RELIABILITY OF COMPUTER FORENSIC EVIDENCE: THE FORENSIC PROCESS & ITS OUTCOME

Professor Donald R. Mason

OBJECTIVES:

After this session, you will be able to:

1. Define and describe “digital evidence”;
2. Identify devices and locations where digital evidence may be found;
3. Identify and describe the basic principles, practices, and tools of digital forensics; and
4. Describe selected trends and challenges in computer forensics.

REQUIRED READING:

[NCJRL PowerPoint].......................................................................................................................1
Digital Evidence
and
Computer Forensics

Don Mason
Associate Director

National Center for Justice
and the Rule of Law
The University of Mississippi School of Law

Objectives
After this session, you will be able to:
- Define and describe “digital evidence”
- Identify devices and locations where digital evidence may be found
- Identify and describe the basic principles, practices, and tools of digital forensics
- Describe selected trends and challenges in computer forensics

Roles of Digital Devices
- Targets
- Tools
- Containers
Roles of Digital Devices

- **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

Computers are Digital Devices

- A computer is like a light switch
  - Switch | Computer | Binary Symbol
  - ON signal present 1
  - OFF no signal present 0

- Each 0 or 1 is a BIT (for BINARY DIGIT)
  - 00000001 = 1
  - 00000010 = 2 (2+0)
  - 00000011 = 3 (2+1)

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

Digital Data

- 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 and Computer Forensics
Technology Assisted Crimes Against Children: Evidentiary and Procedural Matters at Trial, August 1-2, 2011
Copyright © 2011 National Center for Justice and the Rule of Law – All Rights Reserved
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

---

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

---

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 2020: 35 zettabytes will be produced
  - All words ever spoken by human beings, written 7 times
- Compound annual growth rate: 57%

How Much in Real Cases?

- One recent example:
  - 17 terabytes
  - 24+ million images
  - 17,000 movies
  - 4600+ CVIP hits (known CP images)

Sources of Evidence

- Offender’s computer
  - accessed and downloaded images
  - documents
  - chat sessions
  - user log files
  - Internet connection logs
  - browser history and cache files
  - email and chat logs
  - passwords & encryption keys
Sources of Evidence

- Hand-held devices
  - digital cameras
  - PDAs
  - tablets
  - mobile phones
  - GPS devices

Sources of Evidence

- Servers
  - ISP authentication user logs
  - FTP and Web server access logs
  - Email server user logs
  - LAN server logs
  - “Cloud” storage
  - Web pages
  - Social media

Sources of Evidence

- Online activity
  - Internet Protocol addresses
  - Router logs
  - Third party service providers
Computer Forensics

- Obtaining,
- Processing,
- Authenticating, and
- Producing
digital data/records for legal proceedings.
Computer Forensics

- Usually pre-defined procedures followed but flexibility is necessary as the unusual will be encountered
- Was largely "post-mortem"
  - "What's on the hard drive?"
- Rapidly evolving
  - Ex:
    - From "Pull the plug"
      - to
    - "Don't power down before you know what's on it"

Branches, Evolutionary trends

- Computer forensics
- Network forensics
- Live forensics
- Software forensics
- Image forensics
- Mobile device forensics
- "Browser" forensics
- "Triage" forensics
- "Distributed" forensics

Digital Knowledge and Intent Evidence

- Evidence that the CP files were purposely collected
  - CP found in computer's allocated space?
  - In folders assigned to particular "user" of the computer?
  - Files organized, given relevant folder/file titles?
  - Default settings of the computer's software changed?

- Evidence that CP was obtained via Web browsing
  - Evidence in the Index.dat files of web searches for CP?
    - CP found in the Temporary Internet Files?
    - Any CP-related Bookmarks/Favorites saved?

- Evidence that the CP was viewed by a user
  - Any Recent Files/Link Files to the CP?
    - Windows Registry list other devices (scanners, thumb drives, etc.) recently connected to the computer?
    - Any Thumbs.db files containing CP?
    - Any CP videos listed in Windows Media Player/Real Player histories?
Basic Steps

Acquiring (and preserving) evidence without altering or damaging original data

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

Analyzing (searching for) the evidence without modifying it

Popular Automated Tools

Encase
Guidance Software

Forensic Tool Kit (FTK)
Access Data

Skills / Expertise Required

- Technical
  - Data processing and production
- Investigative
  - Understanding computer evidence
  - Building a case
- Legal
  - Maintaining chain of custody
  - Managing digital evidence per the rules
Certifications

- Various offered
  - IACIS's “CFCE”
  - Guidance Software’s “Encase CE”
  - ISFCE’s “CCE”
- Some states require P.I. licenses
- Growing number of schools offering certificate and degree programs
- But no uniform, accepted standards

Acquiring the Evidence

- Seizing computer (“bag and tag”)
- Handling computer evidence carefully
  - Chain of custody
  - Evidence collection (including volatile memory)
  - Evidence identification
  - Transportation
  - Storage
- Making at least two images of each container
  - Perhaps 3rd in criminal case
- 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
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:
3FDSJO90U43JIVJU904FRBEWH

Verification Hash:
3FDSJO90U43JIVJU904FRBEWH

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

Digital Evidence and Computer Forensics
Technology Assisted Crimes Against Children: Evidentiary and Procedural Matters at Trial, August 1-2, 2011
Copyright © 2011 National Center for Justice and the Rule of Law – All Rights Reserved
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.”
  - e4d909c290d0fb1ca068ffaddf22cb70

[MD5 Hash examples](http://www.miraclesalad.com/webtools/md5.php)

---

**“Hashing” an Image**

<table>
<thead>
<tr>
<th>Image 1</th>
<th>MD5</th>
<th>SHA1</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.jpg" alt="Image 1" /></td>
<td>021509c36bc7a8a477118956e78e7a371</td>
<td>77fe03bb70063cfd5dc26b19f5a449e5a97386</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Image 2</th>
<th>MD5</th>
<th>SHA1</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image2.jpg" alt="Image 2" /></td>
<td>eab450e5e8c9a1c17c6effccf9b48</td>
<td>0f1e5f2339b9c6c1d2f88b669cabc</td>
</tr>
</tbody>
</table>
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

Analysis (cont’d)

- 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

Forms of Evidence

- Files
  - Present / Active (doc’s, spreadsheets, images, email, etc.)
  - Archived (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
  - Paragraphs
  - Sentences
  - Words
**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

**How Data Is Stored**

- Track
- Sector
- Clusters are groups of sectors
How Data Is Stored

Files are written to **Clusters**

- Each file may occupy more or less than full clusters
- May write to non-contiguous clusters

**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
Recall how data is stored

How Data Is Stored
- Every file in a computer fills a minimum amount of space
  - In some old computers, one kilobyte (1,024 bytes). In newer computers, 32 KB (32,768 bytes).
  - If file is 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.
EXIF Data

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

Ways of Trying to Hide Data

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

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.

Temporary Internet Files

---

Digital Evidence and Computer Forensics
Technology Assisted Crimes Against Children: Evidentiary and Procedural Matters at Trial, August 1-2, 2011
Copyright © 2011 National Center for Justice and the Rule of Law – All Rights Reserved
“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/
  - E.g., What Judges Should Know About Computer Forensics (2008)

Questions?

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