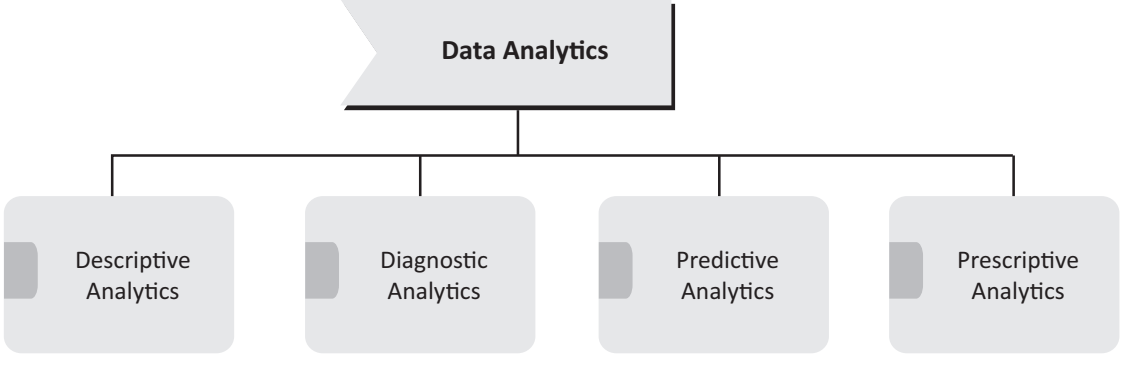
The data must be organized after it is collected so that it can be examined. Data organization can take place on a spreadsheet or other type of software that is capable of taking statistical data.

Step 4: Cleaning the data

The data is first cleaned up to ensure that there is no overlap or mistake. Then, it is reviewed to make sure that it is not incomplete. Cleaning the data helps to fix or eliminate any mistakes before the data goes to a data analyst for analysis.

Data Analytics Types

The following are the four fundamental types of data analytics:



1. **Descriptive Analytics** describes the happenings over time, such as whether the number of views increased or decreased and whether the current month's sales are better than the last one.
2. **Diagnostic Analytics** focuses on the reason for the occurrence of any event. It requires hypothesizing and involves a much more diverse dataset. It examines data to answer questions, such as "Did the weather impact the selling of beer?" or "Did the new ad strategy affect sales?"

3. **Predictive Analytics** focuses on the events that are expected to occur in the immediate future. Predictive analytics tries to find answers to questions like, what happened to the sales in the last hot summer season? How many weather forecasts expect for this year's hot summer?
4. **Prescriptive Analytics** indicates a plan of action. If the chance of a hot summer calculated as the average of the five weather models is above 58%, an evening shift can be added to the brewery, and an additional tank can be rented to maximize production.

Benefits of Data Analytics

1. **Decision-making improves**

   Companies may use the information they obtain from data analytics to guide their decisions, leading to improved results. Data analytics removes a lot of guesswork from preparing marketing plans, deciding what material to make, creating goods, and more. With advanced data analytics technologies, new data can be constantly gathered and analyzed to enhance your understanding of changing circumstances.

2. **Marketing becomes more effective**

   When businesses understand their customers better, they will be able to sell to them more efficiently. Data analytics also gives businesses invaluable insights into how their marketing campaigns work so that they can fine-tune them for better results.

3. **Customer service improves**

   Data analytics provides businesses with deeper insight into their clients, helping them to customize customer experience to their needs, offer more customization, and create better relationships with them.

4. **The efficiency of operations increases**

   Data analytics will help businesses streamline their operations, save resources, and improve the bottom line. When businesses obtain a better idea of what the audience needs, they spend less time producing advertisements that do not meet the desires of the audience.

DATA RECOVERY TOOLS

Data recovery is the process of retrieving data from a storage medium that, for some reason, cannot be accessed normally. This process may be used to recover data from a variety of storage media, such as: hard disk drives, solid-state drives, other flash storage (such as USB drives, and SD cards), or other disk storage (such as CDs, and DVDs). The damage that causes data to be lost typically falls into one of two categories: physical damage (where the hardware is damaged or is malfunctioning), or logical damage (where part of the software and/or file system prevents the data from being accessed by the host operating system.) We'll discuss these different types of storage damage in greater depth a bit later. The term "data recovery" can also be organized into two different contexts: personal data recovery, and forensic data recovery. Personal data recovery is what we normally associate with this topic. It simply refers to the retrieval of data that has been involuntarily lost or made inaccessible due to, for example, damaged storage media. By contrast, forensic data recovery often deals with retrieving data that has been purposely encrypted or hidden to prevent others (such as forensic investigators) from accessing the data.

Data recovery software is a type of software that enables the recovery of corrupted, deleted, or inaccessible data from a storage device. This software reviews, scans, identifies, extracts, and copies data from deleted, corrupted, and formatted sectors or in a user-defined location within the storage device. Data recovery software is primarily used by IT support staff and service providers. Data recovery software generally has access to the core architecture of a hard disk. It can extract data from corrupt storage devices or deleted files/folders by

referring to and accessing the file structure records/entries. Having access and control over file systems and structure, it can also un-format and repair hard drive partitions.

It can be used for both recovering user-stored and system-created data, files, and folders. It can recover data from virtually any storage device including hard disks, flash drives, external storage cards, tape drives, and more. Most data recovery software can perform data recovery on common file systems. Machine learning, also known as artificial intelligence (AI), is gaining traction when it comes to anything related to technology. AI technology is essential to any organization that regularly deals with large data sets, and high-performing companies are always searching for a way to make operations more efficient. Eventually, AI's ability to streamline the processing of large data sets through machine learning and software as a service (SaaS) will replace traditional data centers. Data recovery tools are software applications designed to retrieve data that has been lost, damaged, or corrupted. These tools can be used to recover data from a variety of sources, such as hard drives, USB drives, memory cards, and mobile devices. Data loss can occur due to a variety of reasons, such as hardware failure, human error, malware, or natural disasters. When data is lost, it can be a major setback for individuals and businesses, leading to lost productivity, revenue, and valuable information. Data recovery tools can help mitigate the effects of data loss by providing a way to recover lost data.

There are two main types of data recovery tools: free and paid. Free data recovery tools are typically limited in their functionality and may not be able to recover all types of data. These tools are often useful for simple data recovery tasks, such as retrieving accidentally deleted files. Paid data recovery tools offer more advanced functionality and can often recover data from more complex situations, such as damaged or corrupted files. These tools are typically more expensive than free tools, but they offer a higher success rate for data recovery. There are many data recovery tools available on the market, each with its own set of features and capabilities.

Here are some of the most commonly used data recovery tools:

1. **EaseUS Data Recovery Wizard:** EaseUS Data Recovery Wizard is a popular data recovery tool that can recover data from a variety of sources, including hard drives, memory cards, and USB drives. It offers both free and paid versions, with the paid version offering more advanced features such as deep scanning and advanced filtering.
2. **Recuva:** Recuva is a free data recovery tool that can recover data from hard drives, USB drives, and memory cards. It offers a user-friendly interface and the ability to preview recovered files before they are restored.
3. **Stellar Data Recovery:** Stellar Data Recovery is a paid data recovery tool that can recover data from a variety of sources, including hard drives, SSDs, and mobile devices. It offers advanced features such as disk imaging, which allows users to create a backup of their hard drives before attempting recovery.
4. **Disk Drill:** Disk Drill is a data recovery tool that can recover data from a variety of sources, including hard drives, USB drives, and memory cards. It offers both free and paid versions, with the paid version offering more advanced features such as data protection and a duplicate finder.
5. **R-Studio:** R-Studio is a paid data recovery tool that can recover data from a variety of sources, including hard drives, SSDs, and RAID systems. It offers advanced features such as disk imaging, remote data recovery, and support for virtual machines.

When using data recovery tools, it is important to follow best practices to maximize the chances of successful data recovery. Here are some tips for using data recovery tools effectively:

1. **Stop using the device:** When data is lost, it is important to stop using the device immediately to avoid overwriting the lost data. Continuing to use the device can cause further damage and decrease the chances of successful data recovery.

2. **Identify the cause of data loss:** Before attempting data recovery, it is important to identify the cause of data loss. This can help determine the best approach for data recovery and prevent further data loss.
3. **Use the appropriate data recovery tool:** Different data recovery tools are designed to handle different types of data loss. It is important to choose a data recovery tool that is appropriate for the specific type of data loss being experienced.
4. **Read user reviews:** Before using a data recovery tool, it is important to read user reviews and check the tool's reputation. This can help identify potential issues and ensure that the tool is effective for the specific type of data loss being experienced.
5. **Back up recovered data:** After data has been successfully recovered, it is important to back up the recovered data to prevent future data loss. This can be done by creating a backup on an external hard drive or cloud storage.
6. **Consider professional data recovery services:** In some cases, data recovery tools may not be able to recover lost data. In these situations, it may be necessary to seek professional data recovery services. These services are typically more expensive than using data recovery tools, but they offer a higher success rate for data recovery.

In conclusion, data recovery tools are essential tools for individuals and businesses to recover lost or damaged data. When using these tools, it is important to follow best practices to maximize the chances of successful data recovery. Whether using a free or paid data recovery tool, it is important to choose a tool that is appropriate for the specific type of data loss being experienced and to back up recovered data to prevent future data loss.

There are seven data analysis tools mentioned below in terms of learning, and performance:-

i. **Tableau Public**

It is a free data visualization application that links to any data source you can think of whether it's a corporate Data Warehouse, Microsoft Excel, or web-based information. It also generates data visualizations, maps, dashboards, and so on, all with real-time changes that are shown on the web. These may also be shared on social media or with your customer, and you can download the files in several formats. However, it truly shines when you have an excellent data source. That's when you realize Tableau's ultimate potential. Tableau's Big Data features make it indispensable. Its approach to data analysis and visualization is considerably better than that of any other data visualization software on the market.

ii. **R Programming**

Well, R is the industry's premier analytics tool, and it's extensively used for statistics and data modeling. It can readily alter data and show it in a variety of formats. It has outperformed SAS in several aspects, including data capacity, performance, and results. R may be compiled and run on a broad range of systems, including Windows, UNIX, and macOS. It offers 11,556 packages and lets you explore them by category. Also, R has tools for installing all packages automatically based on user needs, which may be used with Big Data.

iii. **Python**

It's a scripting language that is simple to understand, write, as well as maintain. Furthermore, it's a free open-source tool. Guido van Rossum developed it in the late 1980s and it supports both structured and functional programming methodologies. Python is simple to learn since it is related to Ruby, JavaScript, and PHP. Python also contains excellent machine-learning packages such as Tensorflow, Theano, Scikitlearn, and Keras. Another useful characteristic of Python is that it can be built on any platform, such as a MongoDB database, SQL browser, or JSON. It also excels at handling text data.

iv. Apache Spark

Apache was created in 2009 by the AMP Lab at the University of California, Berkeley. Apache Spark is a large-scale data processing engine that performs applications hundred times quicker when it comes to memory and 10 times faster on disk in Hadoop clusters. It is based on data science, and its design makes data science simple. Spark is also popular for developing data pipelines and machine learning models. Spark also contains the MLlib package, which provides a progressive collection of machine algorithms for recurring data science procedures like classification, collaborative filtering, regression, clustering, and so on.

v. SAS

SAS is basically a data manipulation programming ecosystem and language that is a market leader in analytics. The SAS Institute created it in 1966, and it was expanded upon in the 1980s as well as the 1990s. It is simple to use and administer, and it can analyze data from any source. In 2011, SAS released a significant collection of solutions for customer intelligence, as well as numerous SAS modules for social media, online, and marketing analytics. These are now often used to profile clients and prospects. It can also forecast their actions and manage and improve communications.

vi. Excel

Excel is a popular, basic, and frequently leveraged analytical tool in practically all industries. Whether you are a Sas, R, or Tableau specialist, you will still need to utilize Excel. When analytics on the client's internal data is required, Excel comes in handy. It analyzes the hard work of summarizing the data with a preview of pivot tables, which aids in filtering the data according to the client's needs. Excel includes a sophisticated business analytics feature that aids in modeling skills. It has prebuilt tools such as automated relationship recognition, DAX measure generation, and time grouping.

vii. RapidMiner

It is an extremely capable comprehensive data analysis tool. It's created by the same house that does predictive analysis as well as other advanced analytics such as machine learning, text analysis, visual analytics, and data mining without the use of programming. RapidMiner supports all data source types, including Microsoft SQL, Excel, Access, Oracle, Teradata, Dbase, IBM SPSS, MySQL, Ingres, IBM DB2, Sybase, and others. This tool is quite powerful, as it can provide analytics based on real-world data transformation settings, allowing you to customize the data sets and formats for predictive analysis.

DATA RECOVERY PROCEDURES AND ETHICS

The digital revolution has made our life a lot easier. Digital technology has evolved to become an indispensable part of our daily lives and businesses. Today, in this digital era, information is more accessible to people than ever before. However, despite plenteous benefits, now corporate entities and individual consumers have started recognizing the risks inherent in digital services. The digital economy collects, combine and share data which has come with challenges like loss of the control to personal privacy, compromised data and other cyberspace crimes. Thus, in the digital age, the risk of unethical or even illegal use of consumers' data without their consent can permanently damage consumers' trust in a brand. Therefore, it has become crucial that the digital economy addresses these issues, along with cyber-security threats. Ethics is the key to preventing these types of internal, as well as, external threats.

Digital Ethics

Digital ethics is defined as the field of study concerned with the manner in which digital technology is shaping our political, social ,and moral well-being. In a broader sense, this field deals with the impact of digital Information and Communication Technologies (ICT) on our societies and the environment.

Role of Digital Ethics in Data Storage

The digital advancements have undoubtedly enhanced the business opportunities for companies by enabling them to compete and thrive. However, they must realize and strive toward transparent operations, ethical practices, and protection of privacy. Similar to any other field of work like medicine or accounts, the IT sector also needs a strict set of codes and ethics. The strict guidelines will ensure more stringent legal requirements. Presently, in IT the codes of ethics are not as standard as in other professional careers which makes it difficult to regulate them. CRM tools and other software are immensely useful to collect volumes of data from clients. This has made it necessary to implement ethical guidelines to decide what to do with that data. Companies need to draw a line between what information is ethical to collect and what violates the privacy of clients.

Data Recovery Ethics

When dealing with a massive amount of data, often companies suffer data loss which needs to get recovered. For the purpose of data recovery, companies mostly depend on third parties. Whenever it happens, the data of clients are put at risk if the third-party data recovery service doesn't take the security of clients' data seriously. To make sure that the files are kept safe and confidential during the recovery process, the ethical data recovery company follows ethical safeguard rules and practices, such as secure servers, reliable access protocols, and smart data return policies, etc. Therefore, with great benefits comes great responsibility. It is necessary to establish guidelines and uphold ethical digital practices to build digital trust while reaping the benefits of this revolutionary technology.

There are some common scenarios where data recovery procedures would be necessary:

1. There has been an operating system failure or some critical operating system files have been damaged, causing the device to not be able to boot up properly. In this case, a simple solution would be to use a Live USB to boot up from another operating system so that you can access the data from the storage medium.
2. There has been a hard disk failure and there is physical damage to the storage medium. In this case, you may be able to repair the hardware, but the storage medium is often beyond repair and the focus is more on a one-time recovery in an attempt to salvage any data you can. This will often require the services of a specialized data recovery company.
3. Files have been deleted from a storage medium. As we will discuss later in the presentation, when an operating system "deletes" files, oftentimes the data is not immediately removed from the drive. This allows tools such as file carvers to recover this data.

There are three research methods of data recovery namely, Direct-to-cloud backup, cloud-to-cloud backup, and SaaS backup. With direct-to-cloud, offsite file backups are copied directly to the cloud, bypassing the need for a local device. Cloud-to-cloud backup is the process of copying data from one cloud to another cloud. SaaS backup refers to backing up data created in SaaS applications such as Microsoft 365 or Google G Suite.

Types of Data Recovery

Not all data loss scenarios are the same, so it is important to choose a backup solution that addresses a wide range of restore and recovery needs and reduces data recovery steps:

- **File Restore:** A file restore is exactly what it sounds like—the process of replacing a lost file or files from a backup to its primary location. With SIRIS, an administrator can mount a recovery point, view the protected system's file structure, locate the necessary files, and restore them back to the primary system. If you only need to retrieve a file or a small number of files, this is the ideal restore type.

- **Volume Restore:** When you perform a volume restore on SIRIS, the contents of the chosen recovery point are shared as an iSCSI target. This restore type retrieves files and folders with permissions intact and is used to restore large numbers of files when a bare metal restore is not necessary (i.e., the physical server is intact and operating correctly).
- **Bare metal restores:** This is the process of restoring an entire system image (the protected machine's data, applications, settings, and operating system) from a backup to a new physical server. "Bare metal" refers to the new system's unused, unconfirmed hardware. Bare metal restore is used when a primary server fails, is damaged, or is otherwise rendered inoperable.
- **Local virtualization:** Local virtualization is a feature of BCDR solutions that offers fast recovery of business operations. Local virtualization uses hypervisor technology to boot a virtual server from a snapshot on the backup device. This enables businesses to continue normal business operations while the primary server is restored (using one of the methods above). Local virtualization nearly eliminates costly business downtime. Sometimes, this functionality is known as Instant Virtualization.
- **Cloud virtualization:** Cloud virtualization refers to the process outlined above, but in the cloud rather than on a local backup device. Some BCDR solutions can create a tertiary cloud copy of backup server images. In the event that both the primary and backup servers are inoperable, say because of a fire or flood, business operations can be continued on the cloud backup server image.

Business continuity/disaster recovery (BCDR)

BCDR solutions are designed to enable fast restores that minimize business downtime. To do so, these solutions use snapshot and virtualization technologies to create and store bootable virtual server images on a backup device or in the cloud. In the event of a primary server failure or other outage, business operations are "failed over" to the backup device or cloud while the primary server is being restored, repaired, or replaced. Once the primary server is back up and running, operations are "failed back" to the primary device. BCDR recovery times are typically measured in minutes rather than the hours or even days required of traditional backup tools. BCDR solutions have become popular with businesses of all sizes, but are probably most beneficial for small to medium businesses (SMBs).

GATHERING EVIDENCE- PRECAUTIONS, PRESERVING AND SAFELY HANDLING ORIGINAL MEDIA FOR ITS ADMISSIBILITY

Digital forensics is a process of identifying, preserving, analyzing, and documenting digital evidence; it helps in presenting evidence in a court of law when required. It is to be systematic in a very specific way, that is:

Identification: It is a process to identify where the attacker has stored the data or evidence.

Preservation: The data and evidence are kept secured and preserved; so that they cannot tamper with.

Analysis: This is the process of recreating fragments of data and drawing conclusions based on the evidence found therein.

Documentation: It includes the creation of records of the data, for the recreation of the crime scene.

Presentation: The last step includes summaries and drawing a conclusion based on the data collected.

When collecting digital evidence investigators should maintain a proper and well-documented chain of custody to ensure any evidence collected does not lose its integrity. Different devices should be handled in a specific manner depending on how data is stored on the device.

If mobile devices must be submitted to a lab they should be turned off in order to preserve the cell tower location. This step not only prevents the phone from being used but also prevents remote destruction commands. The

device should be put in a Faraday bag to prevent network interaction from potentially altering data on the device.

Devices that cannot be turned off can instead be placed on airplane mode or disable any Wi-Fi or Bluetooth capabilities. If it is necessary to keep the device powered on, connect it to an external power source such as a portable battery pack. Devices that are found turned off should be left off and their model number, carrier, and unique identifiers should be documented.

Forensic investigators who encounter computers at a scene should prevent any alteration of evidence during collection. They should first document any activity on the computer, components, or devices by screenshotting and recording any information on the screen. If any destructive software is running on the computer, the power must be immediately disconnected to preserve the evidence.

Investigators that have been appropriately trained can also collect digital evidence at the scene. By using tools that help them identify which electronic devices contain evidence related to their case. The ability to preview digital evidence at the scene can save investigators time and resources. Investigators must make duplicate copies of the content contained on devices to maintain the integrity of the primary source of evidence. The data obtained should not be altered or modified.

The simple reasons for collecting evidence are:

- *Future Prevention*: Without knowing what happened, you have no hope of ever being able to stop someone else from doing it again.
- *Responsibility:* The attacker is responsible for the damage done, and the only way to bring him to justice is with adequate evidence to prove his actions. The victim has a responsibility to the community. Information gathered after a compromise can be examined and used by others to prevent further attacks.

Collection Options

Once a compromise has been detected, you have two options:

- Pull the system off the network and begin collecting evidence: In this case, you may find that you have insufficient evidence or, worse, that the attacker left a dead man's switch which destroys any evidence once the system detects that it's offline.
- Leave it online and attempt to monitor the intruder: you may accidentally alert the intruder while monitoring and cause him to wipe his tracks in any way necessary, destroying evidence as he goes.

Obstacles

- Computer transactions are fast, they can be conducted from anywhere, can be encrypted or anonymous, and have no intrinsic identifying features such as handwriting and signatures to identify those responsibly.
- Any paper trail of computer records they may leave can be easily modified or destroyed or maybe only temporarily.
- Auditing programs may automatically destroy the records left when computer transactions are finished with them.
- Investigating electronic crimes will always be difficult because of the ease of altering the data and the fact that transactions may be done anonymously.
- The best we can do is follow the rules of evidence collection and be as assiduous as possible.

Volatile Evidence

The field of computer Forensics Analysis involves identifying, extracting, documenting, and preserving information that is stored or transmitted in an electronic or magnetic form (that is, digital evidence). The volatile data that is held in temporary storage in the system's memory (including random access memory, cache memory, and the onboard memory of system peripherals such as the video card or NIC) is called volatile data because the memory is dependent on electric power to hold its contents. Volatile data is the data that is usually stored in cache memory or RAM. This volatile data is not permanent this is temporary and this data can be lost if the power is lost i.e., when computer loses its connection. During any cyber crime attack, investigation process is held in this process data collection plays an important role but if the data is volatile then such type of data should be collected immediately. Volatile information can be collected remotely or onsite. If there are many numbers of systems to be collected then remotely is preferred rather than onsite. It is very important for the forensic investigation that immediate state of the computer is recorded so that the data does not lose as the volatile data will be lost quickly. If the volatile data is lost on the suspects computer if the power is shut down, Volatile information is not crucial but it leads to the investigation for the future purpose. To avoid this problem of storing volatile data on a computer we need to charge continuously so that the data isn't lost. So that computer doesn't lose data and forensic expert can check this data sometimes cache contains Web mail. This volatile data may contain crucial information.so this data is to be collected as soon as possible. This process is known "Live Forensics".

The final step in evidence assessment specifically deals with the evidence itself. You should identify the stability of the evidence, and collect the most volatile evidence first before moving to nonvolatile evidence. In doing so, you should prioritize the collection and acquisition of evidence so that the evidence that is most likely to contain what you're searching for is examined first. There are different examples of an order of volatility like Registers and cache, Routing tables, Arp cache, Process table, Kernel statistics and modules, Main memory, Temporary file systems, Secondary memory, Router configuration & Network topology.

Methods of Collection

There are two basic forms of collection: freezing the scene and honey potting.

Freezing the Scene

- It involves taking a snapshot of the system in its compromised state. You should then start to collect whatever data is important onto removable non-volatile media in a standard format.
- All data collected should have a cryptographic message digest created, and those digests should be compared to the originals for verification.

Honey Potting

- It is the process of creating a replica system and luring the attacker into it for further monitoring.
- The placement of misleading information and the attacker's response to it is a good method for determining the attacker's motives.

Artifacts

- There is almost always something left behind by the attacker be it code fragments, trojan programs, running processes, or sniffer log files. These are known as artifacts.
- Never attempt to analyze an artifact on the compromised system.
- Artifacts are capable of anything, and we want to make sure their effects are controlled.

Collection Steps

1. Find the Evidence: Use a checklist. Not only does it help you to collect evidence, but it also can be used to double-check that everything you are looking for is there.
2. Find the Relevant Data: Once you've found the evidence, you must figure out what part of it is relevant to the case.
3. Create an Order of Volatility: The order of volatility for your system is a good guide and ensures that you minimize the loss of uncorrupted evidence.
4. Remove external avenues of change: It is essential that you avoid alterations to the original data.
5. Collect the Evidence: Collect the evidence using the appropriate tools for the job.
6. Document everything: Collection procedures may be questioned later, so it is important that you document everything you do. Timestamps, digital signatures, and signed statements are all important.

Controlling Contamination: The Chain of Custody

Once the data has been collected, it must be protected from contamination. Originals should never be used in the forensic examination; verified duplicates should be used. A good way of ensuring that data remains uncorrupted is to keep a chain of custody. This is a detailed list of what was done with the original copies once they were collected. Analysis

- Once the data has been successfully collected, it must be analyzed to extract the evidence you wish to present and to rebuild what actually happened.

Time

- To reconstruct the events that led to your system being corrupted, you must be able to create a timeline.
- Never, ever change the clock on an affected system. Forensic Analysis of Back-ups When we analyze back-ups, it is best to have a dedicated host for the job. We need a dedicated host which is secure, clean, and isolated from any network for analyzing backups.
- Document everything you do. Ensure that what you do is repeatable and capable of always giving the same results.

Reconstructing the Attack

After collecting the data, we can attempt to reconstruct the chain of events leading to and following the attacker's break-in. We must correlate all the evidence we have gathered. Include all of the evidence we've found when reconstructing the attack---no matter how small it is.

Searching and Seizing

There is no one methodology for performing a computer forensic investigation and analysis. There are too many variables for it to be just one way. Some of the typical variables that come to mind include operating systems; software applications; cryptographic algorithms and applications; and hardware platforms. But moving beyond these obvious variables spring other equally challenging variables: law, international boundaries, publicity, and methodology.

There are a few widely accepted guidelines for computer forensic analysis:

- A computer forensic examiner is impartial. Our job is to analyze the media and report our findings with no presumption of guilt or innocence.

- The media used in computer forensic examinations must be sterilized before each use.
- A true image (bit stream) of the original media must be made and used for the analysis.
- The integrity of the original media must be maintained throughout the entire investigation.

Before the Investigation

- For the sake of the first argument, you must have skilled technicians in-house and a top-notch lab with the right equipment, the right computer forensic tools, and so on.
- District attorneys may require more documentation on the chain of evidence handling.
- When you have a case arise, you know what is required and can work the case from the inception in support of these requirements.

Methodology Development

- Define your methodology, and work according to this methodology.
- Here methodology defines a method, a set of rules: guidelines that are employed by a discipline.

The chain of evidence is so important in computer forensic investigations. If resources allow, have two computer forensic personnel assigned to each case every step of the way. Important in the documentation are the times that dates steps were taken; the names of those involved; and under whose authority were the steps taken.

Evidence Search and Seizure

Prior to search and seizure, you already have the proper documents filled as well as permission from the authority to search and seize the suspect's machine.

Step 1: Preparation

You should check all media that is to be used in the examination process. Document the wiping and scanning process. Check to make sure that all computer forensic tools are licensed for use and that all lab equipment is in working order.

Step 2: Snapshot

We should photograph the scene, whether it is a room in a home or in a business. You should also note the scene. Take advantage of your investigative skills here. Note pictures, personal items, and the like. Photograph the actual Evidence. For example, the evidence is a PC in a home office. Take a photograph of the monitor. Remove the case cover carefully and photograph the internals.

Step 3: Transport

If you have the legal authority to transport the evidence to your lab, you should pack the evidence securely. Photograph/videotape and document the handling of evidence leaving the scene to the transport vehicle and from the transport vehicle to the lab examination facility.

Step 4: Examination

You should prepare the acquired evidence for examination in your lab. There are many options to on what tool to use to image the drive. You could use EnCase, the Unix command DD, ByetBack, or also SafeBack. It is wise to have a variety of tools in your lab. Each of these tools has its respective strengths. The important note to remember here is: Turn off virus-scanning software. We must record the time and date of the COMS. Do not boot the suspect machine. When making the image, make sure that the tool you use does not access the file system of the target evidence media. After making the image, seal the original media in an electrostatic-safe container, catalogue it, and initial the container. Finally, the examination of the acquired image begins.

DUPLICATION AND PRESERVATION OF DIGITAL EVIDENCE

Preserving the Digital Crime Scene

- After securing the computer, we should make a complete bitstream backup of all computer data before it is reviewed or processed.
- Bit stream backups are much more thorough than standard backups.
- They involve copying of every bit of data on a storage device, and it is recommended that two such copies be made of the original when hard disk drives are involved.
- Any processing should be performed on one of the backup copies.
- IMDUMP was the first software for taking bit stream back-ups developed by Michael White.

SafeBack

- SafeBack has become a law enforcement standard and is used by numerous government intelligence agencies, military agencies, and law enforcement agencies worldwide.
- SafeBack program copies and preserves all data contained on the hard disk. Even it goes so far as to circumvent attempts made to hide data in bad clusters and even sectors with invalid CRCs.

SnapBack

- Another bit steam backup program, called SnapBack, is also available and is used by some law enforcement agencies primarily because of its ease of use.
- Its prices are several hundred dollars higher than SafeBack.
- It has error-checking built into every phase of the evidence backup and restoration process.
- The hard disk drive should be imaged using specialized bit stream backup software.
- The floppy diskettes can be imaged using the standard DOS DISKCOPY program.
- When DOS DISKCOPY is used, it is recommended that the MS-DOS Version 6.22 be used and the (data verification) switch should be invoked from the command line.
- Know and practice using all of your forensic software tools before you use them in the processing of computer evidence.
- We may only get one chance to do it right.

Computer Evidence Processing Steps

There really are no strict rules that must be followed regarding the processing of computer evidence. The following are general computer evidence processing steps:

1. **Shut down the computer**

   Depending on the computer operating system, this usually involves pulling the plug or shutting down a network computer using relevant commands required by the network involved. Generally, time is of the essence, and the computer system should be shut down as quickly as possible.

2. **Document the hardware configuration of the system**

   Before dismantling the computer, it is important that pictures are taken of the computer from all angles

to document the system hardware components and how they are connected. Labeling each wire is also important so that it can easily be reconnected when the system configuration is restored to its original condition at a secure location.

3. Transport the computer system to a secure location. A seized computer left unattended can easily be compromised. Don't leave the computer unattended unless it is locked up in a secure location.
4. Make bit stream backups of hard disks and floppy disks. All evidence processing should be done on a restored copy of the bit stream backup rather than on the original computer. Bit stream backups are much like an insurance policy and are essential for any serious computer evidence processing.
5. Mathematically authenticate data on all storage devices. You want to be able to prove that you did not alter any of the evidence after the computer came into your possession. Since 1989, law enforcement and military agencies have used a32- bit mathematical process to do the authentication process.
6. Document the system date and time: If the system clock is one hour slow because of daylight-saving time, then file timestamps will also reflect the wrong time. To adjust for these inaccuracies, documenting the system date and time settings at the time the computer is taken into evidence is essential.
7. Make a list of key search words. It is all but impossible for a computer specialist to manually view and evaluate every file on a computer hard disk drive. Gathering information from individuals familiar with the case to help compile a list of relevant keywords is important. Such keywords can be used in the search of all computer hard disk drives and floppy diskettes using automated software.
8. Evaluate the Windows swap file. The Windows swap file is a potentially valuable source of evidence and leads. When the computer is turned off, the swap file is erased. But the content of the swap file can easily be captured and evaluated.
9. Evaluate file slack: It is a source of significant security leakage and consists of raw memory dumps that occur during the work session as files are closed. File slack should be evaluated for relevant keywords to supplement the keywords identified in the previous steps. File slack is typically a good source of Internet leads. Tests suggest that file slack provides approximately 80 times more Internet leads than the Windows swap file.
10. Evaluate unallocated space (erased files). Unallocated space should be evaluated for relevant keywords to supplement the keywords identified in the previous steps.
11. Search files, file slack, and unallocated space for keywords. The list of relevant keywords identified in the previous steps should be used to search all relevant computer hard disk drives and floppy diskettes. It is important to review the output of the text search utility and equally important to document relevant findings.
12. Document file names, dates, and times. From an evidence standpoint, file names, creation dates, and last modified dates and times can be relevant. The output should be in the form of a word-processing-compatible file that can be used to help document computer evidence issues tied to specific files.
13. Identify file, program, and storage anomalies. Encrypted, compressed, and graphic files store data in binary format. As a result, text data stored in these file formats cannot be identified by a text search program. Manual evaluation of these files is required. Depending on the type of file involved, the contents should be viewed and evaluated for their potential as evidence.
14. Evaluate program functionality. Depending on the application software involved, running programs to learn their purpose may be necessary. When destructive processes that are tied to relevant evidence are discovered, this can be used to prove wilfulness.

15. Document your findings. It is important to document your findings as issues are identified and as evidence is found. Documenting all of the software used in your forensic evaluation of the evidence, including the version numbers of the programs used, is also important. Be sure you are legally licensed to USE the forensic software. Screen prints of the operating software also help document the version of the software and how it was used to find or process the evidence.

16. Retain copies of the software used. As part of your documentation process, it is recommended that a copy of the software used to be included with the output of the forensic tool involved. Duplication of results can be difficult or impossible to achieve if the software has been upgraded and the original version used was not retained.

LEGAL ASPECTS OF COLLECTING AND PRESERVING COMPUTER

Forensic Evidence

Definition

- A chain of custody is a roadmap that shows how evidence was collected, analyzed, and preserved in order to be presented as evidence in court.
- Preserving a chain of custody for electronic evidence requires proving that:
  - No information has been added or changed.
  - A complete copy was made.
  - A reliable copying process was used.
  - All media was secured.

Legal Requirements

- When evidence is collected, certain legal requirements must be met. These legal requirements are vast, and complex, and vary from country to country.
- CERT Advisory CA-1992-19 suggests the following text be tailored to a corporation's specific needs under the guidance of legal counsel:
- This system is for the use of authorized users only. Individuals using this computer system without authority, or in excess of their authority, are subject to having all of their activities on this system monitored and recorded by system personnel.
- In the course of monitoring individuals improperly using this system, or in the course of system maintenance, the activities of authorized users may also be monitored.
- Anyone using this system expressly consents to such monitoring and is advised that if such monitoring reveals possible evidence of criminal activity, system personnel may provide the evidence of such monitoring to law enforcement officials.
- The legality of workplace monitoring depends primarily on whether employment policies exist that authorize monitoring and whether that policy has been clearly communicated to employees.
- To prove that the policy has been communicated, employees should sign a statement indicating that they have read, understood, and agreed to comply with corporate policy and consent to system monitoring.

Evidence Collection Procedure

When the time arrives to begin collecting evidence, the first rule that must be followed is Do not rush.

- The investigation team will need a copy of their incident-handling procedure, an evidence collection notebook, and evidence identification tags.
- They may also need to bring tools to produce reliable copies of electronic evidence, including media to use in the copying process.
- In some cases, legal counsel will want photographs of the system prior to search and seizure. Then include a Polaroid camera in the list of tools.

The Incident Coordinator

Policy and procedure should indicate who is to act as incident coordinator.

The Incident Coordinator

- will contact the other members of the response team as outlined in the Incident Response Policy, when an incident is reported.
- will be responsible for ensuring that every detail of the incident-handling procedure is followed, upon arrival at the incident site.
- will assign team members the various tasks outlined in the incident-handling procedure.
- serve as the liaison to the legal team, law enforcement officials, management, and public relations personnel.

Ultimate responsibility for ensuring that evidence is properly collected and preserved and that the chain of custody is properly maintained, belongs to the incident coordinator.

The Evidence Notebook

- One team member will be assigned the task of maintaining the evidence notebook.
- This person will record who, what, where, when, and how of the investigation process. At a minimum, items to be recorded in the notebook include the following task.
    a) Who initially reported the suspected incident along with the time, date, and circumstances surrounding the suspected incident?
    b) Details of the initial assessment leading to the formal investigation.
    c) Names of all persons conducting the investigation.
    d) The case number of the incident.
    e) Reasons for the investigation.
    f) A list of all computer systems included in the investigation, along with complete system specifications. Also include identification tag numbers assigned to the systems or individual parts of the system.
    g) Network diagrams.
    h) Applications running on the computer systems previously listed.
    i) A copy of the policy or policies that relate to accessing and using the systems previously listed.

j) A list of administrators responsible for the routine maintenance of the system.

k) A detailed list of steps used in collecting and analyzing evidence. Specifically, this list needs to identify the date and time each task was performed, a description of the task, who performed the task, where the task was performed, and the results of the analysis.

l) An access control list of who had access to the collected evidence at what date and time.

- A separate notebook should be used for each investigation. It should be bound in such a way that it is obvious if a page or pages have been removed.
- This notebook is a crucial element in maintaining the chain of custody. Therefore, it must be as detailed as possible to assist in maintaining this chain.

Evidence Collection

- Another team member (or members) will be assigned the task of evidence collection.
- To avoid confusion, the number of people assigned to this task should be kept to a minimum.
- This member (or members) should also be highly proficient with copying and analysis tools.
- This person will tag all evidence and work with the person responsible for the evidence notebook to ensure that this information is properly recorded.
- Next, the person will also be responsible for making a reliable copy of all data to be used as evidence.
- The data will include complete copies of drives on compromised or suspect systems, as well as all relevant log files.
- This can be done on-site or the entire system can be moved to a forensics lab, as needs dictate.
- A binary copy of the data is the proper way to preserve evidence.
- A reliable copy process has three critical characteristics.
- The process must meet industry standards for quality and reliability.
- The copies must be capable of independent verification.
- The copies must be tamperproof.
- Once all evidence is collected and logged, it can be securely transported to the forensics lab.
- A detailed description of how data was transported and who was responsible for the transport, along with the date, time, and route, should be included in the log. Storage and Analysis of Data.
- The lab must provide some form of access control; a log should be kept detailing the entrance and exit times of all individuals.
- It is important that evidence never be left in an unsecured area.
- If a defense lawyer can show that unauthorized persons had access to the evidence, it could easily be declared inadmissible.

As analysis of evidence is performed, investigators must log the details of their actions in the evidence notebook. The following should be included at a minimum:

- The date and time of analysis.
- Tools used in performing the analysis.

- Detailed methodology of the analysis.
- Results of the analysis.
- Finally, once all evidence has been analyzed and all results have been recorded in the evidence notebook, a copy of the notebook should be made and given to the legal team.
- If the legal team finds that sufficient evidence exists to take legal action, it will be important to maintain the chain of custody until the evidence is handed over to the proper legal authorities.
- Legal officials should provide a receipt detailing all of the items received for entry into evidence.

COMPUTER IMAGE VERIFICATION AND AUTHENTICATION

Special Needs of Evidential Authentication

- During an investigation, it is decided that evidence may reside on a computer system.
- It may be possible to seize or impound the computer system, but this risks violating the basic principle of innocent until proven guilty, by depriving an innocent party of the use of his or her system.
- It should be perfectly possible to copy all the information from the computer system ina manner that leaves the original system untouched and yet makes all contents available for forensic analysis.
- The courts may rightly insist that the copied evidence is protected from either accidental or deliberate modification and that the investigating authority should prove that this has been done. Thus, it is not the content that needs protection, but its integrity.
- A secure method of determining that the data has not been altered by even a single bit since the copy was taken.
- A secure method of determining that the copy is genuinely the one taken at the time and on the computer in question.
- These elements are collectively referred to as the Digital Image Verification and Authentication Protocol.

DIGITAL IDS AND AUTHENTICATION TECHNOLOGY

- Without an assurance of the software's integrity, and without knowing who published the software, it's difficult for customers to know how much to trust the software. It's difficult to make the choice of downloading the software from the Internet.
- For example (when using Microsoft Authenticode coupled with Digital IDs™ fromVeriSign®), through the use of digital signatures, software developers are able to include information about themselves and their code with their programs.
- When customers download software signed with Authenticode and verified byVeriSign, they should be assured of the content source, indicating that the software really comes from the publisher who signed it, and content integrity, indicating that the software has not been altered or corrupted since it was signed.

Authenticode

- Microsoft Authenticode allows developers to include information about themselves and their code with their programs through the use of digital signatures.
- Through Authenticode, the user is informed

- Of the true identity of the publisher
- Of a place to find out more about the control
- The authenticity of the preceding information
- Users can choose to trust all subsequent downloads of software from the same publisher and all software published by commercial publishers that have been verified by VeriSign.

Public Key Cryptography

- In public key cryptographic systems, every entity has two complementary keys (a public key and a private key) that function only when they are held together.
- Public keys are widely distributed to users, whereas private keys are kept safe and only used by their owners.
- Any code digitally signed with the publisher's private key can only be successfully verified using the complementary public key.
- Code that was successfully verified using the publisher's public key, could only have been digitally signed using the publisher's private key and has not been tampered with.

Certificate Authorities (CA)

- Certification Authorities such as VeriSign are organizations that issue digital certificates to applicants whose identity, they are willing to vouch for. Each certificate is linked to the certificate of the CA that signed it.

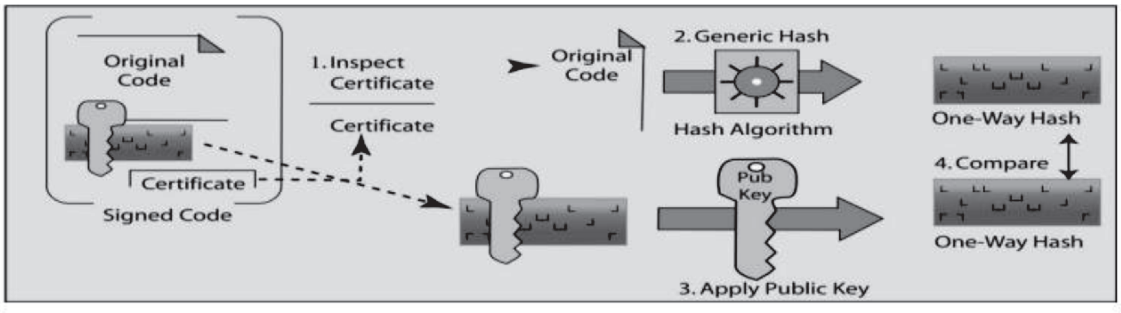
VeriSign has the following responsibilities:

1. Publishing the criteria for granting, revoking, and managing certificates;
2. Granting certificates to applications who meet the published criteria;
3. Managing certificates;
4. Storing VeriSign's root keys in an exceptionally secure manner;
5. Verifying evidence submitted by applicants;
6. Providing tools for enrolment;
7. Accepting the liability associated with these responsibilities;
8. Time-stamping digital signatures.

Digital ID

- A Digital ID/Certificate is a form of electronic credentials for the Internet.
- A Digital ID is issued by a trusted third party to establish the identity of the ID holder.
- The third party who issues certificates is known as a Certificate Authority (CA).
- Digital ID technology is based on the theory of public key cryptography.
- The purpose of a Digital ID is to reliably link a public/private key pair with its owner.
- When a CA such as VeriSign issues a Digital ID, it verifies that the owner is not claiming a false identity.
- When a CA issues you a digital certificate, it puts its name behind the statement that you are the rightful owner of your public/private key pair.

How Authenticode works with VeriSign Digital IDs?



Authenticode: Verisign Digital ID Process

1. Publisher obtains a Software Developer Digital ID from VeriSign.
2. Publisher creates code.
3. Using the SIGNCODE.EXE utility, the publisher.
4. The end user encounters the package.
5. The end user's browser examines the publisher's Digital ID. Using the VeriSign root Public Key, which is already embedded in Authenticode-enabled applications, the end user browser verifies the authenticity of Software Developer Digital ID (which is itself signed by the VeriSign root Private Key).
6. Using the publisher's public key contained within the publisher's Digital ID, the end user browser decrypts the signed hash.
7. The end browser runs the code through the same hashing algorithm as the publisher, creating a new hash.
8. The end user browser compares the two hashes. If they are identical, the browser messages that the content has been verified by VeriSign, and the end user has the confidence that the code was signed by the publisher identified in the Digital ID, and the code hasn't been altered since it was signed.

Time Stamping: Because key pairs are based on mathematical relationships that can theoretically be "cracked" with a great deal of time and effort, it is a well-established security principle that digital certificates should expire.

DOCUMENT A CHAIN OF CUSTODY AND ITS IMPORTANCE

Evidence Collection

The Investigator needs to make sure that evidence must be collected in a systematic and careful manner. The process of evidence collection begins with the preliminary crime scene survey/walk-through, followed by a determination of the evidence collection sequence to be used. There are various methods that can be adopted for evidence collection based on the type of crime scene. The evidence collection sequence may be based on the following information:

- *The scene location:* whether the crime has occurred inside premises or within a vehicle or it is an exterior one.
- *The condition of the evidence:* the condition of evidence (Whether the evidence is fragile or stable) plays an important role in choosing which evidence collection method is to be used.

- Weather conditions which might affect the scene or evidence within.
- Scene management considerations that may alter or contaminate the evidence.

Investigators should use the appropriate equipment when collecting evidence. Equipment that is required for the collection of evidence must be sterile so as to avoid contamination of evidence. Various equipment is used to collect evidence. A few of them are named below:

1. Latex gloves/nitrile gloves (N-DEX, non-latex): - helps in preventing contamination as well as any kind of hazardous exposure to the hands of personnel collecting evidence.
2. Forceps- Forceps and similar tools may have to be used to pick up small items.
3. Tweezers
4. Scalpels
5. Swabs
6. Paper bags
7. Plastic bags
8. Cardboard boxes
9. Wrapping paper
10. Hand tools
11. Thermometer

Evidence Marking and Packaging

Evidence collected from the scene of a crime or received during the investigation of crime scene should be cataloged and packaged before leaving the scene to prevent loss or cross-contamination. Mark the item of evidence when possible. Evidence which cannot be marked, such as soil, hair, and stains, should be placed in an appropriate container or envelope. An important point that is to be kept in mind is that the evidence marked directly might result in interference with the forensic analysis and hence marking should always be done on the outer packaging. When marking evidence directly, include the following:

- Case number
- Item number
- Date recovered or received
- Investigator's initials

Evidence that has been inventoried, marked, and prepared for submittal (or to be returned to the investigating agency) is packaged in an appropriate container and labeled per agency protocol. A trained investigator or evidence collector arrives at the crime scene with all types of packaging materials and tools ready to encounter any type of situation. In order to prevent any change in evidence, the evidence must be packaged carefully. The type of packaging depends on the type of evidence. The evidence must be properly packaged, properly labeled, and sealed with appropriate initials to maintain the chain of custody. The evidence must be packaged in its original condition as it is found at the crime scene. The objects with the trace evidence must be sent as a whole unless it is not possible to transport the whole item such as a wall. As sometimes it takes a long time for a crime lab to process the evidence so it is necessary that the evidence must be packaged in such a manner that the conditions such as evaporation, breakage, etc. should not change its condition. While packaging the chances of cross-contamination must be ended. Each item must be packaged in a separate container. Every

package must be labeled with all the essential details such as Case FIR No., Item No., Type of Evidence (fragile/ stable), etc. After labeling the package must be sealed with evidence tape. Take the entire piece of evidence as it is found on the crime scene, if possible. New and unused packaging materials should be used. Evidence must be sealed using proper methods which prevent tampering. For powders such as drugs or others, ordinary mailing envelopes should not be used because powders will leak out of their corners.

1. Unbreakable Plastic pill bottles with pressure lids or in Manila envelopes, screwcap glass vials, or cardboard pillboxes- Used to store trace evidence such as hair glass fiber, etc.
2. Paper bags and boxes- Used to package larger and/or heavier pieces of evidence.
3. Clean Paint Cans- Used to store Arson evidence.
4. Paper bags or Manila envelopes- Used to store Blood-stained materials/clothing after air drying.
5. Air-tight containers- used to store.

CHAIN OF CUSTODY

After careful collection of the evidence, the next step of the investigator is to submit evidence in the laboratory for examination. In the whole process, the maintenance of the chain of custody is very important. The transfer of property or evidence from a crime scene investigator to any other individual, agency, or location is documented by having a chain of custody. The list of information that is to be included in the chain of custody:

- List of evidence: the item number and a brief description.
- All transfers must include the date and time of the transfer.
- The signature of the individual releasing the evidence to another individual or location.
- The signature of the individual transporting the evidence.
- The signature of the individual receiving the evidence from another individual or location.
- Reason for the transfer as needed.

After all the collected evidence have been packaged properly they should be properly labeled. After labeling the next step is to transport all the packed evidence to the crime lab for forensic analysis or for further evaluation. The chain of custody is a tracking document beginning with detailed scene notes that document where the evidence was received from or collected. The chain of custody is initially established when an investigator takes custody of evidence at a crime scene, or when evidence is received from an officer or detective at, or from, the crime scene. In order to maintain all items, a complete and correct chain of custody must be maintained for all items.

Notes should be prepared which comprise of documentation of recovery location, the date and time of recovery, and also the description of items, condition, and whether any unusual markings or alterations to the item were present during recovery. This is not necessary that the evidence collector only will transport the evidence to the laboratory. Often some other officer transports the evidence to the lab. That's why maintenance of chain of custody log must be maintained indicating the transfer of custody to and from every individual who is involved in transporting or storing the evidence until it gets to the crime lab.

These include:

a) The collecting officer (who collects the evidence from the crime scene),

b) The transportation officer (who transports the collected & packaged evidence from the Crime scene to the laboratory),

c) Any evidence storage officer if the evidence is stored prior to taking it to the lab,

d) Any further transportation officer,

e) Anyone who gets into the evidence for any reason,

f) The laboratory evidence collection person(s),

g) Any other person involved in the whole process,

h) Send all evidence (to the crime lab) by registered or certified mail, return receipt requested, to maintain the chain of custody.

Transfer of Evidence to Property Room

On many occasions, the agencies transfer the evidence to a property room aforementioned to its submission in a crime lab. Property room documentation or secure electronic transfer is used when the investigator submits evidence to the property room. The associated information may include the following:

- Agency case number;
- Type of evidence;
- Officer responsible for the investigation: the name, rank, and identification number of the officer for whom the evidence was recovered. The official laboratory report is addressed to this officer;
- Transporting officer: the name, rank, identification number, and assignment of the investigator;
- Signature or another identifier of responsible officer and date prepared; the date the evidence is submitted to the property room;
- Comment: the address where the incident was located, or where the evidence was recovered.

The list of the evidence/property may include:

- Number each evidence item sequentially;
- Quantity of items included, e.g., 10 spent shell casings;
- Sera ial number of the item, e.g., VCR, handgun;
- Item description;
- Status: e.g., submit for analysis, Hold, or RTC (releasable, return to claimant or owner).

A file system in a computer is the manner in which files are named and logically placed for storage and retrieval. It can be considered as a database or index that contains the physical location of every single piece of data on the respective storage device, such as a hard disk, CD, DVD, or flash drive. This data is organized in folders, which are called directories. These directories further contain folders and files.

For storing and retrieving files, file systems make use of metadata, which includes the date the file was created, date modified, file size, and so on. They can also restrict users from accessing a particular file by using encryption or a password.

Files are stored on a storage media in “sectors”. Unused sectors can be utilized for storing data, typically done in sector groups known as blocks. The file system identifies the file size and position and the sectors that are available for storage. If a structure for organizing files wouldn’t exist, it would not be possible to delete or retrieve files, or to keep two files with the same name since all the files would exist in the same folder. For example, it is because of folders that we are able to name two different image files with the same name, as both exist in two different folders. But if two files are in the same directory, they cannot have the same name.

Most of the applications need a file system to work, hence every partition needs to have one. Programs are also dependent on file systems, which means that if a program is built to be used in Mac OS, it will not run on Windows.

Some commonly used file systems

FAT File System

FAT or File Allocation Table is a file system used by operating systems for locating files on a disk. Due to fragmentation, files may be scattered around and divided into sections. The FAT system keeps a track of all parts of the file. FAT has existed as a file system since the advent of personal computers.

Features

- File Name
  - FAT system in MS-DOS allows file names of 8 characters only
  - FAT file system in Windows supports long file names, with full file path being as long as 255 characters
  - File name should start with alphanumeric characters
  - File names can have any character except "/ = [],? ^""
  - File names can have more than one period and spaces. Characters that come after the last period in the full file name are considered as the file extension.
- FAT file system does not support folder and local security. This means users logged into a computer locally will gain complete access to folders and files that lie in FAT partitions.
- It provides fast access to files. The rate depends upon the size of the partition, file size, type of file, and number of files in the folder.

FAT 32 File System

This is an advanced version of the FAT File system and can be used on drives ranging from 512 MB to 2 TB.

Features

- It is more storage-efficient and supports up to 2TB of size
- Provides a better usage of disk space
- Easier access of files in partitions less than 500 MB or greater than 2GB in size

The figure below shows the partitioning layout in FAT and FAT 32 file systems:

FAT File System

FAT 32 File System

NTFS File System

The NTFS File System stands for New Technology File System.

Features

- Naming
  - File name can be as long as 255 characters
  - File names can have any character other than / " :*
  - They are not case sensitive
- It provides folder and file security. This is done by passing on NTFS permission to files and folders. Security works at the local as well as network levels. Every file and folder in the list has an Access Control List that includes the users, security identifiers, and the access privileges that are granted to the users.
- Files and partition sizes are larger in NTFS than those in FAT. An NTFS partition can be of a size as large as 16 Exabytes, but practically it is limited to 2TB. File size can range from 4 GB to 64 GB.
- It provides up to 50% file compression.
- It is a reliable and recoverable file system that makes use of transaction logs for updating files and folders automatically.
- It provides bad-cluster mapping. This means that it can detect bad clusters or erroneous space in the disk, retrieve the data in those clusters, and then store it in another space. To avoid further data storage in those areas, bad clusters are marked for errors.

NTFS File System

EXT File Systems

Extended file system (EXT), Second Extended file system (EXT2), and Third Extended file system (EXT3) are designed and implemented on Linux. The EXT is an old file system that was used in pioneer Linux systems. EXT2 is probably one of the most widely used Linux file systems. EXT3 also includes the same features as EXT2 but also includes journaling. Here we will talk about the most commonly used EXT2. With the optimizations in kernel code, it provides robustness along with good performance whilst providing standard and advanced Unix file features.

Features

- Supports standard file types in Unix i.e. regular files, device special files, directories, symbolic links
- Can manage file systems created on huge partitions. Originally, file system size was restricted to 2 GB, but with recent work in the VFS layer, this limit has now increased to 4 TB.
- Reserves about 5 percent of blocks for administrator usage, thus allowing the admins to recover from situations of overfilled processes.
- Allows for secure deletion of files. Once data is deleted, the space is overwritten with random data to prevent malicious users from gaining access to the previous data.

What is a file format?

A file format is a layout and organization of data within the file. If a file is to be used by a program, it must be able to recognize and have access to the data in the file. For instance, a text document can be recognized by a program such as Microsoft that is designed to run text files but not by a program that is designed to run audio or video files.

A file format is indicated along with the file name in the form of a file extension. The extension contains three or four letters identifying the format and is separated from the file name by a period.

Steps in the file system forensics process

Carrying out a forensic analysis of file systems is a tedious task and requires expertise every step of the way. Following are the steps that can help analyze a file system for data that may provide evidence in a forensic investigation.

Acquisition

The system should be secured to ensure that all data and equipment stay safe. In other words, all media required for forensic analysis should be acquired and kept safe from any unauthorized access. Find out all files on the computer system including encrypted, password-protected, hidden, and deleted (but not overwritten) files. These files must be acquired from all storage media that include hard drives and portable media. Once acquired, forensic investigators have to make a copy of them so that the original files are kept intact without the risk of alteration.

This can be done in four ways:

- *Disk-to-Image*: This is the most common method as it provides more flexibility and allows to creation of multiple copies.
- *Disk-to-Disk*: Used where disk-to-image is not possible.
- *Logical*: it captures only the files that are of interest to the case. Used when time is limited.
- *Sparse*: It gathers fragments of deleted or unallocated data.

Validation and discrimination

Before you analyze an image, you need to validate it to ensure the integrity of the data.

Hashing algorithms help forensic investigators determine whether a forensic image is an exact copy of the original volume or disk. This validates the integrity of evidence and conforms to its admissibility in the court.

Extraction

Data extraction, which involves the retrieving of unstructured or deleted data and needs to be processed for forensic investigation. Many computer users think that a file, once deleted, will disappear forever from the hard disk. However, this is not true. Deleting files only removes it from the disc contents table. In FAT systems it is called the File Allocation Table, while in NTFS it is called the Master File Table. Data is stored in clusters on the hard disc and consists of a certain number of bits. Parts of files are mostly scattered throughout the disc, and deleting the files makes it difficult to reconstruct them, but not impossible. With increased disk capacity, it now takes longer for all fragments of a file to be overwritten.

In many cases, the criminals may have hidden the data that can turn out to be useful for forensic investigation. Criminals with basic technical knowledge have many options available for hiding data such as disk editor, encryption, steganography, and so on. Recovering and reconstructing this data can be time-consuming, but generally, it produces fruitful evidence.

Extracting data from unallocated space is file carving. It is a helpful technique in digital forensics that finds deleted or hidden files from the media. A hidden file can lie in any area such as slack space, unallocated clusters, or lost clusters of the digital media or disk. For using file carving, a file should have a header that can be located by performing a search that continues till the file footer is located. Data that lies between these two points is extracted and then analyzed for file validation.

Reconstruction

Extracted data can be reconstructed using a variety of available software tools that are based on various reconstruction algorithms such as bottom-up tree reconstruction and inference of partition geometry. Reconstructed data is thoroughly analyzed for further evidence and put forth in the form of a report.

Reporting

In order to keep a track record of every step of the investigation, document every procedural step. Evidence presented without proper documentation may not be admissible in court. This documentation should not only include the recovered files and data but also the physical layout of the system along with any encrypted or reconstructed data. Forensic analysis of time-based metadata can help investigators correlate distinct information quickly and to find notable times and dates of activities related to improper computer usage, spoliation, and misappropriation.

Recovery of Internet Usage Data

Data recovery can be defined as a process of obtaining the information located on a storage device that cannot be accessed by standard means due to its previous deletion or certain damage to the digital medium. Different approaches are used to regain the missing files, yet, only on the condition that their content *is present somewhere within the storage*. For instance, data recovery doesn't cover the situations when a file has never been written to persistent storage, like documents that were created but could not be eventually saved to the hard disk drive due to a power failure. Also, none of the existing restore methods can cope with the cases of permanent erasure which occurs when some other information occupies its storage space – under such circumstances, the lost files can only be retrieved from an external backup.

In general, data recovery techniques are divided into two types: software-based and ones involving the repair or replacement of damaged hardware components in a laboratory setting. A software-based approach is employed in the majority of cases and involves the use of specialized utilities able to interpret the logical structure of the problem storage, read out the required data, and deliver it to the user in a usable form for further copying. Physical repairs are conducted by specialists in the most severe instances, for example, when some mechanical or electrical parts of the drive no longer work properly – in this case, all the measures are directed towards a one-time extraction of the critical content, without the possibility of continued usage of the affected device.

The information remaining on an intact storage can usually be recovered without professional help by means of data-specialized software. However, it is important to keep in mind that *no information is recoverable after being overwritten*. For this reason, nothing should be written to the storage until the last file from it is rescued.

Most data recovery utilities operate using the algorithms of metadata analysis, the method of raw recovery based on the known content of files, or a combination of the two approaches.

Metadata is hidden service information contained within the file system. Its analysis allows the software to locate the principal structures on the storage that keep a record of the placement of files' content, their properties, and directory hierarchy. After that, this information is processed and used to restore the damaged file system. This method is preferred over raw recovery as it allows obtaining files with their original names, folders, dates, and time stamps. If the metadata wasn't seriously corrupted, it may be possible to reconstruct the entire folder structure, depending on the specifics of the mechanisms employed by the file system to get rid of "unnecessary"

items. Yet, such analysis cannot be performed successfully when the crucial parts of metadata are missing. That is why it is extremely important to refrain from using file system repair tools or initiating operations that may result in its modification until the data is restored completely.

As a rule, when the desired result wasn't achieved with the help of metadata analysis, the search for files by their known content it performed. In this case, the "known content" doesn't imply the entire raw content of a file, only particular patterns that are typical for the files of the given format and may indicate the beginning or the end of the file. These patterns are referred to as "file signatures" and can be used to determine whether a piece of data on the storage belongs to a file of a recognized type. Files recovered with this method receive an extension based on the found signature, and new names and get assigned to new folders, usually created for files of different types. The main limitation of this approach is that some files may lack identifiable signatures or have only a signature denoting the start of a file, making it hard to predict where it ends, especially when its parts are not stored consequently.

To get the lost files back with maximum efficiency, data recovery software may use the described techniques concurrently during a single scan launched on storage.

Remote data recovery is performed through a modem or Internet connection by engineers using technology to achieve the same results as if the hard drive had been sent to a lab, yet in a more convenient manner for the customer. Assuming the hard drive is still functioning, remote recovery can be achieved for a single file or for huge volumes of data.

However, many users don't consider remote recovery to be as reliable as sending damaged drives to a lab. They believe recovery can be achieved only by engineers with highly specialized tools in state-of-the-art clean rooms. Users also are concerned about the security of having their computer systems, and their valuable data, connected to a third-party system and any vulnerability that might create.

Depending on the scenario, remote recovery offers the same advantages as in-lab service, with the added benefit of faster recovery times -- often as short as one hour. The initial goal is to either make the original volume mountable -- meaning that the operating system can read and write data to that drive -- or restore the data to its previous location. If this isn't possible, the engineer copies the data to a different location on the customer's system. With no need to dismantle and ship the drive or hardware for service, many concerns about a traditional recovery are eliminated. Security isn't an issue, since each recovery is performed through a connection secured with proprietary communication protocols and encrypted packets.

Remote recovery can solve many data-loss problems because it works for all types of recoveries, including servers, desktops, and laptops, across a wide variety of media, platforms and operating systems. In addition, the pricing structure is similar to traditional in-lab work. With remote recovery, you're not paying more, but you're potentially getting your data back faster.

Requirements of remote recovery

The major requirement of remote service is that the hardware must be working for lost data to be recovered. In these cases, where there is physical damage, the hard drive needs to go into a recovery lab so engineers can use special tools to get the drive running again for long enough to copy the data.

There are other hurdles with remote service for customers already shaken by their data-loss experience. Remote service requires some assistance from the customer so the engineer can connect with the system and produce a file listing. This process might be too intimidating for the less technically inclined user, as many people are afraid to even touch their computers after a data-loss situation out of fear that they might cause more damage.

Another requirement is the need for available staff to work with the remote-recovery engineers. If the circumstances surrounding the data loss are too hectic, sending the drive into a reliable data recovery lab might be a better option.

Remote service requires a stable connection, which can be a challenge in a high-security environment. Many organizations have strict proxy server or firewall policies and may not be able to connect with outside systems. Working with large amounts of data can also be a problem. Although remote recovery is capable of recovering huge volumes of data, sometimes the customer doesn't have the space available for the copied data. For example, working with some Unix or Macintosh systems requires special copy-out destinations to maintain data integrity.

Finally, the unique nature of the service can be a roadblock to using remote data recovery. Since it's a fairly new technology, many users feel that remote recovery is too good to be true and don't believe that it can deliver on the promises it makes. As the following examples show, however, remote data recovery is an option that every data-loss sufferer should always consider.

Data privacy generally means the ability of a person to determine for themselves when, how, and to what extent personal information about them is shared with or communicated to others. This personal information can be one's name, location, contact information, or online or real-world behavior. Just as someone may wish to exclude people from a private conversation, many online users want to control or prevent certain types of personal data collection.

As Internet usage has increased over the years, so has the importance of data privacy. Websites, applications, and social media platforms often need to collect and store personal data about users in order to provide services. However, some applications and platforms may exceed users' expectations for data collection and usage, leaving users with less privacy than they realized. Other apps and platforms may not place adequate safeguards around the data they collect, which can result in a data breach that compromises user privacy.

Why is data privacy important?

In many jurisdictions, privacy is considered a fundamental human right, and data protection laws exist to guard that right. Data privacy is also important because in order for individuals to be willing to engage online, they have to trust that their personal data will be handled with care. Organizations use data protection practices to demonstrate to their customers and users that they can be trusted with their personal data.

Personal data can be misused in a number of ways if it is not kept private or if people don't have the ability to control how their information is used:

- Criminals can use personal data to defraud or harass users.
- Entities may sell personal data to advertisers or other outside parties without user consent, which can result in users receiving unwanted marketing or advertising.
- When a person's activities are tracked and monitored, this may restrict their ability to express themselves freely, especially under repressive governments.

For individuals, any of these outcomes can be harmful. For a business, these outcomes can irreparably harm its reputation, as well as result in fines, sanctions, and other legal consequences. In addition to the real-world implications of privacy infringements, many people and countries hold that privacy has intrinsic value: that privacy is a human right fundamental to a free society, like the right to free speech.

What are some of the challenges users face when protecting their online privacy?

Online tracking: User behavior is regularly tracked online. Cookies often record a user's activities, and while most countries require websites to alert users of cookie usage, users may not be aware of to what degree cookies are recording their activities.

Losing control of data: With so many online services in common use, individuals may not be aware of how their data is being shared beyond the websites with which they interact online, and they may not have a say over what happens to their data.

Lack of transparency: To use web applications, users often have to provide personal data like their name, email, phone number, or location; meanwhile, the privacy policies associated with those applications may be dense and difficult to understand.

Social media: It is easier than ever to find someone online using social media platforms, and social media posts may reveal more personal information than users realize. In addition, social media platforms often collect more data than users are aware of.

Cybercrime: Many attackers try to steal user data in order to commit fraud, compromise secure systems, or sell it on underground markets to parties who will use the data for malicious purposes. Some attackers use phishing attacks to try to trick users into revealing personal information; others attempt to compromise companies' internal systems that contain personal data.

What are some of the challenges businesses face when protecting user privacy?

Communication: Organizations sometimes struggle to communicate clearly to their users what personal data they are collecting and how they use it.

Cybercrime: Attackers target both individual users and organizations that collect and store data about those users. In addition, as more aspects of a business become Internet-connected, the attack surface increases.

Data breaches: A data breach can lead to a massive violation of user privacy if personal details are leaked, and attackers continue to refine the techniques they use to cause these breaches.

Insider threats: Internal employees or contractors might inappropriately access data if it is not adequately protected.

Recover Swap Files/Temporary Files/Cache Files

The biggest problem with data is that once you store it on any form of magnetic media, it stays there forever. When you delete a file, your computer takes a shortcut. Instead of physically destroying the file, the computer simply pretends that the file no longer exists by replacing the first letter of the file name with a special character (hex byte code E5h), which leaves the contents of the file intact.

This process is like taking your name off an apartment building directory to make it look like you no longer live there, but stay in the apartment until someone else moves in. Only when the computer needs the space taken up by the deleted file will it actually overwrite the old file with new data? If your disk has plenty of extra space available, you could go weeks, months, or even years without ever overwriting previously deleted files. (Although, when you defragment your hard disk, your computer will likely overwrite many of these "deleted" files.)

Temp files are the temporary files that store the content of an unsaved in-editing document every minute on Windows for backup purposes. These files are helpful to combat situations of data loss when the system gets crashed, halted, or shut down abruptly, as users can restore content from the temporary files. The Temp folder has two locations on the system hard drive.

- *C:\Username\AppData\Local\Temp*
- *C:\Documents and Settings\ \Application Data\Microsoft (for Windows 7 and XP, when saved on the network drive)*

System users can change or modify this default location of the Temporary files by clicking on the Environment Variables properties of your system or running sysdm.cpl in the Windows Run box.

Ways to Recover Deleted Temp Files

If we talk about any manual way to recover deleted Temp files, then if these are deleted using the simple delete process, the recovery is possible by restoring files through the Recycle Bin. But if it is cleared from the Recycle Bin or Shift+Del action is taken to delete the Temp files, the manual recovery of files is not possible.

However, you can prevent this unpleasant situation by adopting some crucial measures given below.

- Create regular backups of Windows drive, creating a fresh restore point on its storage
- Never save the backup of the Temp files on the same drive/folder
- Avoid using a corrupted hard drive on which the files are lost
- Select a reliable, professional Windows Data Recovery tool for Temp file recovery.

Retrieving deleted files

When a file is deleted, the file system removes the file logically. That is, it removes all the metadata and stamps related to the file. However, the file still resides in the disk as a physical entity until it is overwritten. These physical areas can be very easily explored and read and converted to a readable file using a forensic application. It is observed that data resides on a computer for a very long time and are retrieved to a good extent.

Retrieving cached files

One can find the webpage visited by the suspect or the victim by looking into the cache. The cache file of an application can be spread across the system storage. We can confine only search by using typical keywords related to the case or probable websites.

Retrieving files in unallocated space

In general, a deleted file can be searched sequentially or structurally by looking for file headers or extensions. However, certain tools help us to scan and look for broken headers and use supplementary headers to retrieve data or at least retrieve blocks of a lost file for unallocated space. These retrieved blocks can later be studied and reformed using other tools to retrieve lost files to a great extent. This is also called file carving.

Forensic Toolkit (FTK) is computer forensics software, created by AccessData. It is a court-accepted, digital investigations software that includes many features and capabilities such as full-disk forensic images, decrypting files and cracking passwords, parsing registry files, collecting, process and analyzing datasets, and advanced volatile memory analysis. FTK is recognized as the standard toolkit for cyber defense forensic analysts, incident responders, and other professionals working or collecting forensic evidence. This path will cover the basic tools within the FTK suite - FTK Imager, Registry Viewer, and Password Recovery Toolkit (PRTK.) Then dive into use cases and analysis with FTK Suite.

OpenText™ EnCase™ Forensic is recognized globally as the standard for digital forensics and is a court-proven solution built for deep-level digital forensic investigation, powerful processing, and integrated investigation workflows with flexible reporting options. It is built with a deep understanding of the digital investigation lifecycle and the importance of maintaining evidence integrity. EnCase Forensic empowers any examiner to seamlessly complete any investigation, including investigations of mobile devices. For digital investigations, examiners need to be able to prioritize, collect, and decrypt evidence from a wide variety of devices while maintaining its integrity. The process needs to be quick, efficient, repeatable, and defensible, with the ability to

create intuitive reports. With EnCase Forensic, examiners can be confident the integrity of the evidence will not be compromised. All evidence captured with EnCase Forensic is stored in the court-accepted EnCase evidence file formats.

EnCase Forensic has been used in thousands of court cases and is known for its ability to uncover evidence that may go unnoticed if analyzed with other solutions. Police agencies, federal agencies, and companies across the globe depend on EnCase Forensic for its functionality, flexibility, and track record of court acceptance. New customer-driven enhancements further differentiate EnCase Forensic from all other forensic tools on the market.

FTK can perform forensics analysis on the following file systems: • Microsoft FAT12, FAT16, and FAT32Microsoft NTFS (for Windows NT, 2000, XP, and Vista) • Linux Ext2fs and Ext3fs.

FTK can analyze data from several sources, including image files from other vendors. It can also read entire evidence drives or subsets of data, allowing you to consolidate large volumes of data from many sources when conducting a computer forensics analysis. With FTK, you can store everything from image files to recovered server folders on one investigation drive. FTK also produces a case log file, where you can maintain a detailed record of all activities during your examination, such as keyword searches and data extractions. This log is also handy for reporting errors to Access Data.

At times, however, you might not want the log feature turned on. If you're following a hunch, for example, but aren't sure the evidence you recover is applicable to the investigation, you might not want opposing counsel to see a record of this information because he or she could use it to question your methods and perhaps discredit your testimony. (Chapter 15 covers testimony issues in more detail.) Look through the evidence first before enabling the log feature to record searches. This approach isn't meant to conceal evidence; it's a precaution to ensure that your testimony can be used in court. FTK has two options for searching for keywords. One option is an indexed search, which catalogs all words on the evidence drive so that FTK can find them quickly. This option returns search results quickly, although it does have some shortcomings. For example, you can't search for hexadecimal string values, and depending on how data is stored on the eve-dense drive, indexing might not catalog every word. If you do use this feature, keep in mind that indexing an image file can take several hours, so it's best to run this process overnight. The other option is a live search, which can locate items such as text hidden in unallocated space that might not turn up in an indexed search. You can also search for alphanumeric and hexadecimal values on the evidence drive and search for specific items, such as phone numbers, credit card numbers, and Social Security numbers.

Usage of computer forensics software tools to cross-validate findings in computer evidence-related cases

Computer forensics is the application of investigation and analysis techniques to gather and preserve evidence from a particular computing device in a way that is suitable for presentation in a court of law. The goal of computer forensics is to perform a structured investigation and maintain a documented chain of evidence to find out exactly what happened on a computing device and who was responsible for it.

Computer forensics -- which is sometimes referred to as *computer forensic science* -- essentially is data recovery with legal compliance guidelines to make the information admissible in legal proceedings. The terms *digital forensics* and *cyber forensics* are often used as synonyms for computer forensics.

Digital forensics starts with the collection of information in a way that maintains its integrity. Investigators then analyze the data or system to determine if it was changed, how it was changed, and who made the changes. The use of computer forensics isn't always tied to a crime. The forensic process is also used as part of data recovery processes to gather data from a crashed server, failed drive, reformatted operating system (OS), or other situation where a system has unexpectedly stopped working.

Why is computer forensics important?

In the civil and criminal justice systems, computer forensics helps ensure the integrity of digital evidence presented in court cases. As computers and other data-collecting devices are used more frequently in every aspect of life, digital evidence -- and the forensic process used to collect, preserve and investigate it -- has become more important in solving crimes and other legal issues.

The average person never sees much of the information modern devices collect. For instance, the computers in cars continually collect information on when a driver brakes, shifts and changes speed without the driver being aware. However, this information can prove critical in solving a legal matter or a crime, and computer forensics often plays a role in identifying and preserving that information.

Digital evidence isn't just useful in solving digital-world crimes, such as data theft, network breaches, and illicit online transactions. It's also used to solve physical-world crimes, such as burglary, assault, hit-and-run accidents, and murder.

Businesses often use a multilayered data management, data governance, and network security strategy to keep proprietary information secure. Having data that's well managed and safe can help streamline the forensic process should that data ever come under investigation.

Businesses also use computer forensics to track information related to a system or network compromise, which can be used to identify and prosecute cyber attackers. Businesses can also use digital forensic experts and processes to help them with data recovery in the event of a system or network failure caused by a natural disaster.

As the world becomes more reliant on digital technology for the core functions of life, cybercrime is rising. As such, computer forensic specialists no longer have a monopoly on the field. See how the police in the U.K. are adopting computer forensic techniques to keep up with increasing rates of cybercrime.

Types of computer forensics

There are various types of computer forensic examinations. Each deals with a specific aspect of information technology. Some of the main types include the following:

- **Database forensics.** The examination of information contained in databases, both data and related metadata.
- **Email forensics.** The recovery and analysis of emails and other information contained in email platforms, such as schedules and contacts.
- **Malware forensics.** Sifting through code to identify possible malicious programs and analyzing their payload. Such programs may include Trojan horses, ransomware or various viruses.
- **Memory forensics.** Collecting information stored in a computer's random access memory (RAM) and cache.
- **Mobile forensics.** The examination of mobile devices to retrieve and analyze the information they contain, including contacts, incoming and outgoing text messages, pictures and video files.
- **Network forensics.** Looking for evidence by monitoring network traffic, using tools such as a firewall or intrusion detection system.

How does computer forensics work?

Forensic investigators typically follow standard procedures, which vary depending on the context of the forensic investigation, the device being investigated or the information investigators are looking for. In general, these procedures include the following three steps:

1. **Data collection.** Electronically stored information must be collected in a way that maintains its

integrity. This often involves physically isolating the device under investigation to ensure it cannot be accidentally contaminated or tampered with. Examiners make a digital copy, also called a *forensic image*, of the device's storage media, and then they lock the original device in a safe or other secure facility to maintain its pristine condition. The investigation is conducted on the digital copy. In other cases, publicly available information may be used for forensic purposes, such as Facebook posts or public Venmo charges for purchasing illegal products or services displayed on the Vicemo website.

2. **Analysis.** Investigators analyze digital copies of storage media in a sterile environment to gather the information for a case. Various tools are used to assist in this process, including Basis Technology's Autopsy for hard drive investigations and the Wireshark network protocol analyzer. A mouse jiggler is useful when examining a computer to keep it from falling asleep and losing volatile memory data that is lost when the computer goes to sleep or loses power.

3. **Presentation.** Forensic investigators present their findings in a legal proceeding, where a judge or jury uses them to help determine the result of a lawsuit. In a data recovery situation, forensic investigators present what they were able to recover from a compromised system.

One of the most critical aspects of computer forensics is validating digital evidence because ensuring the integrity of the data you collect is essential for presenting evidence in court. Most computer forensic tools such as ProDiscover, X-Ways Forensics, FTK, and Encase provide automated hashing of image files. For example, when ProDiscover loads an image file, it runs a hash and compares that value to the original hash calculated when the image was first acquired. You might remember seeing this feature when the Auto Image Checksum Verification message box opens after you load an image file in ProDiscover. Computer forensics tools have some limitations in performing hashing, however, so learning how to use advanced hexadecimal editors is necessary to ensure data integrity.

Validating with Hexadecimal Editors Advanced hexadecimal editors offer many features not available in computer forensics tools, such as hashing specific files or sectors. Learning how to use these tools is important, especially when you need to find a particular file—for example, a known contraband image. With the hash value in hand, you can use a computer forensics tool to search for a suspicious file that might have had its name changed to look like an innocuous file. (Recall that two files with exactly the same content have the same hash value, even if they have different names.) Getting a hash value with a full-featured hexadecimal editor is much faster and easier than with a computer forensics tool.

A Computer Forensic Investigation generally investigates the data which could be taken from computer hard disks or any other storage devices with adherence to standard policies and procedures to determine if those devices have been compromised by unauthorized access or not. Computer Forensics Investigators work as a team to investigate the incident and conduct the forensic analysis by using various methodologies (e.g. Static and Dynamic) and tools (e.g. ProDiscover or Encase) to ensure the computer network system is secure in an organization. A successful Computer Forensic Investigator must be familiar with various laws and regulations related to computer crimes in their country (e.g. Computer Misuse Act 1990, the UK) and various computer operating systems (e.g. Windows, Linux) and network operating systems (e.g. Win NT). According to Nelson, B., et al., (2008), Public Investigations and Private or Corporate Investigations are the two distinct categories that fall under Computer Forensics Investigations. Public investigations will be conducted by government agencies, and private investigations will be conducted by a private computer forensic team. This report will be focused on private investigations since an incident occurred at a new start-up SME based in Luton.

LESSON ROUND-UP

- Data recovery tools are an essential tool for individuals to recover lost or damaged data. Such tools perform differently in case of personal data recovery and forensic data recovery.
- Data privacy is also important because in order for individuals to be willing to engage online, they have to trust that their personal data will be handled with care. Organizations use data protection practices to demonstrate to their customers and users that they can be trusted with their personal data.
- Forensic data recovery is the extraction of data from damaged evidence sources in a forensically sound manner. This method of recovering data means that any evidence resulting from it can later be relied on in a court of law.
- Digital forensics tools are either hardware or software designed to aid in the recovery of digital evidence of cyber-attack, and preservation of data or critical systems.
- The chain-of-custody form is critical in evidence gathering, and it comes into play as soon as you arrive at the scene of the incident. Every report, disk, screenshot, and printout is considered evidence; the chain of custody will begin as soon as the evidence is placed in an evidence bag or is tagged as evidence.
- Evidence gathering focuses on collecting all potential evidence, such as might be present in computer/ network logs, on defaced websites, on social media sites, or forensically from a computer hard drive.
- Digital Evidence is any information that is stored or transmitted in the digital form that a party at court can use at the time of trial. Digital evidence can be Audio files, and voice recordings, Address books and contact lists, Backups to various programs, including backups to mobile devices, Browser history, Cookies, Database, Compressed archives (ZIP, RAR, etc.) including encrypted archives, etc.
- Files that are deleted, lost, cached or unallocated can be retrieved using various methods and tools.

GLOSSARY

Data Analytics - Data Analytics is the science of analyzing raw datasets in order to derive a conclusion regarding the information they hold. It enables us to discover patterns in the raw data and draw valuable information from them. Data analytics processes and techniques may use applications incorporating machine learning algorithms, simulation, and automated systems.

Data Recovery - Data recovery is the process of restoring data that has been lost, accidentally deleted, corrupted or made inaccessible.

Chain of Custody - Chain of Custody form the Chain of Custody form (CCF or CoC) is used to record all changes in the seizure, custody, control, transfer, analysis, and disposition of physical and electronic evidence.

Swap File - A swap file is a system file that creates temporary storage space on a solid-state drive or hard disk when the system runs low on memory. The file swaps a section of RAM storage from an idle program and frees up memory for other programs.

Computer forensics – Computer forensics which is sometimes referred to as computer forensic science -- essentially is data recovery with legal compliance guidelines to make the information admissible in legal proceedings. The terms digital forensics and cyber forensics are often used as synonyms for computer forensics.

Data Privacy- Data privacy is the branch of data management that deals with handling personal data in compliance with data protection laws, regulations, and general privacy best practices.

TEST YOURSELF

(These are meant for recapitulation only. Answers to these questions are not to be submitted for evaluation)

1. What is the standard procedure and practice followed for the recovery of data and what ethical norms to be followed in such cases?
2. What are the precautions need to be followed in identification, preservation, analysis and presentation of evidence?
3. Why the issue of data privacy is significant in contemporary world today and what are the challenges faced in protection of such data?
4. Forensic Toolkit can perform forensics analysis on many file systems with different strategies. Discuss.

LIST OF FURTHER READINGS

- Anil Maheshwari, Data Analytics, McGraw Hill Education, First Edition, July 2017
- Jay Liebowitz, Data Analytics and AI, Taylor & Francis, First Edition, August, 2020
- Samaddar, Auerbach P, Data Analytics, Taylor & Francis, First Edition, February, 2019
- Joao Moereira, Andre Carvalho, Tom Horvath, A General Introduction to Data Analytics, Wiley Interscience, 1st Edition, August, 2020
- Goldenfein, J,Images and Biometrics – Privacy and Stigmatisation. In *Monitoring Laws: Profiling and Identity in the World State* (pp. 42-63). Cambridge: Cambridge University Press. doi:10.1017/9781108637657.003,2019
- Van Alsenoy, B., Introduction. In *Data Protection Law in the EU: Roles, Responsibilities and Liability* (pp. 343-346). Intersentia. doi:10.1017/9781780688459.023,2019
- Culnan, M., & Bruening, P., Privacy Notices: Limitations, Challenges, and Opportunities*. In E. Selinger, J. Polonetsky, & O. Tene (Eds.), *The Cambridge Handbook of Consumer Privacy* (Cambridge Law Handbooks, pp. 524-545). Cambridge: Cambridge University Press. doi:10.1017/9781316831960.029,2018