Global Information Assurance Certification Paper

Computer Forensics

Moroni Parra
March 19, 2002

The Scenario

As computers and the Internet continue to become part of our everyday life, the potential for harm caused by computer crime increases. Unfortunately, there's not enough public knowledge about what computer crime is, and how it should be investigated. The result is that such criminal acts become more widespread and costly to our society each year. Computer Forensics is the technology field that attempts to prove thorough, efficient, and secure means to investigate computer crime.

Everyday we witness more and more of this criminal behavior; web sites are routinely defaced, denial of service attacks are more common, and e-commerce sites are broken into. But all of this is only the beginning; this type of activity has been going on for many years now. Information technology operations have been the silent victims of all manner of computer fraud and abuse from inside and out of their network borders. These incidents range all the way from unauthorized accesses and malicious destruction of data, to logic bombs, hacking, and the disclosure of confidential information.

The Computer Security Institute and the FBI have conducted a computer crime survey for several years now. In 1999, this survey showed that only 51% of the respondents could (or would) acknowledge that they had suffered a financial loss due to computer crime. Even more alarming is the fact that only 31% of the respondents could put a dollar figure on their loss. Hence, the $123,779,000 dollars reported by the survey as lost, is merely a lower bound for the survey's participants. This same survey was conducted again in 2000. It showed that 74% of the respondents acknowledged financial losses resulting from computer crime, and 42% reported losses of $265,589,940. The International Computer Security Association has determined that the two major reasons why computer crime is so difficult to be accounted for are: (a) Most computer crimes go undetected by their victims and (b) Out of the attacks that are detected, few are reported

Computer crime is divided in two categories: computer fraud and computer abuse. Computer fraud involves a criminal act, while computer abuse deals with violators of an organization's computer policies. If it can be proved that someone committed computer fraud, and therefore violated the U.S. Criminal Code, such individuals can be fined and/or sent to prison for many years. In contrast, computer abuse can result in a reprimand, demotion, or termination of employment. Computer crime is a dangerous and damaging activity. To address the task of investigating computer fraud and abuse, a relatively new field called Computer Forensics is emerging.

Once the computer crime has been committed and detected, it is necessary to gather and preserve evidence to use in a legal process. Such evidence is both physical and logical. It may consist of hardware components and media (which contain data) or just data alone. The physical side of computer forensics involves what is called search and seizure of computer evidence. Here, an investigator travels to the scene of a computer crime, and searches for, and takes into custody computer hardware and media that are involved in the crime. On the other hand, the logical side of computer forensics deals with the extraction of raw data from any relevant information resource. This is referred to as information discovery and normally involves an investigator combing through log files, searching the Internet, retrieving data from a database, etc.

An investigator must be able to extract information from the evidence at hand, but without causing changes to the original state of this evidence; also the original state of the evidence must be preserved throughout an investigation - from the moment the evidence is located, to the moment the investigation is closed.

An important tool used by investigators to safeguard evidence, is something called the chain of custody. Essentially, this is a means of accounting for who has touched a given piece of evidence, when they touched it, and what they did to the evidence. It's a way of demonstrating that evidence hasn't been damaged or tampered with while in the care of the investigator. .

Forensic Tools

Computer evidence is inherently complex and volatile in its own unique way. It is complex because it can be derived from any computing resource, at any level of operation (i.e. machine language – ones and zeros). Computer evidence is also volatile, since it can be digitally altered or destroyed with ease, and often without detection. To cope with these issues, the skills and tools that the computer fraud and abuse investigator deploys must be tailored to fit the job.

An effective computer fraud and abuse investigator must have a fundamental understanding of information systems. The investigator must be familiar with good systems administration practices, and possess skills and knowledge relevant to computer security. He/she must understand how computers, operating systems, databases, and computer networks function, and must have strong understanding of the various concepts at work in these areas like computer organization, distributed computing, database architecture and administration, network architecture and protocols, etc.

The tools needed to investigate computer crime consist of both hardware and software. The best place to work on these type of investigations are laboratories adapted to preserve evidence. This must be a highly secure environment – physically and logically – where computer evidence is processed and stored.

Regarding the software tools needed for investigating computer crime, a data archiving program to manage the tape backup and CD reader/writer systems, and a case management system, which will be a key component throughout the investigation. It provides the investigator with a means of storing case notes and information about all of the players and items in a given investigation. Ideally, all of the tracking data should be stored and interacted with in a secure manner: when case data are transmitted or archived they should be strongly encrypted, and access to case data should be through a means using strong authentication.

The Methodology

The methodology implemented while investigating a computer crime is divided in two:

- Search and Seizure
- Information discovery

In search and seizure area of forensic science, the investigator goes to the computer crime scene, and faces the task of recovering and processing physical evidence. In contrast, information discovery involves the investigator accessing data sources on *un-seized* materials (e.g., log files, databases, etc.), in an effort to locate and process information that may prove or disprove something. Often, a case may require search and seizure as well as information discovery.

One might be tempted to assume that information discovery necessarily follows the stages for search and seizure: e.g., after locating a computer at a crime scene, an investigator performs information discovery activities on that computer to look for logical evidence on its hard drive. Although sound in concept, this is incorrect in practice. The last stage of search and seizure, Process Evidence, is where an investigator would actually examine the data on a crime scene computer's hard drive. The distinguishing characteristic here is that information discovery does not take place on seized computers, components, or media. Rather, it is the retrieval of logical evidence on un-seized materials.

The process of search and seizure begins by formulating a plan. The most sound and expeditious means of tracking evidence and handling case data, is to use a software system designed for that purpose. It is a requisite that any investigative processes used to carry out computer forensics, also exhibit the characteristics of a scientific methodology. For example, a valid process should consist of rational, well-conceived steps that can be repeated in all investigations. In addition, such steps should help safeguard against inconsistent and biased results, by providing a framework of reason within which investigative activities can take place. This is an outline of the steps t follow:

1. Identify and research a problem
2. Formulate a hypothesis
3. Conceptually and empirically test the hypothesis
4. Evaluate the hypothesis with regards to the test results - devise and execute new tests if the results are inconclusive
5. If the hypothesis is acceptable, evaluate its impact

Deploying a formal method for investigating computer crime is obviously resource intensive. By selectively using the method in its entirety on only high profile cases, the process of investigation can be streamlined. However, the risk is that serious weaknesses can be introduced into evidence gathered by an investigator.

The processes of search and seizure, and information discovery could be different for different categories of urgency assigned to a case. Within each category, however, these processes would be carried out in exactly the same way.

Search and Seizure involves the recovering and processing of physical computer evidence from a computer crime scene. Although mostly just good, common sense, the following six rules should always be in the mind of the investigator throughout the stages of search and seizure forensic work:

1. Do not alter original evidence
2. Do not execute programs on a crime scene computer (especially the operating system)
3. Do not allow a suspect to interact with a crime scene computer
4. Always back up a crime scene computer; if a crime scene computer is on, do not turn it off until any valuable data in temporary memory have been saved
5. Document all investigative activities
6. Regarding the storage of computer evidence: if you are comfortable there, then the computer and components will be comfortable there

In particular, these rules of thumb are helpful for establishing a protocol by which evidence is accounted for, gathered, handled and stored. Not only is this essential for tracking and managing evidence that may originate from a computer crime scene, but such a protocol also protects against defense attorneys wanting to get your evidence thrown out of court because of potential mishandling

Data Recovery

Data recovery in a forensic analysis is an important component in understanding what has happened in the past, as burglary tools, data files, correspondence, and other clues can be left behind by interlopers. This is the information that other people have thrown away

Every operating system works in it's very unique way, and some data recovery techniques that work with one of them may not work with another. For instance, the UNIX Internet FAQ (ftp://rtfm.mit.edu/) has made the following statement since 1993

> For all intents and purposes, when you delete a file with "rm" it is gone...However, never say never. It is theoretically possible *if* you shut down the system immediately after the "rm" to recover portions of the data. However, you had better have a very wizardly type person at hand with hours or days to spare to get it all back.

Nevertheless, it's actually quite simple to view the data on the disk, deleted or not, by simply looking at the raw disk. Since the data is still there, the easiest way to examine it is by utilizing the standard UNIX tools — like strings, grep, text pagers, and so forth Unfortunately, such tools have no way of discerning what data is allocated and unallocated. Now, that doesn't mean that they cannot be useful, especially if you know what you're looking for. Let's say, for instance that an intruder deleted all your system log files (which might start with the month, day, and time) from the first week of January, you could type this to see them:

```
strings /dev/raw/disk/device | egrep '^Jan 0[1-7] [0-9][0-
9]:[0-9][0-9]:[0-9][0-9]'| sort | uniq -c > date-file
```

Through the investigation of destroyed filesystems, it has been demonstrated that modern UNIX filesystems do not scatter the contents of a file randomly over the disk. Instead, they are remarkably successful in avoiding file fragmentation, even after years and years of intensive use.

Clearly, file contents with little fragmentation are easier to recover than file contents that are scattered all over the disk. But good filesystem locality has more benefits. It allows deleted information to survive much longer than you would expect.

Now, having said that, in the case above we are specifically talking about a UNIX system. This wouldn't work if we were dealing with a PC, nevertheless there are other mechanism to retrieve "deleted" information in this type of situation. You would essentially have to toggle the delete flag and hope that no one has overwritten your data.

In the case of Linux, its "ext2" filesystem does not delete the location of the first 12 data blocks stored in the inode when a file is removed. This means that such "deleted" data could be restored directly by using icat on the inode number that we are referring to. Now, as with any other data recovery method, there's no guarantee that the data will still be there. When a file is deleted in Linux, the inode's *dtime* is updated. Using that information you can recover the data from the most recent 20 inode numbers that were deleted. The following is a piece of Bourne shell code recently deleted data from a Linux system:

```
for inode_num in `ils /dev/device |

sort -n +7 -t |

tail -20 |

awk -F | {print $1}'`

      do

      icat /dev/device $inode_num > $inode_num.result

      done
```

The Coroner's Toolkit

There are several useful tools available to recover deleted data, especially for UNIX systems. One of these tools is The Coroner's Toolkit (TCT); a tool may be used to track digital data. The TCT is a suite of freeware tools, and it was originally written by Dan Farmer, a researcher for Earthlink Networks, and Wietse Venema a researcher at IBM Corp.'s T.J. Watson Research Center.

TCT is a standard tool, or rather a collection of tools that are designed to assist in a forensic examination of a computer. It's designed for Unix systems, but it can also get some data collection and analysis from non-Unix disks and media. It has been pointed out that though it may be a time-consuming process, data files will be recovered and reconstructed to try and piece together evidence. As it has been mentioned in previous

paragraphs, once you delete a file, contrary to popular belief, it doesn't just go away. You have an index that can then be used to trace information.

One aspect of TCT is something called grave-robber, a program that controls a number of other tools, all working to capture as much information as possible about a potentially compromised system and its files. TCT is one of those tools that has been used many times in court cases to prove when a computer crime has been committed.

Lazarus is a program included in The Coroner's Toolkit. Lazarus is a rather strange, but at the same time, simple program that produces unusual results. Its goal is to give some "shape" to unstructured data so that a user can view it and manipulate it too. It achieves this goal via a few simple heuristics. The results are predicated on two lemmas:

- The UNIX FFS never starts writing file data except on well-defined boundaries. If we choose an input block size that is consistent with this, we will never miss an opportunity for dividing up a file appropriately — 1024 bytes should succeed for this goal.
- UNIX filesystems like to write files in contiguous blocks when possible for performance reasons. (The UNIX filesystem always keeps itself relatively defragmented, unlike many PC filesystems.)

With these basic rules, a sort of primitive digital X-ray device can be created. The map of the disk that is created essentially makes the drive transparent — you can peer into the disk and see the data by content type, but the highly useful filesystem abstraction is lost.

Lazarus starts by reading in a block of data from its input stream and roughly determining what sort of data – text or binary – the block is. This is done by examining the first 10 percent of the bytes in the block – if they are mostly unprintable characters, then it is flagged as a binary block; otherwise, it is flagged as text data. If the block has been flagged as text, then Lazarus checks the data against a set of regular expressions to attempt to determine what it is with finer detail. For instance, if it sees "From: foo@bar.com," it further marks the text as mail; "

The program unrm is also part of TCT. This program which can emit all the unallocated blocks on a filesystem. It does this by reading the list of free blocks in a filesystem, going to each logical block, and seeing if it contains any blocks or fragments of unallocated data. Fortunately, the free list covers all blocks in the partition, ignoring disk abstractions such as cylinder group maps, boot blocks, and the like, so you're pretty much guaranteed to get all the data blocks.

Unrm can be a very powerful tool if you're looking for something that you know is deleted. For example, assuming you accidentally deleted your password file (a file composed of lines of seven fields separated by semicolons, the third and fourth fields being numeric), you could probably recover most of it by using unrm and a bit of editing:

```
unrm /dev/raw/disk/device | egrep '^.*:.*:[0-9]*:[0-
9]*:.*:.*:' | sort -u > unrm-password-file
```

Contrary to the popular belief that it's hard to recover information, it's actually starting to appear that it's very hard to remove something even if you want to. The persistence of data is remarkable. Getting data "back" is just a matter of having the right set of tools and the "know how" to get the job done. The unrm/lazarus combination is a fine, so to speak, trash can analyzer. While the results can be spotty for simple single file "undeletion," these tools turn out to bee very useful for forensic analyst.

To demonstrate data persistency, experiments like the following have been run: Windows 95 was installed on a disk and used for a while. Then the same disk was used to install a firewall running Solaris. Finally the disk was once more converted to be used by a Linux system. Having done that, the programs lazarus and unrm were run on this disk. It turned out that files and data from the prior two installations were clearly visible. Now, this is a great proof of data persistence. Forensic data is everywhere on computers. It is necessary to continue researching different methods for retrieving it.

Summary

Throughout this document we have talked about Computer Forensic and its importance in the process of uncovering computer crime. We have talked about computer fraud and computer abuse. We have discussed the importance of having a methodology to execute an investigation and the importance of preserving evidence intact. Finally we have analyzed different programs and methods that will help a Forensic Scientist to uncover hidden activity of intruders. And after all that has been written in this document, I can only say that we have barely scratched the surface of what the fascinating world of Computer Forensics really is. It is a relatively new discipline that is growing at a fast pace. I believe that we will see more development and more tools related to this field of computer science in the near future.

Sources

Wietse Venema, "File Recovery Techniques", December 2000
http://www.ddj.com/documents/s=878/ddj0012h/0012h.htm

Dan Farmer, Wietse Venema, "Computer Forensic Analysis Class" Aug 1999
http://www.porcupine.org/forensics/handouts.html

Dan Farmer, Wietse Venema "CouBeing Prepared for Intrussion", April 2001
http://www.ddj.com/documents/s=868/ddj0104f/0104f.htm#rs1

Carnegie Mellon Software Engineering Institute "Using the Coroner's ToolKit"
http://www.cert.org/security-improvement/implementations/i046.02.html

Farmer, Venema, "Computer Forensics Analysis Class Handouts", August 1999
http://www.fish.com/forensics/class.html

Timothy Wright, "The Field Guide for investigating Computer Crime" (Part 5), Sep 2001
http://www.securityfocus.com/infocus/1248

Wietse Venema, "Strangers in the Night", November 2000
http://www.ddj.com/documents/s=879/ddj0011g/0011g.htm

Farmer and Venema, "The Coroner's Toolkit (TCT)" August 1999
http://www.porcupine.org/forensics/tct.html

Laura Rohde, "Forensic Tools may play Role in Investigation", Sep 2001
http://www.cnn.com/2001/TECH/industry/09/12/tech.forensics.idg/index.html

Last Updated: September 19th, 2017

Upcoming Training

Baltimore Fall 2017 - SEC401: Security Essentials Bootcamp Style	Baltimore, MD	Sep 25, 2017 - Sep 30, 2017	vLive
Rocky Mountain Fall 2017	Denver, CO	Sep 25, 2017 - Sep 30, 2017	Live Event
SANS Copenhagen 2017	Copenhagen, Denmark	Sep 25, 2017 - Sep 30, 2017	Live Event
SANS Baltimore Fall 2017	Baltimore, MD	Sep 25, 2017 - Sep 30, 2017	Live Event
Community SANS New York SEC401*	New York, NY	Sep 25, 2017 - Sep 30, 2017	Community SANS
SANS DFIR Prague 2017	Prague, Czech Republic	Oct 02, 2017 - Oct 08, 2017	Live Event
Community SANS Sacramento SEC401	Sacramento, CA	Oct 02, 2017 - Oct 07, 2017	Community SANS
Mentor Session - SEC401	Minneapolis, MN	Oct 03, 2017 - Nov 14, 2017	Mentor
Mentor Session - SEC401	Arlington, VA	Oct 04, 2017 - Nov 15, 2017	Mentor
SANS October Singapore 2017	Singapore, Singapore	Oct 09, 2017 - Oct 28, 2017	Live Event
SANS Phoenix-Mesa 2017	Mesa, AZ	Oct 09, 2017 - Oct 14, 2017	Live Event
SANS Tysons Corner Fall 2017	McLean, VA	Oct 14, 2017 - Oct 21, 2017	Live Event
CCB Private SEC401 Oct 17	Brussels, Belgium	Oct 16, 2017 - Oct 21, 2017	
SANS Tokyo Autumn 2017	Tokyo, Japan	Oct 16, 2017 - Oct 28, 2017	Live Event
Community SANS Omaha SEC401	Omaha, NE	Oct 23, 2017 - Oct 28, 2017	Community SANS
SANS vLive - SEC401: Security Essentials Bootcamp Style	SEC401 - 201710,	Oct 23, 2017 - Nov 29, 2017	vLive
San Diego Fall 2017 - SEC401: Security Essentials Bootcamp Style	San Diego, CA	Oct 30, 2017 - Nov 04, 2017	vLive
SANS Seattle 2017	Seattle, WA	Oct 30, 2017 - Nov 04, 2017	Live Event
SANS San Diego 2017	San Diego, CA	Oct 30, 2017 - Nov 04, 2017	Live Event
SANS Gulf Region 2017	Dubai, United Arab Emirates	Nov 04, 2017 - Nov 16, 2017	Live Event
SANS Miami 2017	Miami, FL	Nov 06, 2017 - Nov 11, 2017	Live Event
Community SANS Vancouver SEC401*	Vancouver, BC	Nov 06, 2017 - Nov 11, 2017	Community SANS
Community SANS Colorado Springs SEC401**	Colorado Springs, CO	Nov 06, 2017 - Nov 11, 2017	Community SANS
SANS Sydney 2017	Sydney, Australia	Nov 13, 2017 - Nov 25, 2017	Live Event
SANS Paris November 2017	Paris, France	Nov 13, 2017 - Nov 18, 2017	Live Event
SANS London November 2017	London, United Kingdom	Nov 27, 2017 - Dec 02, 2017	Live Event
Community SANS St. Louis SEC401	St Louis, MO	Nov 27, 2017 - Dec 02, 2017	Community SANS
Community SANS Portland SEC401	Portland, OR	Nov 27, 2017 - Dec 02, 2017	Community SANS
SANS San Francisco Winter 2017	San Francisco, CA	Nov 27, 2017 - Dec 02, 2017	Live Event
SANS Khobar 2017	Khobar, Saudi Arabia	Dec 02, 2017 - Dec 07, 2017	Live Event
SANS Austin Winter 2017	Austin, TX	Dec 04, 2017 - Dec 09, 2017	Live Event