INTRODUCTION TO DIGITAL EVIDENCE & FORENSICS/WHAT IS CYBER CRIME?

Professor Donald R. Mason

OBJECTIVES:

After this session you will be able to:

1. Define “cyber crime”;
2. Define and describe “digital evidence”;
3. Identify devices and locations where digital evidence may be found;
4. Define basic computer and digital forensics; and
5. Identify and describe the basic practices, principles, and tools used in digital forensics.

REQUIRED READING:

Donald R. Mason, Introduction to Cyber Crime, Digital Evidence, and Computer Forensics (Mar. 2010) [NCJRL PowerPoint].................................................................1
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
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
  - Threats, harassment
  - Child pornography
- Computer as **Container**
  - From drug dealer records to how to commit murder

Murder!

- Studied currents
- Researched …
  - Bodies of water
    - including San Fran Bay
  - How to make cement anchors
  - Tide charts

Had 5 home computers
“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

Computer & Internet Uses

- Remote Computing
- Research
- Commerce
- Recreation
- Communication

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
- Personal digital devices
- Laptops
- Cell Phones

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

Digital Evidence

- **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

*According to report by technology research firm IDC
How Much Data?

- 1 Byte (8 bits): A single character
- 1 Kilobyte (1,000 bytes): A paragraph
- 1 Megabyte (1,000 KB): A small book
- 1 Gigabyte (1,000 MB): 10 yards of shelved books
- 1 Terabyte (1,000 GB): 1,000 copies of Encyclopedia
- 1 Petabyte (1,000 TB): 20 million four-door filing cabinets of text
- 1 Exabyte (1,000 PB): 5 EB = All words ever spoken by humans

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
Vehicle “black boxes”
- Event data recorders
- Sensing and diagnostic modules
- Data loggers
Chicago’s 911 Network

Room in Virtual World

Ex: Second Life
Cell Site Location Data

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 – Comments**

**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: 3FDSJ09U43JIVJU904FRBEWH
Verification Hash: 3FDSJ09U43JIVJU904FRBEWH

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\WELLESLEY.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: 88F7BA9EBEB33EEDC2AF312D0399FEC
Verification Hash: 88F7BA9EBEB33EEDC2AF312D0399FEC

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>FAT32</td>
<td>0</td>
<td>26700030</td>
<td>12.7GB</td>
</tr>
</tbody>
</table>

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."
  e4d909c290d0fb1ca068ffadef22cb0

What happens when you rename a file?
Or rename the extension?

“Hashing” an Image

MD5
021509c96bc7a647718950e78e7a371
SHA1
77fe03b7c0663c05ddc268b19f5a449e5a97386
(single pixel changed using Paint program)

MD5
ea8450e5e8cfa1c17c6effcc95b484
SHA1
01f57f3309b96c16d5872f5c1decdf6b88e69cbc
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 $\sigma$ in place of first letter
    - "taxes.xls" appears as "$\sigma$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

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

Ways of Trying to Hide Data

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

Password Protection

Ex: Secrethelper

![Password Protection Example](image)
**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 – Example**

- StenographyOriginal.png (200 × 200 pixels, file size: 88 KB)
- StenographyRecovered.png (200 × 200 pixels, file size: 19 KB)
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

- 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/
  - E.g., What Judges Should Know About Computer Forensics (2008)