La recuperación de la información y la informática forense: Una propuesta de proceso unificado

Ana Haydée Di Iorio[1], Rita Evelina Sansevero[2], Martín Castellote[3], Ariel Podestá[4], Fernando Greco[5], Bruno Constanzo[6], Julián Waimann[7]

[1] Ingeniera en Informática, Docente e Investigadora en Universidad FASTA, diana@ufasta.edu.ar

[2] Ingeniera en Informática, Docente e Investigador de la Facultad de Ingeniería de la Universidad FASTA, resansevero@gmail.com

[3] Ingeniero en Informática, Docente e Investigador de la Facultad de Ingeniería de la Universidad FASTA. castellotemartin@yahoo.com.ar

[4] Ingeniero Informático, Docente e Investigador de la Facultad de Ingeniería de la Universidad FASTA. arielpodesta@gmail.com

[5] Ingeniero en Informática, Docente e Investigador en Universidad FASTA, fmartingreco@gmail.com

[6] Técnico en Informática. Auxilliar de Investigación Alumno, Facultad de Ingeniería de la Universidad FASTA. Bru.constanzo@gmail.com

[7] Técnico en Informática. Auxilliar de Investigación Alumno, Facultad de Ingeniería de la Universidad FASTA.julianw@ufasta.edu.ar

Abstract. La recuperación de la información en el ámbito de la informática forense puede ser considerado hoy en día un proceso crítico. Una de las mayores problemáticas es la falta de un proceso unificado que guíe a los expertos forenses en esta tarea tan compleja. Se presenta en éste trabajo los avances obtenidos hasta el momento en el proyecto de investigación denominado PURI "Proceso Unificado de Recuperación de la Información" cuyo objetivo es formalizar un proceso marco que abarque las fases, tareas, herramientas y actividades aplicables a diversos entornos y dispositivos.

Keywords: Informática Forense – Peritaje Informático – Recuperación de la Información.

1 Introducción

Con el advenimiento de la Sociedad de la Información, las nuevas tecnologías irrumpen en practicamente todos los aspectos de nuestra cotidianeidad. Es así que, la recuperación de la información digitalizada pasa a ser considerado un aspecto crítico, tanto en ámbitos privados como públicos. Vinculado este tema a la resolución de conflictos judiciales, una de las mayores problemáticas es la falta de un proceso unificado que guíe a los expertos forenses en esta tarea tan compleja, ya sea por la variedad de plataformas tecnológicas que pueden encontrarse, como por la diversidad de formas que puede tomar una evidencia digital.

En este contexto, complejo y velozmente cambiante, existen dificultades a sortear en la recuperación de información que incrementan la complejidad de la tarea: defierentes tecnologías, diversidad de métodos de almacenamiento, localización de la información, leyes heterogéneas, tecnologías que naturalmente eliminan evidencias, mecanismos internos de protección de la información, falta de herramientas específicas, criptografía, aplicaciones que cubren solo una parte del proceso y de efectividad desconocida, y falta de guías y de mecanismos de validación.
La finalidad de la informática forense es el hallazgo de información de valor almacenado o transmitido en forma binaria, con el fin de ayudar a determinar el origen de incidentes.
La correcta extracción de la información es crucial en la obtención de evidencias y es justamente el objeto de análisis de este proyecto.
Un proceso clásisco de recuperación de la información comprende al menos las siguientes etapas básicas: adquisición, validación, análisis, interpretación, documentación y presentación de las pruebas; por lo tanto, se trata de un proceso compuesto de varias etapas de distinta naturaleza que, por consiguiente, presentan diferentes problemáticas y dificultades a sobrellevar.
Durante la realización del proyecto "Proceso Unificado de Recuperación de Información" que en adelante denominaremos PURI, se han estudiado las técnicas y herramientas disponibles en el mercado con el fin de generar un proceso que proponga alternativas al profesional forense.
En el presente documento, se presenta el estado actual del proyecto y se introducen los resultados obtenidos en la etapa de validación del proceso unificado propuesto.

2 La Informática Forense

La automatización de los cálculos en la información, la irrupción de la informática, y la expansión de la capacidad de los dispositivos, ha generado que el volumen de información almacenada crezca exponencialmente. El valor de estos datos, que forman parte de la realidad de nuestras vidas, ha obligado en algunos casos a darles carácter de críticos.
La Informática Forense nace como una rama de las ciencias forenses, una disciplina auxiliar a la justicia, que consiste en la aplicación de técnicas que permiten adquirir, validar, analizar y presentar datos que han sido procesados electrónicamente y guardados en un medio computacional.
Las tareas de informática forense pueden llevarse a cabo tanto en procesos judiciales, como en cuestiones extra judiciales, sin embargo, la importancia de contar con un proceso unificado que auxilie en estas tareas está relacionada con la existencia de un aval científico que le permita a un oficial de justicia confiar en las tareas desarrolladas dentro de un proceso judicial.
El Dr. Julio Téllez Valdés define al Delito Informático como "una conducta típica, antijurídica y culpable en que se tiene a las computadoras como instrumento o fin", es decir, como instrumento para cometer cualquiera de los delitos ya tipificados, o como un fin en si mismo. Por ejemplo, en una estafa a través de sistemas informáticos, la

informática es el medio; en cambio, en el caso de la distribución de virus informáticos, la informática es el fin.
En Argentina, la ley de delitos informáticos, ley 26.388 es sancionada en Junio del año 2008.
Es así que, a pesar del carácter correctivo que demarcan las bases de la informática forense, orientadas en forma casi exclusiva al marco legal, hoy en día su alcance es mucho más amplio, considerando que la presencia de las tecnologías de la información y la comunicación en la sociedad es una tendencia irreversible.

3 El Proyecto

El proyecto PURI nace en la conjunción de dos cátedras docentes de la facultad de Ingeniería de la Universidad FASTA, la cátedra de Sistemas Operativos y la de Informática y Derecho.
Tiene como objetivo general la generación de un proceso unificado para recuperar información y la presentación de propuestas de desarrollo de nuevas técnicas y herramientas, a partir de la detección de carencias.
Los objetivos específicos derivados del objetivo general propuesto son:

- El estudio y análisis de las fases a seguir para la recuperación de la información según las buenas prácticas sugeridas por organismos internacionales.
- La generación de un proceso unificado de recuperación de la información, donde para cada una de las fases se indique: el objetivo, las tareas que la componen, las técnicas disponibles y las herramientas.
- La propuesta de nuevas técnicas a utilizar en áreas carentes.
- La propuesta de desarrollo de nuevas herramientas y el desarrollo de prototipos funcionales

Para alcanzar los objetivos mencionados, se definen entonces cinco etapas bien identificadas, que conforman el plan de trabajo del proyecto con una duración estimada en veinticuatro meses.
En la primer etapa se propone realizar una investigación del tipo exploratorio para recopilar y analizar toda la documentación obtenida sobre recuperación de información con el fin de conocer el estado del arte. Luego, se procede a sintetizar y formalizar este conocimiento en un proceso unificado divido por fases, cada una con sus objetivos, técnicas y tareas específicas. El proceso luego se validará con al menos dos casos de uso típicos en distintas técnologías base.
En esta etapa se prevé también el estudio de las técnicas y herramientas disponibles en el mercado con el fin de generar un proceso accesible al profesional forense actual.
Una vez validado el proceso, en la etapa siguiente se prevé estudiar, analizar y proponer, sobre los nichos carentes, el desarrollo de nuevas técnicas y herramientas, sobre las que se trabajará en el desarrollo y validación del prototipo.
Finalizado el proyecto se espera obtener el diseño de un proceso, las propuestas de las técnicas y herramientas efectivas al efecto de las fases carentes de ellas y los prototipos correspondientes.

Actualmente, el Proyecto se encuentra finalizando la segunda etapa, es decir, en la fase de validación del proceso obtenido en la primer etapa.
En la redacción propuesta de proceso se sugieren ciertas herramientas hoy existentes en el mercado. Se han considerado las siguientes variables para determinar que herramientas sugerir: las pruebas de efectividad técnica realizadas por el equipo, el tipo de licencia (es decir, si se trata de software libre o propietario, si es open source o no), la compañia de respaldo, y el grado de madurez de la técnica.

4 El Proceso Propuesto

El objetivo general del proyecto es crear un proceso unificado de recuperación de información a partir de la estructuración y organización de las tareas, técnicas y herramientas que lo componen, respetando las buenas prácticas propuestas por los organismos internacionales, de manera que se convierta en un elemento de guía y consulta para los profesionales forenses, tanto en el ámbito comercial como en el judicial.
Este proceso se define como una secuencia de fases compuestas por etapas que involucran tareas a llevar a cabo aplicando técnicas implementadas por herramientas concretas que permiten ejecutar dichas tareas.
Este modo de definir el proceso, brindará una visión detallada y abarcadora de todo lo concerniente a esta actividad.

A continuación se presenta el Proceso Unificado de Recuperación de la Información propuesto:

Fase de Adquisición
Esta fase comprende toda actividad vinculada con la generación de una réplica exacta de todo el contenido digital alojado en el dispositivo original. El motivo de realizar una copia de tal información viene originado por distintas razones que se mencionan a continuación:
1) La ciencia forense debe respetar tres principios básicos: No contaminación, actuar metódicamente y Mantener la cadena de evidencia. Justamente en las ciencias informáticas la no contaminación se garantiza a traves de la copia bit a bit del original, de esa manera, al trabajar sobre la copia se resguarda el original y se garantiza la no contaminación de la evidencia.
2) Además, el proceso de recuperación de información demanda cierto tiempo, durante el cual el dispositivo quedaría inutilizado para otras actividades. Al trabajar sobre la copia, se podría entregar el original al dueño.
3) Eventualmente el dispositivo que almacena la información en cuestión puede no ser siempre el más indicado para llevar a cabo las pruebas requeridas, debido a problemas de accesibilidad o velocidad. Esta es otra razón que fundamenta la obtención de una copia exacta de los datos a fin de trabajarlos eficientemente en un entorno apropiado.
Esta fase de adquisición comprende etapas que de acuerdo al entorno en el que se deba llevar a cabo la recuperación de la información, aplicará o no involucrarlas en el

proceso. Es así que se procedió a dividir la adquisición de dispositivos móviles de otros dispositivos por sus características altamente diferenciadoras a todo nivel, tanto físico (hardware), cómo lógico (software).
A continuación se enumeran dichas etapas.

1. Adquisición de medio de almacenamiento persistente
2. Adquisición de medio de almacenamiento volátil
3. Adquisición de tarjetas SIM.
4. Adquisición de Tarjetas de Memoria Extraíble Persistente del dispositivo móvil
5. Adquisición de Memoria ROM interna del dispositivo móvil
6. Adquisición de Memoria Volatil (RAM) del dispositivo móvil

Estas etapas, a su vez, comparten una secuencia de tareas que básicamente consisten en: 1.Bloquear el dispositivo para evitar escrituras indeseadas. Este proceso puede ser por hardware o software.

2.Capturar y resguardar una copia fiel del contenido del dispositivo. Esta copia debe ser a nivel bit con el fin de que sea una réplica exacta del original.

3.Comprimir o dividir la imágen obtenida. Esto con el fin de poder almacenarla en un medio con menor capacidad que el original.

4.Validar la copia con los datos originales, para asegurar que la copia a sido bien realizada. Usualmente mediante algún algoritmo de hash.

Fase de Preparación

Esta fase involucra todos los procedimientos necesarios para generar el entorno de pruebas preciso para llevar a cabo en primer lugar la inspección, y eventualmente la recuperación de la información.
Como primera etapa, la fase de preparación contempla la restauración de la imagen. Esto significa que si la misma se encontrara dividida, encriptada o comprimida deberá realizarse el proceso contrario, a fin de lograr el original.
A continuación se deberá validar que la restauración ha sido exitosa mediante un algoritmo de hash, como se mencionó previamente.
Si la imagen que se obtuvo es de un sistema de archivos de un determinado sistema operativo, entonces será útil generar una máquina virtual que tome dicha imagen como su disco principal. Al hacerlo se debería realizar una copia a fin de no alterar la imagen original.
Opcionalmente puede montarse la imágen a fin de tratar los datos contenidos en la misma como un dispositivo de almacenamiento conectado al equipo de trabajo.
Finalmente esta etapa contempla la identificación de tipos de sistemas de archivos y sistemas operativos contenidos en los medios de almacenamiento originales.

Fase de Análisis

Esta fase comprende el fuerte del trabajo en donde se analiza el contenido adquirido en busca de vestigios de lo que se quiere hallar. El objetivo final de la fase de análisis en el caso de un proceso judicial o pre-judicial es encontrar la denominada Evidencia Digital, es decir, aquello que relaciona el hecho ocurrido con el “imputado” y la “victima”. Entonces, se piensa en la evidencia digital como en un tipo de evidencia física que está construida de campos magnéticos y pulsos electrónicos que pueden ser recolectados y analizados con herramientas y técnicas especiales.

La fase de análisis comprende las siguientes etapas:

1. Extracción lógica
2. Extracción física
3. Análisis de relaciones

La extracción lógica representa la recuperación de información eliminada a partir del sistema de archivos. Por esa razón se denomina “lógica”, ya que no se accede en forma directa a los bloques, sino a través del Sistema de Archivos, y del Sistema Operativo como intermediario. La mayoría de los sistemas operativos no eliminan la información en el momento en el que un Usuario solicita el borrado de un archivo determinado, sino que , de alguna manera, dejan registrado que el espacio que ocupaba dicho archivo ahora se encuentra disponible. De esta manera, por ejemplo, si fuese posible hallar tal espacio entonces sería posible reconstruir la información original.

Esta etapa de extracción lógica contempla las siguientes tareas:

- Recuperación de archivos eliminados
- Extracción de información a examinar por tipo de archivo.
- Extracción de metadatos del archivo presentes en el sistema de archivos.
- Extracción de metadatos propios del archivo
- Extracción de archivos protegidos con contraseña
- Extracción de archivos comprimidos
- Extracción de archivos encriptados
- Determinar el tipo de archivo encriptado
- Búsqueda de determinado tipo de archivo oculto. Esta tarea está vinculada a la búsqueda en particular de determinado tipo de archivo, que puede estar oculto en un nombre de archivo con otra extensión
- Búsqueda de información en el área de paginado del Sistema Operativo.
- Búsqueda de Información de Configuración. Referida a la información que puede obtenerse del registry de Windows u otros archivos de configuración de las distintos Sistemas Operativos.
- Búsqueda de Información en Procesos en Memoria. Referida a la obtención de información existente en los dump de memoria, en el caso de examinación de equipos que se encontraban en funcionamiento al momento de la adquisición.

La extracción física comprende la búsqueda de la información directamente en el espacio de datos omitiendo todo tipo de estructura de sistema de archivos. Con lo cual se da caso omiso a los metadatos y se aplican diferentes técnicas sobre el contenido puro del bloque en el dispositivo de almacenamiento.

En esta etapa podrían estar incluidas, de ser necesario, las siguientes tareas:

- Búsqueda de palabras en el bloque del dispositivo
- Extracción de archivos en espacio desalojado (marcado como libre) y no fragmentado
- Extracción de archivos en espacios desalojados, que puedan estar fragmentados
- La etapa de análisis de relaciones trata justamente de identificar relaciones entre conjuntos de archivos, con el fin de obtener una conclusión. Esto involucra puntualmente la Identificación de relaciones entre conjunto de

archivos vinculados a una actividad en particular (ej: archivos relacionados a la navegación por internet) y la verificación de aplicaciones instaladas, entre otros.

5 Situación Actual y Conclusiones

El proceso propuesto presentado se validó para las plataformas: Linux Ubuntu, Android, Windows y Mac OSX. Es decir, se realizaron tres validaciones en plataformas de escritorio y una para plataforma de dispositivos móviles.
A partir de esta validación se detectaron áreas carentes, tanto de técnicas como de herramientas que apliquen técnicas que han sido propuestas por algunos autores.
Con este mapa de la realidad actual, se procederá a seleccionar las áreas en las que se trabajará en el próximo año.
Como ya se ha mencionado, las ciencias forenses deben cumplir con tres principios básicos: evitar la contaminación, actuar metódicamente y controlar la cadena de evidencia. El método es el que permite garantizar el trabajo realizado, su confrontación en un juicio oral, de ser necesario, y la trazabilidad.
Por las características propias de las tecnologías de la información y la comunicación, su gran dinamismo y diversidad, era necesario contar con algún proceso que sea lo suficientemente amplio, como para adaptarse a cualquier tecnología, y a su vez, tuviera guías concretas de implementación en plataformas específicas con herramientas actuales y a disposición.
Entendemos que esta primer propuesta de un proceso unificado de recuperación de la información, su difusión y uso, pueden ser el puntapié para que este proceso siga madurando y fortaleciéndose.

Referencias Bibliográficas

1. CANO M. Jeimy J, Computación Forense. Ed. Alfaomega, p 1-10 Año 2009.
2. Forensic Examination of digital Evidence: A Guide for law enforcement, NIJ Report, US Department of Justice, Office of Justice Programs, disponible en http://www.ojp.usdoj.gov/nij (accedido el 30 de Mayo de 2012)
3. ACPO (Association of Chief Police Officers) – England, Wales and North Ireland. Good Practice Guide for Computer-Based Electronic Evidence. Oficial release version. Disponible en www.acpo.police.uk (accedido el 28 de Mayo de 2012)
4. CARVEY Harlan, Windows Forensic Analysis. Ed. Syngress. Año 2009.