National District Attorneys Association
National Center for Prosecution of Child Abuse

Computer Forensics for Prosecutors

February 18-19, 2013 ● Portland, Oregon

Detective Michael Smith
Computer Crimes & Computer Forensics
Linn County Sheriff's Office
Voice: 541-812-9900
Email: msmith@cc.linn.or.us

Special Thanks To

- Lt. Joshua Moulin, Southern Oregon High Tech Crimes Task Force
- Lt. Jason Rampolla, Park Ridge (NJ) Police Department
- Richard Kahlan, Computer Forensic Specialist, USDOJ CEOS
- Daniel Hisaki, Windows Azure and SkyDrive Forensics, Microsoft Corporation
- Katherine Smythe, Software Engineer, Google Android Team
  - For their willingness to collaborate and share ideas in the digital world

Objectives

- Be able to identify sources of technical investigations
- Understand common terms related to computer hardware
- Understand how the Internet works and how IP addresses are assigned
- Understand how data is written, stored and deleted from storage devices
- Understand what are backdoors and how they work
- Understand the content of a computer forensics report

Sources of Investigations

- Walk-in complaints from citizens
- Cybertips from The National Center for Missing and Exploited Children – passed on from the ICAC Task Force
- Referrals from other Law Enforcement Agencies
- Child Protection System undercover operations
- Referrals from Cloud Storage providers – passed on from NCMEC

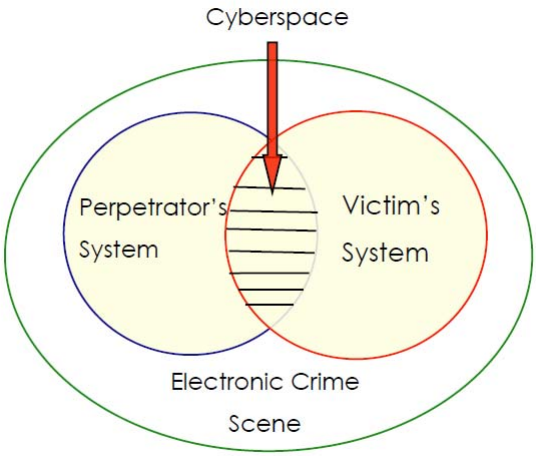
Computer Forensics Defined

- "Pertaining To The Law"
- Coined in 1991 in the first training session held by the IACIS in Portland
- Described as the autopsy of a computer hard disk drive
- Now extends to cloud storage and mobile devices
- New technologies offer new possibilities

Examination & Documentation

- Digital Evidence can be:
  - The Fruits of the Crime
  - The Instrumentality
  - The Evidence
- Your Electronic Crime Scene just changed again!
  - Mobile Devices
  - Cloud Storage
  - Encryption

Where Is The Crime Scene?



What Type Of Investigation Is Needed?

- Tier 1 – On-Scene Preview Of Digital Evidence
  - Seizure of evidence, documentation, interviews
  - Encryption, P2P evidence, wireless / storage
  - RAM capture, Forensic Scan, zSearch, Bitlocker
  - Live scanning of cloud storage
  - Bluetooth sniffing and capture
- Tier 2 – Evidentiary Forensic Analysis
  - Acquisition, analysis for indictment and plea agreements
  - Coordination with vendors (Microsoft, Apple, Amazon) to preserve chain of evidence
  - Case-specific forensic analysis
  - Evidence to corroborate statements, CVIP submission

What Type Of Investigation Is Needed?

- Tier 3 – Requests from DA/Defense
  - Analysis to answer concerns and requests of DA
  - Analysis offered to Defense to exculpate their client
  - Opportunity to close door on defenses, move plea forward
  - Vendors to provide documentation for evidence chain
- Tier 4 – Trial Prep Forensics and Analysis
  - Includes all seized digital evidence for case
  - Defeating known/plausible defenses, complete analysis report, preparation of demonstrative evidence, meeting with DA, prep of expert witness questions/testimony
  - Verify that vendor participation is scrubbed from report

Basics To Understand

- Common types of digital storage media
- How data is stored?
- Hashing, how it works, and why it is important

Identifying Digital Evidence

Computer Forensics Defined

- Collection,
- Preservation,
- Examination,
- Documentation, and
- Presentation

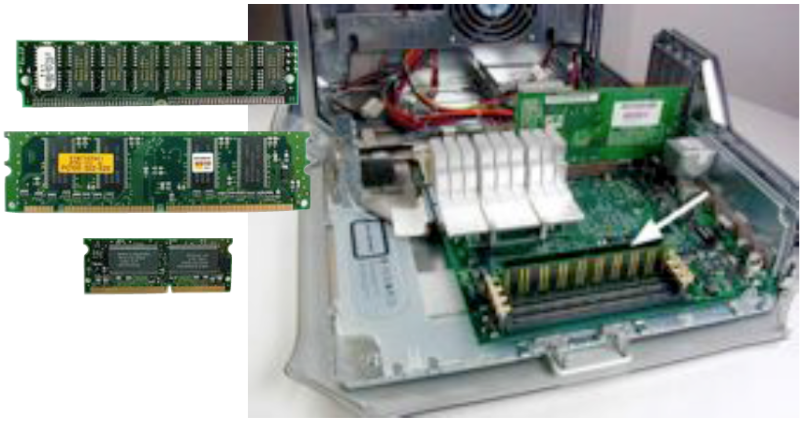
… of Computer Related Evidence

Digital Evidence

What does it look like?

- USB Drives
- Memory Cards
- External Hard Drives
- Computers
- Mobile Devices
- GPS Devices
- Cloud Storage
- RAM (not this ram!)

Digital & Electronic Evidence: RAM



Digital & Electronic Evidence: Wireless Devices

- Be prepared to investigate wireless devices
- Understand how your own devices may interact wirelessly with suspect devices
- Wireless devices can contain evidence of crimes
- Evidence on wireless devices is generally volatile, and gone once power is lost

Evidence Of Wireless Devices

Understanding Data : Data Sizes

- Bit (b) is a single zero or one
- Byte (B) is eight bits in sequence together
- Kilobytes (KB) is 1024 bytes, sometimes shown as 1000 bytes
- Megabytes (MB) is 1,048,576 bytes, sometimes shown as a million bytes
- Gigabytes (GB) is 1,073,741,824 bytes, sometimes shown as a billion bytes
- Terabytes (TB) is 1,099,511,627,766 bytes, sometimes shown as a trillion bytes

How Data Is Written

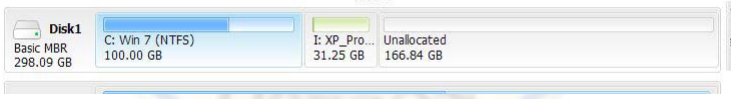
- Data is written and read in 1's and 0's on the drive
- The hard drive is equipped with platters which spin at generally 7200 or 10000 rpm
- Mechanical arms move back and forth over the platters while they spin and write or retrieve data
- The data is written as the mechanical arm changes the magnetic coating on the platter's surface as either + or – (a 1 or 0)
- Solid state drives work the same way but do not use platters and have no moving parts

How Digital Data Is Stored

- Data is written in binary code, or 1's or 0's
- These 1's and 0's are grouped together in blocks of 8 and called bytes
- For example a sequence of "1010011" represents the letter "S". The sequence "1001111" is the letter "O"

Understanding Unallocated Space

- Allocated Space: Physical space on the hard drive that has been assigned and is being used by the file system at a specific moment in time. This includes:
  - Visible files
  - Hidden files
  - Slack space

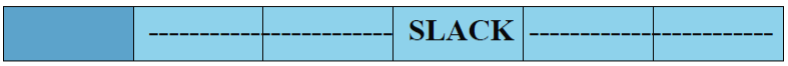


Slack Space

- File slack can be an excellent source of evidence
- Computers write data one sector at a time but must allocate a minimum number of sectors for each file. These sectors are allocated even if you don't use them
- It's like a video tape… If you say that a video tape can only have one show on it at a time, you would allocate a 2 hour video tape per show. Now if you record a ½ hour program, you still have 1½ hours of tape left

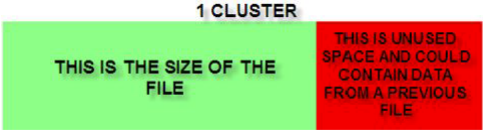
Slack Space

- If there was a program on the tape before you recorded the new ½ hour show, you would see it at the end minus the first ½ hour. This is slack space.



½ hour program

1½ hour of old program



Slack Space Recovery

- Often if data resides in slack space it can be forensically recovered
- Evidence from slack space will normally not have dates/times associated with it because that information may have been overwritten
- It is possible to get enough of a document or image to prosecute an individual

Understanding Unallocated Space

- Unallocated Space = Physical space on the hard drive that has not been assigned by the file system at a specific moment in time and is considered available for use. This includes:
  - Deleted files
  - Space that has not been assigned to a file
- Until something else is placed in its spot on the drive, the file will remain and can be recovered with forensic methods

Methods Impacting Deleted Files

- Running system utilities such as defrag can rearrange data and overwrite unallocated space and slack space
- Using secure erase features such as Norton Secure Erase or other third party applications that are designed to "shred" data
- Although this class is primarily about Windows computers, it should be noted that Mac computers have functionality built in to securely erase data

Terminology - Forensic Image

- It is no longer recommended to call forensic images a "mirrored image"
- Mirroring would imply that the duplicate looks exactly like the original. Although the content is the same, it looks nothing like the original
- "Forensic Image" is the most appropriate and recommended term

Hashing

- Hashing is a very important tool for forensics
- Hashing is like a digital fingerprint for a file. It is mathematically derived from the contents of the item being hashed
- The odds of two files with different content sharing the same MD5 hash value is more than 1 in 340 undecillion (or 1 followed by 36 zeros)
- Hashing in used in forensics for many things:
  - Known File Filters
  - Narrow search scope
  - Exclude items to be searched
  - Find known images of child pornography
  - Compare files to determine if they have been altered
  - Ensure integrity of a forensic image process

Hashing

- There are several algorithms such as MD5 (Message Digest 5), SHA1 (Secure Hash Algorithm) and others
- MD5 is a 128 bit 32 character algorithm and is the most commonly used hashing algorithm
- There are other hashing algorithms available but digital forensics primarily focuses on MD5 and SHA1
- Hashing has been used in many other areas such as download confirmation and encryption
- Hashing is used by cloud storage providers to rapidly scan uploaded files to detect any illegal content

What Affects A Hash Value?

- Any changes to the contents of the file
  - One pixel in a picture
  - Add / remove one character in a document
- Changing the filename or file extension has no effect on the hash value
- Sophisticated CP traders modify files to change hashes, and avoid detection
- This can be circumvented by image analysis software used by Microsoft and Google in their cloud drives. Interested agencies can request for free copies for internal use

What's A Backdoor?

- A method to bypass data encryption or security
  - Does not require the password or passphrase to be known
  - Saves time, cost and effort to access encrypted or secured data
  - Allows data to be accessed, copied and even modified without tipping off the owner
- Currently available for major encryption software – Microsoft Bitlocker, FileVault, BestCrypt, TrueCrypt, etc
- Currently implemented by major cloud storage provider to comply with NCMEC requirements

Legality Of Backdoor Access

- The Patriot Act allows for the use of backdoors for counter-terrorist investigations.
- Requests for backdoor access can be initiated as part of a counter-terrorism investigation
- "Fruit of the poisonous tree" can be circumvented
  - The use of backdoors cannot be detected or proven
  - Vendors are legally and commercially prevented from acknowledging their backdoors. Defense will not be able to prove their existence
  - The files can be described as "forensically obtained"

Legality Of Backdoor Access

- Users of mobile devices and cloud storage sign off on their rights to data scanning. There is no opt-out option
- All cloud stored content are automatically hash-scanned and image-analyzed by their service providers and infringing content reported to NCMEC
- Mobile content are automatically scanned when they are synced with cloud storage like Apple iCloud or Dropbox. Mobile devices that are not cloud-synced can be accessed by their respective vendors
- Evidence procured this way are not considered fruit of the poisonous tree

Foreign Backdoor Access

- Vendors are allowed to provide backdoor access to qualified foreign LEA or governments, but all requests for non-pornographic content must be cleared with the State Department
- Verify with the vendors that the foreign LEA / government have been pre-qualified for backdoor access before discussing this method with their agents
- If a foreign LEA colleague requests backdoor access via your department, please ensure it is only for illegal pornography content, not:
  - content that has political or commercial interest
  - content that are of personal interest
  - content that can be used to prosecute a US citizen

End Of Part 1

- Detective Stu Pitt will take over for Part 2:
  - Tier 1 : On-Scene Preview
  - Tier 1 : Defeating Passwords
  - Tier 1 : Collection & Preservation
  - Tier 2 : Evidentiary Forensic Analysis
  - Tier 2 : Examination & Documentation
  - Tier 3 : Requests from DA / Defense
  - Tier 4 : Trial Forensics Examination
- Tomorrow, Detective Laughlin Foo will conduct Part 3:
  - Mobile Devices
  - Cloud Storage
  - Forensic Principles

Those who are interested in last year's presentation can download it at
http://www.ndsaa.org/Computer_Forensics_for_Prosecutors.pdf