Digital Evidence Locations & Introduction to Computer Forensics

Don Mason
Associate Director

Objectives

After this session, you will be able to:
- Define “digital evidence” and identify types
- Describe how digital evidence is stored in computers
- Identify devices and locations where digital evidence may be found
- Define basic computer and digital forensics
- Identify and describe the essential principles, tools, and trends in digital forensics

Special Acknowledgments

- Justin T. Fitzsimmons
  Senior Attorney, NDAA National Center for Prosecution of Child Abuse
- Sergeant Josh Moulin
  Commander, Southern Oregon High-Tech Crimes Task Force
Advancing Technology

Computer / Digital Technology
- Personal computers, at work and at home
- Digital cameras
- Web cams
- Camera and video cam cell phones
- Document and image scanners
- Digital recording and duplicating devices
- Large digital storage capacities
- Portable media

How Digital Evidence is Stored
- Data is written in binary code – 1’s and 0’s
- These 1’s and 0’s are grouped together in blocks of 8, called “bytes.”
- For example, the sequence “10001111” represents the letter “O”.

---

Digital Evidence Locations; Computer Forensics Introduction
Handling Child Pornography Cases, February 4-5, 2010
Copyright © 2010 National Center for Justice and the Rule of Law – All Rights Reserved
How Data is Stored

Clusters are groups of sectors

Track
Sector

Digital Evidence

Information of probative value that is stored or transmitted in binary form and may be relied upon in court

Digital Evidence

- Information stored in binary format but convertible to
  - e-mail, chat logs, documents
  - photographs (including video)
  - user shortcuts, filenames
  - web activity logs
- Easily modified, corrupted, or erased
- Correctly made copies indistinguishable from original
The Internet

- World Wide Web (the Web)
- E-mail
- Instant messaging (IM)
- Webcam/Internet Telephone (VoIP)
- Peer-to-peer (P2P) networks
- Legacy Systems
  - Newsgroups
  - Telnet and File transfer (FTP) sites
  - Internet Relay Chat (IRC)
  - Bulletin boards

Web 2.0

- Interactive Internet communities
- Social networks
- Blogs
- "Wikis"
- Video or photo sharing sites
- Online role-playing games
- Virtual worlds
Cloud Computing

- Basically, obtaining computing resources from someplace outside your own four walls, and paying only for what you use
  - Processing
  - Storage
  - Messaging
  - Databases
  - etc.

Ex: Google docs
What Kinds of Computers Can Be on the Internet?

- Mainframes
- Laptops
- Personal Digital Devices
- Cell Phones
- Personal Computers

Internet Connectivity

- Home PCs
  - Telephone Dial-in Line
  - Cable Modem Connection
  - DSL Line
- Internet Service Provider (ISP)
  - Internet Addressing
    - Every network / host (and each home computer connected to the Internet) has a unique numeric Internet protocol (IP) address *num1.num2.num3.num4*
      - e.g., 172.20.53.229
    - Nearly all hosts and networks also have corresponding domain names that are easier for humans to remember and use
      - e.g., www.ncjrl.org or oag.state.gov.us
Why Addressing Matters

- The Internet is a packet-switched network
- The component parts of a communication (i.e., the packets) sent to another host may travel by different paths
- Each packet makes one or more “hops” along the network on the way to its destination

What’s in a Packet?

- An IP data packet includes:
  - routing information (where it came from, where it’s going)
  - the data to be transmitted
- Replies from the receiving host go to the packet’s source address
  - here, 172.31.208.99

Packet Switching

ISP.COM, a/k/a 172.31.208.99

AGENCY.GOV, a/k/a 10.135.6.23
Packet Switching
ISP.COM, a/k/a 172.31.208.99

Packet Switching
AGENCY.GOV, a/k/a 10.135.6.23

Packet Switching
ISP.COM, a/k/a 172.31.208.99

Packet Switching
AGENCY.GOV, a/k/a 10.135.6.23
Packet Switching

ISP.COM, a/k/a 172.31.208.99

AGENCY.GOV, a/k/a 10.135.6.23

Computer & Internet Uses

- Remote Computing
- Research
- Commerce
- Recreation
- Communication
Why It Matters How Computers, Networks, and the Internet Work

- Immense amount of digital data created, transmitted, stored
- Some created by humans
- A lot necessarily created by machines “in the background”

Digital Evidence

- User-created
  - Text (documents, e-mail, chats, IM’s)
  - Address books
  - Bookmarks
  - Databases
  - Images (photos, drawings, diagrams)
  - Video and sound files
  - Web pages
  - Service provider account subscriber records

- Computer-created
  - Dialing, routing, addressing, signaling info
  - Email headers
  - Metadata
  - Logs, logs, logs
  - Browser cache, history, cookies
  - Backup and registry files
  - Configuration files
  - Printer spool files
  - Swap files and other “transient” data
  - Surveillance tapes, recordings
### Forms of evidence

- **Files**
  - Present / Active (doc's, spreadsheets, images, email, etc.)
  - Archive (including as backups)
  - Deleted (in slack and unallocated space)
  - Temporary (cache, print records, Internet usage records, etc.)
  - Encrypted or otherwise hidden
  - Compressed or corrupted
- **Fragments of Files**
  - Paragraphs
  - Sentences
  - Words

### Sources of Evidence

- **Offender’s computer**
  - accessed and downloaded images
  - user log files
  - Internet connection logs
  - browser history and cache files
  - email and chat logs
- **Hand-held devices (embedded computer systems)**
  - digital cameras
  - PDAs
  - mobile phones

### More Sources of Evidence

- **Servers**
  - ISP authentication user logs
  - FTP and Web server access logs
  - Email server user logs
  - LAN server logs
- **Online activity**
  - IP addresses of chat room contributors
Digital Devices / Locations Where Digital Evidence May be Found

Mainframes, Desktops, Laptops

Hard Drives
Solid State Drives

Removable Media

USB Storage Devices
Vehicle “black boxes”
  – Event data recorders
  – Sensing and diagnostic modules
  – Data loggers
“True GPS”

- "Handset solution"
  - The data is "inside the box"
  - Involves search of the device

GPS devices

Computer Forensics
Computer Forensics

- "preservation, identification, extraction, documentation, and interpretation of computer media for evidentiary and/or root cause analysis"
- Usually pre-defined procedures followed but flexibility is necessary as the unusual will be encountered
- Was largely "post-mortem" but is evolving

Computer/Digital Forensics

- Sub branches / activities / steps
  - Computer forensics
  - Network forensics
  - Live forensics
  - Software forensics
  - Mobile device forensics
  - "Browser" forensics
  - "Triage" forensics

Basic Computer Forensics

- **Seizing** computer evidence
  - bagging & tagging
- **Imaging** seized materials
- **Searching** the image for evidence
- **Presenting** digital evidence in court
Myth v. Fact

Myth
- A computer forensic analyst can recover any file that was ever deleted on a computer since it was built.

Fact
- The analyst can recover a deleted file, or parts of it, from unallocated file space until the file system writes a new file or data over it.

Myth v. Fact

Myth
- Metadata ("data about data") is the all knowing, all seeing, end all piece of info on a file.

Fact
- Metadata does contain useful information about a file but it is limited.
  - E.g.: Author
  - MAC times
  - File name, size, location
  - File properties
  - Might contain revisions, comments, etc.

Metadata – basic examples
**Metadata – Track changes**

Metadata is data about data. It includes information about the file itself, such as its creation date, author, and other details. This data can be useful for tracing the origin and history of a file. For example, if a file contains sensitive information, the metadata can help identify the source and auditor.

**Metadata – Comments**

Comments are additional notes or information that can be added to a file. They can be used to record important details about the file, such as its purpose or any limitations. Comments can be viewed within the file itself, allowing users to see and edit them as needed.

**EXIF data**

- **Exchangeable Image File Format**
- **Embeds data into images containing camera information, date and time, and more**
Basic steps – 3 A’s

Acquiring evidence without altering or damaging original

Authenticating acquired evidence by showing it’s identical to data originally seized

Analyzing the evidence without modifying it

Acquiring the evidence

- Seizing the computer: Bag and Tag
- Handling computer evidence carefully
  - Chain of custody
  - Evidence collection
  - Evidence identification
  - Transportation
  - Storage
- Making at least two images of each evidence container
  - Perhaps third in criminal case – for discovery
- Documenting, Documenting, Documenting

Preserving digital evidence

The “Forensic Image” or “Duplicate”

A virtual “snapshot” of the entire drive

- Every bit & byte
- “Erased” & reformatted data
- Data in “slack” & unallocated space
- Virtual memory data
Write Blockers

Hard drives are imaged using hardware write blockers

Authenticating the evidence

- Proving that evidence to be analyzed is exactly the same as what suspect/party left behind
  - Readable text and pictures don’t magically appear at random
  - Calculating hash values for the original evidence and the images/duplicates
    - MD5 (Message-Digest algorithm 5)
    - SHA (Secure Hash Algorithm) (NSA/NIST)

What is a Hash Value?

An MD5 Hash is a 32 character string that looks like:

Acquisition Hash: 3FDSJO90U43JIIVJU904FRBEWH
Verification Hash: 3FDSJO90U43JIIVJU904FRBEWH

The Chances of two different inputs producing the same MD5 Hash is greater than:

1 in 340 Unidecillion: or 1 in 340,000,000,000,000,000,000,000,000,000,000,000
Hashing Tools – Examples

- http://www.fileformat.info/tool/md5sum.htm

- Also, AccessData’s FTK Imager can be downloaded free at http://www.accessdata.com/downloads.html

MD5 Hash

- 128-bit (16-byte) message digest – a sequence of 32 characters
  - “The quick brown fox jumps over the lazy dog”
  - 9e107d9d372bb6826bd81d3542a419d6
- “The quick brown fox jumps over the lazy dog.”
  - e4d909c290d0f1ba068ffaddf22cb0

More Examples of Hash Values
File "F:\Wellesley\WELLESLEY01" was acquired by Detective Papargiris at 02/21/02 06:40:56PM. The computer system clock read: 02/21/02 06:40:56PM.

Evidence acquired under DOS 7.10 using version 3.19.

File Integrity:
Completely Verified, 0 Errors.

Acquisition Hash: 88F7BA9EB338EEDC2AF312D039BFEC
Verification Hash: 88F7BA9EB338EEDC2AF312D039BFEC

Drive Geometry:
Total Size: 12.7GB (26,712,000 sectors)
Cylinders: 28,266
Heads: 15
Sectors: 63

Partitions:
<table>
<thead>
<tr>
<th>Code</th>
<th>Type</th>
<th>Start Sector</th>
<th>Total Sectors</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>0C</td>
<td>FAT32X</td>
<td>0</td>
<td>2670030</td>
<td>12.7GB</td>
</tr>
</tbody>
</table>
What happens when you rename a file?
Or Rename The Extension
“Hashing” an image

MD5
021509c96bc7a6a47718950e78a7a371
SHA1
77fe03b7c0663cf55dc268b19f5a449e5a97386

(single pixel changed using Paint program)

MD5
ea8450e5e8cf71a1c17c6effced95b484
SHA1
01f57f330f06c16d58725f51decdfeeb88b69c6c

Analyzing the evidence

- Working on bit-stream images of the evidence; never the original
- Prevents damaging original evidence
- Two backups of the evidence
  - One to work on
  - One to copy from if working copy altered
- Analyzing everything
  - Clues may be found in areas or files seemingly unrelated

Popular Automated Tools

Encase
Guidance Software

Forensic Tool Kit (FTK)
Access Data
Validation of Computer Forensics Tools

Subjecting EnCase to Daubert analysis
1. Subject to testing criteria
   - NIST 2004 study [link]
   - Lab-specific testing
2. Subject to peer review and publication
   - Featured in a number of articles and forensics/incident response books
3. High known or potential rate of error?

Validation of Computer Forensics Tools
4. General acceptance within the scientific community

Case law/judicial notice of prior Daubert hearings in other jurisdictions

Use in law enforcement and corporate/private sectors
Taught in academic institutions

EnCase and Legal Challenges

Analysis (cont.)

- Existing Files
  - Mislabeled
  - Hidden
- Deleted Files
  - Trash Bin
  - Show up in directory listing with $ in place of first letter
    - "taxes.xls" appears as "$axes.xls"
- Free Space
- Slack Space
- Swap Space

Free Space

- Currently unoccupied, or "unallocated" space
- May have held information before
- Valuable source of data
  - Files that have been deleted
  - Files that have been moved during defragmentation
  - Old virtual memory

Slack Space

- Space not occupied by an active file, but not available for use by the operating system
- Every file in a computer fills a minimum amount of space
  - In some old computers, this is one kilobyte, or 1,024 bytes. In most new computers, this is 32 kilobytes, or 32,768 bytes
  - If you have a file 2,000 bytes long, everything after the 2000th byte is slack space
**How “Slack” Is Generated**

- File A (In Memory)
- File A saved to disk, on top of File B
- File A over-writes File B, creating slack
- File A (Now On Disk)
- Remains of File B (slack)

*Slack space:* the area between the end of the file and the end of the storage unit

**Recall how data is stored**

**Other sources mined for Transient Data**

- Browser cache, history, cookies
- Residual chat data
- Activity logs
- Registry & registry backup files
Sources of Digital Gold
- Internet History
- Temp Files (cache, cookies etc...)
- Slack/Unallocated space
- Buddy Lists, chat room records, personal profiles, etc...
- News Groups, club listings, postings
- Settings, file names, storage dates
- Metadata (email header information)
- Software/Hardware added
- File Sharing ability
- Email

Selected “Trends” in Digital Forensics

“Browser” Forensics

“Triage” Forensics

Browser Forensics

Web browsers (e.g. Microsoft Internet Explorer, Mozilla Firefox, Safari, Opera) maintain histories of recent activity, even if not web related
Internet History

Computers store Internet history in a number of locations including:
- Temporary Internet files
- Windows Registry
- Browser / Search Term history
- Cookies

This information is browser specific.
“Triage” Forensics

- Also occasionally referred to as “rolling” forensics, or “on-site preview”
- Image scan
- Especially useful in “knock & talk” consent situations or in screening multiple computers to determine which to seize
- Caveat: Not all agencies are equipped or trained yet to do this.

“Triage” Forensics - Steps

- Attach/Install write-blocking equipment
- Turn on target device
- Scan for file extensions, such as:
  - .doc
  - .jpg (.jpeg)
  - .mpg (.mpeg)
  - .avi
  - .wmv
  - .bmp
“Triage” Forensics - Steps

- Pull up thumbnail views - 10-96 images at a time
- Right click on image, save to CD or separate drive.
- Determine file structure or file path.

“Triage” Forensics

- Increasingly important, as the number and storage capacities of devices rapidly grow.
- But does NOT enable a comprehensive forensically sound examination of any device on the scene.

Resources

- https://blogs.sans.org/computer-forensics/
- http://craigball.com/