[Volumen 33 · Febrero 2003]

Ágora sic Divulgación

SUMARIO

Retos de la informática forense

José María Sierra
Julio César Hernández
Arturo Ribagorda
Departamento de Informática
UNIVERSIDAD CARLOS III DE MADRID

El advenimiento y posterior desarrollo de las tecnologías de la información ha desembocado en numerosos avances para la mayoría de los ámbitos sociales. Este importante avance puede observarse en los múltiples cambios en la forma en la que la información es generada, transmitida y almacenada. Por esta razón, y como ocurre en otros ámbitos, se hace necesaria la adaptación de determinadas disciplinas para que estas puedan tener efectividad en el nuevo entorno. La informática forense intenta satisfacer una necesidad, cada vez más importante, de reconstruir determinados hechos en función de un conjunto de evidencias digitales.

Retos de la informática forense

Introducción

No existe ninguna duda de que la comentada mejora tecnológica está permitiendo un desarrollo superior para las empresas, la educación, la investigación, etc. y que por lo tanto sus beneficios sociales son evidentes. Sin embargo, es también cierto que puede observarse igualmente un desarrollo tecnológico en la delincuencia. Este hecho se ve plasmado en la utilización de potentes ordenadores, cifrado de comunicaciones, almacenamiento cifrado de información y en definitiva tecnología de vanguardia que les permite ocultar sus operaciones, una mayor coordinación entre sus integrantes y, en último caso, poder eludir sus responsabilidades ante la justicia.

Parece razonable que sea necesario un tiempo de adaptación para que, tanto las fuerzas de seguridad del estado[1] como la judicatura, puedan combatir los denominados delitos tecnológicos. Asimismo, cabe apuntar que el tratamiento informatizado de informaciones digitales posee determinadas características intrínsecas que no pueden ser dejadas de lado puesto que condicionan su tratamiento. Este hecho provoca situaciones totalmente distintas a las que existían en tiempos pretéritos ya que:

- cuando se sustraen informaciones, éstas siguen estando allí donde se tomaron.
- la copia realizada puede ser totalmente idéntica al original de las mismas.
- los datos pueden arrebatarse por conductos digitales sin interacción física con el repositorio de los mismos.
- la información puede ser ocultada, o incluso cifrada, de forma que no pueda ser detectada.
- la información puede ser distribuida en pocos segundos a numerosos lugares de todo el mundo.
- y, por último, la información puede eliminarse sin que exista forma alguna de recuperarla.

De esta manera, y aunque puede que los objetivos de los delitos no hayan variado, es evidente que han aparecido nuevas formas en las que los delincuentes intentan alcanzarlos. A este respecto es habitual que cuando los delincuentes son detenidos la mayoría de sus datos se encuentren almacenados en su ordenador y, que el análisis exhaustivo de éste sea un paso clave para el descubrimiento de pistas y pruebas de gran importancia para el proceso judicial.

Figura 0: ilustración original de "El principito" de Saint-Exupèry

En particular, este artículo versa sobre la cada vez más trascendental informática forense[2], sus técnicas y herramientas son de utilidad en la recuperación de evidencias digitales, las cuales no sólo pueden ser utilizadas en los procesos judiciales, sino que además son usadas para el estudio de ordenadores que han sufrido algún incidente de seguridad. En este último caso, las técnicas de informática forense representan un paso fundamental en los planes de actuación ante incidentes, ya que son la única forma de estudiar las causas internas de los mismos y, por ende, de corregir y prevenir su repetición.

Una definición aceptada de informática forense es la siguiente "El proceso de identificación, preservación, análisis y presentación de evidencias digitales". Además, y como dificultad añadida al citado proceso, en caso de que el análisis forense vaya a ser utilizado con propósitos legales, las técnicas y herramientas utilizadas deben ser extremadamente escrupulosas con el mantenimiento de la integridad de las evidencias, evitando cualquier tipo de modificación de las mismas, aunque éstas fueran accidentales. Por lo tanto, y debido a las comentadas características de la información digital, son muchas las restricciones asociadas a las técnicas de informática forense y por ello, en el siguiente apartado se detallarán las más importantes.

De todo lo comentado sobre las características de las

[1] A este respecto existen desde hace ya algún tiempo departamentos especializados en este tipo de delitos. Estos grupos poseen una formación avanzada en estas técnicas y constantemente actualizan sus conocimientos.

[2] Término traducido del inglés *Computer Forensics*, quizá mucho más familiar para el lector.

informaciones digitales, se desprende que la tarea del investigador forense tiene tintes de quimera, pues las evidencias pueden cobrar formas muy diversas, estar cifradas con algoritmos computacionalmente seguros o, directamente, haber sido borradas de forma segura.

Sin embargo, existen otra serie de factores que juegan a favor de éste, el más importante de los cuales es la diferencia entre la visión suministrada por sistema operativo y la "realidad física" de las informaciones contenidas en los dispositivos de almacenamiento. En suma que en muchos casos, como apuntaba Saint-Exupèry, lo fundamental es invisible a los ojos.

Características del problema

Si bien resulta usual intuir en los estudiantes de Seguridad Informática (principalmente al comienzo de sus estudios) cierta predilección por el reto que supone la ruptura de las medidas de seguridad en los sistemas informáticos, cabría suponer que ante un reto de, al menos, igual calibre (como sin duda es la informática forense), estos mismos alumnos se sintieran atraídos, retados por su dificultad y tomaran con entusiasmo el estudio de sus técnicas. A pesar de ello y, como seguro han intuido, esta atracción no se produce (por razones que se escapan a la ciencia), pero sin embargo el desafío que supone la informática forense está ahí y en este apartado mostraremos algunas de las causas que la condicionan.

II

De igual manera que continuamente aparecen nuevas herramientas software para la protección de los sistemas informáticos frente a las amenazas que los acechan, es incluso más copiosa la producción de herramientas software (una parte muy significativa de ellas de carácter gratuito) para la ruptura de los citados sistemas. Esta gran diversidad hace inabordable el conocimiento del funcionamiento interno de las mismas, el cual es el único medio de poder encontrar pistas no evidentes sobre las acciones llevadas a cabo con ellas. Así, si bien la utilización de las herramientas sólo exige de un conocimiento superficial de su funcionamiento, el forense informático precisa de un conocimiento más allá de su uso. Por ello, es habitual que para el análisis forense sea a menudo necesaria la actuación de varios expertos, factor que se ve multiplicado si contemplamos el hecho de que una misma herramienta puede comportase y situar las pistas en ubicaciones distintas según sea el sistema operativo sobre el que se utiliza.

El sistema operativo analizado es un punto también a tener en cuenta. Así, sistemas como Windows, cuyo funcionamiento provoca una abstracción más completa del almacenamiento físico de los datos, permiten un margen de maniobra mayor para el forense informático. Esto se debe a que los usuarios de los sistemas Windows disponen de una visión del almacenamiento sesgada por el sistema operativo de forma que por ejemplo, datos que han sido eliminados por el sistema operativo (y por lo tanto no pueden percibirse mediante las herramientas de éste), se encuentran en realidad aún almacenados y pueden ser recuperados por el forense mediante otras vías. Este sistema operativo supone también una ventaja, ya que su comportamiento no está claramente descrito y surgen vulnerabilidades que pueden ser utilizadas para recuperar información del sistema. Un asombroso ejemplo puede ser la vulnerabilidad referida a la previsualización de documentos ("*Thumbnail View*")[3], la cual permite ver parte del contenido de un documento aunque este haya sido eliminado o se encuentre cifrado[4] (ver **figura 1**).

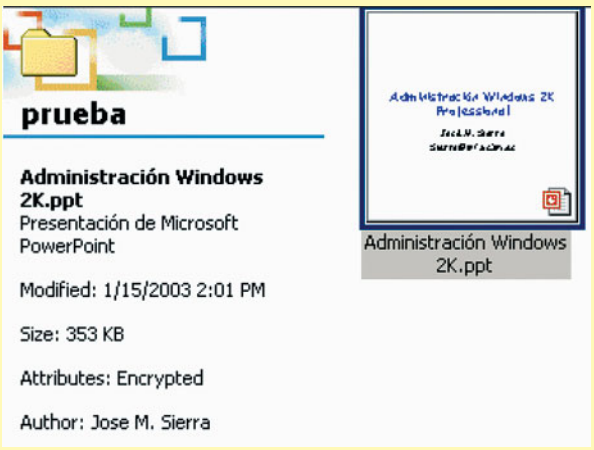


Figura 1: previsualización de un fichero cifrado

En los sistemas basados en Unix, esta situación varía considerablemente, sobretodo cuando se trata del sistema Linux. Esto se debe a varias razones, primera que sus usuarios tienen mayores conocimientos de informática y, por consiguiente, conocen la operativa del almacenamiento de información, y por otro lado, que se trata de sistemas operativos donde existe una menor abstracción entre el hardware del sistema y los usuarios. El forense informático debe hacer un esfuerzo suplementario en estos casos.

La posibilidad de que los datos almacenados se encuentren cifrados constituye otro reto para el investigador forense. El cifrado de datos, en suma el cifrado de informaciones digitales de cualquier formato y contenido, es una técnica al alcance de cualquier usuario de un ordenador[5]. De hecho, los propios sistemas operativos permiten, de manera transparente para los usuarios, el cifrado de informaciones y de igual forma, existen numerosas aplicaciones, algunas de reconocido prestigio (PGP, BestCrypt, Crypt2000, etc.), que facilitan su utilización bajo cualquier sistema operativo. A este respecto la labor del forense informático debe centrarse en el análisis de la implementación que las aplicaciones hacen de acreditados algoritmos criptográficos (Rijndael, TwoFish, RC6, etc., son algunos ejemplos)[6]. En ocasiones vulnerabilidades detectadas son corregidas en posteriores versiones y, son éstas uno de los pocos puntos de apoyo que la informática forense puede tener frente al cifrado de datos.

También inicialmente diseñadas para la seguridad, las herramientas de borrado seguro pueden dificultar de manera muy significativa el análisis forense. Estas aplicaciones sobreescriben varias veces (con bits pseudoaleatorios) las partes del disco donde se encuentra la información que se desea eliminar (las herramientas BCWipe y SecureClean son ejemplos de este tipo de herramien-

[3] http://support.microsoft.com/default.aspx?scid=kb%3Ben-us%3B298608

[4] Claro está que cifrado mediante el sistema EFS (*Encrypted File System*) de Microsoft.

[5] Aunque en este artículo no se mencionen, ya que se centra en la informática forense, el cifrado de datos es un pilar fundamental de la seguridad informática y el hecho de que su uso esté a disposición de cualquier usuario es, desde nuestro punto de vista, beneficioso para la sociedad de la información.

[6] Este tipo de algoritmos han sido validados por la comunidad científica internacional – en el caso de los tres reseñados se trata de algoritmos finalistas del concurso NIST para elegir el estándar estadounidense de cifrado AES– y por lo tanto la búsqueda de debilidades en ellos carece de sentido en una investigación forense.

[7] Son precisamente este tipo de debilidades de gran componente técnico las que deben ser conocidas y aprovechadas por el forense informático para contrarrestar el uso de estas herramientas. http://www.securiteam.com/windowsntfocus/5ZP0M0U60G.html

tas). Recientemente han sido detectadas ciertas limitaciones presentes en un gran número de herramientas de borrado seguro, relacionadas con el tratamiento de los "Streams" NTFS[7]. (**figura 2**)

La esteganografía, que en resumen permite ocultar informaciones dentro de otras, es otra técnica que puede dificultar el análisis forense,. Esta técnica, para la que al igual que en el caso del cifrado existen numerosas aplicaciones a disposición del usuario, es utilizada habitualmente para ocultar imágenes (en casos de contenido ilícito) dentro de otras de forma que, aunque aparentemente el fichero contiene una determinada imagen inocua, es posible extraer de ese mismo fichero otra imagen que se encuentra oculta.

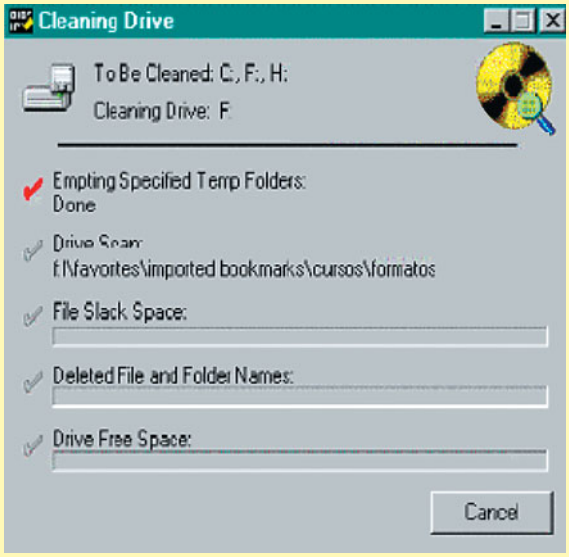


Figura 2: herramienta Secure Clean

El volumen de datos a analizar también representa una dificultad añadida; así, y aunque las capacidades de cómputo han ido creciendo de manera muy significativa, del mismo modo lo han hecho tanto las aplicaciones como los dispositivos de almacenamiento. Esto obliga al forense informático a realizar búsquedas restringidas a porciones de los datos a su disposición, pues las búsquedas completas exigirían una cantidad considerable de tiempo, que en ocasiones no está disponible.

Por último, es necesario apuntar la necesidad de cuestionar los datos que puedan extraerse en un análisis forense. Esto es debido a la ductilidad de las informaciones digitales, ya que las pistas dejadas en una determinada máquina no tienen por que ajustarse a la realidad. De hecho, para un delincuente puede que la mejor forma para evitar ser detectado no sea sólo eliminar aquellas evidencias digitales que puedan involucrarle, sino incluso dejar indicios falsos que apunten en cualquier dirección menos en la correcta.

Precisamente para la falsificación de las pistas de auditoría de los sistemas existen diversas aplicaciones que permiten manipularlas fácilmente. En el entorno Unix están disponibles las herramientas denominadas "RootKits", que permiten tomar control de la máquina y entre otras utilidades la falsificación de todo tipo de informaciones del sistema operativo. En el entorno Windows existe un abanico más heterogéneo de herramientas que va desde virus avanzados a programas troyano. Además existen diversos programas que permiten la eliminación de las pistas de auditoría generadas por navegadores de Internet (*PurgeIE* y *Evidence Eliminator*, son ejemplos de este tipo de herramientas). (**figura 3**)

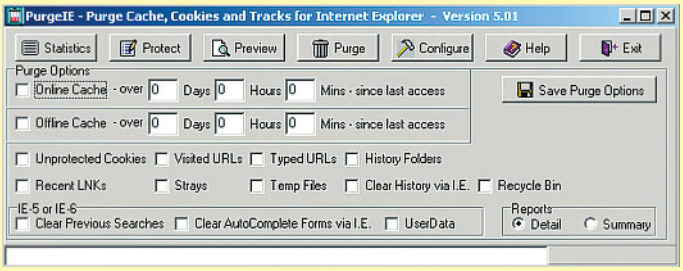


Figura 3: herramienta PurgeIE

En resumen, el trabajo del forense informático debe ser lo bastante bueno como para compensar todas estos inconvenientes, excavando entre las pruebas hasta encontrar aquellas que puedan ser irrefutables. En cualquier caso, y este es un problema de aún más difícil resolución, en caso de que el análisis forense precise relacionar las evidencias encontradas en una máquina con una determinada persona, ni tan siquiera el hecho de que sólo esa persona tenga acceso físico a la máquina puede asegurarnos la comentada asociación. Esto se debe a que siempre que la máquina sea accesible desde Internet, cabe pensar que la seguridad de la máquina haya sido vulnerada y que ésta no sea más que una plataforma utilizada por un tercero para enmascarar sus actuaciones.

El proceso

Si bien las metodologías han sido denostadas en ocasiones por su formalidad y manera estricta y secuencial de entender los procesos, en este caso no podemos evitar utilizarla. Esto se debe a que si bien nos será de gran utilidad para realizar un escrupuloso análisis de una determinada máquina, resulta inexcusable su aplicación si además, como ocurre en muchas ocasiones, el análisis forma parte de un proceso judicial; ya que es preciso controlar y documentar de manera detallada todos los pasos realizados. Así, cuando va a realizarse el tipo de análisis que nos ocupa es necesario "ponerse los guantes" y acometer con esmero cada una de las fases del proceso.

De esta forma la aplicación de la metodología y el estudio de la documentación generada en un análisis forense debe indefectiblemente conducir a los mismos resultados, con independencia del experto forense que realice la investigación.

La International Organization of Computer Evidence (IOCE)[8] entre otros principios que deben ser tenidos en cuenta para la protección de evidencias digitales cita:

– Cualquier labor realizada no debe modificar en ningún caso la evidencia.

– En caso de que sea necesario operar sobre el original de la evidencia digital, la persona debe estar capacitada para realizar tal labor sin que la evidencia digital sea afectada.

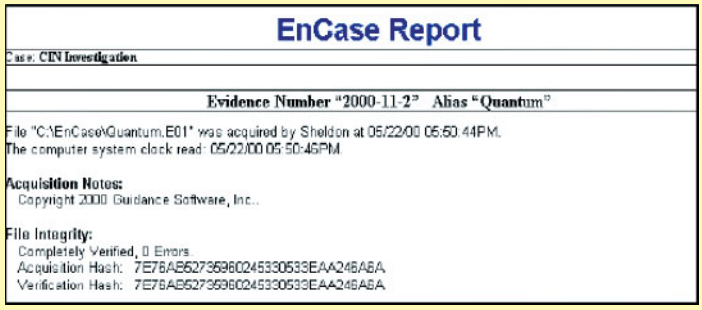
– Cualquier actividad relacionada con la evidencia digital debe ser documentada minuciosamente, protegerse y estar disponible para que pueda realizarse su revisión y comprobación.

Para la puesta en marcha del proceso, el forense informático puede apoyarse en determinadas herramientas que además de automatizar tareas, también le ayudan a secuenciar sus pasos y a documentar cada uno de ellos. Así, y aunque en los últimos años han aparecido

[8] http://www.ioce.org/

muchas herramientas de este tipo, existen dos que son claramente más utilizadas; se trata de EnCase[9] para sistemas Windows y The Coroner's Toolkit[10] para sistemas Unix. Estas dos herramientas poseen cualidades que las hacen las más aconsejables y aunque responden a filosofías distintas ambas pueden ser de mucha utilidad.

La primera de las fases debe ser realizada con el mayor cuidado, pues errores de manipulación de las evidencias pueden provocar la invalidez del todo el proceso. Así, la primera labor a ejecutar es una copia digital completa de todos los datos identificados (incluyendo tanto las partes ocupadas del disco, como zonas libres –pues estas últimas pueden aún contener datos útiles–). Además de las dos herramientas citadas existen otras (SafeBack[11] es un ejemplo) que permiten realizar una copia de los datos en estas condiciones. Una vez realizada la copia, es necesario aplicar una función de integridad sobre original y copia de forma que no pueda albergarse duda alguna sobre la autenticidad de la copia (la siguiente figura muestra el resultado obtenido con EnCase). A partir de ese momento sobre la evidencia digital original no se realizará ninguna tarea. (**figura 4**)


EnCase Report

Case: CIN Investigation

Evidence Number "2000-11-2" Alias "Quantum"

File "C:\EnCase\Quantum.E01" was acquired by Sheldon at 05/22/00 05:50:44PM.
The computer system clock read: 05/22/00 05:50:46PM

Acquisition Notes:
Copyright 2000 Guidance Software, Inc.

File Integrity:
Completely Verified, 0 Errors
Acquisition Hash: 7E78AB5273696024533053 3EAA246A8A
Verification Hash: 7E78AB5273696024533053 3EAA246A8A

Figura 4: resultado mostrado por la herramienta EnCase

Así, y sobre la copia digital exacta de la evidencia, se realizará la siguiente fase de análisis. A continuación, y aunque la heterogeneidad de los análisis obliga al forense informático a tomar decisiones dependiendo de cada caso, entre otras tareas será necesario estudiar las evidencias a la búsqueda de datos que puedan ser recuperados (para lo cual es fundamental que el conocimiento en profundidad del sistema de almacenamiento). Otra labor a realizar es la búsqueda de informaciones particulares, estas búsquedas pueden basarse en fechas, tamaño, formato, contenido, etc.; la velocidad y eficacia de las búsquedas facilitará el trabajo de análisis.

Es muy interesante conocer si los programas básicos del sistema operativo han sido manipulados. A este respecto existen herramientas automatizadas que comparan éstos con los programas originales, facilitando en gran medida esta tarea.

Por otro lado, en muchas ocasiones los archivos con imágenes son el objetivo fundamental del análisis. Es habitual que la búsqueda de los mismos se realice con independencia de la extensión que posean y basándola en patrones conocidos que están presentes en este tipo de archivos. Una vez localizados, también será ineludible su estudio para descartar que se hayan utilizado técnicas esteganográficas para ocultar en ellos otras informaciones.

También dentro de la fase de análisis de la evidencia debe realizarse un trabajo de búsqueda de informaciones especiales. Entre otras son muy interesantes las contraseñas de las distintas aplicaciones que requieran autenticación (correo electrónico, agendas, mensajería instantánea, páginas web, etc.), los nombres de usuario (*login*) y datos personales.

Por último, la elaboración del informe es de suma importancia en el resultado final del proceso. Éste debe ser sumamente claro, de forma que su compresión sea asequible incluso para personas con escasos conocimientos en el área. El informe debe incluir la metodología seguida, las pruebas de que la cadena de custodia se ha seguido correctamente (preservando la validez de las evidencias) y todas las circunstancias relevantes que acontecieron durante su desarrollo. También, y aunque se sinteticen las ideas principales y conclusiones del mismo, es necesario que contenga los resultados de manera detallada, incluyendo –a ser posible directamente– las informaciones producidas por las herramientas utilizadas durante el proceso de análisis.

Resumen

En los últimos años hemos asistido a un crecimiento muy importante del interés por la informática forense. Este crecimiento es respuesta a la imparable digitalización de la información en todos los órdenes de la sociedad, y a la necesidad de reconstruir determinados hechos en función de evidencias digitales. Así, en pocos años este área ha pasado de ser utilizada como parte de los planes de actuación ante incidentes de seguridad, al tiempo que se ha convertido en un procedimiento indispensable para numerosos procesos judiciales.

Es por ello, y por la pericia necesaria para poder realizarla, por la que la comentada demanda apenas puede verse satisfecha por el escaso número de profesionales que pueden desempeñarla. Además, resulta lógico pensar que este área tiene por delante una dilatada vida ya que, la dialéctica entre la ocultación de rastros digitales y la informática forense no tiene un vencedor claro. ❖

José María Sierra
Profesor Titular Interino
Julio Cesar Hernández
Ayudante de Universidad
Arturo Ribagorda
Catedrático de Universidad
Departamento de Informática
UNIVERSIDAD CARLOS III DE MADRID
[sierra, jcesar, arturo]@inf.uc3m.es

[9] En la actualidad está ya disponible las versión 4 de la herramienta. http://www.guidancesoftware.com/

[10] La herramienta puede funcionar sobre FreeBSD 2-4.*, OpenBSD 2.*, BSD/OS 2-3.*, SunOS 4-5.*, Linux 2.* y puede descargarse desde http://www.fish.com/tct o http://www.porcupine.org/forensics/tct.html

[11] http://www.forensics-intl.com