

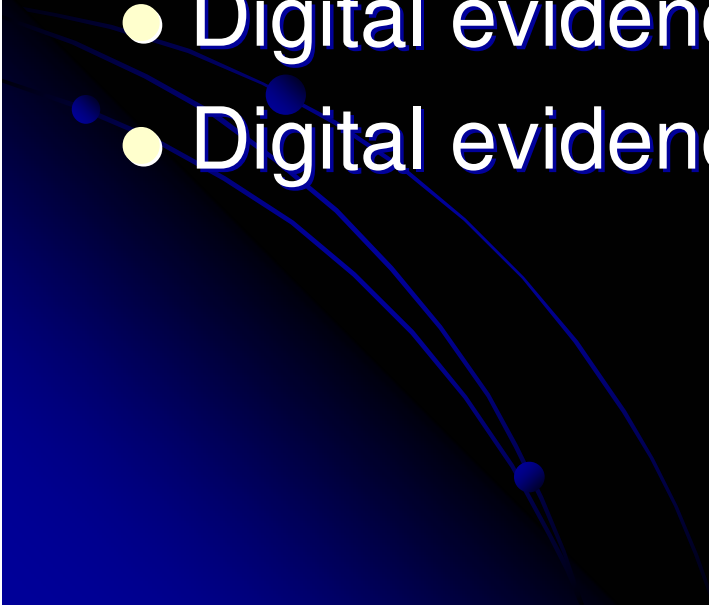
Digital Evidence

Harley Kozushko
Graduate Seminar
Fall 2003

Introduction

- Digital Evidence – encompasses any and all digital data that can establish that a crime has been committed or can provide a link between a crime and its victim or a crime and its perpetrator
- This presentation will explore collection, preservation and identification of digital evidence.

Overview

- Introduction to Intrusion Detection Systems
 - The rules and guidelines surrounding the gathering and use of digital evidence.
 - Digital evidence on the target machine.
 - Digital evidence on the network.
- 

Purpose

- The purpose of this presentation is to provide a reference to the recovery, collection, preservation and identification of digital evidence.



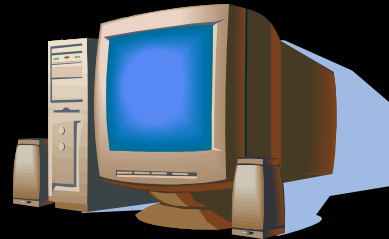
Intrusion Detection

- A brief overview
 - Intrusion detection systems collect information from a variety of system and network sources, then analyze the information for signs of intrusion and misuse.

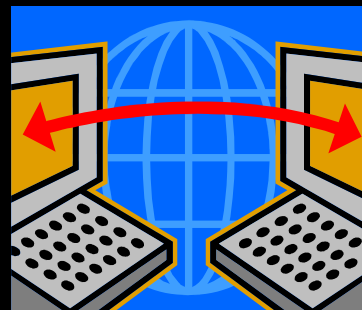


Intrusion Detection – 2 Types

- Host-Based



- Network-Based

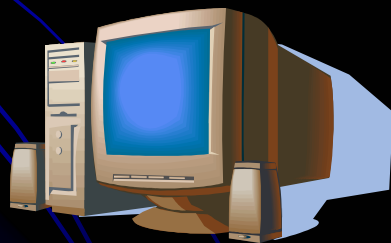


Intrusion Detection – Host-Based

- Host-based intrusion detection systems
 - The system is used to analyze data that originates on computers (hosts).
 - Examines events like what files are accessed and what applications were executed.
 - Logs are used to gather this data
 - Resides on every system and usually reports to a central command console.
 - Uses signatures or predefined patterns that have been defined as suspicious by the security officer.

Intrusion Detection – Host Based cont.

- Used primarily for detecting insider attacks.
 - For example an employee who abuses their privileges, or students changing their grades.
- Audit policy
 - Defines which end-user actions will result in an event record being written to an event log.
 - For example accesses of mission-critical files.



Intrusion Detection – Network-Based

- Network-Based

- The system is used to analyze network packets.
- Used to detect access attempts and denial of service attempts originating outside the network.
- Consists of sensors deployed throughout a network.
- Sensors then report to a central command console

Intrusion Detection – Network-Based

- Uses packet content signatures
 - Based on the contents of packets.
 - Patterns are detected in the headers and flow of traffic.
- Encryption prevents detection of any patterns.



Issues with Network-Based Intrusion Detection Systems

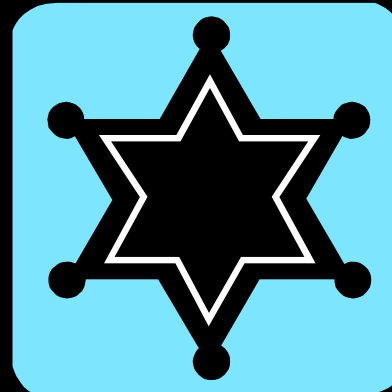
- Because network-based intrusion detection can relate information like IP addresses which can be spoofed, a valid user could be wrongly attributed to perpetrating the misuse of the system.
- This makes network-based intrusion detection data invalid, which also makes most, but not all network digital evidence invalid.
- However, host-based intrusion detection data can be valid digital evidence.

Issues with Data Forensics

- Because data forensics is relatively new, laws dictating the validity of evidence are sketchy and not widely-known.
- Evidence is needed to fully prosecute the attacker.
- This evidence has to come from the security administrator who must ensure the validity of the evidence.
- The security administrator must know the rules that govern the admissibility of evidence in the United States.

Digital Evidence

- Evidence must pass the test of admissibility and weight.
 - Admissibility is a set of legal rules applied by a judge.
 - These rules are extensive.



Digital Evidence

- Weight is a measure of the validity and importance.
 - Essentially whether the judge or jury believes the evidence.
 - There are few guidelines except what is convincing and well presented.
 - Evidence must be authentic, accurate, and complete for it to pass any standard of weight.



Computer Forensic Evidence

- Computer evidence is just like any other evidence in the sense that it must be:
 - Authentic
 - Accurate
 - Complete
 - Convincing to Juries
 - In conformity with common law and legislative rules (admissible)

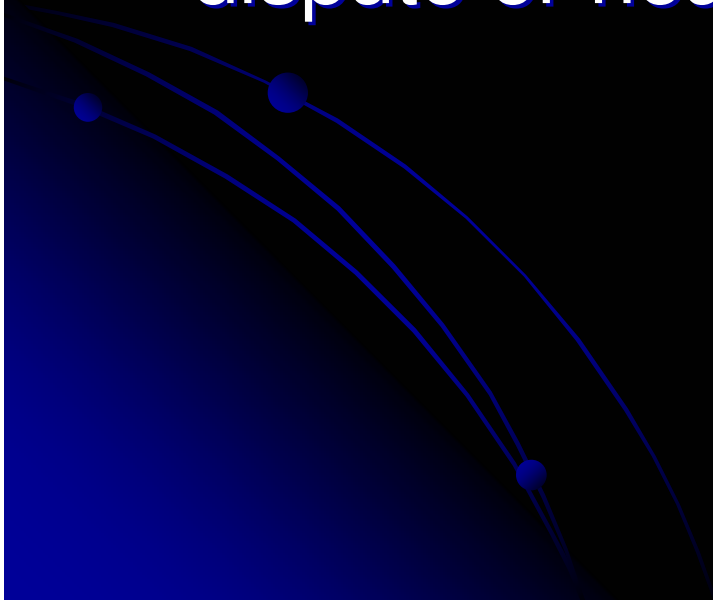
Computer Forensic Evidence

- Authenticity: Does the material come from where it purports?
- Reliability: Can the substance of the story the material tells be believed and is consistent? Are there reasons for doubting the correct working of the computer?



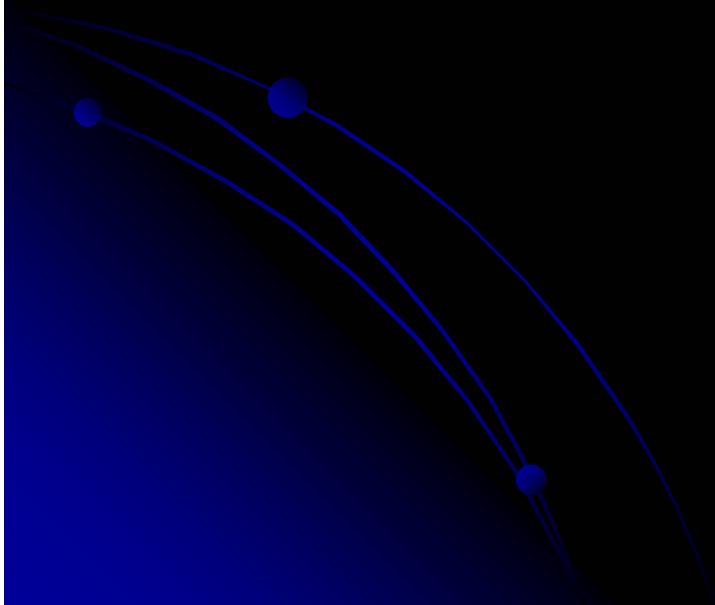
Computer Forensic Evidence

- **Completeness:** Is the story that the material purports to tell complete? Are there other stories that the material also tells that might have a bearing on the legal dispute or hearing?



Computer Forensic Evidence

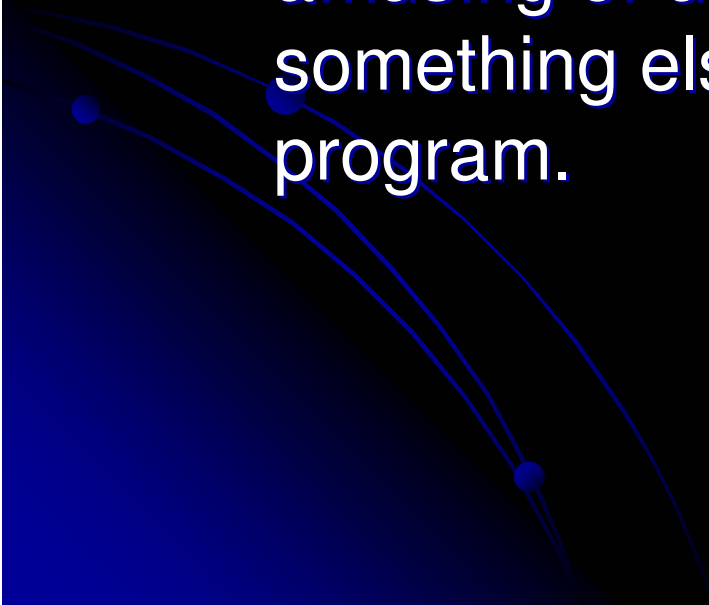
- Freedom from interference and contamination: Are the levels acceptable as a result of forensic investigation and other post-event handling?



Classification of Digital Evidence

- Digital evidence can be classified, compared, and individualized in several ways:
 - Contents – investigators use the contents of an e-mail message to classify it and to determine which computer it came from.
 - Also, swap files and slack space contain a random assortment fragments of digital data than can often be classified and individualized.

Classification of Digital Evidence

- Function – investigators examine how a program functions to classify it and sometimes individualize it.
 - A program that appears to do something amusing or useful but actually does something else, is classified as a Trojan horse program.
- 

Classification of Digital Evidence

- Characteristics – file names, message digests, and date stamps can be helpful in classifying and individualizing digital evidence.



Rules of Evidence

- The five properties that evidence must have in order to be useful:
 - Admissible
 - Authentic
 - Complete
 - Reliable
 - Believable



Rules of Evidence

- Admissible – evidence must be able to be used in court.
 - Failure to comply with this rule is equivalent to not collecting the evidence in the first place, except the cost is higher.



Rules of Evidence

- Authentic – evidence must be tied to the incident in order to prove something.
 - The evidence must be shown to relate to the incident in a relevant way.

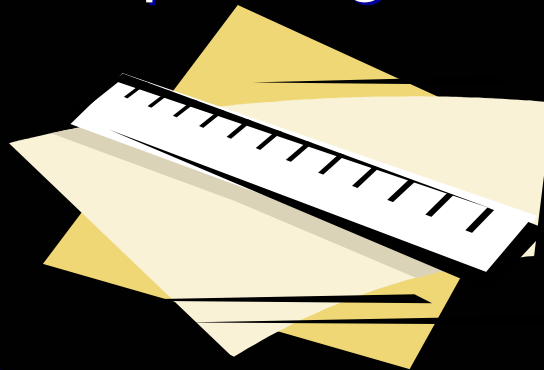


Rules of Evidence

- Complete – It's not enough to collect evidence that just shows one perspective of the incident.
 - Not only should you collect evidence that can prove the attacker's actions, but also evidence that could prove their innocence.

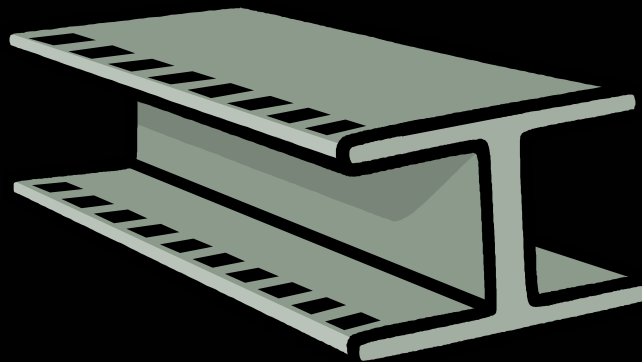
Rules of Evidence

- Complete cont.
 - For instance, if you can show the attacker was logged in at the time of the incident, you also need to know who else was logged in, and why you think they didn't do it.
 - This is called exculpatory evidence, and is an important part of proving a case.



Rules of Evidence

- Reliable – Your evidence collection and analysis procedures must not cast doubt on the evidence's authenticity and veracity.



Rules of Evidence

- Believable - The evidence you present should be clearly understandable and believable by a jury.
- There's no point in presenting a binary dump of process memory if the jury has no idea what it all means.



Do's

- Using the preceding five rules, some basic do's and don'ts can be derived:
 - Minimize handling/corruption of original data
 - Account for any changes and keep detailed logs of your actions
 - Capture as accurate an image of the system as possible
 - Be prepared to testify
 - Ensure your actions are repeatable
 - Work fast

Do's

- Minimize handling/corruption of original data
 - Once you've created a master copy of the original data, don't touch it or the original itself – always handle secondary copies.
 - Any changes made to the originals will affect the outcomes of any analysis later done to copies.
 - For example no programs that modify the access times of files should be run.

Do's

- Account for any changes and keep detailed logs of your actions
 - When evidence alteration is unavoidable it is absolutely essential that the nature, extent, and reasons for the changes be documented.



Do's

- Capture as accurate an image of the system as possible.
- Capturing an accurate image of the system is related to minimizing the handling or corruption of the original data.
 - Differences between the original system and the master copy count as a change to the data.
 - You must be able to account for these differences.



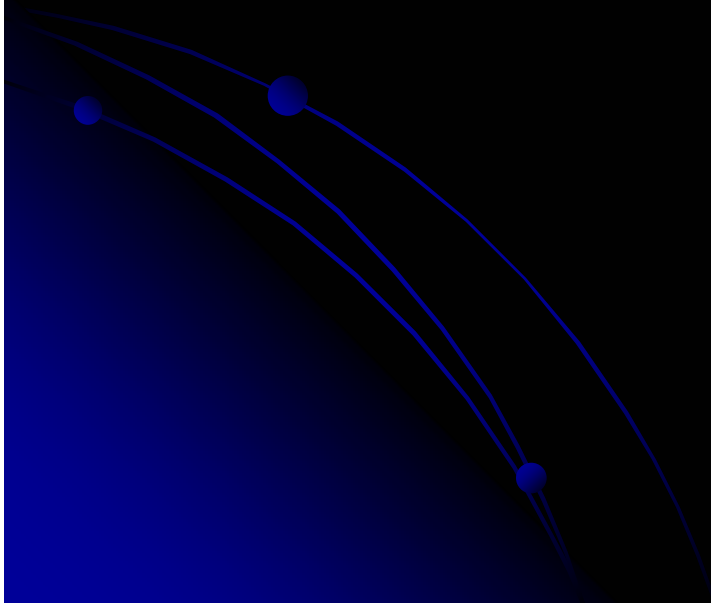
Do's

- Be prepared to testify.
 - Without the collector of the evidence being there to validate the documents created during the evidence-collection process, the evidence becomes hearsay, which is inadmissible.



Do's

- Ensure your actions are repeatable.
 - No one is going to believe you if they can't replicate your actions and reach the same results.
 - This rules out an trial and error actions.



Do's

- Work Fast

- The faster you work, the less likely the data is going to change. Volatile evidence may vanish entirely if you don't collect it in time.
- Automation of certain tasks makes collection proceed even faster.



Don'ts


- Don't shutdown before collecting evidence.
- Don't run any programs on the affected system.



Don'ts

- Don't shutdown before collecting evidence.
 - There is the possibility of loss of volatile evidence and the attacker may have trojaned the startup and shutdown scripts, Plug and Play may alter the system configuration and temporary file systems may be wiped out.

Don'ts

- Don't run any programs on the affected system.
 - There is the possibility of inadvertently triggering something that could change or destroy evidence.
 - Any programs used should be on read-only media and should be statically linked.
- 

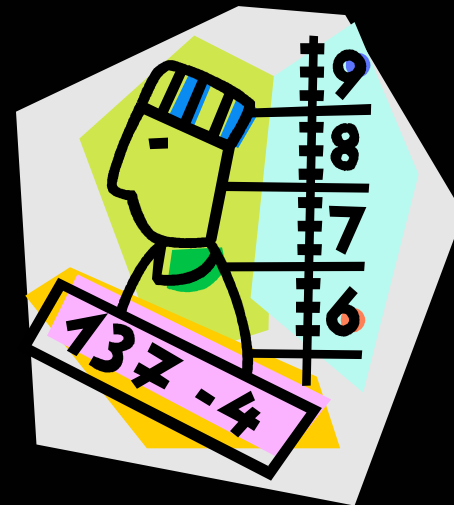
General Procedure

- When collecting evidence there is a general four step procedure to follow which include:
 - Identification of evidence
 - Preservation of evidence
 - Analysis of evidence
 - Presentation of evidence



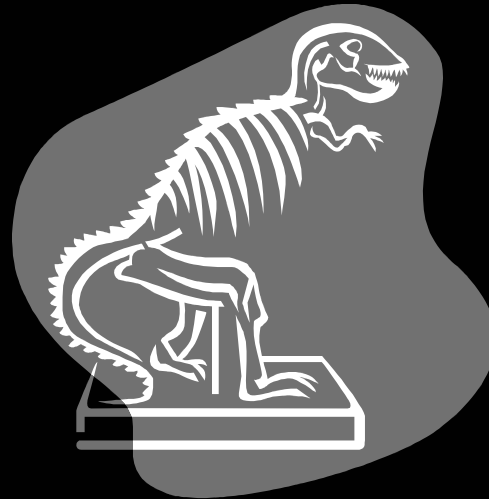
General Procedure

- Identification of evidence:
 - You must be able to distinguish between evidence and junk data.
 - For this purpose you should know what the data is, where it is located, and how it is stored.



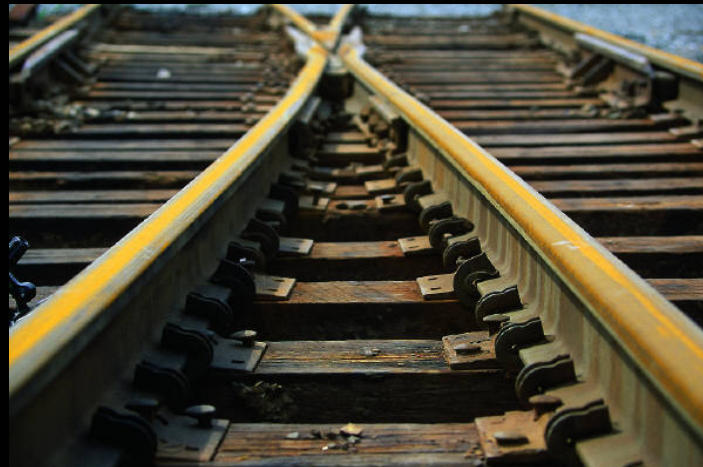
General Procedure

- Preservation of evidence:
 - The evidence you find must be preserved as close as possible to its original state.
 - Any changes made during this phase must be documented and justified.



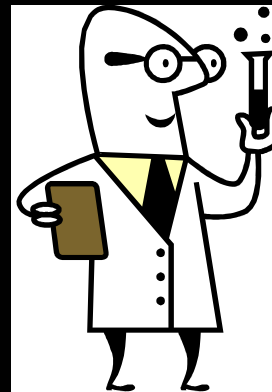
General Procedure

- Preservation of evidence cont.
 - All procedures used in the examination should be auditable, that is, a suitably qualified independent expert appointed by the other side of a case should be able to track all the investigations carried out by the prosecution's experts.



General Procedure

- Analysis of evidence:
 - The stored evidence must then be analyzed to extract the relevant information and recreate the chain of events.



General Procedure

- Presentation of evidence:
 - Communicating the meaning of your evidence is vitally important – otherwise you can't do anything with it.
 - The manner of presentation is important, and it must be understandable by a layman to be effective.

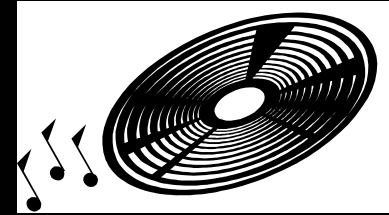


Records

- Through every step of the procedure, it is crucial to record and document everything that is done and everything that is used.
 - This ensures that the procedure is repeatable.

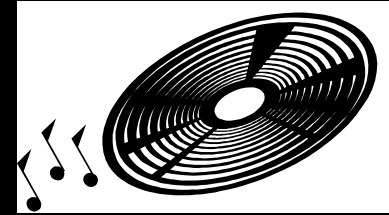


Records



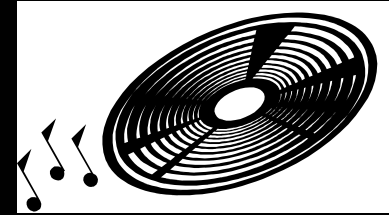
- What to record:
 - Who initially reported the suspected incident along with the time, date, and circumstances surrounding the suspected incident.
 - Details of initial assessment leading to the formal investigation.
 - Name of all persons conducting the investigation.

Records



- More of what to record:
 - The case number of the incident.
 - Reasons for the investigation.
 - A list of all computer systems included in the investigation, along with complete system specifications.
 - Network diagrams.
 - Applications running on the computer systems previously listed.

Records



- More of what to record:
 - A copy of the policy or policies that relate to accessing and using the systems previously listed.
 - A list of administrators responsible for the routine maintenance of the system.
 - A detailed list of steps used in collecting and analyzing evidence.
 - An access control list of who had access to the collected evidence at what date and time.

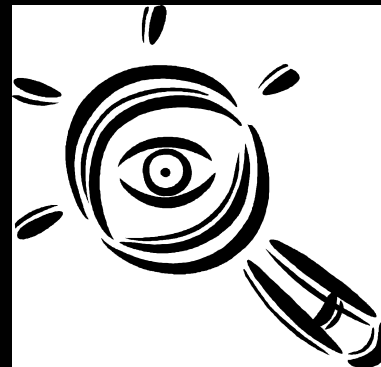
Collection Steps

- Step by step guide for collecting evidence:
 - Find the evidence
 - Find the relevant data
 - Create an order of volatility
 - Remove external avenues of change
 - Collect the evidence
 - Document everything



Collection Steps

- Find the evidence:
 - Determine where the evidence you are looking for is stored.
 - Use a checklist to double check that everything you are looking for is there.



Collection Steps

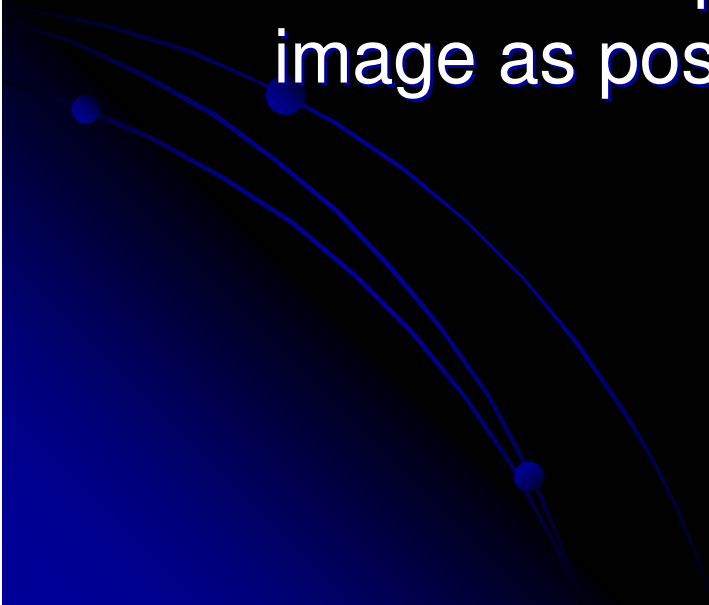
- Find the relevant data:
 - Once you've found the data, you need to figure out what part of it is relevant to the case.
 - In general you should err on the side of over-collection.

Collection Steps


- Create an order of volatility:
 - Now that you know exactly what to gather, work out the best order in which to gather it.
 - Ensures that you minimize loss of uncorrupted evidence.



Collection Steps

- Remove external avenues of change:
 - It is essential that you avoid alterations to the original data.
 - Preventing anyone from tampering with the evidence helps you to create as exact an image as possible.
- 

Collection Steps

- Collect the evidence:
 - Collect the evidence using the appropriate tools for the job.
 - As you go, re-evaluate the evidence you've already collected. You may find that you missed something important.
- 

Collection Steps

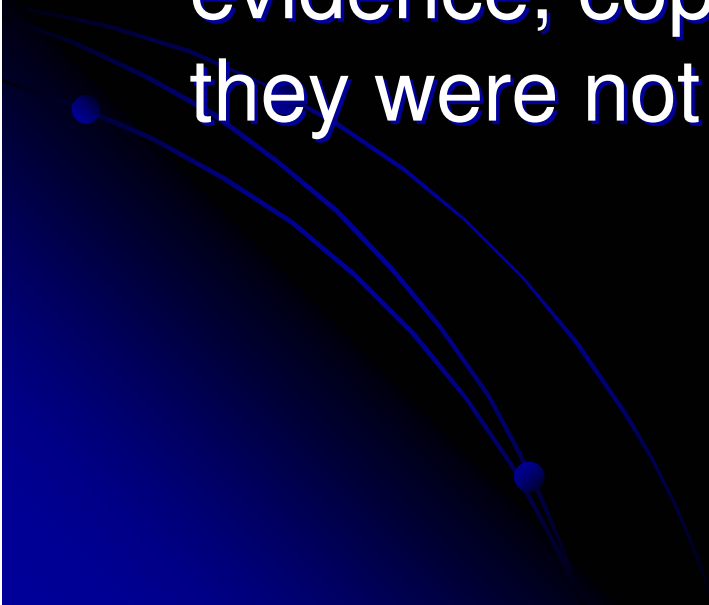
- Document everything:
 - Your collection procedures may be questioned later, so it is important that you document everything that you do.



Digital Evidence vs. Physical Evidence

- It can be duplicated exactly and a copy can be examined as if it were the original.
 - Examining a copy will avoid the risk of damaging the original.
- With the right tools it is very easy to determine if digital evidence has been modified or tampered with by comparing it with the original.

Digital Evidence vs. Physical Evidence

- It is relatively difficult to destroy.
 - Even if it is “deleted,” digital evidence can be recovered.
 - When criminals attempt to destroy digital evidence, copies can remain in places they were not aware of.
- 

Collecting and Preserving Digital Evidence

- The focus of digital evidence is on the contents of the computer as opposed to hardware.
- Two kinds of copies:
 - Copy everything.
 - Just copy the information needed.
- When there is plenty of time and uncertainty about what is being sought, but a computer is suspected to contain key evidence, it makes sense to copy the entire contents.

Collecting and Preserving Digital Evidence

- When collecting the entire contents of a computer, the general concept is the same in most situations:



Collecting and Preserving Digital Evidence

- All related evidence should be taken out of RAM.
- The computer should be shut down.
- Document the hardware configuration of the system.
- Document the time and date of the CMOS.
- The computer should be booted using another operating system that bypasses the existing one and does not change data on the hard drive(s).
- A copy of the digital evidence from the hard drive(s) should be made.

Collecting and Preserving Digital Evidence

- When collecting the entire contents of a computer, a bit stream copy of the digital evidence is usually desirable.
- In short, a bit stream copy copies what is in slack space and unallocated space, whereas a regular copy does not.

Agenda for Duplication and Preservation of Evidence

- Make bit stream back-ups of hard disks and floppy disks cont.
 - Tools to accomplish this:
 - Encase
 - DD
 - Byte back
 - Safeback
 - Note the tool used
- When making the bit stream image, note and document how the image was created.
 - Also note the date, time, and the examiner

Empirical Law

- Empirical Law of Digital Collection and Preservation:
 - If you only make one copy of digital evidence, that evidence will be damaged or completely lost.



Computer Image Verification

- At least two copies are taken of the evidential computer.
- One of these is sealed in the presence of the computer owner and then placed in secure storage.
- This is the master copy and it will only be opened for examination under instruction from the Court in the event of a challenge to the evidence presented after forensic analysis on the second copy.

Collecting and Preserving Digital Evidence

- Collecting evidence out of RAM on a Unix machine is not a simple task.
- The 'ps' command is used to list programs that a machine is running but one must specify that one wants to see all the processes.
- "ps -aux"

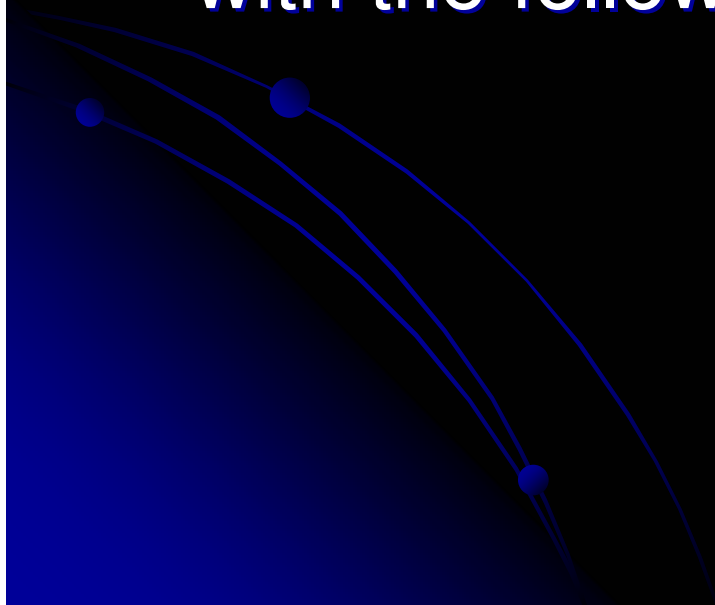


Collecting and Preserving Digital Evidence

- Some types of Unix allow one to save and view the contents of RAM that is associated with a particular program using the “gcore” program.
- There are also programs that provide a list of files and sockets that a particular program is running – “lsof”
- Investigators can use the “dd” command to make a bit stream backup

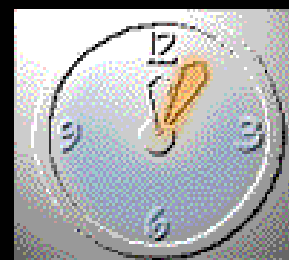
Collecting and Preserving Digital Evidence

- Whenever digital evidence is copied onto a floppy disk, compact disk, tape or any other form of storage media, an indelible felt-tipped pen should be used to label it with the following information:



Collecting and Preserving Digital Evidence

- Current date and time and the date/time on the computer (any discrepancy should be noted).



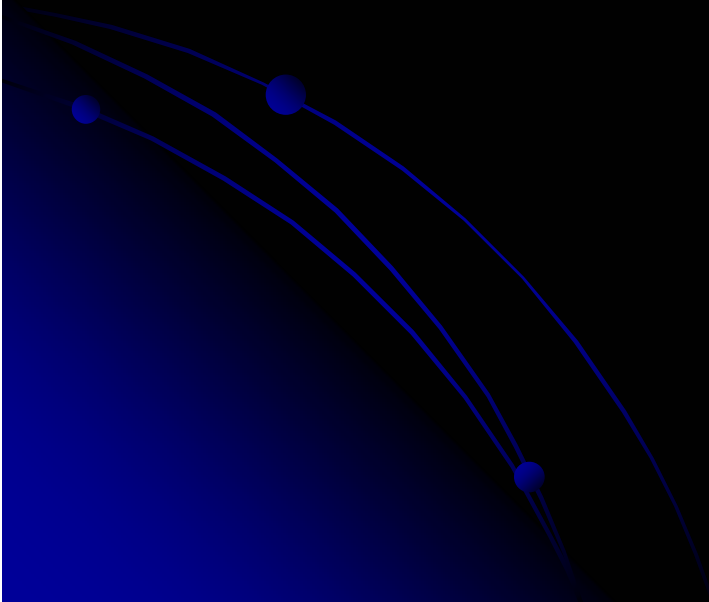
Collecting and Preserving Digital Evidence

- The initials of the person who made the copy.
- The name of the operating system.



Collecting and Preserving Digital Evidence

- The program(s) and/or command(s) used to copy the files.
 - Retain copies of software used.
- The information believed to be contained in the files.



Collecting and Preserving Digital Evidence

- Since the evidence has been collected, it is important to ensure the integrity of the evidence.



Controlling Contamination

- The chain of custody.
 - Once the data has been collected, it must be protected from contamination.
 - Originals should never be used in forensic examination – verified duplicates should be used.



Controlling Contamination

- Chain of Custody: analysis.
 - Once data has been successfully collected, it must be analyzed to extract the evidence you wish to present and rebuild exactly what happened.
 - You must make sure to fully document everything you do – your work will be questioned and you must be able to show that your results are consistently obtainable from the procedures you performed.

Controlling Contamination

- Time

- To reconstruct the events that led to your system being corrupted, you must be able to create a timeline.
- Log files use time stamps to indicate when an entry was added, and these must be synchronized to make sense.

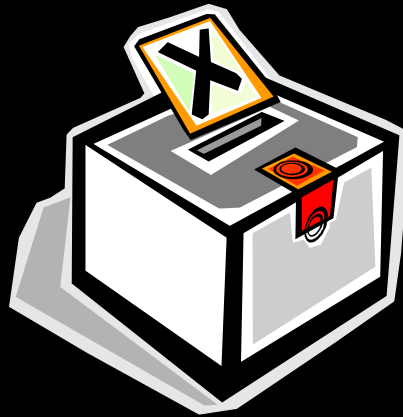


Controlling Contamination

- Forensic Analysis of backups:
 - When analyzing backups it is best to have a dedicated host for the job.
 - This examination host should be secure, clean, and isolated from any network.
 - Document everything you do, ensure that what you do is repeatable and capable of always giving the same results.

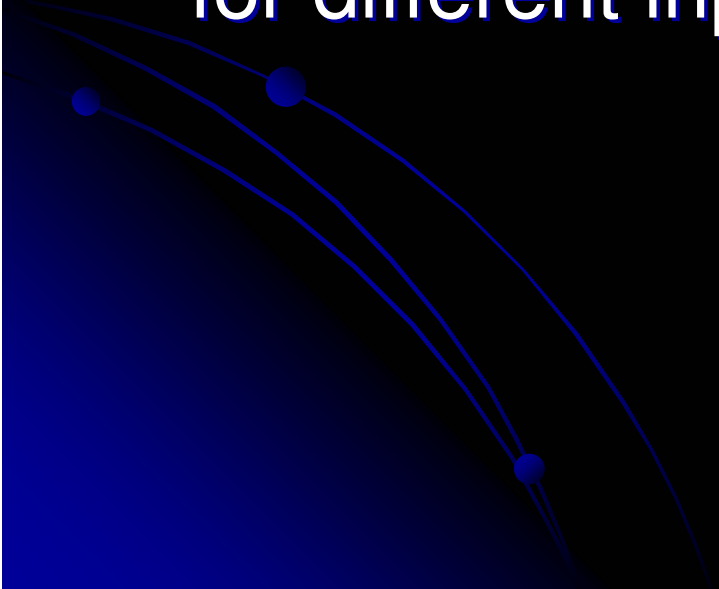
Message Digests

- A message digest algorithm can be thought of as a black box, that accepts a digital object and produces a number.



Message Digests

- A message digest always produces the same number for a given input.
- Likewise, a good message digest algorithm will produce a different number for different inputs.



Message Digests

- Therefore, an exact copy will have the same message digest as the original but if a file is changed even slightly it will have a different message digest from the original.



Message Digests

- The MD5 algorithm can be used for calculating message digests.
- The algorithm uses the data in a digital object to calculate a combination of 32 numbers and letters.



Message Digests

- It is highly unlikely that two files will have the same message digest unless the files are duplicates.
 - It is conjectured that the difficulty of coming up with two messages having the same message digest is on the order of 2^{64} operations, and that the difficulty of coming up with any message having a given message digest is on the order of 2^{128} operations.

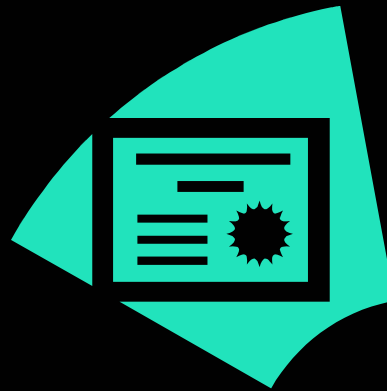
Message Digests

- Message digests provide a method of near individualization and therefore, are sometimes referred to digital fingerprints.
- Message digests are also useful for determining if a piece of digital evidence has been tampered with.



Message Digests

- In essence, the message digest speaks for the integrity of the file.



Host-Based Intrusion Detection

- As it pertains to host-based intrusion detection, the primary source of digital evidence is in log files.



Legal Requirements

- U.S. Code Title 28, Section 1732
 - Provides that log files are admissible as evidence if they are collected in the regular course of business.
- Rule 803(6) of the Federal Rules of Evidence
 - Logs, which might be considered hearsay, are admissible as long as they are collected in the course of regularly conducted business activity.

Legal Requirements

- This means that you are much safer to log everything all the time and deal with the storage issues, than to turn on logging only when an incident is suspected.



Legal Requirements

- Another factor in the inadmissibility of log files is the ability to prove that they have not been subjected to tampering.
- Whenever possible, digital signatures should be used to verify log authenticity.
- Other protective measures include storing logs on a dedicated logging server and/or encrypting log files.

Log Files

- Computer log files are created routinely and contain information about acts and events made at specific times by, or from information transmitted by, a person with knowledge.
- Some computer-generated information has been seen as so reliable that it has been accepted as direct evidence.
 - Direct evidence is usually something tangible that is presented to prove a fact.

Log Files

- It is important to keep these logs secure and to back them up periodically.
- Because logs are automatically time stamped, a single copy should suffice, although you should digitally sign and encrypt any logs that are important, to protect them from contamination.



Log Files

- acct – contains every command typed by every user on the computer.
- loginlog – records failed logins
- syslog – main system log file that contains a wide range of messages from many applications
- sulog- records every attempt to log in as the administrator of the computer (root).

Log Files

- Utmp – contains a record of all users currently logged into a computer. The “who” command accesses this file.
- Wtmp – contains a record of all of the past and current logins and records system startups and shutdowns. The “last” command accesses this file.

Log Files

- Xferlog – contains a record of all files that were transferred from a computer using the file transfer protocol.



Digital Evidence on the Internet

- How to apply the following procedures and guidelines to digital evidence on the network.



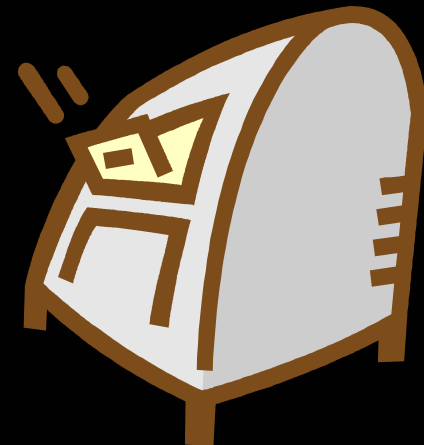
Digital Evidence on the Internet

- For an example, e-mail and internet mail (hotmail):



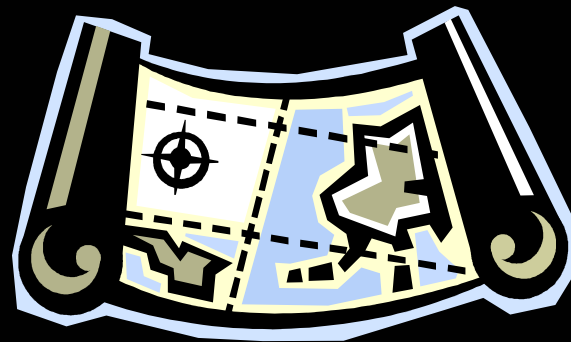
Digital Evidence on the Internet

- Message Transfer Agents (MTA) – computers which are the equivalent of post offices for electronic mail.
- This MTA adds the current time and name of the MTA along with some technical information to the header.
 - Called a received header.



Digital Evidence on the Internet

- Therefore, to track an email message back to the sender you simply retrace the route that the e-mail traveled by reading through the e-mail's received headers.



Digital Evidence at the Transport and Network Layer

- There are services on the Internet that can help you learn more about the sender of an e-mail message:
 - finger
 - telnet
- The most direct method of finding contact information for a given host is to search the Whois databases (<http://whois.arin.net>)

Digital Evidence at the Transport and Network Layer

- A program called traceroute provides a list of routers that information passes through to reach a specific host.
- This program is very useful for determining which computers were involved in the transport of information on the Internet.
 - Intermediate routers sometimes have relevant digital evidence in log files.

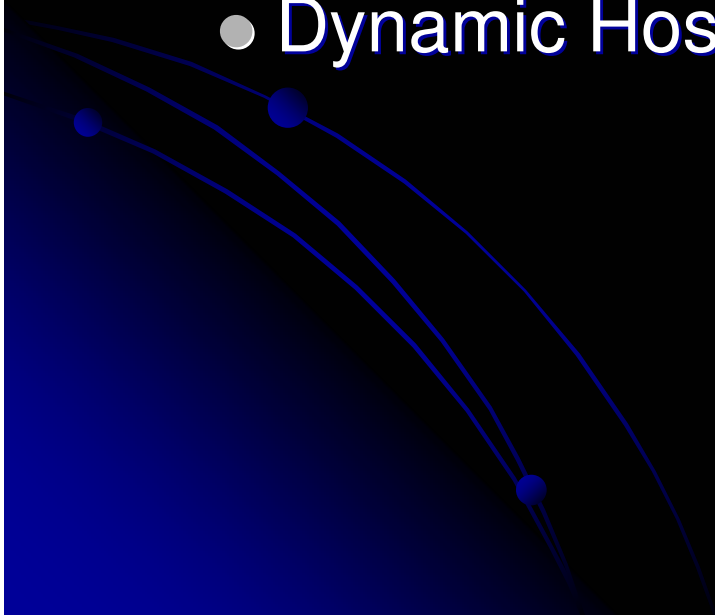
Digital Evidence at the Transport and Network Layer

- Recall that the whois databases contain contact information for the people who are responsible for each router.



Digital Evidence at the Transport and Network Layer

- Protocols for assigning IP addresses.
- To combat IP spoofing two protocols are used:
 - Bootstrap Protocol (BOOTP)
 - Dynamic Host Configuration Protocol (DHCP)



Digital Evidence at the Transport and Network Layer

- BOOTP and DHCP are quite similar
 - Both require hosts to identify themselves using a MAC address before obtaining IP addresses.
- When a computer boots up, it sends its MAC address to the BOOTP or DHCP server which recognizes the MAC address and sends back an IP address.

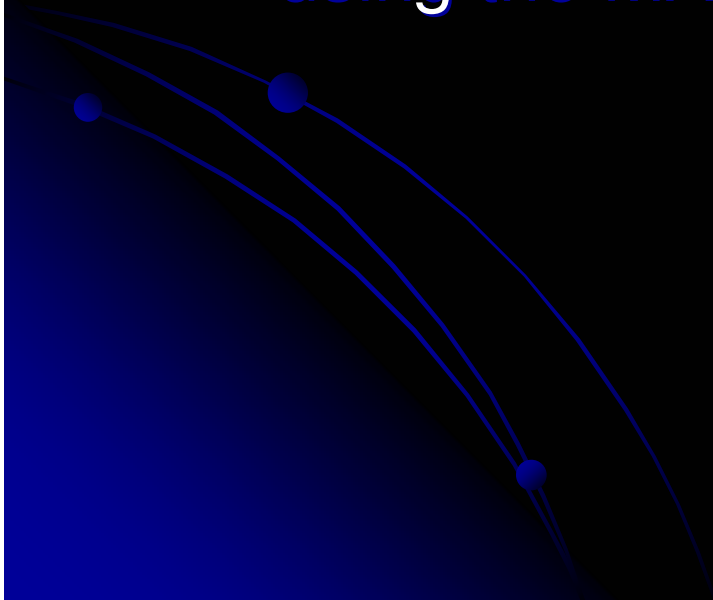
Digital Evidence at the Transport and Network Layer

- The server can be configured to assign a specific IP address to a specific MAC address thus giving the effect of static IP addresses.



Digital Evidence at the Transport and Network Layer

- The criminal could reconfigure his computer with someone else's IP address to hide his identity.
 - Investigators can still identify the computer using the MAC address.

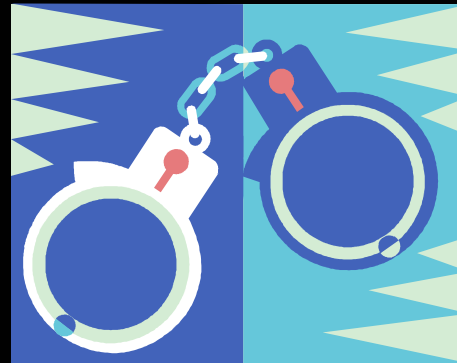


Digital Evidence at the Transport and Network Layer

- The criminal's computer must use the Address Resolution Protocol to device data from a router which requires the actual MAC address of the computer.
- The router would have an entry in its ARP cache showing a particular computer using someone else's IP address.

Digital Evidence at the Transport and Network Layer

- Therefore, there is a good chance that the criminal's computer and the criminal himself will be located and caught.



Summary

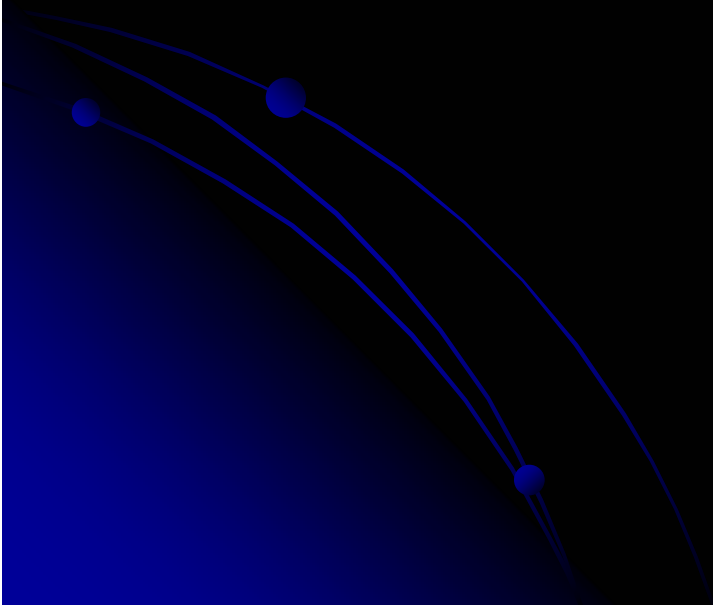
- Data forensics is best applied to digital evidence that resides on host computers rather than on the network.
- Collection and preservation of evidence is key to the use of digital evidence in a court of law.
- The entire process needs to be documented extensively.
- A good forensic toolkit can provide the tools for digital evidence collection and preservation.

Conclusions

- There are set methods for collecting evidence off of a host computer.
 - These methods will provide the administrator with prosecution support.
- A good forensics toolkit and a trained administrator can make collection and preservation of digital evidence a quick, routine task.
- It is a good idea to get trained by professional, or to seek professional help during the process.

Current Project

- Data Forensics Toolkit
 - Primarily host-based toolkit
 - Bootable Linux toolkit



Where to Get More Information

- Casey, Eoghan. Computer Evidence and Computer Crime: Forensic Science, Computers, and the Internet. Cambridge: Cambridge University Press, 2000.
- Vacca, John R. Computer Forensics Computer Crime Scene Investigation. Massachusetts: Charles River Media, 2002.
- Kurose James F. and Keith W. Ross. Computer Networking. United States of America: Pearson Education,
- R. Rivest, "The MD5 Message Digest Algorithm," RFC 1321, Apr. 1992.
<http://www.rfc-editor.org/rfc/rfc1321.txt>.