Digital Forensics

Raj Jain
Washington University in Saint Louis
Saint Louis, MO 63130
Jain@cse.wustl.edu

Audio/Video recordings of this lecture are available at:

http://www.cse.wustl.edu/~jain/cse571-17/

Overview

1. What is Digital Forensics?
2. Digital Artifacts
3. Network Forensics
4. Sources of Network-based Evidence
5. Mobile Device Digital Artifacts

These slides are based partly on the "Introduction to Digital Forensics Course" developed by University of Illinois's NSF-supported Digital Forensics Curriculum development initiative.

What is Digital Forensics?

- *Forensics*: The application of science to legal problems and investigations.
- *Digital forensics*: A branch of forensics involving the recovery and investigation of digital evidence.
- *Digital evidence*: Data stored or transmitted using a digital device: Computer, Storage, Network, Mobiles.
  - Computer Forensics
  - Storage Forensics
  - Network Forensics
  - Mobile Forensics

Evidence Integrity

- Should ensure evidence is not damaged or altered.
- Any alteration must be precisely documented.
- Avoid loss of evidence.
- Avoid introducing artifacts that may confuse investigation.
- Don't do examination on the original device.
  Make a forensic duplicate. NIST Computer Forensics Tool Testing (CFTT) certified tools.
- Verify integrity by computing crypto hashes
- Turn-off computer/Mobile? Depends.
  May loose evidence such as open decrypted files

Network Forensics: Example

- Justin Ross Harris was charged with murder of his toddler son, who died after being left in a hot car.
  - Internet browsing history on Justin's work computer showed he had searched for information on how child deaths occur in cars and how hot it needs to be to kill them.
  - He was eventually found guilty.

Storage Forensics: Example

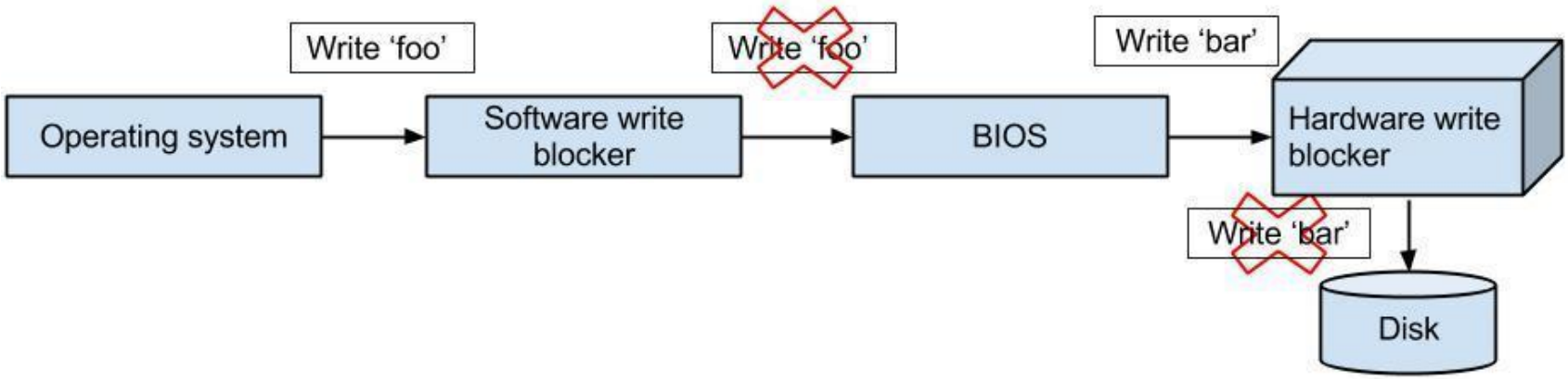
- "BTK" serial killer Dennis Rader murdered 10 people in Kansas across 3 decades.
- He sometimes sent mocking messages to the police.
- In 2005 he sent a message asking the police if they could trace the source of documents on a floppy disk; via a newspaper ad, they told him no.
- He sent a message on a floppy disk, which also included a deleted Word document whose metadata referred to the Christ Lutheran Church in Wichita (where he was president of the church council) and named "Dennis" as the last person to modify the file.
- Within 10 days of mailing the disk, he was arrested.

Digital Artifacts

- Operating system: Event logs, Registry data.
- File system: access times, modification.
- Disk: deleted files, hidden partitions.
- Internet: browser history, e-mail.
- Media: photos, videos, audio.
- Documents: Office, PDFs, RTF, XML.
- Databases: MySQL, Oracle.
- Application data: instant messaging.

Write Blockers

- **Software write blockers** can block the message from the OS.
- **Hardware write blocker** is required to block the message from the BIOS.



Metadata Files

- Files that contain data about the file system itself.
  - $MFT: master file table (file records).
  - $MFTMirr: mirror of $MFT.
  - $LogFile: transaction log to recover from failure.
  - $BadClus: bad clusters.
  - . Root directory.
  - $Secure: special access control list database.

File Recovery vs. Carving

- **Recovery**: file system metadata are intact; use them to find file ("undelete").
- **Carving**: pulls the ***raw*** bytes from the media. Recovery of files ***without*** file system metadata.
  - Header-footer carving: All JPEG files start with 0xFFD8 and ends with 0xFFD9.
    Tools: Scalpel, Foremost, EnCase.
  - File-structure-based carving: Use internal file metadata
    Tools: Foremost, PhotoRec
  - Content-based carving: statistical signatures indicating language or file content.

Memory Carving

- Snapshots (images) of RAM can be taken by some tools (dd, dev, mem).
  - Just like a disk.
- Crash dumps.
- Sleep images are left behind (maybe unencrypted) by some OSes.
  - E.g., /private/var/vm/sleepimage.
- RAM is highly fragmented.

Computer Forensics

- Windows Registry.
- Event logs.
- Link files.
- Recycle Bin
- Application Metadata: E.g., Microsoft Office
  - List of last 10 authors.
  - A list of revisions, along with who made them, if "track changes" is on.
  - Hidden annotations and comments.
  - Created, last written, and accessed times.
- Browser History and Cache
- Prefetch Files: Stored in C:\Windows\Prefetch*.pf

Network Forensics

- E-mails and IM sessions; browser activity; routinely kept packet logs; /var/log/messages
- Depending on the method of capture, packet capture may be treated as recordings of events, i.e., there are **privacy and legal implications**, so you must be careful.

Sources of Network-based Evidence, 1

- **Wireless access points**.
  - Easy to eavesdrop on traffic that passes through.
  - Management and control frames are not encrypted.
- **Switches** (devices that connect computers together).
  - Contain tables that store, e.g., mappings between physical ports and devices' MAC addresses.
  - Traffic can be captured from switches as well.
- **Routers** (devices that connect segments of networks together).
  - Contain routing tables that map router ports with networks.
  - Routers may also function as packet filters.

Sources of Network-based Evidence, 2

- **DHCP servers** (which give out IP addresses via a temporary leasing system; every machine needs an IP address to connect to the Internet).
  - DHCP server logs contain IP address to MAC address mappings and the times the IP addresses were leased.
- **DNS servers and name servers** (which convert domain names to IP addresses).
  - Logs may reveal connection attempts from internal to external websites, e-mail, or SSH servers, with time stamps.
- **Authentication servers** (which, e.g., validate username & password credentials).
  - Log successful/failed login attempts for all devices within their authentication domains.

Sources of Network-based Evidence, 3

- **Network intrusion detection/prevention systems (NIDS/NIPS)** (which monitor what's coming in and going out of a network).
  - May detect attacks in progress.
  - Can be tuned to give more granular data.
- **Firewalls** (which block incoming and/or outgoing traffic).
  - Firewalls can be configured to keep logs, which can serve as digital evidence.
- **Web proxies** (intermediary servers that support clients' access to Web content).
  - An enterprise's Web proxy can store Web surfing logs for the entire organization.

Sources of Network-based Evidence, 4

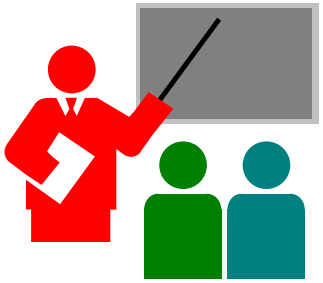
- **Application servers**.
  - E.g., database, Web, e-mail, chat, and voicemail servers. These can be very important.
- **Local network diagrams** and applications.
  - You can review these to find out which servers are useful for a particular investigation.

Flow Record Processing System

- A *flow record processing system* captures, stores, and then processes information about the flows of interest.
- It can include these components:
  - **Sensor**: The device that monitors flows and extracts important bits to a flow record. Does actual capture.
  - **Collector**: A server that receives and stores flow records from sensors. (Sensors have limited storage.)
  - **Aggregator**: Aggregates data from multiple collectors.
  - **Analysis**: Software tools are used for analysis.
- E.g., Wireshark includes all 4 components in 1 tool.

Mobile Device Digital Artifacts

- Internal memory (NAND).
  - GPS, text messages, call records.
  - Photos, contacts.
  - Email, browser history.
- External memory (SIM cards).
  - Subscriber data.
  - Extra storage for internal memory.
- Service provider logs.
  - Call records, cell usage.
  - Location parameters.
  - Text/multimedia messages.
  - Intercepted data.



Summary

1. Digital Forensics involves recovering digital evidence from computer, storage, network, or mobile devices
2. Recovering files without metadata is called 'File Carving'
3. Computer forensics requires looking at registry, logs, link files, recycle bin, and application meta data
4. Network forensics requires looking at emails, instant messages, traffic, browser histories
5. Network evidence can be collected by access points, switches, routers, DHCP servers, DNS Servers, Authentication Servers, NIDS, Firewalls, and web proxies
6. Mobile device artifacts includes contents of internal memory, SIM card, Service provider logs

Lab 25: Autopsy

- Autopsy is a forensic tool commonly used to analyze disk images
- Read about the software
  - https://www.sleuthkit.org/autopsy/v2/
- Download the software
  - Mac: http://macappstore.org/autopsy/
  - Windows: https://www.sleuthkit.org/autopsy/download.php
- You will analyze a Remix OS image, read about this OS at: https://en.wikipedia.org/wiki/Remix_OS
- Download remix img: _http://files.pine64.org/os/remix/remix-v2.0-20160415-pine64-32GB.zip
- Unzip the file to get the image file

Lab 25 (Cont)

- Launch autopsy and create a new case
- Add the downloaded image to source data
- Select "all ingest" module and let the program download all files (It takes a while)
- Expand the view tree to see all the files that exist in this image

1. Generate a timeline on the image and answer the following questions (Submit screenshots to support your answer):
   a. At what timeline most of the activities where recorded ?
   b. Name a file that is modified and give the date
   c. Name a file that is created and give the date
   d. Name a file that is accessed and give the date
   e. Name a file that is and give the date changed

Lab 25 (Cont)

f. Is there any web activities recorded, what type?

g. Is there any MISC activities recorded, what type?

h. Generate a snapshot record and submit a timeline snapshot

2. Answer the following "View Tree" questions (Submit screenshots to support your answer):

a. How many image files can you see? HTML? Audio? Executable?

b. Locate notice.html, when was this file modified? What is its MD5 hash?

c. How many system file have been deleted ?

Lab 25 (Cont)

d. Locate bt_config.new in the deleted file, when was the file deleted? Assuming that you have a recovery software, can you still recover it? Why or why not?
e. Locate .booting in the deleted file, when was the file deleted? Assuming that you have a recovery software, can you still recover it? Why or why not?

3. Answer the following "Results Tree" questions (Submit screenshots to support your answer):
   a. How many contacts file are recorded?
   b. How many email messages are recorded?
   c. How many email addresses are recorded?
   d. Locate one of the email messages recorded
      - Who are the sender and receiver ?
      - What date was it sent?
      - What is the subject of the message ?

Acronyms

- ASCII — American Standard Code for Information Interchange
- BIOS — Basic Input/Output System
- CFTT — Computer Forenscis Testing Tool
- CS — Computer Science
- DHCP — Dynamic Host Control Protocol
- DNS — Domain Name Server
- GPS — Geo Positioning System
- IM — Instant Message
- IP — Internet Protocol
- JPEG — Joint Picture Expert Group
- MAC — Media Access Control
- MFT — Master File Table
- MO — Missouri
- MySQL — My Structured Query Language
- NAND — Not And
- NIDS — Network Intrusion detection system

Acronyms (Cont)

- NIPS Network Intrusion Prevention system
- NIST National Institute of Science and Technology
- NSF National Science Foundation
- OS Operating System
- RAM Random Access Memory
- RTF Rich Text Format
- SIM Subscriber Information Module
- SSH Secure Shell
- XML Extensible Markup Language

Scan This to Download These Slides

Raj Jain
http://rajjain.com

Related Modules

CSE571S: Network Security (Spring 2017),
http://www.cse.wustl.edu/~jain/cse571-17/index.html

CSE473S: Introduction to Computer Networks (Fall 2016),
http://www.cse.wustl.edu/~jain/cse473-16/index.html

Wireless and Mobile Networking (Spring 2016),
http://www.cse.wustl.edu/~jain/cse574-16/index.html

CSE571S: Network Security (Fall 2014),
http://www.cse.wustl.edu/~jain/cse571-14/index.html

Audio/Video Recordings and Podcasts of
Professor Raj Jain's Lectures,
https://www.youtube.com/channel/UCN4-5wzNP9-ruOzQMs-8NUw