Introduction To Digital Forensics

Module 1

Objective Questions

- What is forensic science?
- What is Digital Forensics?
- What is digital evidence?
- Where is digital evidence found?
- How can traditional forensic science be applied to digital evidence?
- What are some basic examples of digital evidence?
- What are the basics of Digital Forensic investigations?

Background - What is Forensic Science?

- Concept of applying science to investigations - both criminal & civil in nature
- Long established history
  - Toxicology analysis techniques developed in the 1700s
  - Fingerprint analysis developed in mid 1800s, started gaining widespread use in early 1900s
  - DNA analysis pioneered in the 1980s
- Can be used to reconstruct events based on findings
- Uses scientific method to analyze available evidence / validate hypothesis

What is Digital Forensics?

- Subset of Forensic Science
- Deals with collection & analysis of digital evidence
  - How do you apply scientific methods to digital data?
- Overarching Term - multiple sub-fields
  - Disk Forensics
  - Memory Forensics
  - Mobile Forensics
  - Network Forensics
  - Intrusion Forensics
- Formed out of the need for a standardized method
  - *“Necessity is the mother of invention”*

History of Digital Forensics

- Early on, investigators encountered computer systems
  - Information in those systems relevant to an investigation
  - Needed to know how to obtain the information
  - Often needed to work with system administrators
- People realized there was a need for skills & training
  - Often systems were used by people committing criminal acts
  - Investigators started meeting and discussing their needs
  - Sometimes difficult to get agency buy-in
- Relevant data stored in electronic form - more and more prevalent
- Mid 90s - Becomes more and more evident something is needed
- Groups / Organizations formed in an effort to standardize the field
- Today - more standardized methods

Digital Evidence - What is it?

- *“Any data stored or transmitted using a computer which can support or refute a theory of how an offense occurred, or that addresses critical elements of the offense, such as intent or alibi” (adapted from Chisum, 1999)*
- Data = literal binary data
  - 1s & 0s
  - Combined, they can represent anything - text, images, videos, etc.
- Different sources require different handling
  - Files vs. Emails vs. Volatile Memory

Evidence - Direct vs. Circumstantial

- Direct evidence
  - Establishes a fact
  - Example: Video footage shows a suspect present at a crime scene
- Circumstantial evidence
  - Requires an inference of fact
  - Example: Dirt found on a suspect's shoes matches dirt at a crime scene
    - Puts them at the scene, but not at a specific point in time
  - Multiple circumstantial pieces of evidence can be combined to see the big picture
- Digital Evidence
  - Can be direct or circumstantial
  - Often times somewhere in the middle, which encompasses additional challenges
  - "Putting someone at the keyboard" can be difficult
    - Who was logged in - who had access to the account

Example Evidence: File

- Can be found many places
  - Hard drive, removable media, communications, volatile memory
- Files contain information (content)
- Files have metadata (data about the data)
  - Modified, Accessed, Created date/times (MAC)
  - Possible author information
  - Device origin information
  - User account origin information
  - Original creation location/path

Example Evidence: Communication

- What do you use to communicate on a daily basis?
  - SMS / iMessage
  - E-Mail
  - Social media direct messaging
  - Phone calls
- Where does that content exist?
- What artifacts are generated?
- Where can those artifacts be found? For how long?

Example Evidence - Event Records

- Logging can be a huge source of evidentiary value
- Windows event logs
- Server logs
- Domain authentication logs
- Web server access logs
- Firewall logs
- Custom application logs

Digital Evidence - Everywhere

- Think about civil / criminal offenses

- Now try and think about one which doesn't include some aspect of digital evidence…….

Digital Evidence - Awareness

- Organizations becoming more aware of situations / incidents
  - Computer intrusions
  - Fraud (internal / external)
  - IP Theft
  - Physical incidents / altercations
- Everything has some kind of associated digital evidence
  - Just because you aren't using a digital device, doesn't mean there isn't one
- At this point, our lives are so incredibly dependant on digital devices & technology, we sometimes don't even have awareness of its grasp

Digital Forensics - Point of View

- The approach to the field can differ
  - Your point of view depends on your end goal
- What are you doing?
  - Are you conducting an internal investigation?
  - Have you been retained to investigate at another organization
  - Are you law enforcement investigating a crime?
- Your organizational needs can define your approach
- In this class, approach will primary be technical
  - Will try and be approach agnostic
  - We will address differences as they come up
- Does this mean we don’t concern ourselves with legal stuff?
  - No

Legal Authority

What you need to look at other people's stuff

Disclaimer!

- I am not a lawyer
- Do not interpret anything here as legal advice
- Always consult your in-house counsel with questions

Legal Authority - What is it?

- Essentially permission to look at other people's data
- Usually comes in one of two forms
  - Court ordered
  - Consent

Court Ordered Authority

- Both civil and criminal applications
- Civil - litigation often comes with court ordered e-discovery
  - Might be agreed upon by both parties
  - Could be ordered by court at request of opposition
- Criminal - different levels of authority, different burden to obtain
- Subpoena - least invasive
  - Generally limited to user / subscriber information
  - Network metadata / net-flow
- Search warrant - most invasive
  - Signed by judge after demonstrating probable cause
  - Allows for full content of particular data
  - Full disk image of a computer / Full content network monitoring (PCAP)

Consent Authority

- More than likely, this is what you will operate under
- External consultant
  - You receive authority when your client hires you - they give you permission to collect and analyze their data
- In-house response
  - You will be operating as part of your organization
  - Most organizations will have use-agreements - login banners
  - You should not have any expectation of privacy when you use company devices / networks
- If you are unsure what authority and the extent you are allowed to operate under, consult your in-house counsel

Evidence Response

- Need to understand what to collect and how to collect
  - Crazy coincidence - you are in the right place!
- How to handle evidence once collected?
  - Chain of custody?
  - What if it ends up going to LE? Court? Don't make assumptions...
- How quick do you need to respond and collect?
  - Too quick - might overlook things, make mistakes, miss opportunities
  - Too slow - might miss volatile artifacts
  - Must strike the right balance

Forensics Principles

- Forensic science - large body of proven investigative techniques & methods
- Forensic - characteristic of evidence which satisfies suitability for admission as fact, along with some level of ability to persuade based upon proof
- Certainty - Must use with great care
  - More often than not, cannot be certain of what happened
  - Generally, can only present possibilities of varying likelihood
  - Key is to increase the likelihood & confidence of a conclusion
- Conclusions - based on fact
- “Forensically Sound” - used to describe something which embodies the generally accepted principles of forensic science
  - Many times within Digital Forensics, you will create your own tools / techniques
  - Ensure they are forensically sound by following accepted methods

Forensically Sound

- “In order to be useful in an investigation, digital evidence must be preserved and examined in a forensically sound manner. Some think that a method of preserving or examining digital evidence is only forensically sound if it does not alter the original evidence source in any way. This is simply not true.”

- Example: Examining DNA evidence (blood, tissue, etc.) requires destroying a small amount of the evidence in order to conduct the analysis, yet when done properly and according to accepted practices, this is still regarded as sound forensics.

Forensically Sound

- Example: Imaging a failing hard drive
  - Pre/post MD5 values might not match - evidence not completely invalidated - just need to document what happened
- Example: Live data acquisition - disk / memory
  - Running a tool on a system creates its own artifacts within the collected evidence
  - However, this has become a generally accepted practice in Digital Forensics, especially with memory when there isn't an alternative way
- Bottom Line:
  - Alter as little as possible
  - Document as much as possible
  - Asses your results to determine if they were skewed
    - Validation is key

Evidence Exchange

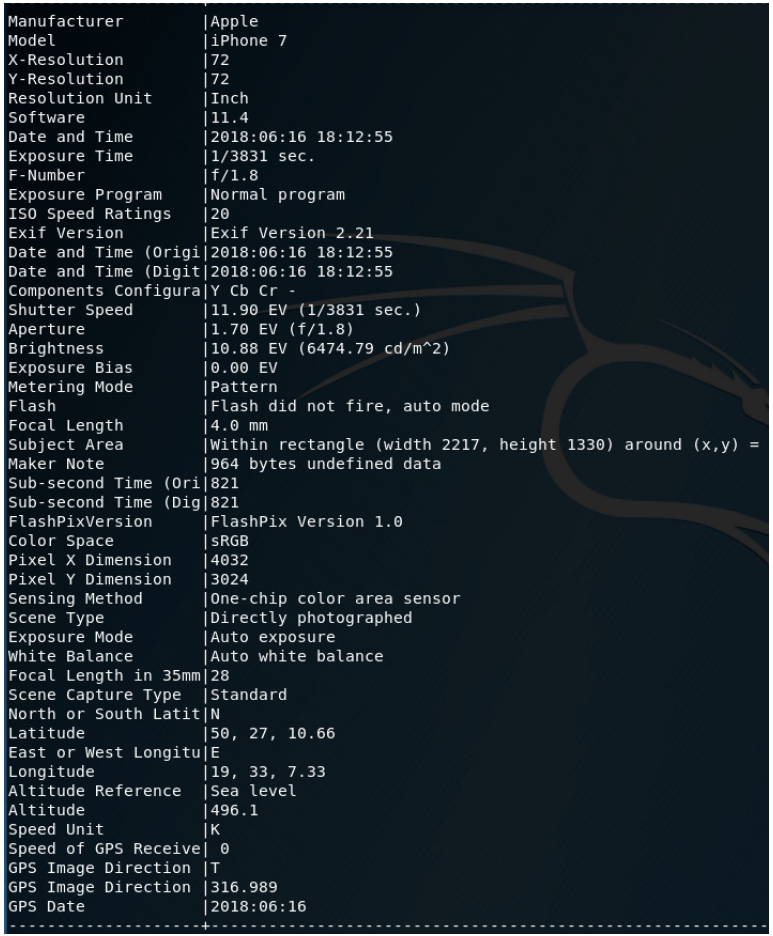
- Locard's Exchange Principle
  - Contact between two items results in an exchange
  - DNA of an attacker might be found under the fingernails of a victim
  - Blood from a victim ending up on the attacker
- Computer Intrusion - Evidence exchanged
  - File system activity
  - Registry artifacts
  - Event / Network logs
  - Theft of info - attacker takes evidence back
  - Network connections - socket artifacts generated in memory

Evidence Characteristics

- Class vs. Individual
- Class - traits common to similar items
  - Shoe print found at scene of crime - make/model/size - not specific shoe
  - EXIF data which identifies a specific make and model of a device which took a photo
- Individual - characteristics unique to each specific item
  - A fingerprint found at the scene of a crime is matched to a specific individual
  - Specific anomalies on a camera lense or scanner bed might be used to link back to the device a specific digital image

EXIF Data

- Exchangeable Image File format
- Standard for embedding additional metadata fields in multimedia files
- Possible for massive amount of metadata to be present
- Easy to access
  - Kali linux has the exif command pre-installed
  - **`root@kali:~#exif /home/user/file.jpg`**



```
Manufacturer        |Apple
Model               |iPhone 7
X-Resolution        |72
Y-Resolution        |72
Resolution Unit     |Inch
Software            |11.4
Date and Time       |2018:06:16 18:12:55
Exposure Time       |1/3831 sec.
F-Number            |f/1.8
Exposure Program    |Normal program
ISO Speed Ratings   |20
Exif Version        |Exif Version 2.21
Date and Time (Origi|2018:06:16 18:12:55
Date and Time (Digit|2018:06:16 18:12:55
Components Configura|Y Cb Cr -
Shutter Speed       |11.90 EV (1/3831 sec.)
Aperture            |1.70 EV (f/1.8)
Brightness          |10.88 EV (6474.79 cd/m^2)
Exposure Bias       |0.00 EV
Metering Mode       |Pattern
Flash               |Flash did not fire, auto mode
Focal Length        |4.0 mm
Subject Area        |Within rectangle (width 2217, height 1330) around (x,y) =
Maker Note          |964 bytes undefined data
Sub-second Time (Ori|821
Sub-second Time (Dig|821
FlashPixVersion     |FlashPix Version 1.0
Color Space         |sRGB
Pixel X Dimension   |4032
Pixel Y Dimension   |3024
Sensing Method      |One-chip color area sensor
Scene Type          |Directly photographed
Exposure Mode       |Auto exposure
White Balance       |Auto white balance
Focal Length in 35mm|28
Scene Capture Type  |Standard
North or South Latit|N
Latitude            |50, 27, 10.66
East or West Longitu|E
Longitude           |19, 33, 7.33
Altitude Reference  |Sea level
Altitude            |496.1
Speed Unit          |K
Speed of GPS Receive| 0
GPS Image Direction |T
GPS Image Direction |316.989
GPS Date            |2018:06:16
--------------------+-----------------------------------------------------
```

Evidence Authentication

- How do you ensure evidence maintains integrity once collected?
- Two primary issues to address
  - Who was in possession of the evidence from the time it was collected, to when it is introduced?
    - Chain of custody
  - Has the evidence been altered in any way since it was initially collected?
    - Pre/Post hash values

Chain of Custody

- Typically a form which accompanies a piece of collected evidence, documenting any and all people who have had possession of the evidence from collection to the current point in time
- A court may require everyone on the chain to testify to ensure its accuracy
- When there is a break in the chain, it introduces a **possibility** to invalidate the evidence
- Minimize if possible - reduce chance of problem
    - Working copy of evidence for analysis

Evidence Integrity

- How do you ensure collected data has not been modified?
- If only there was some way to “fingerprint” a portion of data, and generate a value which could be used to easily identify if the source data had changed....
- It would be really cool if even a slight modification of the source data would trigger a significant and noticable change in the resultant fingerprint data….

Evidence Integrity - Hashing

- Hash / Message Digest / Cryptographic Hash Value
- Hashing operates as a “black box” algorithm
  - We send it an input value -> we receive the output value (hash)
  - Not actually black box - algorithms are publically known
- Identical input value with a specific algorithm will always yield same result
- Several different algorithms
  - MD5, SHA1, SHA256
- Some functions like MD5 have been touted as ‘non-secure’
  - Hash collisions have been identified
- Still used in the Digital Forensics world
  - Just because something isn’t cryptographically secure doesn't negate all of its value

Hashing

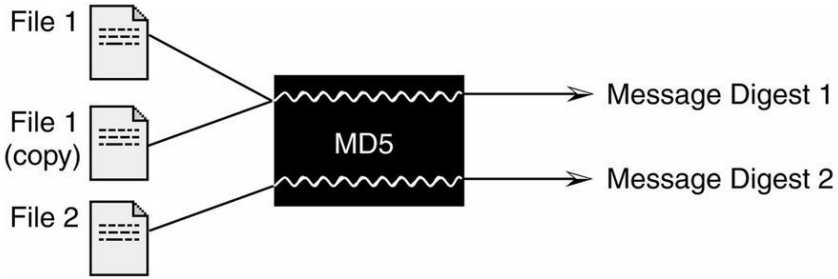


FIGURE 1.3 Black box concept of the message digest.

Hashing - Input Change

- Even a small input change manifests as a significant change in the output of a hashing algorithm
- Makes it very easy to see if files are identical by glancing at their hash values

```
bash-3.2$ echo "This is a test" | md5
ff229413369560 98ae9a564289d1bf1b
bash-3.2$ echo "This is a test." | md5
02bcabffffd16fe0fc250f08cad95e0c
```

Hashing Uses

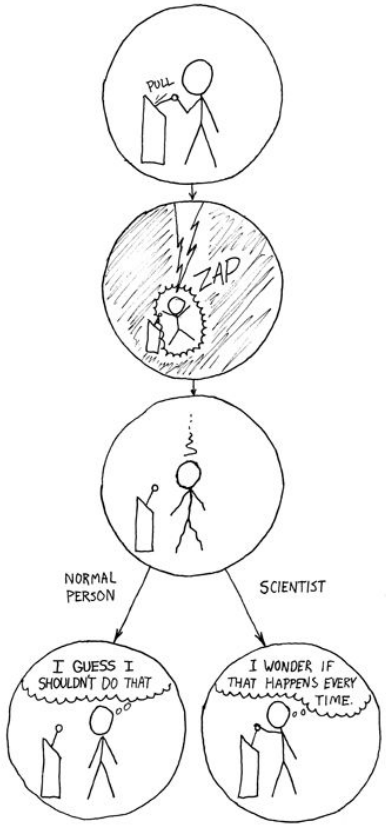
- Evidence integrity
  - During evidence acquisition, a hash value of the source data is recorded
  - At any point going forward, the collected data can be re-hashed
  - Values can be compared to determine if integrity has been maintained
- Duplicate files
  - After hashing a group of files, it becomes trivial to identify duplicates
- Searching
  - Known search criteria files can be hashed
  - Files on a system can be hashed, and then search for search criteria hash values
- Filtering
  - Remove known file types by hash
  - National Institute of Standards and Technology (NIST) - National Software Reference Library (NSRL) - https://www.nist.gov/itl/ssd/software-quality-group/national-software-reference-library-nsrl

Preserve Evidence Integrity

- How do we ensure we don't alter the evidence during collection?
  - Physical source drive - write blocker
- Live Response
  - Software mechanisms
  - Sometimes unavoidable modifications happen
    - Memory collection

Scientific Method

- “An important aspect of the scientific method is that any experiments or observations must be repeatable in order to be independently verifiable.”
- Documentation is key
  - “If you didn’t document it, you didn’t do it”
- Document what you acquire
  - Details of the source evidence and how you acquired it
- Document what you examine / analyze
  - What tools / methods you used
- Document what you found
  - Where you found it



Understanding Technology

- In order to properly examine and analyze digital evidence, it requires an understanding of the systems and devices which create and store it
- “A lack of understanding of how computers function and the processes that sophisticated tools have automated makes it more difficult for digital investigators to explain their findings in court and can lead to incorrect interpretations of digital evidence.”
- Pushing buttons in forensic software and not knowing how to verify the results will lead to problems
    - Casey Anthony - Recommended Reading #1

Computer Science Matters

- Binary & Endianess / Byte Order
  - 510 = 00000001 11111110 (256 + 254 big endian)
  - 510 = 11111110 00000001 (254 + 256 little endian)
  - Little endian - (intel) - register with smaller part of number read first
- Who likes reading through mountains of data in binary? Anyone?
  - Fortunately, we have hex & ASCII
  - Decimal 12 = 1100 in binary = ‘C’ in Hex
    - 1100 = (8*1)+(4*1)+(2*0)+(1*0)
- Forensic analysis often utilizes Hex
  - Easier to page through data

Digital Investigation Stages

- Preparation
- Survey
- Documentation
- Preservation
- Examination & Analysis
- Reconstruction
- Reporting Results

Preparation

- Think about what your end goals are
  - How will you achieve them?
- Research - understand what you are going to encounter
  - Workstations / Servers / Mobile Devices / Networks / Cloud Resources
- Ask questions if you need to
- Be prepared
  - Have hardware and software for the unexpected
- Plan for a large amount of data needing storage
  - Drive sizes are only going up
- Double check your legal authority / Scope of Work

Survey

- Understand what you have in front of you
- Determine what you need to deal with
  - Identify hardware components - computers, flash drives, mobile devices
  - Locate network storage - work with staff if possible
- Don’t discount less obvious sources of evidence
  - Gaming systems
  - Memory cards installed in other devices
  - Non-standard devices (Raspberry Pi)
- Follow all the cables if need be

Documentation

- As mentioned, documentation is essential at every step
- Identify what you encounter, how you encounter it, etc.
- Document what steps you took with each device
- Include the information about each device
  - Make, model, serial #, physical description, discovered state, location
- Initiate chain of custody
- Don’t hide mistakes
  - Everyone makes them - own it and explain it
  - If you try and cover it up, you lose your credibility - and you don’t ever get it back
- Take notes during analysis - even if something doesn’t work
  - May not be needed for final report - but may be helpful for you later on, especially if you end up needing to testify about it years later

Preservation

- Once relevant / responsive evidence is identified, what next?
  - Need to properly preserve the evidence
- Different evidence requires different preservation techniques
  - Hard drive / RAID array
  - Volatile memory - Determine if needed
  - Network data
  - Mobile devices
  - Loose media (USB drives, memory cards, optical discs, etc.)
- Notes / “Desk Litter”
  - Post-It notes on around a workspace with passwords written down? 👍
- If situation warrants, seal items with evidence tape
  - Initial / sign, date/time

Examination & Analysis

- Once collected, need to understand what you have
- Filtering & Reduction can be a good first step
  - Exclude non-relevant data (NIST NSRL) or non relevant data types
  - Focus on user-created content
  - Filter by time frame
- Look for file anomalies
  - Does file signature match extension?
  - Is there hidden encrypted data? (entropy score)
  - Is there evidence a file was moved from a different origin source?
- Identify deleted / partially overwritten files

Reconstruction

- Remember, finding technical artifacts is only half the challenge
  - Still need to reconstruct what happened
- Functional reconstruction
  - Understanding how systems functioned and determine their capabilities
  - If a large file was thought to be downloaded to disk within a certain time frame, was their network connection fast enough to support the theory?
- Relational reconstruction
  - Determine relationships between evidence - useful for network analysis & intrusion forensics
- Temporal reconstruction
  - Timestamps are critical during forensics analysis - allows for putting an overall picture together
  - Taking pieces of evidence from different sources and collating them together, sorting by time
  - Allows for an overall master timeline of activity surrounding an incident

Reporting

- What good is your analysis if you can't convey the results?
- Need to be able to both summarize, and articulate details
- Sample report structure:
  - Introduction - Brief background of case and short summary of findings
  - Evidence Summary - Identify what evidence was obtained and analyzed
  - Examination Summary - Identify what tools/methods used and any important discoveries
  - File System Examination - Inventory of any relevant files/directories with supporting evidence
  - Analysis - Describe any reconstruction methods used and their outcomes
  - Conclusions - Summary of conclusions based on facts identified with supporting evidence
  - Glossary - Define any key complex technical terms or concepts
  - Appendix - Additional evidence used to reach conclusions (file lists, etc.)

Next Steps & Reminders

- Ensure you read the required readings for this week
- When you are ready, attempt the quiz
- Start looking at Module 1 - Forensic Analysis Assignment
  - Unknown File Analysis
- You may post general questions in the forums
  - Please refrain from posting specific questions/answers - Use your judgement
  - Helping each other learn the material is the end goal
    - “What is a good exif viewer for OSX?” - Totally fine
    - “Did everyone else get ‘C’ for #7? - Totally NOT fine

Questions?