Cyber Crime, Digital Evidence Locations, and Computer Forensics

Don Mason
Associate Director

Objectives
After this session, you will be able to:
- Define "cyber crime"
- Define and describe "digital evidence"
- Identify devices and locations where digital evidence may be found
- Define basic computer and digital forensics
- Identify and describe the basic practices, principles, and tools used in digital forensics

Advancing Technology
Always Something New

Roles of Digital Devices

- **Targets**
- **Tools**
- **Containers**

New Crimes, New Techniques

- **Computer as Target**
  - Unauthorized access, damage, theft
  - Spam, viruses, worms
  - Denial of service attacks
- **Computer as Tool**
  - Fraud, ID theft
  - Threats, harassment, bullying
  - Child pornography
- **Computer as Container**
  - From drug dealer records to how to commit murder
“Cyber Crime”
- “Computer crime”
- “Network crime”
- “Computer-related crime”
- “Computer-facilitated crime”
- “High tech crime”
- “Internet crime” or “Online crime”
- “Information age crime”

Any crime in which a computer or other digital device plays a role, and thus involves digital evidence

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 code but convertible to, for example:
  - e-mail, chat logs, documents
  - photographs (including video)
  - user shortcuts, filenames
  - web activity logs
- Easily modified, corrupted, or erased
- But correctly made copies are indistinguishable from the original
How Data Is Stored

- Track
- Sector
- Clusters are groups of sectors

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
Computer and Internet Uses

- Remote Computing
- Research
- Commerce
- Recreation
- Communication

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.
What Kinds of Computers Can Be on the Internet?

- Mainframes
- Personal computers
- Personal digital devices
- Laptops
- Cell Phones

Internet Connectivity

- Home PCs
- Internet Service Provider (ISP)
- Telephone dial-in line
- Cable modem connection
- DSL line
- Network
- High-speed data link
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
- 172.31.208.99
- 10.135.6.23
- 0111001010101011
  1011011000100101
  0100...
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
Data Generated in 2006

- 161 billion gigabytes (161 exabytes)
- 12 stacks of books each reaching from the Earth to the Sun
- 3 million times all the books ever written
- Would need more than 2 billion iPods to hold it

Projections for 2006-2010

- Six fold annual information growth
- In 2010: 988 exabytes to be created and copied
  - More than 73 stacks of books taller than 93 million miles!
- Compound annual growth rate: 57%

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
Digital Devices / Locations Where Digital Evidence May be Found

Computer Hardware

Challenges
- Increasing ubiquity and convergence of digital devices
- Increasing data storage capacity
- Shrinking devices and media
- Growing use of solid state devices
More Digital Devices

And Still More

More
Vehicle “black boxes”
- Event data recorders
- Sensing and diagnostic modules
- Data loggers

More

More
Ex: Second Life

Join the Internet’s largest user-created, 3D virtual world community.

Cell Site Location Data
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**

<table>
<thead>
<tr>
<th>Myth</th>
<th>Fact</th>
</tr>
</thead>
<tbody>
<tr>
<td>A computer forensic analyst can recover any file that was ever deleted on a computer since it was built.</td>
<td>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.</td>
</tr>
</tbody>
</table>

---

**Myth v. Fact**

<table>
<thead>
<tr>
<th>Myth</th>
<th>Fact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metadata (&quot;data about data&quot;) is the all knowing, all seeing, end all piece of info on a file.</td>
<td>Metadata does contain useful information about a file but it is limited.</td>
</tr>
</tbody>
</table>

- **E.g.**:
  - Author
  - MAC times
  - File name, size, location
  - File properties
  - *Might* contain revisions, comments, etc.
EXIF Data

- Exchangeable Image File Format
- Embeds data into images containing camera information, date and time, and more

Basic Steps

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 3rd in criminal case – for discovery
- Documenting, Documenting, Documenting
Preserving Digital Evidence
The “Forensic Image” or “Duplicate”

A virtual “clone” 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: 3FD3J09U43JJIVJU904FRBEWH
Verification Hash: 3FD3J09U43JJIVJU904FRBEWH

The Chances of two different inputs producing the same MD5 Hash is greater than:
1 in 340 Unidecillion: or 1 in 340,000,000,000,000,000,000,000,000,000,000,000,000

File “F:\Wellesley\WELLESLE.E01” 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: 88F7BA968E31EE9CD31F32DD395BFEC
Verification Hash: 88F7BA968E31EE9CD31F32DD395BFEC

Drive Geometry:
Total Size 12.7GB (20,712,000 sectors)
Cylinders: 28,266
Heads: 15
Sectors: 63

Partitions:
Code Type Start Sector Total Sectors Size
0C FAT32X 0 26700030 12.7GB
Hashing Tools – Examples

- [http://www.fileformat.info/tool/md5sum.htm](http://www.fileformat.info/tool/md5sum.htm)

Also, AccessData’s [FTK Imager](http://www.accessdata.com/downloads.html) can be downloaded free at [http://www.accessdata.com/downloads.html](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.”
  - e4d909c290d0fb1ca068ffaddf22cbd0


![This is a demonstration of hash values](image)
What happens when you rename a file?
Or rename the extension?
“Hashing” an Image

MD5
021509c96bc7a6a47718950e78a7a371
SHA1
77fe03b7ec063cf55dc268b19f5a449e9a97386

(single pixel changed using Paint program)

MD5
ead450e5e8cf1ac17cbeffcc09fb484
SHA1
01f577330fb6c16d58725c1decdfbeb8eb9cbe

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
Analysis (cont.)

- Existing Files
  - Mislabeled
  - Hidden
- Deleted Files
  - Trash Bin
  - Show up in directory listing with σ in place of first letter
    - "taxes.xls" appears as "σaxes.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 RAM)
- File A saved to disk, on top of File B
- File A (Saved To Disk)
- File A overwrites File B, creating slack
- File B (“Erased,” On Disk)
- Remains of File B (Slack)

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

Ways of Trying to Hide Data

- Password protection schemes
- Encryption
- Steganography
- Anonymous remailers
- Proxy servers

Password Protection

Ex: Secrethelper
**Encryption**

- Sometimes used as security measure to prevent others from accessing file data.
  - Example: "Pretty Good Privacy"
  - Scrambles file data so that it is unusable.

**Steganography**

SteganographyOriginal.png (200 × 200 pixels, file size: 88 KB)  
SteganographyRecovered.png (200 × 200 pixels, file size: 19 KB)
Another example

Selected “Trend”

“Triage” Forensics

“Triage” Forensics

- "Rolling" forensics, or "on-site preview"
- Image scan
- Especially useful in "knock & talk" consent situations, screening multiple computers to determine which to seize, or probation or parole monitoring
- Not all agencies equipped or trained yet to do this.
“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.

“When is enough enough?”

“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.
Resources

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

Questions?

662-915-6898
drmason@olemiss.edu
www.ncjrl.org