I. Define technology uses and digital evidence

II. Provide an overview of “traditional” computer forensics

III. Describe advances in computer forensics
   1. Network Forensics
   2. Live Forensics
   3. Remote Forensics
   4. Software Forensics
   5. Image Forensics
   6. Browser Forensics
   7. Triage Forensics

IV. Discuss the new era of Mobile Device Forensics

I. DEFINING DIGITAL EVIDENCE AND HOW PREDATORS CREATE IT
WHAT IS DIGITAL FORENSICS?

Branch of forensic science encompassing the recovery and investigation of digital evidence found in digital devices.

INSIDE V. OUTSIDE THE BOX

What is stored on your computer

- Hard drive & other memory
- Docs & pics
- Outlook emails
- Internet cache
- CD’s and floppy disks
- iPods
- Cell phones
- External hard drives

What is not stored on your computer

- Online email accounts (e.g., Gmail or Yahoo)
- Internet shopping accounts
- Social networking accounts
- Text message backups
- Cell site location data
- Using pen/trap for Internet “DRAS” information
- Subscriber account records
- Contents of Websites
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)

- 00000001 = 1
- 00000010 = 2 (2+0)
- 00000011 = 3 (2+1)

An 8-bit sequence = 1 byte = a keystroke

- 01000001 = A

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 exploitation

- Computer as **Container**
  - From drug dealer records to how to commit murder

CHALLENGES

- Increasing ubiquity and convergence of digital devices
- Increasing data storage capacity
- Shrinking devices and media
- Growing use of solid state devices
DIGITAL EVIDENCE
- Information of probative value that is stored or transmitted in binary form and may be relied upon in court

**User-created**
- Text (docs., e-mail, chats, IM's)
- Address books
- Bookmarks
- Databases
- Images (photos, drawings)
- Video and sound files
- Web pages
- Service provider account subscriber records

**Computer-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

INTERNET DISTRIBUTION
- E-mail
  - File transfer protocol (FTP)
  - Bulletin and message board systems
  - Newsgroups
  - Chat rooms

- Web sites
  - Webcams
  - E-groups
  - Instant messaging
  - Peer-to-Peer Networks
INTERACTIVE DISTRIBUTION

- Interactive Internet communities
- Social networks
- Blogs
- Wikis
- Video or photo sharing sites
- Online role-playing games
- Virtual worlds

II. COMPUTER FORENSICS

The collection and analysis of data from computer media to locate digital evidence in a manner that will make the evidence admissible in a court of law

COMPUTER FORENSICS

- The application of computer investigation and analysis techniques to gather digital evidence
  - AKA Cyber Forensics
- Generally, the process entails pre-defined procedures followed, but flexibility is necessary
THE TRIPLE-A PROCESS

- Acquire
  - Without altering or damaging original data – locate, obtain, and preserve

- Analyze (search)
  - Without modifying it

- Authenticate
  - Show it’s the same as data seized for legal proceedings

ACQUISITION

In a “forensically sound manner”

- Photograph, secure, and seize (“bag & tag”)
- Handle evidence carefully (collection; chain of custody; storage)
- Document all steps – image drive, load it into a forensic software suite, and index and analyze it

Evolving from Dead Box forensics

  - “Cut the power immediately”
  - “Don’t power down before you know what’s on it”

ANALYSIS

- Working on bit-stream images of the evidence; never the original
  - Prevents damaging original evidence
  - Two backups

- Analyzing everything
  - Clues may be found in areas or files seemingly unrelated
**ANALYSIS TOOLS**
- Password crackers
- Hard drive tools – Fdisk / Linux
- Viewers – QVP / Disk view
- ThumbsPlus
- Unerase tools
- CD-R utilities
- Text search tools
- Drive imaging – FTK (forensic tool kit); EnCase; X-Ways, Prodiscover; Autopsy (bootable CDs for imaging include Deft and CAINE)

**AUTHENTICATION**
- Proving evidence to be analyzed / offered is exactly same as what was seized
  - Readable text and pictures don’t magically appear at random
  - Calculating hash values for original evidence and duplicates
    - Hash value = a mathematical algorithm of the original media (Essentially analogous to a file’s "fingerprint" or DNA)
    - MD5 (Message-Digest algorithm 5)
    - SHA (Secure Hash Algorithm) (NSA/NIST)

**AUTHENTICATION**
Preserving Evidence = “Forensic Image” of entire hard drive

- A virtual copy of the entire drive
  - Every bit & byte
  - “Erased” & reformatted data
  - Data in “slack” & unallocated space
  - Virtual memory data
MD5 Hash

- 128-bit (16-byte) message digest – a sequence of 32 characters
- Analogous to a file’s “fingerprint” or DNA
- Like a unique file name

Basic illustration:

“The quick brown fox jumps over the lazy dog”
9e107d9d372bb6826bd81d3542a419d6

“The quick brown fox jumps over the lazy dog.”
e4d909c290d0fb1ca068ffaddf22cbd0

“Hashing” Data Sets

MD5 Hash values:
Acquisition Hash: 3FDSJO90U43JIVU904FRBEWH
Verification Hash: 3FDSJO90U43JIVU904FRBEWH

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
"HASHING" AN IMAGE

<table>
<thead>
<tr>
<th>MD5</th>
<th>SHA1</th>
</tr>
</thead>
<tbody>
<tr>
<td>021509c96bc7a5a477f8950e78a7a371</td>
<td>77fe03b07c0063cf35dc268b1f5a449e5a97386</td>
</tr>
<tr>
<td>M5</td>
<td>SHA1</td>
</tr>
<tr>
<td>ea8450e5e8c8f1af1c17c6effcc95b484</td>
<td>01f573330b06c16e55725fc1decdf6b88969cbc</td>
</tr>
</tbody>
</table>

OTHER WAYS OF AUTHENTICATING

- Witness with knowledge
- Process
- Chain of custody, etc.

Prima facie showing, per Rule 901, that evidence is what it is purported to be

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

File B (Now on disk) → Remains of File A (Slack)

Slack space: The area between the end of the file and the end of the storage unit
(2) Is it powered on?
   (a) If no, continue process
   (b) If yes, turn off computer

(3) Remove hard drive from target system

(4) Attach hard drive to forensic system

(5) Make imaged copy and analyze it

---

**“DEAD BOX” STRENGTHS**

- Simple – clearly defined and straightforward stages of the acquisition, which can be verified at any time
- Reliable – no risk of altering / contaminating evidence as the data on the hard disk is not modified
- Thorough – analysis of every byte on the hard disk is acquired including unallocated, slack space and metadata

---

**“DEAD BOX” WEAKNESSES**

- Over-inclusive with data acquisition (can be invasive)
- Logistical problems with moving extensive amounts of computer equipment
- Encryption (an exact copy of an encrypted file is of no use)
- Increasing storage capacity of target devices
- Understanding that some evidence is lost by immediately powering down device
  - Decryption keys and network data may be stored as volatile data and lost when the computer is turned off
III. “NEW DIGITAL FORENSICS”

NEW BRANCHES

NETWORK FORENSICS
SOFTWARE FORENSICS
REMOTE FORENSICS
IMAGES FORENSICS
TRIAGE FORENSICS
COMPUTER FORENSICS

NETWORK FORENSICS

- Captures, records, and analyzes network events to discover the source of security attacks or other incidents; reveals specifics about who's using a network and how they're using it.

Typically used by companies hiring network forensic monitors – two methods:
- “Catch it as you can” – all packets passing through a certain point are captured; requires large storage abilities
- “Stop, look and listen” – each packet is initially analyzed and only certain info is saved for future analysis; requires less storage but may require a faster processor
**NETWORK FORENSICS**

- **Strengths:**
  - Useful for security planning
  - Can ID system abuses before they occur
    - Detect clues of corporate espionage or unauthorized use
    - Find hackers and analyze how the attack was carried out to avoid repeat attacks

- **Weakness:**
  - Usually requires a high level of skilled resources to be effective

**LIVE FORENSICS**

- **a live analysis of programs on a suspect’s computers to acquire RAM, unencrypted files, and other pertinent data**

- **Must stay within SW and minimize modified data**

- **Functions:**
  - Determine if possible encryption;
  - Perform a live image of hard disks (limited guidelines though);
  - Assess the scope of the local network and identify other devices attached to the network

- **Should be seen as an improvement to the process, not a total replacement of the traditional method**

- **Strengths:**
  - Avoids logistical problems: seizing a large number of devices for later analysis, which costs considerable resources

- **Weaknesses:**
  - Increasing target size
  - Difficulty in locating target (unidentified machines)
  - Analyzing drives of a large group of machines consumes considerable resources
REMOTE FORENSICS

= the ability to access a computer or network from a remote distance and copy data from it

- Tools vary in configurations and capabilities
- Two choices for civil firms:
  - USB drives – ship a remote collection software "toolkit" loaded onto a thumb drive or hard drive to custodian
  - Web-based tools – use Internet / network connection; the data collected is generally stored locally on a USB drive; allow investigators immediate access

Strengths:
- Allows you to preview suspect's drive while in use; perform a live acquisition; encrypt connection; and copy RAM
- Captures volatile system info
- Analyzes current running processes
- Cost savings for clients (civil)

Weaknesses:
- Data transfer speeds can cause problems
- Gaining access permission may be difficult
- Heavy traffic could cause delays / errors

SOFTWARE FORENSICS

= analyzes computer program authorship by using measurements from source code; ascertain the circumstances that led to the defect in code or a malicious application

- Four types:
  - Author Discrimination (one or more)
  - Author Characterization (programming style)
  - Author Intent Analysis (deliberate intent (malice) v. accidental error)
  - Author Identification
- Crime examples: malicious code; plagiarism; and computer fraud
SOFTWARE FORENSICS

- Case Study
  - The Internet Worm—written by Robert Morris and released onto the Internet on November 1988
    - Analysis & Conclusions:
      - The code was not well written; not portable; probably not checked using lint
      - The code contained little error handling behavior, suggesting that the author was sloppy and performed little testing
      - The code seemed to have been written over a long period of time
      - Etc.
    - This indicates the amount of knowledge that can be extracted from source code

IMAGE FORENSICS

- extracts and analyzes traces from digital content to determine if image is doctored or similar to others
  - Helps alleviate the work of manually examining a large number of digital pictures in order to identify potential evidence
  - EX: Content-based image retrieval techniques allows for queries for contraband images and queries for images related to some known images

SPOTTING FAKE PHOTOS

- Clone location—finding cloned regions by using algorithms of 6x6 blocks
Strength:
- Analysis of image alterations can provide very useful information

Weaknesses:
- No real comparative study exists to evaluate the actual competency of DIF tools
- Rise of counter-forensics exposes the limitations of DIF tools

**IMAGE FORENSICS**

**BROWSER FORENSICS**
- Uses programs to help rebuild cached browser pages and present the data in the same format as it was viewed by the suspect
- Can be seen as not a separate discipline, but rather a subset of identification through programs
- Program Example: NetAnalysis: Browser Artefact Recovery and Analysis – developed by digital forensics practitioner working for a police Digital Forensics Unit in the UK
  - The Program is a powerful tool to search for, filter through, and identify target information

**TRIAGE FORENSICS**
- A procedural model for the investigation of digital crime scenes including both traditional crime scenes and the more complex battlefield crime scenes
- AKA “rolling forensics” or “on-site preview forensics”
- Allows screening of multiple computers to determine which to seize or monitor
**TRIAGE FORENSICS**

- **Strengths:**
  - Avoids the loss of valuable info that results from computer shutdown
  - Saves time
  - Increasingly important as the number and storage capabilities of devices grows
  - Useful in consent situations

- **Weaknesses:**
  - Not all agencies are equipped or trained yet to do this
  - Does not enable a comprehensive forensically sound examination

**EX: osTRIAGE**

- Primarily built for child exploitation investigations
- Can be used for any computer investigation
  - Provides details: most every USB device ever plugged into the computer (make, model, and serial #), full browser history, registry keys, FirstFolder, TypedPaths, and many more, extracts passwords from p2p, email, chat, and other sources, full network details, ARP cache, DNS cache, open ports, running processes, installed software, and on and on.
  - Allows for adding additional items of interest beyond what is included.

**CLOUD FORENSICS**

- A sub-network of digital forensics that involves retrieving evidence which are not under the control of the user and may not be identified and located easily due to the dynamic nature of the cloud

- Unanswered Questions:
  - Is there a reasonable expectation of privacy on the cloud?
  - Does the SCA apply to cloud stored data?
  - How does the warrant requirement apply to cloud-stored data?
  - Authenticating cloud data for trial will turn on the validity of the system used to logically allocate data to individual users
MOBILE DEVICE FORENSICS

- The science of recovering digital evidence from a mobile phone under forensically sound conditions using accepted methods

- If the device is found on, do not switch it off, if the device is off, do not switch it on
- Forensic tests can be carried out once the phone is seized
- Cell site location analysis can be carried out without the seized phone

A DIFFERENT STORY

Special Challenges
- Preventing loss of data
- Stopping overwriting or wiping
- Removing from network
- Finding passwords and working around locks
- Diversity of makers, models, hardware
- Phones
- SIM cards
- Removable media
- Diversity of platforms and operating systems
- CDMA, GSM, et al.
- iOS, Android, et al.
- Array of diverse tools that don’t work the same with all devices
- Data recovery may be via
  - Physical acquisition (write protection; “chip-off”)
  - Logical analysis (capturing active files, etc.)
  - Manual extraction (e.g., screen shots with camera)
IS IT “FORENSICS”?

It is “forensics” - follows computer forensics process
- Identification
- Preservation
- Collection
- Examination
- Analysis
- Reporting

It is not “forensics” – not the same as computer forensics process
- Works on original evidence
- Writes to phone
- Deleted data may not be recoverable
- Process documentation may be more important for authentication
- “Hashing” may not be possible
- Some advocate “cell phone data recovery”

RECOVERABLE DATA

- Contacts
- Recent call history
- Ringtones
- Apps
- Custom wallpapers
- Voice mail
- Text messages (SMS)
- Calendar
- Memos
- GPS info
- Multimedia (MMS) files
  - Audio
  - Images
  - Videos
- E-mail
- Browser history
- Chats
- Documents

RECOVERABLE DATA

- Depends on
  - Phone model
  - Service provider plans and capabilities
  - Acquisition tools
  - Analysis software
FINAL OBSERVATIONS

- Mobile devices increasingly important
- Mobile devices contain wealth of probative information
- Multiple tools must be used as needed
- Results must often be verified manually

QUESTIONS?

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