Hacker Highschool
SECURITY AWARENESS FOR TEENS

LESSON 8
DIGITAL FORENSICS

Hacker Highschool
SECURITY AWARENESS FOR TEENS

ISECOM
INSTITUTE FOR SECURITY AND OPEN METHODOLOGIES

"License for Use" Information

The following lessons and workbooks are open and publicly available under the following terms and conditions of ISECOM:

All works in the Hacker Highschool project are provided for non-commercial use with elementary school students, junior high school students, and high school students whether in a public institution, private institution, or a part of home-schooling. These materials may not be reproduced for sale in any form. The provision of any class, course, training, or camp with these materials for which a fee is charged is expressly forbidden without a license including college classes, university classes, trade-school classes, summer or computer camps, and similar. To purchase a license, visit the LICENSE section of the Hacker Highschool web page at www.hackerhighschool.org/license.

The HHS Project is a learning tool and as with any learning tool, the instruction is the influence of the instructor and not the tool. ISECOM cannot accept responsibility for how any information herein is applied or abused.

The HHS Project is an open community effort and if you find value in this project, we do ask you support us through the purchase of a license, a donation, or sponsorship.

Table of Contents

"License for Use" Information....2
Contributors....4
8.0 Introduction....5
8.1 Forensic Principles....6
8.1.0 Introduction....6
8.1.1 Avoid Contamination....6
8.1.2 Act Methodically....6
8.1.3 Chain of Evidence....6
8.1.4 Conclusion....6
8.2 Stand-alone Forensics....7
8.2.0 Introduction....7
8.2.1 Hard Drive and Storage Media Basics....7
8.2.2 Encryption, Decryption and File Formats....8
8.2.3 Finding a Needle in a Haystack....10
8.2.3.1 find....10
8.2.3.2 grep....10
8.2.3.3 strings....11
8.2.3.4 awk....11
8.2.3.5 The Pipe "|"....11
8.2.4 Making use of other sources....11
8.3 Network Forensics....13
8.3.0 Introduction....13
8.3.1 Firewall Logs....13
8.3.2 Mail Headers....13
Further Reading....14

Contributors

Simon Biles, Computer Security Online Ltd.
Pete Herzog, ISECOM
Chuck Truett, ISECOM
Marta Barceló, ISECOM
Kim Truett, ISECOM

ISECOM
INSTITUTE FOR SECURITY AND OPEN METHODOLOGIES

/// computer security online

8.0 Introduction

Forensics concerns the application of a methodical investigation technique in order to reconstruct a sequence of events. Most people are now familiar with the concept of forensics from TV and films, "CSI (Crime Scene Investigation)" being one of the most popular. Forensic science was for a long time – and still is really – most associated with Forensic Pathology – finding out how people died. The first recorded description of forensics was on just this subject In 1248, a Chinese book called *Hsi DuanYu* (the Washing Away of Wrongs) was published. This book describes how to tell if someone has drowned or has been strangled.[1]

Digital forensics is a bit less messy and a bit less well known. This is the art of recreating what has happened in a digital device. In the past it was restricted to computers only, but now encompasses all digital devices such as mobile phones, digital cameras, and even GPS[2] devices. It has been used to catch murderers, kidnappers, fraudsters, Mafia bosses and many other decidedly unfriendly people.

In this lesson, we are going to cover two aspects of forensics (all computer based I'm afraid – no mobile phone stuff here).

1. What people have been up to on their own computer.

This covers ...

- ... the recovery of deleted files.
- ... elementary decryption.
- ... searching for certain file types.
- ... searching for certain phrases.
- ... looking at interesting areas of the computer.

2. What a remote user has been doing on someone else's computer.

This covers ...

- ... reading log files.
- ... reconstructing actions.
- ... tracing the source.

This lesson is going to focus on the tools available under Linux. There are tools that are available under Windows, as well as dedicated software and hardware for doing forensics, but with the capability of Linux to mount and understand a large number of alternate operating and file systems, it is the ideal environment for most forensic operations.

1 Apparently it is something to do with marks left around the throat, and the level of water penetration into the lungs.
2 Global Positioning System – a thing which tell you where you are in the world using a number of orbiting satellites.

8.1 Forensic Principles

8.1.0 Introduction

There are a number of basic principles that are necessary regardless of whether you are examining a computer or a corpse. This section is a quick summary of these principals.

8.1.1 Avoid Contamination

On TV you see forensic examiners dressed up in white suits with gloves, handling all evidence with tweezers and putting it into sealed plastic bags. This is all to prevent "contamination". This is where evidence is tainted, for example, by fingerprints being added to the handle of a knife by someone picking it up (think *The Fugitive* if you have seen it ... Look what trouble it got him into !)

8.1.2 Act Methodically

Whatever you do, when (if ?) you get to court, you will need to justify all the actions that you have taken. If you act in a scientific and methodical manner, making careful notes of what it is that you are doing and how you do it, this justification becomes much easier. It also allows for someone else to follow your steps and verify that you haven't made a mistake which may cast the value of your evidence in doubt.

8.1.3 Chain of Evidence

You must maintain something called the "Chain of Evidence". This means that at any point in time from the seizure of the evidence until it's final presentation in court, that you can account for who has had access to it, and where it has been. This rules out the possibility that someone has tampered with it, or falsified it in some way,

8.1.4 Conclusion

Keep these things in mind, and even if you are not going to take your work to court, you will be able to maximize your abilities as a forensic examiner.

8.2 Stand-alone Forensics

8.2.0 Introduction

This section is about the forensic examination of an individual machine. For want of a better term, we will call it "stand-alone forensics". This is probably the most common part of computer forensics - its main role is to find out what has been done using a particular computer. The forensic examiner could be looking for evidence of fraud, such as financial spreadsheets, evidence of communication with someone else, e-mails or an address book, or evidence of a particular nature, such as pornographic images.

8.2.1 Hard Drive and Storage Media Basics

There are several components that make up an average computer. There is the processor, memory, graphics cards, CD drives and much more. One of the most crucial components is the harddisk (hard drive). This is where a majority of the information that the computer requires to operate is stored. The Operating System (OS) such as Windows or Linux resides here, along with user applications such as word processors and games. This is also where significant amounts of data is stored, either deliberately, through the action of saving a file, or incidentally, through the use of temporary files and caches. This allows a forensic examiner to reconstruct the actions that a computer user has carried out on a computer, which files have been accessed and much, much more.

There are several levels at which you can examine a harddisk. For the purposes of this exercise, we are only going to look at the file system level. It is worth noting though, that professionals are capable of looking in a great level of detail at a disk to determine what it used to contain – even if it has been overwritten many times.

The file system is the computer's implementation of a filing cabinet. It contains drawers (partitions), files (directories) and individual pieces of paper (files). Files and directories can be hidden, although this is only a superficial thing and can easily be overcome.

Working through the following Exercises should give you a far better understanding of the basics of disk storage.

Exercises:

For each of the following terms about storage media, search for information and learn how they work. Understanding how equipment functions normally is your first step toward forensics.

1. Magnetic/Hard/Physical Disk: This is where your computer stores files. Explain how magnetism is used on a hard disk.

2. Tracks: What are referred to as "tracks on a hard disk?

3. Sectors: This is a fixed space that data fits into. Explain how.

4. Cluster/Allocation unit: Explain why when a file is written to a hard disk that it may be assigned more space than it needs. What happens to that empty space? Looking up the term "file slack" should help you.

5. Free/"Unallocated" Space: This is what you have left after files are deleted. Or are those files really gone? Explain how a file is deleted on the computer. Looking for tools on "secure

delete" may help you. Knowing how you are supposed to securely delete a file so it's really gone is a great way to learn why such tools are needed.

6. Hash, also known as an MD5 hash: Explain what this hash is and what it's used for.

7. BIOS: This stands for "Basic Input/Output System". What is this and where is it stored on a PC?

8. Boot Sector: This works with partition tables to help your PC find the operating system to run. There are many tools for working with partitions, with the standard one being called fdisk. Knowing how these tools work is your first clue to understanding partitions and the boot sector.

9. Cyclical Redundancy Check (CRC): When you get a "read error" message from your hard disk, this means that the data failed a CRC check. Find out what the CRC check is and what it does.

10. File Signature: Often times a file has a small 6-byte signature at the start of the file which identifies what kind of file it is. Opening a file in a text-editor is the easiest way to see this. Open 3 files of each of the following file types in a text editor: .jpg, .gif, .exe, .mp3. What was the first word at the top of the file for each?

11. RAM (Random-Access Memory): This is also known as "memory" and it is a temporary location to read and write information. It is much, much faster than writing to the hard disk. It's also gone when power is lost to the computer. Explain how RAM works. Knowing your computer may have anywhere from 64 to 512 Mb of RAM, search for information about a computer that has more RAM than that.

12. Currently, the largest RAM disk (a super fast hard disk emulated in RAM) is 2.5 Tb (Terabyte). How many times larger than your PC is that?

8.2.2 Encryption, Decryption and File Formats

A lot of the files that you will come across will not be immediately readable. Many programs have their own proprietary file formats, while others use standard formats – for example the standard picture formats - gif, jpeg, etc. Linux provides an excellent utility to help you to determine what a given file is. It is called **file.**

Command Line Switch	Effect
-k	Don't stop at the first match, keep going.
-L	Follow symbolic links
-z	Attempt to look inside compressed files.

An example of the use of the file command is shown below:

```
[simon@frodo file_example]$ ls
arp.c                   nwrap.pl
isestorm_DivX.avi       oprp_may11_2004.txt
krb5-1.3.3              VisioEval.exe
krb5-1.3.3.tar          Windows2003.vmx
krb5-1.3.3.tar.gz.asc
[simon@frodo file_example]$ file *
arp.c:                              ASCII C program text
```

```
isestorm_DivX.avi:                RIFF (little-endian) data, AVI
krb5-1.3.3:                       directory
krb5-1.3.3.tar:                   POSIX tar archive
krb5-1.3.3.tar.gz.asc:            PGP armored data
nwrap.pl:                         Paul Falstad's zsh script text
executable
oprp_may11_2004.txt:              ASCII English text, with very long
lines, with CRLF line terminators
VisioEval.exe:                    MS-DOS executable (EXE), OS/2 or MS
Windows
Windows2003.vmx:                  a /usr/bin/vmware script text
executable
[simon@frodo file_example]$
```

From this you can start to make some attempts to read a certain type of file. There are a number of file conversion utilities available to you under Linux, and even more available on the Internet, as well as a number of file viewers for various formats. Sometimes it may require more than one step to get to a place where you can really work with the data – try to think laterally!

Occasionally, you will come across files which have been encrypted or password protected. The complication that this presents varies, from encryption that is easily broken to stuff that would even give the NSA (or GCHQ or whatever your local government agency happens to be) a headache. There are again a number of tools available on the Internet that you can use to try to break the encryption on a file. It pays to examine the area surrounding the computer that you are dealing with. People aren't very good at remembering passwords, it may well be written down somewhere nearby. Common choices for passwords also involve : pets, relatives, dates (marriage, date of birth), telephone numbers, car registrations, and other simple combinations (123456, abcdef, qwerty etc.). People are also reluctant to use more than one or two passwords for everything, so if you can reverse engineer a password on one file or application, try it on the others. It is highly likely to be the same.

Exercises:

For these Exercises, we will learn about password cracking. While it is legal to crack your own passwords if you forget them, it is not legal in some countries to figure out how something else is encrypted, in order to protect the other material from being cracked.

DVD movies are encrypted to prevent them from being stolen off the DVD and sold. While this is an excellent use of encryption, it is illegal for anyone to research how that encryption is used. This leads to your first exercise:

1. What is "DeCSS" and how does it relate to DVD encryption? Search on "decss" to learn more.

2. Knowing that something is password protected means learning how to open that file. This is known as "cracking" the password. Find information about cracking various types of passwords. To do this search for "cracking XYZ passwords" where XYZ is the password type you are looking for. Do this for the following password types:

   a. MD5

b. Adobe PDF

c. Excel

3. If the encryption method is too strong to be broken, it may be necessary to perform a "dictionary attack" (sometimes known as "brute force"). Find out what a dictionary attack is.

8.2.3 Finding a Needle in a Haystack

Commercial forensic software includes powerful search tools that allow you to search for many combinations and permutations of factors. Without these expensive commercial tools you need to be a little more resourceful. Linux provides you with plenty of scope to construct similar tools using standard utilities. The following text details the use of **find, grep** and **strings**, and then describes the use of the **pipe** to combine them.

8.2.3.1 find

```
find [path...][expression]
```

find is used to locate files meeting certain criteria within the operating system. It is not designed for looking within the files. There must be a million permutations of expressions that can be combined to search for a file.

Exercise:

1. Read the manual page for find. Complete the "Effect" for each "Expression" in the table below. (Hint: Where a number is given as an argument, it can be specified as follows: +n – for **greater** than **n**; -n – for **less** than **n**; n – for **exactly n**.)

Expression	Effect
-amin n	File last accessed n minutes ago
-anewer	
-atime	
-cnewer	
-iname	
-inum	
-name	
-regex	
-size	
-type	
-user	

8.2.3.2 grep

grep is an immensely powerful tool. It is used to find certain lines within a file. This allows you to quickly find files that contain certain things within a directory or file system. It also allows for

searching on regular expressions. There are search patterns that allow you to specify criteria that the search must match. For example: finding all strings in the dictionary that start with "s" and finish with "t" to help with doing a crossword.

```
grep ^s.*t$ /usr/share/dict/words
```

Exercises:

1. Read the manual page for grep.
2. Look up regular expressions for grep on the Internet. Try to construct a regular expression that looks for all words that are four letters long and contain an "a".

8.2.3.3 strings

strings is another useful utility. This will search through a file of any type for human readable strings. This can return a great deal of information about a specific file, often providing information about the application that created it, authors, original creation time and so on.

Exercise:

1. Read the manual page for strings.

8.2.3.4 awk

awk is a programming language designed for working with strings. It is used to extract information from one command to feed into another. For example, to take just the running programs from the ps command, you would use the following:

```
ps | awk '{print $4}'
```

Exercise:

1. Read the manual page for awk.

8.2.3.5 The Pipe "|"

All of the above tools are easily combined using the UNIX "pipe" command. This is shown with the "|" symbol. This allows you to take the output of one command and feed it down a pipe to another command. To find all files in the current directory that are mpg files, use the following:

```
ls | grep mpg
```

Exercises:

1. Using the pipe, the ls command and grep, find all files in the current directory that were created this month.
2. Using the ps command and awk, print a list of all the running process names.

8.2.4 Making use of other sources

There are many other interesting ways of examining how a computer has been used. Nearly every application that gets run will record some additional data beyond the files that it

directly takes in, or files that it puts out. This could include temporary files for processing, lists of last accessed files or the history of a web-browser.

Exercises:

1. What is browser cache? Find the location where your web browser stores its cache.

2. What are browser cookies? Find the location where your web browser stores its cookies.

3. Search for information about web browser cookies. What kinds of cookies are there and what kind of information is stored in them?

4. Your computer uses temporary directories where it writes files by default for the user. This is often times known as Application Data. Find the temporary directories you have available on your computer. While may be called tmp or temp, often times, there are many more that you don't know about. Try FIND on files written with today's date as a great way to find temporary files. Do those files disappear when you reboot the computer?

8.3 Network Forensics

8.3.0 Introduction

Network forensics is used to find out where a computer is located and to prove whether a particular file was sent from a particular computer. While network forensics can be very complicated, we will cover some of the basics that can be applied to everyday life.

8.3.1 Firewall Logs

Who's connecting to me? The firewall is a utility which can choke connections between two points in a network. Many types of firewalls exist. Regardless of the type and job of the firewall, it is the firewall logs which give you the details. Only by using the logs, can you find patterns of attacks and abuse to your firewall.

Exercises:

1. Visit the website http://www.dshield.org. This website takes firewall logs from all over the world to find patterns of network attack attempts. This helps security professionals be sure to verify if the networks they are protecting are vulnerable to those particular attacks before they happen. Read through the website and explain how that pie graph of the world is made and what it means.

2. On the same website, read through the "Fight back" section and the response e-mails they receive. Explain the purpose of this.

8.3.2 Mail Headers

E-mails come with information of every computer they pass through to get to you. This is kept in the headers. Sometimes even more information is in the headers. To view the headers however is not always so simple. Various mail clients will all have different ways to view this. The real trick to reading headers, though, is to know they are backwards. The top of the list is you. Then it travels goes with each line until the very last line is the computer or network that the mail was sent from.

Exercises:

1. A great resource focused on network forensics for fighting SPAM is http://www.samspade.org. Visit SamSpade.org and go to the section called "The Library". Using this section you should be able to explain how to read e-mail headers. You should also read about forged e-mail headers and e-mail abuse. Explain the various ways e-mail can be used to cause harm.

2. Determine how to look at your e-mail headers in the e-mails you receive. Are there any particular fields in those headers that seem foreign to you? Look them up. You should be able to explain what each field means in that header.

Further Reading

The following links are in English.

http://www.honeynet.org/papers/forensics/
http://www.honeynet.org/misc/chall.html - Some forensic Exercises.
http://www.porcupine.org/forensics/ - The classics
http://www.computerforensics.net/
http://www.guidancesoftware.com/corporate/whitepapers/index.shtm#EFE
http://www.forensicfocus.com/
http://www.securityfocus.com/infocus/1679
http://www.linuxsecurity.com/feature_stories/feature_story-139.html
http://www.linuxsecurity.com/feature_stories/feature_story-140.html
http://www.securityfocus.com/incidents
http://staff.washington.edu/dittrich/talks/blackhat/blackhat/forensics.html
http://www.openforensics.org/
http://fire.dmzs.com/
http://www.sleuthkit.org/
http://www.fbi.gov/hq/lab/fsc/backissu/oct2000/computer.htm