Digital Evidence Locations and Computer Forensics

Don Mason
Associate Director

Objectives
After this session, you will be able to:

- Define and describe “digital evidence”
- Identify devices and locations where digital evidence may be found
- Define “computer forensics” and describe the basic practices, principles, and tools used in digital forensics

Advancing Technology
Computers Are Digital Devices

- A computer is like a light switch
  
<table>
<thead>
<tr>
<th>Switch</th>
<th>Computer</th>
<th>Binary Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>ON</td>
<td>signal present</td>
<td>1</td>
</tr>
<tr>
<td>OFF</td>
<td>no signal present</td>
<td>0</td>
</tr>
</tbody>
</table>

- Each 0 or 1 is a BIT (for BINARY DIGIT)
  
  - 0 0 0 0 0 0 1 = 1
  - 0 0 0 0 0 1 0 = 2 (2+0)
  - 0 0 0 0 0 1 1 = 3 (2+1)

- An 8-bit sequence = 1 byte = a keystroke

$$01000011 = \text{A}$$
Diagram of a Hard Drive or Floppy

How Data Is Stored

Clusters are groups of sectors

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

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

Data Generated in 2010
- 1200 trillion gigabytes (1.2 zettabytes)
- 89 stacks of books each reaching from the Earth to the Sun
- 22 million times all the books ever written
- Would need more than 750 million iPods to hold it
- 90 trillion emails sent in 2009

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%
Projections for 2006-2010

- Six fold annual information growth
- In 2020: 35 zettabytes will be produced
  - All words ever spoken by human beings, written 7 times
- 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
Challenges

- Increasing ubiquity and convergence of digital devices
- Increasing data storage capacity
- Shrinking devices and media
- Growing use of solid state devices
Internal Drives

Removable Media

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

More
More

GPS devices

Evidence Containers
"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 Basic Examples]

---

### Metadata – Track Changes

![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: 3FDSJ090U43JIVJU904FRBEWH
Verification Hash: 3FDSJ090U43JIVJU904FRBEWH

The Chances of two different inputs producing the same MD5 Hash is greater than:
1 in 340 Unidecillion: or 1 in 340,000,000,000,000,000,000,000,000,000,000,000,000,000
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: 8F7BA9EBEB33EEDECD2AF312DD39DFE6C
Verification Hash: 8F7BA9EBEB33EEDECD2AF312DD39DFE6C

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>

Hashing Tools – Examples
- [http://www.fileformat.info/tool/md5sum.htm](http://www.fileformat.info/tool/md5sum.htm)

Also, AccessData’s FTK Imager 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."
  e4d909c290d0fb1ca068ffaddf22c2d0

This is a demonstration of hash values.
What happens when you rename a file?

Or rename the extension?
“Hashing” an Image

MD5
021509c96bc7a6a47718950e78e7a371
SHA1
77fe03b07c0063cf35dc268b19f5a449e5a97386

(single pixel changed using Paint program)

MD5
ea8450e5e8cf1a1c17cd6effccdd95b484
SHA1
01f57f330fb05c16d5872f5c1decdfeeb88b69cbc

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 α in place of first letter
    - "taxes.xls" appears as "gaxes.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 2,000\textsuperscript{th} byte is slack space.

How “Slack” Is Generated

{\begin{center}
\begin{tikzpicture}

\node[align=center] (a) at (0,0) {File A \textbf{(In RAM)}};
\node[align=center] (b) at (1,0) {File A saved to disk, on top of File B};
\node[align=center] (c) at (2,0) {File A \textbf{(Saved To Disk)});
\node[align=center] (d) at (3,0) {File A overwrites File B, creating slack};
\node[align=center] (e) at (4,0) {File B \textbf{("Erased," On Disk)});
\node[align=center] (f) at (5,0) {Remains of File B \textbf{(Slack)});

\draw[->] (a) -- (b);
\draw[->] (b) -- (c);
\draw[->] (c) -- (d);
\draw[->] (d) -- (e);
\draw[->] (e) -- (f);

\end{tikzpicture}
\end{center}

\textit{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.

Encoded

Decoded
Steganography

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

Another example

What do you see?

- F-22s
- What else?
  - Embedded 121-page extract of a terrorist training manual
  - The F-22 image, the "carrier" file, is 2.25MB bitmap file (.bmp).
  - The "payload," the training manual extract, is a text file (.txt) that is only 227KB. So the payload easily fits in.
And another example

![Train Image]

The seemingly innocuous image of the train contains this 39,958 byte simulated child pornography image.
- Hidden using InPlainView, an application that employs the Least Significant Bit (LSB) Image Encoding technique.
- Technique manipulates the least significant bits, or smallest units, of the color components of selected bytes that represent the color of each pixel in image.
- Technique works because the change in color of each pixel is so slight it cannot be detected by the human eye.

What do you see?

![Image with Obstruction]

Recent Example

**Arrests of alleged spies draws attention to long obscure field of steganography**

By David Montgomery
Washington Post Staff Writer
Wednesday, June 30, 2010; 11:49 PM
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