Hacker Highschool
SECURITY AWARENESS FOR TEENS

LECCIÓN 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.

All works copyright ISECOM, 2004.

Información sobre la "Licencia de Uso"

Las lecciones y cuadernos de trabajo siguientes son de acceso público y están disponibles bajo las siguientes condiciones de ISECOM:

Todos los trabajos del proyecto "Hacker Highschool" son proporcionados para su uso no comercial con estudiantes de escuelas primarias, secundarias, bachilleratos y ciclos formativos dentro de las actividades académicas propias de la institución. Dichos materiales no pueden ser reproducidos con fines comerciales de ningún tipo. La impartición con estos materiales de cualquier clase, curso o actividad de formación para el que sea necesario pagar un importe, queda totalmente prohibida sin la licencia correspondiente, incluyendo cursos en escuelas y universidades, cursos comerciales o cualquier otro similar. Para la compra de una licencia visite la sección "LICENSE" de la página web del proyecto "Hacker Highschool" en www.hackerhighschool.org/license.
El proyecto HHS es una herramienta de aprendizaje y, como tal, la formación final debe proceder realmente de la influencia del instructor y no basarse únicamente en el uso de la herramienta.

ISECOM no puede aceptar bajo ningún concepto responsabilidad alguna sobre la forma de aplicar, ni sus consecuencias, de cualquier información disponible dentro del proyecto. El proyecto HHS es un esfuerzo de una comunidad abierta, por lo que si encuentra útil este proyecto le invitamos a patrocinarlo a través de de la compra de una licencia, una donación o un patrocinio.

Todos los Derechos Reservados ISECOM, 2004.

Índice

Contribuciones

Simon Biles – Computer Security Online Ltd.

Pete Herzog – ISECOM

Chuck Truett, ISECOM

Marta Barceló, ISECOM

Kim Truett, ISECOM

Guiomar Corral - Enginyeria La Salle

Jaume Abella - Enginyeria La Salle

ISECOM
INSTITUTE FOR SECURITY AND OPEN METHODOLOGIES

LA SALLE
Enginyeria
Arquitectura
Universitat Ramon Llull

8.1. Introducción

Forensics o el análisis forense está relacionado con la aplicación de técnicas de investigación metódicas para poder reconstruir una secuencia de eventos. La mayoría de personas conocen el concepto de análisis forense por la televisión y las películas, como por ejemplo la serie "CSI (Crime Scene Investigation)" que es una de las más conocidas. La ciencia del análisis forense ha estado durante mucho tiempo, e incluso todavía lo está, relacionada con la Patología Forense (averiguar las causas de la muerte de personas). La primera descripción detallada que se ha encontrado del análisis forense es respecto a la Patología Forense y data del 1248, en un libro chino llamado Hsi DuanYu (the Washing Away of Wrongs). Este libro describe como distinguir si alguien ha muerto ahogado o estrangulado.[1]

Digital forensics o el análisis forense digital es un poco menos confuso pero también es menos conocido. Es el arte de recrear qué ha pasado en un dispositivo digital. Al principio, estaba restringido solo a los ordenadores, pero ahora comprende cualquier tipo de dispositivo digital como teléfonos móviles, cámaras digitales e, incluso, dispositivos GPS[2]. Se ha utilizado para capturar asesinos, secuestradores, infractores, jefes de la Mafia y muchas otras clases de gente poco amistosa.

En esta lección, cubriremos dos aspectos del análisis forense (lo sentimos, todos basados en ordenadores – no hay temas de móviles aquí).

1. Así pues, veremos lo que la gente puede llegar a tener en sus propios ordenadores.

Esto cubre...

- … la recuperación de ficheros borrados
- … decodificación elemental
- … la búsqueda de cierto tipo de ficheros
- … la búsqueda de ciertas frases
- … la búsqueda en áreas interesantes del ordenador

2. También veremos lo que un usuario remoto ha estado haciendo en el ordenador de otra persona.

Esto cubre...

- … la lectura de ficheros log
- … la reconstrucción de acciones
- … la búsqueda en áreas interesantes del ordenador
- … el rastreo de la fuente

1 Aparentemente está relacionado con las marcas dejadas alrededor de la garganta, y el nivel de penetración del agua en los pulmones.

2 Global Positioning System – Sistema que puede comunicar tu posición en cualquier parte del mundo utilizando satélites orbitales.

Esta lección se centrará en las herramientas disponibles en Linux. Existen también herramientas disponibles para Windows así como software y hardware dedicado para el análisis forense, pero con la capacidad de Linux para montar y entender un gran número de sistemas de ficheros y sistemas operativos.

8.2. Principios del Forensics

8.2.1. Introducción

Existen un gran número de principios básicos que son necesarios independientemente de si estás examinando un ordenador o un cadáver. Esta sección resume la mayoría de estos principios.

8.2.2. Evita la contaminación

En televisión salen los examinadores forenses ataviados con batas blancas y guantes, cogiendo todas las pruebas con pinzas y poniéndolas en bolsa de plástico selladas. Todo ello es para prevenir la "contaminación". Aquí es donde las evidencias se pueden echar a perder, por ejemplo, si alguien coge un cuchillo y deja sus huellas digitales en la hoja del cuchillo (¿te acuerdas de la película del Fugitivo?… Piensa en los problemas que llegó a tener!).

8.2.3. Actúa metódicamente

En cualquier cosa que hagas, si tuvieras que ir a un juicio, necesitarías justificar todas las acciones que hayas tomado. Si actúas de una manera científica y metódica, tomando cuidadosas notas de todo lo que haces y cómo lo haces, esta justificación es mucho más fácil. También permite a cualquier otra persona poder seguir tus pasos y verificar que tú no has cometido ningún error que pueda poner en duda el valor de tu evidencia.

8.2.4. Cadena de Evidencias

Siempre debes mantener lo que se denomina la "Cadena de Evidencias". Esto significa que, en cualquier momento del tiempo, desde la detección de la evidencia hasta la presentación final en el juicio, puedes justificar quién ha tenido acceso y dónde ha sido. Esto elimina la posibilidad de que alguien haya podido sabotearlo o falsificarlo de alguna manera.

8.2.5. Conclusiones

Ten siempre presentes estas cosas, incluso si tu trabajo no se va a presentar a ningún tribunal. Así podrás maximizar tus habilidades como analista forense.

8.3. Análisis forense individualizado

8.3.1. Introducción

Esta sección trata sobre el análisis forense en una sola máquina. Para asignarle un nombre mejor, lo llamaremos "stand-alone forensics" o análisis forense individualizado. Probablemente, esta es la parte más común del análisis forense de ordenadores y su papel principal se basa en descubrir qué se ha estado haciendo con un ordenador en particular. El analista forense podría estar buscando una evidencia de fraude, como por ejemplo hojas de balance financieras, evidencia de comunicación con alguien, correos electrónicos o una agenda de direcciones, o evidencia de un tema particular, como imágenes pornográficas.

8.3.2. Fundamentos sobre discos duros y medios de almacenaje

Los elementos que componen un ordenador común son, entre otros: el procesador, la memoria, la tarjeta gráfica, la unidad de CD, etc. Uno de los más importantes es el disco duro. Aquí es donde se guarda la mayor parte de la información que el ordenador necesita. El sistema operativo (OS) como Windows o Linux reside aquí, junto con las aplicaciones de usuario como el procesador de texto y los juegos. Aquí también se guarda gran cantidad de información, ya sea de forma deliberada al guardar un fichero, o de forma incidental con del uso de ficheros temporales y cachés. Esto es lo que permite a un analizador forense reconstruir las acciones que el usuario ha llevado a cabo con el ordenador, los ficheros a los que ha accedido y mucho, mucho más.

El disco duro puede ser examinado a varios niveles, pero para el propósito de esta lección nos centraremos únicamente en el nivel del sistema de ficheros (file system). Sin embargo, debemos tener en cuenta que los profesionales son capaces de analizar el disco a un gran nivel de detalle para determinar lo que contenía; incluso si ha sido sobrescrito varias veces.

El sistema de ficheros (file system) es la implementación en un ordenador de un gabinete de ficheros. Este contiene cajones (las particiones), carpetas (los directorios) y papeles (los ficheros). Los ficheros y los directorios pueden estar ocultos, aunque sólo es una característica superficial que puede ser fácilmente desactivada.

Los siguientes ejercicios nos permitirán tener un mejor conocimiento de las bases del almacenaje de información en los discos.

Ejercicios

Para cada uno de los términos siguientes relacionados con los medios de almacenaje, busca información y aprende cómo funcionan. Tu primer paso en el análisis forense consiste en conocer cómo funciona un equipo de manera normal.

1. Disco físico/duro/magnético: Aquí es donde básicamente tu ordenador almacena los ficheros para que permanezcan allí cuando se apaga el ordenador. Explica cómo se utiliza el magnetismo en un disco duro.

2. Pistas: ¿Qué son las 'pistas' de un disco duro?

3. Sectores: Esto es un espacio fijo donde se almacenan los datos. Explica cómo.

4. Cluster/Unidad de asignación: Explica porqué, cuando se escribe un fichero en el disco duro, puede ser que ocupe más espacio del que necesita. ¿Qué pasa con el espacio vacío? Si buscas el término "file slack" te ayudará a entender el concepto.

5. Espacio libre/sin asignar: Esto es lo que queda una vez que has suprimido archivos. ¿Pero esos ficheros realmente han desaparecido? Explica cómo se borra un fichero del ordenador. Si buscas las herramientas relacionadas con el borrado o eliminación segura (secure delete), éstas te pueden ayudar.

6. Hash, también conocido como MD5 hash: Explica qué es hash y para qué se utiliza.

7. BIOS: Son las siglas de "Basic Input/Output System". ¿Qué es y donde está en el PC?

8. Sector de arranque: Éste trabaja con las tablas de partición para ayudar a que tu PC encuentre el sistema operativo que debe ejecutar. Hay muchas herramientas para trabajar con las particiones, siendo la estándar la herramienta llamada fdisk. Conocer cómo funcionan estas herramientas es tu primer indicio para entender cómo funcionan las particiones y el sector de arranque.

9. Cyclical Redundancy Check (CRC): También se le conoce como Código de Redundancia Cíclica. Cuando el disco duro te informa de un mensaje de error de lectura o "read error", esto significa que los datos fallaron en el chequeo de CRC. Averigua qué es el chequeo de CRC y qué hace.

10. Firma de fichero: A veces un fichero tiene una firma pequeña de 6 bytes al inicio del fichero que identifica el tipo de fichero. La forma más fácil para verla es abrir el fichero con un editor de texto. Abre 3 ficheros de cada una de las siguientes extensiones con un editor de texto: .jpg, .gif, .exe, .mp3. ¿Cuál es la primera palabra al inicio de cada uno de los ficheros?

11. RAM (Random-Access Memory): También se conoce como simplemente "memoria" y es el emplazamiento temporal para leer y escribir información, ya que es mucho más rápido que escribir en el disco duro. El problema es que se pierde toda la información cuando se apaga el ordenador. Explica cómo trabaja la RAM. Sabiendo que tu ordenador debe de tener entre 64 y 512 Mb de RAM, busca información sobre algún ordenador que tenga más RAM que esa.

Actualmente, el disco RAM mayor (un disco duro superrápido emulado en RAM) es de 2.5 Tb (Terabyte). ¿Cuántas veces es mayor que tu PC?

8.3.3. Encriptación, Desencriptación y Formatos de Ficheros

Muchos de los ficheros que te encontrarás no son descifrables de manera inmediata. Muchos programas tienen sus propios formatos de fichero, mientras que hay otros que utilizan formatos estándar – por ejemplo, los formatos de fotografías estándar - gif, jpeg, etc. Linux proporciona una utilidad excelente para ayudarte a empezar a determinar el tipo de cada fichero. Es lo que se denomina **file**.

Command Line Switch	Efecto
-k	No para a la primera coincidencia, continúa
-L	Sigue links simbólicos
-z	Intenta mirar en el interior de archivos comprimidos.

Un ejemplo de la utilización del comando **`file`** se muestra a continuación:

```
[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]$
```

A partir de aquí, puedes intentar leer cierto tipo de ficheros. Hay ciertas utilidades de conversión de ficheros disponibles para Linux e, incluso, hay más disponibles en Internet, así como también visualizadores de ficheros para varios formatos. A veces se puede requerir más de un paso para llegar hasta donde realmente se puede trabajar con la información ¡Intenta pensar de forma lateral!

De manera ocasional, te podrás encontrar ficheros que han sido encriptados o protegidos con alguna contraseña. La complicación que esto presenta varía según la encriptación proporcionada por ciertas aplicaciones, pero puede dar grandes quebraderos de cabeza incluso a entidades como la NSA (o GCHQ o cualquier agencia estatal o local). También existe un gran número de herramientas disponibles en Internet, que puedes utilizar para romper la encriptación de un fichero. Solo hace falta que mires alrededor del ordenador con el que estás trabajando, para ver que las personas no son muy buenas recordando contraseñas y seguramente habrán dejado escrito en algún papel su contraseña. Además, es muy común que las contraseñas estén relacionadas con sus mascotas, familiares, fechas (aniversarios, nacimientos…), números de teléfono, matrículas y otras combinaciones sencillas (123456, abcdef, qwerty...). Además, las personas son reticentes a utilizar más de una o dos contraseñas para todo, por lo que si consigues una contraseña de algún fichero o aplicación, prueba la misma con otros ficheros porque seguramente será la misma.

Ejercicios

Aunque es legal crackear tus propias contraseñas si las has olvidado, en algunos países no es legal intentar resolver cómo se han encriptado los ficheros para protegerlos de ser crackeados.

Las películas DVD también están encriptadas para prevenir que se puedan extraer del DVD y se vendan. Aunque es una manera excelente de encriptación, no es legal buscar métodos para averiguar cómo se ha utilizado la encriptación. Esto lleva a tu primer ejercicio:

1. ¿Qué es "DeCSS" y qué relación tiene con la encriptación de DVDs? Busca información sobre "decss" para aprender más.

2. Saber que algo está protegido mediante contraseñas significa aprender cómo abrir ese fichero. Esto es lo que se conoce como "crackear" la contraseña. Busca información sobre cracker distintos tipos de contraseñas. Para hacerlo busca "cracking XYZ passwords" donde XYZ es el tipo de contraseña que estás buscando. Hazlo para el siguiente tipo de contraseñas:

   a. MD5

   b. Windows Administrator

   c. Adobe PDF

   d. Excel

3. Si el método de encriptación es demasiado fuerte para romperlo, seguramente será necesario realizar un "ataque de diccionario" o "dictionary attack" (también llamado de "fuerza bruta" o "brute force"). Encuentra qué es un ataque de diccionario y mira si puedes encontrar una herramienta que realice un "ataque de diccionario" contra el fichero de password de UNIX: el "/etc/passwd".

8.3.4 Buscando una aguja en un pajar

Los programas comerciales existentes sobre análisis forenses incluyen herramientas de búsqueda de grandes prestaciones, permitiéndole buscar según diferentes combinaciones y permutaciones de factores. Pero para no depender de estas herramientas, en su mayoría de gran coste, haremos uso de nuestro ingenio para realizar ese análisis. Nos ayudaremos de las herramientas básicas que nos proporciona Linux. Las siguientes tablas detallan el uso de los comandos **find**, **grep** y **strings**, y explica como utilizarlas en combinación, una con otra.

8.3.4.1 Find

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

find se usa para encontrar archivos que reúnen ciertos criterios dentro del sistema operativo, no está diseñado para mirar dentro de cada archivo. Debe haber un millón de permutaciones de expresiones que se pueden combinar para buscar un archivo.

Ejercicio:

1. Lee la página del manual del comando find. En la siguiente tabla completa el "Efecto" para cada una de las "Expresiones". (Indicación: Cuando se pasa un número como argumento se puede especificar de las siguientes formas: **+n** para indicar **mayor** que **n**; **-n** para indicar **menor** que **n**; **n** para indicar **igual** que **n**).

Expresión	Efecto
-amin n	Archivos accedidos hace menos de *n* minutos
-anewer	
-atime	
-cnewer	
-iname	
-inum	
-name	
-regex	
-size	
-type	
-user	

8.3.4.2 Grep

grep es una herramienta inmensamente potente. Se utiliza para encontrar ciertas líneas dentro de un archivo. Esto te permite encontrar rápidamente ficheros que contienen cierta información dentro de un directorio o de un sistema de ficheros. También permite buscar ciertas expresiones comunes. Existen algunos patrones de búsqueda que permiten especificar ciertos criterios que la búsqueda debe encontrar. Por ejemplo: encontrar todas las palabras de un diccionario que empiezan por "s" y terminan por "t" para ayudarte a resolver un crucigrama.

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

Ejercicios:

1. Lee la página del manual del comando **grep**.

2. Busca por Internet expresiones comunes para el comando **grep**. Intenta construir una expresión regular que permita encontrar todas las palabras que tengan cuatro letras y contengan una "a".

8.3.4.3 Strings

strings es otra utilidad muy útil. Esta herramienta busca dentro de un fichero cualquier tipo de expresiones o frases que puedan ser leídas ("human readable strings"). Esto nos puede dar gran cantidad de información sobre un fichero específico, y proveer información sobre la aplicación que lo creó, autores, fecha de creación, etc.

Ejercicio:

1. Lee la página del manual del comando **strings**.

8.3.4.4 Awk

awk es un lenguaje de programación diseñado para trabajar con frases. Se utiliza para extraer información de un comando y usarla como parámetros de otro. Por ejemplo, para extraer del comando **ps** únicamente el nombre de los programas que se están ejecutando, deberíamos usar:

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

Ejercicio:

1. Lee la página del manual del comando **awk**.

8.3.4.5 El pipe "|"

Todos los comandos anteriores pueden combinarse usando el comando "pipe" (enlace) de UNIX, el cual se representa con el símbolo "|". Esto nos permite coger el resultado de un comando para proveer de parámetros a otro. Por ejemplo, para encontrar en el directorio actual todos los ficheros *mpg*, utilizaremos:

```
ls | grep mpg
```

Ejercicios:

1. Usando el comando **ls**, el comando **grep** y el comando **pipe**, encuentra en el directorio actual todos los ficheros que fueron creados este mes.

2. Usando el comando **ps** y el comando **awk**, muestra el nombre de todos los programas que se están ejecutando.

8.3.5 Haciendo uso de otras fuentes

Existen muchas más vías para examinar la manera en que se ha utilizado un ordenador. Casi todas las aplicaciones que se ejecutan en un ordenador registran datos adicionales más allá de los que se necesitan para correr la aplicación. Se pueden incluir los archivos temporales que utiliza, los archivos temporales de Internet, etc.

Ejercicios:

1. Qué es el "browser cache"? Encuentra el lugar donde el explorador que utilizas guarda su cache.

2. Qué son los "browser cookies"? Encuentra el lugar donde el explorador que utilizas guarda sus "cookies".

3. Busca información acerca de las "cookies" del explorador. ¿Qué tipo de "cookies" son y qué tipo de información hay en ellas guardas?

4. Tu ordenador usa directorios temporales donde guardar archivos por defecto para el usuario. Se conocen como "Datos de aplicación" (Application Data). Busca los directorios temporales que tienes disponibles en tu ordenador. A menudo se suelen llamar como "tmp" o "temp", pero hay muchos más que desconoces, que también son temporales. Una buena manera de encontrarlos es haciendo un **FIND** de los archivos que contengan la fecha de hoy. Desaparecen estos archivos al reiniciar el ordenador?

8.4 Network Forensics

8.4.0 Introducción

Network forensics (forensics de red) se usa par buscar donde se encuentra un ordenador y para saber si un archivo en particular se ha enviado desde un ordenador en concreto. El network forensics es bastante complicado, pero se van a tratar los temas básicos que se puedan aplicar al día a día.

8.4.1 Firewall Logs

¿Quién se conecta a mi ordenador? El firewall es una utilidad que hace que se bloqueen las conexiones entre dos puntos de la red. Existen diferentes tipos de firewalls. Sin tener en cuenta el tipo ni la tarea del firewall, descubriremos que los archivos "logs" serán los que nos ayuden más. Solamente utilizando los logs, podremos encontrar la manera o conducta con la que nos atacan nuestro firewall, para así combatirles.

Ejercicios:

1. Visita la web http://www.dshield.org. Esta web contiene los archivos logs de multitud de firewalls de todo el mundo, con los que se puede contrastar los diferentes ataques que sufren y los patrones que siguen dichos ataques. Esto

ayuda a los profesionales a verificar si su red estará bien protegida si sufre algún tipo de ataque como el que ya se ha realizado en otro lugar. Haz una leida por la web y explica cómo se ha realizado este gràfico de proporciones y su significado.

2. En la misma web, lee el apartado "Fight back" y las respuestas de los emails recibidos. Intenta explicar su propósito.

8.4.2 La cabecera de los mails

Los e-mails contienen además de la información propia, una información acerca de cada ordenador por el que ha ido pasando hasta llegar al ordenador tuyo. Esta información se encuentra en la cabecera. A veces, hay más información en las cabeceras, pero no es tarea fácil ver esta información. Los clientes de correo tienen diferentes maneras de ver dicha información. Lo realmente interesante es saber como ha sido escrita esta información para atrás. En lo más alto de la lista esta tu ordenador, y según van avanzando las líneas, te vas alejando de tu ordenador hasta llegar finalmente al ordenador que te ha enviado el correo.

Ejercicios

1. Visita la web http://www.samspade.org y ves a la sección "The Library". Después de leer esta sección deberías ser capaz de saber leer las cabeceras de los e-mails. Échale un vistazo también a los e-mails falsificados y al abuso de e-mails. Explica los diferentes métodos para poder hacer daño mediante el uso del correo electrónico.

2. Determina como averiguar las cabeceras de los e-mails que recibes propios. ¿Tienen algún campo en particular que no te pertenezca? Analízala e intenta explicar cada uno de los campos que hay en ella.

8.5 Lecturas de interés

Los siguientes links están en inglés:

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