Digital Evidence 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
- Identify and describe the basic principles, practices, and tools of digital forensics
- Describe selected trends and challenges in computer forensics

From the “old days” to ...
Evolving technology in …

The “Digital age” with …

Convergent, “Smart” Devices
“Post-PC Era”?  

Cellular phone a “computer”?  
- Yes, as defined in Computer Fraud and Abuse Act  
- Ultimately, does it make any difference whether a device capable of storing digital evidence is deemed to be a “computer”?

Computers = *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 0 1 | = 1 |
  | 0 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  
  
  | 0 1 0 0 0 0 0 1 | = A |
Digital Devices

The Investigative Future is Here

- Criminal Connectivity:
  - iPads
  - Kindles
  - iTouches
  - E-Readers
  - Appliances!
- From homes, offices, coffee shops, airplanes, cars, buses, trains, … almost anywhere

Always Something New
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
Digital Evidence

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

Two types

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/Network-created

- Email headers
- Metadata
- Activity 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

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
- 1 Zettabyte (1,000 EB, or 1 billion TB) = 250 billion DVDs, 36 million years of HD video, or the volume of the Great Wall of China

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
- 107 trillion emails sent in 2010
Projection

- In 2020: 35 zettabytes will be produced
  - All words ever spoken by human beings, written 7 times

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

- Servers
  - Internet Protocol addresses
  - ISP authentication user logs
  - FTP and Web server access logs
  - Email server user logs
  - Subscriber account information
  - LAN server logs
  - “Cloud” storage
  - Web pages
  - Social media

Sources of Evidence

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

"inside the box, outside the box"
Inside the Box
What the computer owner actually has possession of

- Computer’s hard drive and other memory
  - Documents
  - Pictures
  - Outlook Emails
  - Internet Cache
- CD’s and floppy disks
- iPods
- Cell Phones
- External Hard Drives

Outside the Box
What is not stored on the owner’s computer

- Online Email Accounts (Gmail and Yahoo)
- Internet Shopping Accounts
- Social Networking Accounts
- Backups of text messages
- Cell Site Location Data
- Using Pen/Trap for Internet “DRAS” information
- Subscriber account records
- Contents of Websites
Outside the Box
What is not stored on the owner's computer

Variety of “Boxes”

Computer Hardware

Digital Evidence and Computer Forensics
Copyright © 2012 National Center for Justice and the Rule of Law – All Rights Reserved
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
And Still More

Remember this news item?

More
Vehicle “black boxes”
- Event data recorders
- Sensing and diagnostic modules
- Data loggers
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”
Terms, Branches, 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 Unidecillion: or 1 in 340,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000

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.”
  e4d909c290d0fb1ca068ffaddf22cbd0


“Hashing” an Image

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  in place of first letter
    - “taxes.xls” appears as “œaxes.xls”
- Free Space
- Slack Space

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

- 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.

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
Pop Quiz

How can you reliably “destroy” data?

Jackhammer hard drive shredder

Slack Space

Space not occupied by an active file, but not available for use by the operating system

How “Slack” Is Generated

File B (Draft in RAM) → File B saved to disk, on top of File A → File B (Saved to disk) → File B overwrites part of File A, creating slack → Remains of File A (Slack) → Slack space: The area between the end of the file and the end of the storage unit → File B (Now on disk)

File B (Draft in RAM) → File B saved to disk, on top of File A → File B (Saved to disk) → File B overwrites part of File A, creating slack → Remains of File A (Slack) → Slack space: The area between the end of the file and the end of the storage unit → File B (Now on disk)

File B (Draft in RAM) → File B saved to disk, on top of File A → File B (Saved to disk) → File B overwrites part of File A, creating slack → Remains of File A (Slack) → Slack space: The area between the end of the file and the end of the storage unit → File B (Now on disk)
Selected Developments 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

- “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

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.

Tool Example: osTriage

- "Live response tool"
- Developed by F.B.I. SA in SLC
- Free to U.S. law enforcement
- Validated by F.B.I. November 2011
- 43 MB software package
- Run from USB storage (e.g., thumb drive or external hard drive)

osTriage – Reasons to Use

- Increasing use and ease of “virtualization”
  - May be multiple additional “computers”
- Increasing use of free & low cost encryption
- Loss of valuable info when computer is rebooted
- Loss of visibility of network storage
- Saves time
osTriage – Capabilities

- Display comprehensive details
  - User accounts
  - Physical and logical hard drives
  - Mapped networked drives
  - NIC information
  - Every USB device ever inserted into machine
  - Browser history
  - “Flash cookies”
  - Applications running (e.g., P2P or encryption)

osTriage – Capabilities

- Searches drives, finds images/videos, displays thumbnails
- Allows easy copying of contraband images, videos to USB storage device
- Compares images/videos to SHAs
- Checks files names against keyword list
- Has built-in image viewer
- Supports viewing any EXIF data and thumbs.db

osTriage – Capabilities

- Extracts saved passwords
- Extracts list of recently opened files
- Writes nothing to computer being scanned
- Allows for custom searches
- Looks inside archives for key word filenames
- Gathers and saves volatile data before shutdown
osTriage - Limitations

- Cannot find and display data no longer there (e.g., cleared browser history)
- Doesn’t look for deleted files
- Doesn’t look at file headers to identify images or videos
- Does not substitute for full, forensically sound examination of device, if needed

Ways of Trying to Hide Data

- Password protection schemes
- Encryption
- Steganography
- Anonymous remailers
- Proxy servers
- Changing File Extensions

Password Protection

- Computer/BIOS Passwords
- Encryption Programs
- Archive Passwords
- Document Passwords
Changing File Extensions

Encryption

- Sometimes used as security measure to prevent others from accessing file data.
  - Examples: "Pretty Good Privacy" and "Truecrypt"
    - 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)

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 Remember the Cloud

Questions?

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