Digital Forensics
Lecture 01- Disk Forensics

An Introduction to

Akbar S. Namin
Texas Tech University
Spring 2017

Digital Investigation Foundations

- Digital Investigations and Evidence
  - Investigation of some type of digital device that has been involved in an incident or crime
  - Committed a physical crime or executed a digital event that violated a policy or law
    - E.g., a suspect used the internet to conduct research about a physical crime
    - E.g., an attacker gains unauthorized access to a computer, a user downloads contraband materials, or a user sends a threatening email.
  - An investigator's job: When the violation occurred and who or what caused it to occur

Digital Investigation Foundations

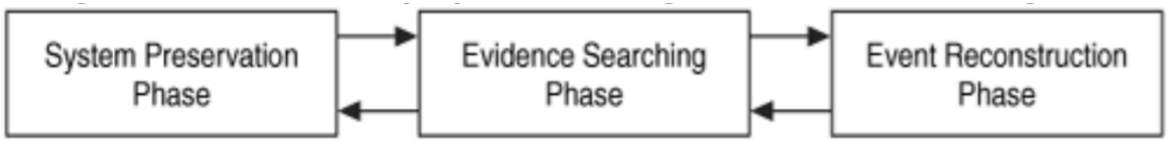
- Digital Investigations
  - A process where we develop and test hypotheses that answer questions about digital events.
- Use scientific methods: develop a hypothesis using evidence and test the hypothesis by looking for additional evidence that shows the hypothesis is impossible.
- Digital evidence
  - A digital object that contains reliable information that supports or refutes a hypothesis

Digital Investigation Foundations

- Forensic
  - The American Heritage Dictionary: “An adjective and relating to the use of science or technology in the investigation and establishment of facts or evidence in a court of law”
- A Digital Forensic Investigation
  - A process that uses science and technology to analyze digital objects and that develops and tests theories, which can be entered into a court of law, to answer questions about events that occurred.
  - A more restricted form of digital investigation

Digital Investigation Foundations

- Digital Crime Scene Investigation Process
  - There is no single way
  - A typical approach:
    - Three major phases
      - System prevention
      - Evidence searching
      - Event reconstruction



  - This process can be used when investigating both live and dead systems
    - A live analysis: occurs when we use the OS or other resources of the system being investigated to find evidence
      - We risk getting false information because the software could maliciously hide or falsify data
    - A dead analysis: occurs when we are running trusted applications in a trusted OS to find evidence
      - More ideal, but is not possible in all circumestances

Digital Investigation Foundations

- System Preservation Phase
  - Try to preserve the state of the digital crime scene
  - The purpose: reduce the amount of evidence that may be overwritten
  - The process continues after data has been acquired from the system because we need to preserve the data for future analysis
  - The goal: reduce the amount of evidence that is overwritten, we want to limit the number of processes that can write to the storage devices
- Preservation Techniques
  - For a dead analysis: we terminate all processes by turning the system off, and make duplicate copies of all data
  - For a live analysis, suspect processes can be killed or suspended
  - The network connection can be unplugged
  - Important data should be copied from the system in case it is overwritten while searching for evidence
  - When data are saved, a cryptographic hash should be calculated to later show that the data have not changed (e.g., MD5, SHA)

Digital Investigation Foundations

- Evidence Searching Phase
  - Looking for data that support/refute hypotheses about the incident
  - Typically starts with a survey of common locations based on the type of incident
    - E.g., if we are investigating Web-browsing habits, we should look at the Web browser cache, history file, and bookmarks.
    - E.g., if we are investigating a linux intrusion, we look for signs of a rootkit or new user accounts.
  - While the investigation proceeds, we develop hypotheses
    - An iterative process

Digital Investigation Foundations

- Search Techniques
  - Mostly done in a file system and inside files
  - A common search common: search for files based on their names or patterns
  - Another common technique: search for files based on a keyword in their content
  - A third one: search for files based on their temporal data, (i.e., last accessed, written time)
  - Search for known files by comparing the MD5 or SHA hash of a file's content with a hash database such as the National Software Reference Library (NSRL)
    - (http://www.nsrl.nist.gov)
  - Hash databases can be used to search for files based on signatures in their content

Digital Investigation Foundations

- Event Reconstruction Phase
  - Use the evidence that we found
  - This phase requires knowledge about the applications and the OS that are installed on the system
  - We may have found several files that violate a corporate policy or law, but that does not answer questions about events
    - We should determine what application downloaded an application
    - Is there any evidence that a Web browser downloaded them, or a malware has done it?

Digital Investigation Foundations

- General guidelines
  - PICL (Preservation, Isolation, Correlation, and Logging)
  - Preservation
    - Do not modify any data that could have been evidence
    - Copy important data
    - Calculate MD5 or SHA hashes of important data
    - Use a write-blocking device
    - Minimize the number of files created during a live analysis
      - They can overwrite evidence in unallocated space
    - Be careful when opening files
      - You could be modifying important data (e.g., last access time)

Digital Investigation Foundations

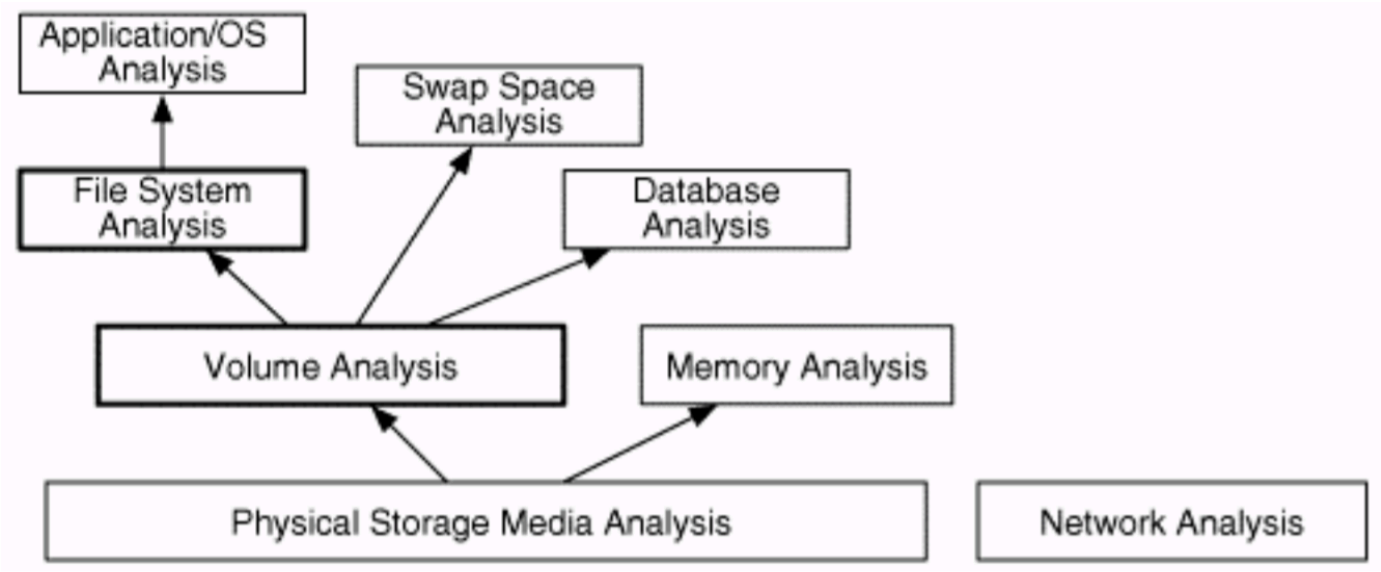
- General guidelines
  - PICL (Preservation, Isolation, Correlation, and Logging)
  - Isolation
    - Isolate the analysis environment from both the suspect data and the outside world
    - The reason: you do not know what it might do
    - Isolation is implemented by viewing data in a virtual environment (e.g., vmware)
    - Isolate from the the outside world
      - If tampering is done , you do not transmit anything

Digital Investigation Foundations

- General guidelines
  - PICL (Preservation, Isolation, Correlation, and Logging)
  - Correlate
    - Correlate data with other independent sources
    - It helps reduce the risk of forged data
    - E.g., timestamps can be easily changed
      - If time is important, try to find log entries, network traffic, or other events
  - Log
    - Helps identify what searches you have not yet conducted and what your results were
    - Specially it is important when doing live analysis
      - Document what you do

Digital Investigation Foundations

- Data Analysis
  - Analysis types
    - Basically two independent analysis areas:
      - Based on storage devices (disk forensics)
      - Based on communication devices (network forensics)
    - A different analysis areas:



Digital Investigation Foundations

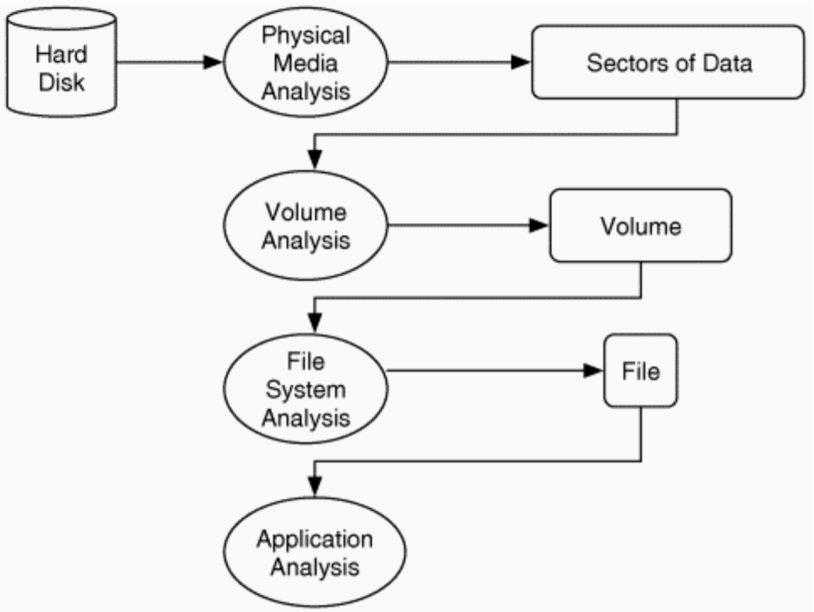
- Data Analysis
  - Physical storage media analysis
    - The analysis of the physical storage medium
    - E.g., hard disks, memory chips, and CD-ROMs
    - Reading magnetic data from in between tracks or other techniques that require a clean room
    - A stream of 1s and 0s
  - Memory
    - Organized by processes
    - Volatile storage
  - Volumes
    - Storage devices that are used for non-volatile storage
    - A volume is a collection of storage locations that a user or application can write to and read from
    - Two major concepts:
      - Partitioning
      - Assembly

Digital Investigation Foundations

- Data Analysis
  - Volumes
    - Two major concepts:
      - Partitioning
        - » Divide a single volume into multiple smaller volumes
      - Assembly
        - » Combine multiple volumes into one larger volume
    - File systems are the most common contents
      - A collection of data structures that allow an application to create, read, and write files
      - The results of file system analysis could be file content, data fragments, and metadata associated with files

Digital Investigation Foundations

- Data Analysis
  - Application analysis
    - To understand what is inside a file
    - The picture:
      - A disk that is analyzed to produce a stream of bytes
      - Volumes are analyzed at the file system Layer to produce a file



Digital Investigation Foundations

- Overview of Toolkits
  - Christine Siedma's Electronic Evidence Information site
    - (http://www.e-evidence.info)
  - Jacco Tunnissen's Computer Forensics, Cybercrime, and Steganography site
    - http://www.forensics.nl
  - A list of open source forensics tools
    - http://www.opensourceforensics.org
  - EnCase by Guidance Software
    - http://www.encase.com
    - A Windows-based tool
    - Can analyze many file system formats (e.g., FAT, NTFS, HFS+, UFS, Reiser, JFS, CD-ROMs, DVDs)
    - Allows listing the files, recovering deleted files, conducting keyword searches, viewing all graphic images, make timelines of file activity, etc.
    - It has a scripting language called EnScript (it helps automate many tasks)

Digital Investigation Foundations

- Overview of Toolkits
  - Forensics Toolkit (FTK) by AccessData
    - http://www.accessdata.com
    - Windows-based
    - Can acquire and analyze disk, file system, and application data
    - Supports FAT, NTFS, Ext2/3 file systems
    - Application-level analysis
    - Sophisticated searching abilities
  - ProDiscover by Technology Pathways (ProDiscover)
    - A Windows-based analysis tool
    - Can analyze FAT, NTFS, Ext2/3, and UFS file systems
  - SMART by ASR Data
    - Linux-based analysis tool
    - Can analyze FAT, NTFS, Ext2/3, UFS, HFS+, JFS, Reiser, CD-ROMs

Digital Investigation Foundations

- Overview of Toolkits
  - The Sleuth Kit / Autopsy (TSK)
    - http://www.sleuthkit.org
    - Unix-based command line analysis tool
    - Based on the Coroner's Toolkit (TCT) (http://www.porcupine.org)
    - Can analyze FAT, NTFS, Ext2/3, UFS file systems
    - Can list files and directories, recover deleted files, make timelines of file activity

Disk Forensics

- Reference
- File System Forensic Analysis (Brian Carrier)