3° Conferencia Nacional de Informática Forense

6 y 7 de Junio de 2019,

Facultad de Ciencias Exactas Físicas y Naturales (FCEFyN)
Universidad Nacional de Córdoba (UNC)

INFOCONF 2019

Actas

3° Conferencia Nacional de Informática Forense

INFOCONF 2019

6 y 7 de Junio de 2019,

Facultad de Ciencias Exactas Físicas y Naturales (FCEFyN)
Universidad Nacional de Córdoba (UNC)

Actas de la 3ra Conferencia Nacional de Informática Forense / Ignacio Martín Gallardo ... [et al.]. -
1a ed. - Córdoba : Universidad Nacional de Córdoba. Facultad de Ciencias Exactas, Físicas y Naturales, 2019.
Libro digital, PDF

Archivo Digital: descarga y online
ISBN 978-950-33-1553-8

1. Seguridad Informática. 2. Forense. I. Gallardo, Ignacio Martín
CDD 005.4

Comité organizador

Presidencia

Mg. Ing. Pablo RECABARREN (UNC – FCFEYN)

Ing. Luis Bosch (UNC – FCFEYN)

Mg. Ing. Eduardo Casanovas (IUA-UNDEF)

Dra. Maria Alejandra Juana Hillman (MPF)

Comité Academico

Presidente

Mg. Ing. Miguel Solinas (UNC – FCFEYN)

Miembros

Esp. Ing Javier Jorge (UNC)
Ing. Juan Ignacio Alberti (UFASTA)
Ing. Hugo Carrer (UNC)
MCs Ing. Eduardo Esteban Casanovas (IUA)
Ing. Martín Castellote (UFASTA)
Esp. Ing. Alicia Castro (UNSL)
Abog. José María Cifuentes (MPBA)
Ing. Bruno Constanzo (UFASTA)
MCs. Ing. Hugo Curti (UFASTA)
Esp. Ing. Ana Di Iorio (UFASTA)
Ing. Maximiliano Eschoyez (UNC)
Ing. Franco Filippi (MPFCBA)
Dr. Juan Fraire (UNC)
Lic. Ana Cecilia Gallardo (MPFCBA)
Abog. María Laura Giménez (MPBA)
Ing. Fernando Greco (MPBA)
Esp. Ing. Walter Heffel (ENERSA)

Ing. Juan Ignacio Iturriaga (UFASTA)
Esp. Abog. Sabrina Lamperti (UFASTA)
Ing. Sebastián Lasia (UFASTA)
Dr. Ing. Marcelo Martín Marciszack (UTN-FRC)
Abog. Franco Martini (MPFCBA)
Dr. Fernando Menzaque (UNC)
Dr. Ing. Orlando Micolini (UNC)
MCs. Abog. Luciano Monchiero (MPFCBA)
Dr. Damian Morero (UNC)
Abog. Franco Pilnik (MPFCBA)
Ing. Ariel Podestá (UFASTA)
Ing. Gonzalo Ruiz de Angeli (UFASTA)
Ing. Santiago Trigo (UFASTA – MPBA)
Ing. Luis Ventre (UNC)

Auspiciantes

INFOCONF 2019

Ciberseguridad y Ciberdefensa (WCCD)

Gestión de Datos de Seguridad / Seguridad y Salud / Sistemas de Control Industrial / Seguridad en Redes / Seguridad en la Nube / Seguridad en Sistemas Embebidos / Tecnologías Móviles Seguras / Desarrollo de Software Seguro / Análisis y Respuesta a Incidentes de Seguridad / Políticas de Seguridad

Seguridad en dispositivos móviles Android

Alegre Andrés Santiago, Martinez Campos Silvana Pompeya

Universidad Católica de Salta, Facultad de Ingeniería e Informática
{asantiago.alegre, mcpompeya}@gmail.com

Resumen. El presente trabajo tiene por finalidad mostrar el funcionamiento del sistema operativo Android en equipos celulares, un detalle de los riesgos actuales en materia de seguridad y robo de información y un compendio de buenas prácticas para evitar ser víctimas de delitos informáticos. Este trabajo se realiza en el marco del proyecto titulado "Aplicaciones de metodologías, procesos y técnicas forenses digitales a nuevas tecnologías" el cuál se desarrolla en el marco de la Facultad de Ingeniería de la Universidad Católica de Salta.
En el siguiente trabajo se presenta los siguientes objetivos:
Conocer los conceptos básicos de seguridad.
Conocer el funcionamiento y el proceso de inicialización de un dispositivo Android. Conocer los riesgos y amenazas asociados al uso de equipos móviles con sistemas operativos Android. Conocer las mejores prácticas para reforzar la seguridad en dispositivos móviles.

Palabras Clave: Dispositivos móviles, ataques, seguridad, amenaza, Android.

1 Introducción

La evolución de los dispositivos móviles ha aumentado considerablemente en los últimos años, han pasado de ser simples celulares a computadores de mano. Por esta razón, las actividades cotidianas de las personas como revisar e-mail, redes sociales, sitios de interés, noticias, incluso el banco online, han pasado a utilizarse en dichos equipos. Éstos, se han convertido en una extensión de la vida diaria de las personas, tanto personal como laboralmente.

Ante el marcado aumento en el uso de dispositivos móviles, el mercado destinado a estas terminales ha crecido fuertemente; los servicios ofrecidos y las aplicaciones cada día son más numerosas. Se puede estar conectado prácticamente desde cualquier lugar, a la hora que se desee. Pero estos beneficios no son gratis debido a que llama la atención de personas y cibercriminales que han visto un gran mercado y que en los últimos años su explotación ha estado en aumento, sacando provecho a través de los diferentes delitos informáticos. Debido a esto los dispositivos móviles están expuestos a un sinnúmero de peligros. La penetración del sistema operativo Android en los usuarios de dispositivos móviles en América Latina ha sido muy fuerte, por su flexibilidad de adaptarse a cualquier equipo hardware.

2 Conociendo al gigante

Android es un sistema operativo basado en Linux y de código abierto, para dispositivos móviles, que inicialmente se desarrolló por Android Inc. Empresa que más tarde compró Google en 2005. Fue presentado el 5 de noviembre de 2007 con la fundación de la Open Handset Alliance [1]. Es esta empresa la que incorporó el primer dispositivo en utilizar el sistema operativo Android que recibió el nombre de HTC Dream y al ser de código abierto fue adoptado por una gran cantidad de fabricantes que hoy lideran el mercado. Debido a que cualquier fabricante puede adaptar Android para sus dispositivos, la penetración que tiene en el mercado es enorme y continúa creciendo convirtiéndose en el sistema operativo para dispositivos móviles más utilizado según los datos recopilados por StatCounter [2] hasta Agosto del 2018 en donde Android sobrepasó a Windows en un 5.73 %.

Cada nueva versión de Android introduce mejoras y nuevas funcionalidades. También se corrigen bugs y errores detectados. Se identifican por una numeración. A los dos primeros dígitos de esta se le da el nombre de un postre popular en inglés. A cada versión generalmente se le agregan modificaciones, surgiendo pequeñas actualizaciones menores. Todas las versiones se pueden ver a continuación, ordenadas por las más recientes.

Versión/Nombre	**Fecha**
Android 9.0 Pie	6 de Agosto de 2018
Android 8.0 Oreo	21 de Agosto de 2017
Android 7.0 Nougat	Agosto de 2016
Android 6.0 Marshmallow (Malvavisco)	Octubre de 2015
Android 5.0 Lallipop	Noviembre de 2014
Android 4.4 KitKat	Noviembre de 2013
Android 4.3 JellyBean (Michel)	Julio de 2013
Android 4.2 JellyBean (Gummy bear)	Noviembre de 2012
Android 4.1 JellyBean (Gomita confitada)	Julio de 2012
Android 4.0 Ice Cream Sandwich (Sandwich de helado)	Octubre de 2011
Android 3.0/3.1/3.2 Honeycomb (Panal de miel)	Febrero de 2011
Android 2.3 Gingerbread (Pan de jengibre)	Diciembre de 2010
Android 2.2 Froyo (Yogur helado)	Mayo de 2010
Android 2.0 Éclair	Octubre de 2009
Android 1.6 Donut	Septiembre de 2009
Android 2.5 Cupcake	Abril de 2009
Android 1	Septiembre de 2008

Table 1. Versiones de Android hasta el 2018

3 Arquitectura del sistema operativo Android

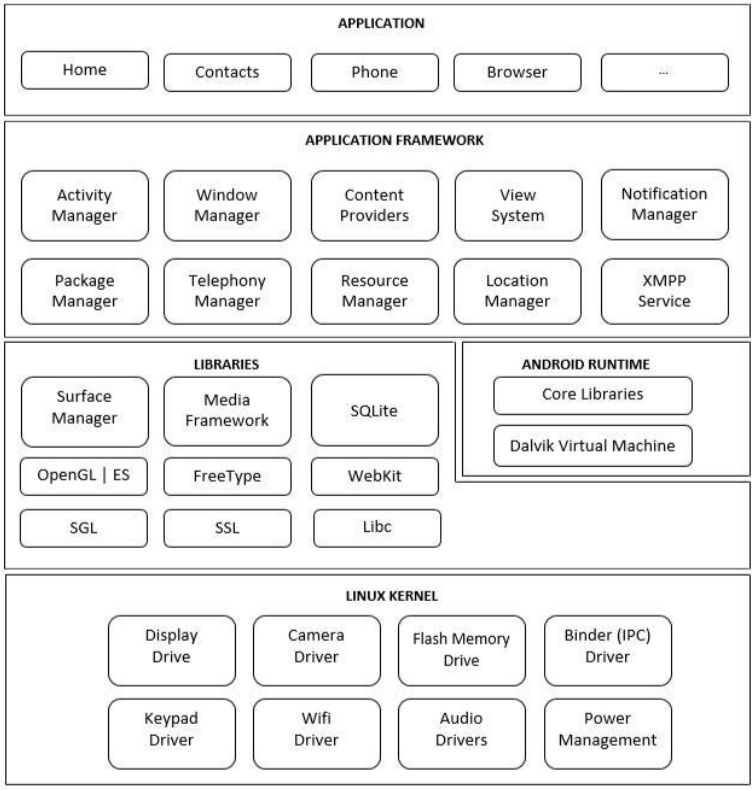


Fig. 1. Arquitectura de Android

Cómo se muestra en la Figura 1 la arquitectura está formada por cuatro (4) capas. Ellas son las siguientes:

1. **Aplicaciones:** Este nivel contiene, tanto las incluidas por defecto de Android como aquellas que el usuario vaya añadiendo posteriormente, ya sean de terceros o de su propio desarrollo. Todas estas aplicaciones utilizan los servicios, las API y librerías de los niveles anteriores.
2. **Framework de aplicaciones:** Los desarrolladores tienen completo acceso a los mismos APIs del framework usados por las aplicaciones base. Esta capa está diseñada para simplificar la reutilización de componentes; cualquier aplicación puede publicar sus capacidades y cualquier otra aplicación puede luego hacer uso de esas

capacidades (sujeto a reglas de seguridad del framework). Este mismo mecanismo permite que los componentes sean reemplazados por el usuario.

3. **Librerías:** La siguiente capa la componen las bibliotecas nativas de Android, también llamadas librería. Incluye un conjunto de librerías en C/C++ usadas en varios componentes de Android, están compiladas en código nativo del procesador y muchas utilizan proyectos de código abierto. Normalmente están hechas por los fabricantes, quienes también se encarga de instalarlas en los dispositivos. El objetivo de las librerías es proporcionar funcionalidad a las aplicaciones para tareas que se repiten con frecuencia. Entre las librerías más importantes ubicadas aquí se encuentran OpenGL, Bibliotecas multimedia, Webkit, SSL, FreeType, SQLite, entre otras.
4. **Tiempo de ejecución de Android**: Al mismo nivel que las librerías de Android se sitúa el tiempo de ejecución. Está constituido por las Core Libraries, que son librerías con multitud de clases Java y la máquina virtual Dalvik. Dalvik ha sido escrito de forma que un dispositivo puede correr múltiples máquinas virtuales de forma eficiente. Cada aplicación Android corre su propio proceso, con su propia instancia de la máquina virtual Dalvik.
5. **Kernel:** El núcleo de Android está formado por el sistema operativo Linux con un Kernel versión 2.6. Esta capa proporciona servicios como la seguridad, el manejo de la memoria, el multiproceso, la pila de protocolos y el soporte de drivers para dispositivos. Dicha capa actúa como capa de abstracción entre el hardware y el resto de la pila. Por lo tanto, es la única que es dependiente del hardware.

4 Proceso de inicialización de Android

Android como cualquier otro sistema operativo también tiene su proceso de arranque. Hay que tener en cuenta que el procesador que utilizan la mayoría de los dispositivos móviles es ARM (se basan en el modelo RISC y están licenciados por la compañía británica ARM Holdings) y el hardware es distinto también, por lo tanto, el kernel está configurado y adaptado para el hardware del dispositivo. A continuación, se procede a describir la Figura 2, proceso de Inicialización de Android:

Boot ROM: Cuando se inicializa el dispositivo y arranca el procesador, este tiene la dirección de memoria en donde se encuentra el Bootloader y procederá a cargar dicho programa en la RAM del dispositivo para comenzar su ejecución.

Bootloader: Es el primer programa que corre cuando se enciende un dispositivo móvil y es el encargado de gestionar el arranque del sistema. También se asegura de que todos los componentes de hardware estén en correcto funcionamiento y en ese caso se encargará de hacer iniciar Android en el caso de un inicio normal o en el caso de que sea solicitado por el usuario podrá iniciar con Recovery, Fastboot u otros.

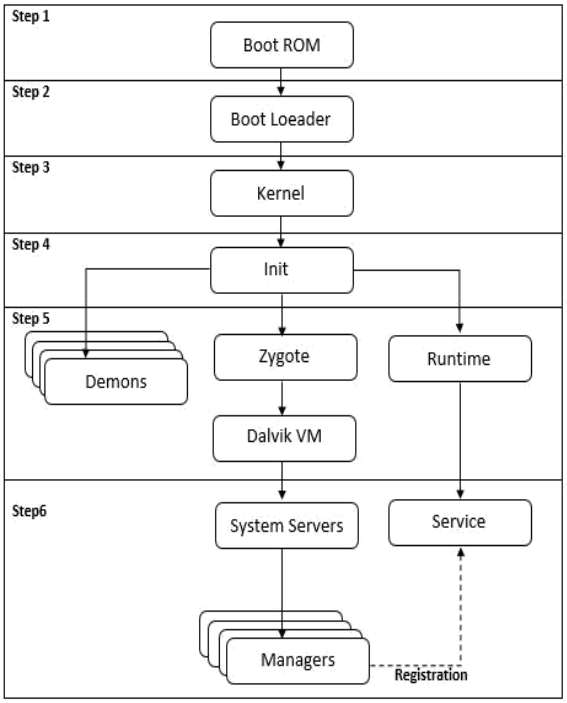


Fig. 2. Proceso de Inicialización de Android

Kernel: El Kernel de Android inicializa igual que cualquier sistema linux: montará los filesystem necesarios, inicializa la memoria, dispositivos, el cache, cargará los drivers y cuando termine todas estas tareas accederá a la partición raíz para lanzar el proceso "init" el cual será el primero en ejecutar del sistema operativo que dará inicio a Android.

El proceso Init: Es el primero de todos los que corre el sistema operativo de base, se puede decir que es el proceso raíz. Se encuentra en la base del system "/" proceso encargado de montar los principales filesystems del sistema operativo para luego dar lugar a la ejecución del archivo "init.rc" el cual se encargará de cuatro clases bien definidas de operaciones: acciones, servicios, comandos y opciones, que servirán de base para el funcionamiento de los siguientes procesos en jerarquía de Android. Al realizar todas estas tareas lo primero que se podrá visualizar en el teléfono es el logo de Android o su correspondiente logo de inicialización.

Zygote y Dalvik: El servicio de Zygote es inicializado desde Init y es el encargado de inicializar las máquinas virtuales Dalvik. Crea máquinas virtuales por cada proceso nuevo que se inicia. Dalvik es la máquina virtual que se generará para cada proceso y tiene un funcionamiento muy similar al de las VM de Java.

Servicios del Sistema o Servicios*: Completada* la etapa anterior, Zygote se encargará de inicializar los servicios de sistema. Entre los principales servicios se encuentran: telefonía, teclado, batería, alarmas, sensores, administrador de ventanas, y otros agentes

y servicios de google. Cuando finaliza el inicio de estos servicios, el sistema está listo para interactuar con el usuario.

***Boot completo**:* Este es el último paso en la etapa de inicialización, cuando todas las etapas concluyeron se dispara una acción de BROADCAST denominada "ACTION_BOOT_COMPLETED" la cual indica la finalización del proceso.

5 Riesgos y amenazas asociados al uso de equipos móviles con sistemas operativos Android

Ataques como malware, códigos QR falsos, phishing, fraudes informáticos, pérdida o robo del dispositivo, conexiones inalámbricas inseguras, entre otros hacen que cada vez sea más necesario adoptar medidas de protección tanto desde el ámbito tecnológico como de educación y concientización. Para empezar, nos centraremos en el malware y los diferentes objetivos que este puede llegar a tener. Puede tener objetivos muy variados, siendo los más comunes obtener datos personales y beneficio económico. Su modo de funcionamiento puede ser automático o controlado remotamente. Los principales tipos de malware son según el informe *Tendencias 2013: Vertiginoso crecimiento de malware para móviles.* [4]

- **Virus:** Es un programa malicioso que infecta a otros archivos del sistema con la intención de modificarlos o hacerlos inservibles. Cuando un archivo ha sido infectado, también se convierte en portador del virus y, por lo tanto, en una nueva fuente de infección. Para que un virus se propague, este archivo debe ser ejecutado por el usuario.
- **Gusano:** Es un programa malicioso autor replicable que aprovechará las vulnerabilidades de la red para propagarse.
- **Troyano:** Es un pequeño programa oculto en otra aplicación. Su objetivo es pasar inadvertido por el usuario e instalarse en el sistema cuando el usuario ejecuta la apli-cación. Una vez instalado, sin el consentimiento del usuario puede realizar diversas acciones instantáneamente o estar fijadas para realizarse en un futuro.
- **Spyware:** Es una aplicación que recoge información sobre una persona u organización sin su consentimiento. Generalmente, el objetivo final de esta información recopilada es venderla a empresas de publicidad.
- **Keylogger:** Es una aplicación encargada de almacenar todas las acciones del móvil. Por lo tanto, puede capturar información confidencial, como el número de la tarjeta de crédito o las contraseñas u otro tipo de información sensible.
- **Hijacker:** Es un programa que realiza cambios en la configuración del navegador **w**eb. Un ataque típico es cambiar la página de inicio por una página de publicidad.
- **Dialer:** Es un programa que de manera oculta realiza llamadas a teléfonos con tarifas especiales. De esta manera, el atacante puede obtener beneficios económicos.

La mayoría de las amenazas para dispositivos móviles no presentan sospecha de que el dispositivo de la víctima está infectado. Android ha sido diseñado con dos capas de

seguridad, la primera de ellas es restringida al usuario y en donde se ejecutan y almacenan todos los procesos y archivos vitales para el sistema y la segunda capa de seguridad es donde reside la información personal del usuario, las aplicaciones y configuraciones que ha descargado e instalado. Así cuando un Smartphone es infectado, las acciones de estos solo tienen efecto sobre los datos y configuraciones del usuario. Es por esto que técnicas como el rooteo siempre es una tarea muy peligrosa porque deja una puerta abierta a la entrada de virus a los niveles más seguros del sistema.
A continuación en la Tabla 2 [5] se muestran algunos comportamientos que pueden indicar que un dispositivo está infectado.

Síntomas	**Posibles causas**
Cargos extraños en la cuenta o gastos excesivos de saldo.	Troyanos SMS: Gran cantidad de malware para Android está diseñado para suscribir a la víctima a números de mensajería SMS Premium.
Comportamientos como la apertura y cierre de aplicaciones, envío de mensajes y llamadas internacionales no realizadas.	Malware: Algunos códigos maliciosos solicitan varios permisos para lograr obtener el control del teléfono inteligente.
Aparición de aplicaciones no instaladas por el usuario.	Malware: Códigos maliciosos como Android/MarketPay descargan e instalan aplicaciones sin el consentimiento de la víctima.
Consumo excesivo de datos.	Malware/otros: Un consumo desmesurado de datos en redes 3G/4G podría indicar la presencia de una infección.

Tabla 2. Síntomas de un dispositivo infectado.

Si el usuario percibe algunos de estos síntomas es posible que su teléfono se encuentre infectado con código malicioso.

- **Phishing:** El phishing es una amenaza informática a través de la cual los cibercriminales suplantan a una entidad de confianza, como un banco u otra empresa, para robar información de la víctima. Por lo general, el phishing se suele propagar median-te un correo electrónico en el que se le indica al destinatario que deberá entregar determinados datos; de lo contrario, se amedrenta al usuario indicándole que su cuenta o servicio podría presentar problemas. También existen ataques de phishing que requieren interacción telefónica para poder concretar el robo, es decir, utilizan el teléfono celular como una herramienta más para poder llevar a cabo el robo de dinero.
- **Fraudes electrónicos (scam):** Los fraudes electrónicos o scam también son amenazas que suelen llegar a través de correo electrónico. A diferencia del phishing, los scam utilizan premios falsos para seducir a la potencial víctima. En el texto es común observar que los atacantes solicitan datos personales del usuario y una suma

de dinero que debe depositarse antes de poder cobrar el premio (dicha suma suele ser considerablemente más alta que el monto solicitado a pagar). Al tratarse de amenazas que utilizan el correo electrónico y sitios web para operar, ambos ataques pueden afectar a usuarios de dispositivos móviles de igual modo que a una persona que utiliza un computador de escritorio o notebook.

- **Smishing:** También existen ataques de phishing diseñados y adaptados específicamente para usuarios de equipos móviles. Este tipo de amenaza llega a través de mensajes de texto (SMS) y se le clasifica como Smishing (conjunción de SMS y Phis-hing).
- **Vishing:** Conjunción de Voice y Phishing. Son fraudes informáticos cometidos mediante la tecnología Voz sobre IP, es decir, a través de un llamado telefónico. En estos casos, el cibercriminal solicita información sensible como el número de la tarjeta de crédito de la víctima, a la que se le pide que lo ingrese a través del teclado del dispositivo móvil.
- **Robo o extravío físico del equipo:** Debido a que los dispositivos móviles acompañan al usuario prácticamente a cualquier lugar, la posibilidad de que se pierdan o sean robados es bastante alta. Frente a este tipo de situaciones, la pérdida del equipo no es el mayor problema, la preocupación pasa por la información sensible almacenada en el dispositivo que puede ser obtenida por terceros. Asimismo, en el caso de que el usuario no cuente con un respaldo de los datos, recuperar esa información podría complicarse, más si son documentos de autoría propia.
- **Ingeniería social:** En términos técnicos, la ingeniería social se refiere a un ataque en el que se utilizan las habilidades sociales para obtener información, ya sea personal como sobre los sistemas informáticos. Las personas que practican la ingeniería social (estafadores) utilizan la interacción social ya sea para engañar a sus víctimas para que entreguen la información, o manipularlos para que confíen en él o ella y compartan información con el atacante.

Por lo general, el atacante fingirá ser otra persona, Este tipo de ataque no es una amenaza específica contra la seguridad informática, sino más bien una cuestión general de seguridad personal que se traslada también a la seguridad móvil.

- **Conexiones inalámbricas wifi y Bluetooth inseguras**: Una de las particularidades de los dispositivos móviles es la capacidad de conexión a diferentes redes inalámbricas para navegar por Internet o para compartir recursos como impresoras, archivos, etc. En esta línea, una de las tecnologías más utilizadas por los usuarios son las conexiones inalámbricas wifi. Debido a su facilidad de uso e implementación en varios lugares como hoteles, casas, oficinas, cafés, restaurantes, entre otros, es común que las personas se conecten a este tipo de redes con el objetivo de consultar el correo electrónico, visitar sitios web y redes sociales, realizar transferencias bancarias, etc. Aunque esto supone una comodidad, una conexión wifi insegura puede poner en riesgo la integridad de la información del usuario. En este sentido, aquellas redes que no están protegidas por una contraseña y que no implementan ningún tipo de cifrado son susceptibles a ataques como sniffing o robo de paquetes, más si las redes son de acceso público. En dichos escenarios, un cibercriminal podría interceptar el tráfico de red para robar información sensible como credenciales

de acceso. Asimismo, un atacante puede montar una red wifi fraudulenta que utilice el mismo nombre de punto de acceso (AP) que otra conexión con el objetivo de realizar acciones maliciosas. Con res-pecto a la tecnología Bluetooth, esta facilita la conexión a Internet utilizando otro dispositivo como también la posibilidad de compartir archivos entre usuarios que se encuentran físicamente cercanos. Pese a que algunos teléfonos vienen con esta función desactivada, mantener activado el Bluetooth no solo consume más batería, sino también expone al usuario a riesgos de seguridad como la posibilidad de que un código malicioso utilice una conexión de este tipo para propagarse de un equipo hacia otro. Además, desde que se implementó esta tecnología, se han descubierto vulnerabilidades que atentan en contra de la seguridad del usuario.

- **Códigos QR maliciosos:** Los QR (Quick Response) son un tipo de código de barras en dos dimensiones que facilitan el acceso a sitios web y otro tipo de información. Debido a su facilidad de uso y a la proliferación de diversos teléfonos inteligentes, muchas empresas de marketing e incluso algunos supermercados, utilizan esta tecnología para que las personas puedan acceder rápidamente a varios recursos como la compra de productos en línea. Con tan solo apuntar la cámara de un dispositivo móvil se puede leer este tipo de código de barras, sin embargo, con la misma facilidad, un atacante puede utilizar un código QR con propósitos maliciosos como dirigir al usuario hacia la descarga de malware o sitios de phishing.
- **Bring Your Own Device (BYOD):** El fenómeno BYOD (BringYourOwnDevice o Traiga su propio dispositivo) consiste en la utilización de dispositivos como notebooks, teléfonos inteligentes y tabletas que son propiedad del usuario en ambientes corporativos. Pese a que esta tendencia permite una mayor flexibilidad al facilitar que un empleado pueda trabajar desde cualquier lugar y utilizando el equipo de su elección, también atenta en contra de la Seguridad de la Información en caso que no se adopten los resguardos necesarios.

6 Mejores prácticas para reforzar la seguridad en dispositivos móviles

Android es el sistema operativo móvil que cuenta con el mayor market share, por lo tanto, es el más afectado por malware. La importancia de la información y la existencia de otras amenazas hacen necesario la implementación de medidas de seguridad [5]:

- **Redes VPN:** Posibilitan que las personas puedan conectarse a redes corporativas a través de una conexión a Internet. Protegen la información que es transmitida del computador del usuario hacia sitios y servicios remotos como correo electrónico, redes sociales, portales bancarios y otros. Se la debe utilizar siempre que sea posible pero en entornos corporativos es fundamental.
- **Verificación de HTTPS en dispositivos móviles:** Otra práctica que permite mitigar ataques como el robo de contraseñas a través de una conexión wifi insegura es el uso del protocolo HTTPS y un certificado válido. Es fundamental cuando se utilizan ser-vicios como portales bancarios, redes sociales, correo electrónico, y cualquier otro sitio que requiera de credenciales de acceso o datos sensibles para funcionar.

- **Prevención de códigos maliciosos y otras amenazas:**
  - Activar el bloqueo de la tarjeta SIM y establecer una clave de acceso para prevenir que terceros puedan acceder a información confidencial. Esta es más segura si está conformada por más de cuatro caracteres y los números no son consecutivos ni fáciles de adivinar (los bloqueos por patrones y reconocimiento facial son más inseguros). Se debe implementar una solución de seguridad para dispositivos móviles que permita protegerlo de códigos maliciosos, mensajes indeseados y del borrado de la información de forma remota en caso de robo o extravío.
  - Mantener el sistema operativo móvil y las aplicaciones actualizadas para evitar la explotación de vulnerabilidades corregidas.
  - Descargar aplicaciones provenientes exclusivamente de tiendas o repositorios oficiales. También es fundamental leer el contrato de licencia que acompaña al software.
  - Al momento de instalar una aplicación nueva, es importante fijarse en los permisos que solicita dicho software para poder funcionar. En este sentido, se hace necesario establecer qué funciones cumple un programa y en base a esto, los permisos estrictamente necesarios para que la aplicación pueda ejecutarse correctamente.
  - Realizar una copia de seguridad de la información almacenada en el equipo.
  - Desactivar el uso de conexiones inalámbricas wifi, Bluetooth, Infrarrojo y de la tecnología de transmisión NFC, con el objetivo de prevenir amenazas que se propagan por los mismos.
  - No seguir enlaces provenientes de mensajes, correos electrónicos, redes sociales, etc. Asimismo es recomendable ingresar a sitios web de esta naturaleza escribiendo la dirección directamente en la barra del navegador.
  - Anotar el número de identificación IMEI (International Mobile Equipment Identity) en un lugar que no sea el mismo dispositivo móvil. Dicho código se utiliza para identificar a los teléfonos celulares y sirve para bloquear el acceso de un equipo a la infraestructura móvil por parte del operador en caso de robo o pérdida.
- **Configuración de Parámetros de seguridad en Android:** a continuación, se mencionan una serie de parámetros de configuración que pue-den ser modificados para aumentar el nivel de seguridad de un dispositivo que ejecuta Android. Es posible que algunas opciones aparezcan de forma distinta de acuerdo al modelo, marca y versión del sistema operativo móvil.

  (a) **Menú seguridad:** lo primero que se debe modificar son algunos parámetros de la opción “Seguridad”. A través de esta sección, es posible activar el bloqueo de pantalla (1) y tarjeta SIM (2), ocultar las contraseñas escritas (3), visualizar las aplicaciones con mayores privilegios de ejecución (4), limitar la posibilidad de instalar aplicaciones provenientes de orígenes desconocidos o repositorios no oficiales (5) e instalar certificados de seguridad (6) en caso de ser necesario (comúnmente en entornos corporativos). Todas estas opciones pueden ser accedidas a través de: **Menú principal – Ajustes - Seguridad**. A continuación, se muestra una captura de referencia:

Fig. 3. Parámetros de seguridad de Android

Es importante destacar que en la opción que muestra aquellas aplicaciones con mayores privilegios (6), solo deberán aparecer programas que realmente lo necesitan como soluciones de seguridad, software de rastreo por GPS, entre otros.

Fig. 4. Opciones de cifrado de Android

(b) **Activar cifrado de memoria y tarjeta SD**: otra opción que incluyen algunas versiones de Android es la posibilidad de cifrar la información almacenada en el equipo. De este modo y sin la contraseña necesaria, será muy difícil que un ter-

cero pueda acceder a los datos sin la respectiva clave. Para activar dicha característica se debe ingresar de esta forma: **Menú principal -Ajustes -Memoria**. La siguiente imagen muestra las opciones que se deben activar:

La "Encriptación de almacenamiento" (1) cifra aquella información almacenada en la memoria interna del dispositivo. Por otro lado, la opción "Encriptación de memoria del teléfono" (2) cifra los datos guardados en la tarjeta SD del equipo.

(c) **Actualizaciones automáticas**: para facilitar la tarea de actualizar las aplicaciones es posible configurar Android para que este procedimiento sea realizado de forma automática. Para configurar esta opción se debe acceder al **Menú principal - Google Play -Ajustes**.

Fig. 5. Activar actualizaciones automáticas en Google Play

Al activar las actualizaciones automáticas es posible especificar que dicho procedimiento sea realizado solo a través de redes wifi o mediante cualquier conexión como datos (puede incurrir en gastos adicionales dependiendo del plan contratado por el usuario).

Fig. 6. Configuración para redes VPN

(d) **Configuración de redes VPN**: A continuación, se mencionan los pasos necesarios para poder establecer una red VPN en Android, sin embargo, los datos de con-figuración deben ser proveídos por la empresa prestadora del servicio. Ingresar a **Menú principal** -**Ajustes** -**Más** -**VPN**.

- **Política Screen-lock (Bloqueo de Pantalla):** Poder configurar una pantalla de bloqueo mediante un PIN, patrón de puntos, contraseña o biometría es una característica común en los dispositivos móviles para aumentar la seguridad y la privacidad del usuario proporcionando un mecanismo de autenticación para evitar que un atacante pueda tener acceso directo a los contenidos del dispositivo. Una política de bloqueo de pantalla está configurada definiendo cual es el nivel mínimo de longitud del PIN o contraseña aceptable implementando entropía. La entropía es una medida de la seguridad de la contraseña y su resistencia contra ataques de fuerza bruta.

  Una política adecuada de bloqueo de pantalla para dispositivos móviles debe incluir lo siguiente:

  - Requisito de bloqueo de pantalla: Los usuarios deben tener un PIN o contraseña activada en el dispositivo.
  - Alta entropía: Especificar una longitud mínima, número de letras, dígitos, símbolos, etc.
  - Intentos fallidos: Si se hacen demasiados intentos de inicio de sesión fallidos, limpie el dispositivo.
  - Temporizador de Bloqueo de pantalla: Tiempo de inactividad antes de bloqueo de pantalla se active.

  Esto evitará que usuarios no autorizados tengan acceso directo a las aplicaciones y archivos, y prohíbe la fuerza bruta y ataques de diccionario, de esta manera se mejora significativamente la confidencialidad e integridad de los datos del usuario cuando se configura una política de contraseñas. La longitud del PIN, la cantidad de puntos de un Patrón, la cantidad de caracteres y tipos de una contraseña determinarán cuan seguro será el método.

- **Bloqueo remoto, tracking y borrado de la información:** Los dispositivos que se utilizan en una organización con frecuencia tienen más probabilidades de contener algún tipo de datos sensibles que los que son de uso personal. Una manera eficaz de proteger el dispositivo y la confidencialidad e integridad de los datos, además de realizar un seguimiento geo posicional del dispositivo y una política de limpieza es un bloqueo remoto. Esto permitirá que un usuario o un administrador pueda activar de forma remota el bloqueo del dispositivo, conocer la ubicación o iniciar un borrado que elimine todo el contenido del dispositivo. La función de bloqueo remoto tiene como fin prohibir el acceso directo al dispositivo y el borrado remoto intenta asegurar que los datos ya no están en riesgo.
- **Control de aplicaciones:** Android presenta la opción de habilitar o deshabilitar el permiso para que se puedan instalar aplicaciones desde fuentes que no sean su

tienda oficial. Ello implica que posee mayores riesgos que las demás plataformas, al hacer "más fácil" que el usuario cometa un error; instalando una aplicación maliciosa, que se hace pasar por otra cosa, desde una fuente externa. Sin embargo, los repositorios oficiales no están completamente libres de peligros por lo que siempre es recomendable realizar una revisión de las aplicaciones antes de instalarlas independientemente del medio desde donde se obtengan. Si a la hora de instalar una aplicación que hace uso del led como una linterna esta solicitara permisos por ejemplo para usar el GPS, para realizar llamadas o acceder a datos de navegación, el usuario ya tiene una buena base de información para sospechar que la aplicación posee intencionalidades diferentes a las que informa.

7 Conclusión.

En este trabajo se llevó a cabo un recorrido básico sobre las características de seguridad que se debería de tener en cuenta ante dispositivos móviles. Hoy en día la mayoría de las organizaciones están en peligro, no importa cuál sea la dimensión, pero peligro al fin. Existen tantas medidas de seguridad como modelos de dispositivos hay; pero no estamos lo suficientemente capacitados o preparados para afrontar la perdida de información valiosa, perdida de materia prima fundamental que hace a la vitalidad de nuestra organización; por eso importante es ser precavidos, conocer sobre el tema y tomar las medidas necesarias para mitigar los riesgos que conlleva la fuga de datos. Por eso es tan importante la educación y el uso responsable de dispositivos tanto para fines personales como de trabajo.

Referencias

1. O. H. Alliance, 12 Noviembre 2007. [En línea]. Available: http://www.openhandsetalliance.com/.
2. Statcounter, «Statcounter - GlobalStat,» Septiembre 2018. [En línea]. Available: http://gs.statcounter.com/os-market-share.
3. A. Goujon, «welivesecurity by ESSET,» 29 Noviembre 2012. [En línea]. Available: https://www.welivesecurity.com/la-es/2012/11/29/tendencias-2013-vertiginoso-crecimiento-malware-moviles-parte-i/.
4. K. Yaghmour, de Embedded Android, 0'REILLY, 2013, p. 33.
5. A. ESSET, «Academia ESSET,» [En línea]. Available: https://www.academiaeset.com/default/store/14041-seguridad-en-dispositivos-moviles. [Último acceso: Agosto 2018].

Blockchain, formalidad contractual en la era digital

Eduardo Casanovas[1], José Ignacio Casanovas[1], Serafín Fernandez[1]

[1] Universidad de la Defensa Nacional – Facultad de Ingeniería CRUC – IUA
Av Fuerza Aérea 6500-Córdoba
ecasanovas@iua.edu.ar, {casanovasjoseignacio,
serafinfernandez2208}@gmail.com
http://www.iua.edu.ar

Resumen. En la actualidad es complejo asegurar a) la fiabilidad de la información, b) la transparencia y trazabilidad de las operaciones, c) la inmutabilidad y no manipulación de datos una vez guardados. Para hacer frente a estos desafíos existe una alternativa eficiente llamada blockchain que se puede ejecutar a través de contratos inteligentes.
En el presente trabajo les mostraremos las cuestiones básicas del uso de esta tecnología y los requerimientos necesarios para la programación de los contratos inteligentes. Además describiremos las ventajas y desventajas de decidir realizar un desarrollo sobre la Blockchain pública o privada.
Esta tecnología innovadora consta de mecanismos robustos basados en la transparencia absoluta de las operaciones realizadas, sin necesitar la participación de actores de intermediación que lo garanticen. Estamos frente a un verdadero cambio de paradigma, pues la confianza que se requiere al realizar cualquier acto entre partes quedará garantizada íntegramente por la tecnología, transformando toda transacción existente en una vinculación peer to peer.
Nuestra finalidad será exponer el marco teórico y fundamentos sobre los cuales esta arquitectura se maneja, para poder entender las ventajas y analizar aplicaciones sobre las que se puede implementar esta solución.

1 Introducción

Blockchain (BC) funciona como una arquitectura de base de datos distribuida que mantiene una lista continuamente creciente de registros protegidos de manipulación.
Blockchain existe en una red de ordenadores o "nodos", que pueden ser públicos o privados, según el tipo de desarrollo que se quiere implementar. Cuando se realiza una nueva transacción, esta es validada antes de ser incluida en el bloque siguiente de la cadena.

Cuando se describe a Blockchain se la compara con un libro de contabilidad pública y se destaca que esta tecnología no solo registra todas las transacciones que han ocurrido sino que todos los nodos intervinientes tienen una copia exacta de esos registros. Sus ventajas más importantes son la fiabilidad, transparencia, inmutabilidad y seguridad de la información sin la necesidad de intermediarios. Por otro lado, su desventaja más importante es el conocido "Ataque del 51", en donde para "forzar" un cambio en la red, más del 51% de su poder de cómputo debe estar de acuerdo y hacerlo

al mismo tiempo. Aunque es un problema lejano en las inmensas redes populares que usan esta arquitectura, existe como posibilidad.

Esta tecnología disruptiva permite a extraños que no tienen confianza entre sí realizar operaciones, dando nacimiento al paradigma de la formalidad contractual en la era digital.

2 Usos de esta tecnología

Un ejemplo claro de uso, son las operaciones de cambio de valores en el ciberespacio. El más famoso es Bitcoin, una criptomoneda (moneda digital), concebida en 2009. Gracias a Blockchain, las criptomonedas no dependen de la confianza en un emisor central sino en las características propias de la tecnología.
El mundo de las **finanzas** está siendo revolucionado por las criptomonedas, que avanzan hacia alcanzar las dos principales características que las diferencian del dinero tradicional:

- ser un medio de pago generalmente aceptado
- estar respaldado en algo que garantice su valor

La garantía del valor la puede suplir la tecnología innovadora y la aceptación generalizada gana terreno a pasos agigantados con adeptos a aceptarlas como medios de pago alrededor del mundo. Son entonces los propios Bancos quienes validan su potencialidad disruptiva, al convertirse en los principales inversores de las fintech emergentes.

En la actualidad, el uso de Blockchain no solo se enfoca en el negocio de las criptomonedas, sino que hay muchos modelos de negocio que aprovechan sus ventajas.

Estamos ante una revolución digital, pero lo que notamos es que las innovaciones siempre avanzan primero dentro de cada campo como si se tratara de compartimentos estancos. Esas innovaciones se convierten en disrupciones cuando se combina su uso en funcionalidades completamente nuevas. Se avanza en internet y derivamos en internet de las cosas. Se avanza en redes sociales y buscadores que usan internet y derivamos en Marketing digital. Se avanza en blockchain y derivamos en su uso para **contratos inteligentes**. El término inteligente se debe simplemente a que están implementados a través de Blockchain.

Como expertos en seguridad informática y consultoría en marketing digital proponemos la combinación de contratos inteligentes aplicados al mundo del marketing digital. Un ecosistema que consideramos altamente permeable a la implementación de un avance de esta magnitud. Todos los actores intervinientes manejan terminología y conocimientos relacionados a la era digital. De esta forma la aplicación de blockchain en la formalización contractual puede servir de ejemplo para el resto de los actos jurídicos que impliquen un vínculo entre partes. El marketing digital requiere como cualquier actividad de contratos entre dueños de agencias y sus empleados, entre las agencias y sus clientes. La industria es particularmente dinámica y una solución que reduzca tiempos y costos mientras formaliza relaciones, resulta clave para la eficientización de las prestaciones de servicio. Un uso fundamental es el de aplicación

de blockchain para proteger los derechos de autor. En un ecosistema competitivo donde la creatividad es uno de los factores de éxito, es indispensable garantizar su genuina autoría. Otra aplicación de la blockchain es para compartir y **gestionar el uso de propiedad intelectual**, la tecnología Blockchain agiliza los derechos de propiedad, aportando transparencia a todos y una trazabilidad verificable desde el origen.

Otros de los usos son:

En **manejo de identidades,** para el seguimiento y la gestión de identidades digitales seguras y eficientes

En **"Internet de las cosas"(IoT)** para resolver la escalabilidad, la privacidad y problemas de fiabilidad

En **registros médicos** para proporcionar servicios bajo las instrucciones del hospital que la use, manteniendo el control total.

En el **ámbito policial, político, inmobiliario** y todo aquél sector que implique actos entre partes.

3 Requerimientos para su programación

Dijimos que el contrato es inteligente cuando se implementa a través de blockchain. En este segmento abordaremos la forma en la que esto se lleva efectivamente a cabo.
El contrato inteligente es un programa informático y como tal se construye con un lenguaje de programación.

Solidity es un lenguaje de programación de alto nivel cuya sintaxis es similar a otro de los lenguajes de programación más usados hoy en día. Este lenguaje está diseñado y compilado en código de bytes (bytecode) para crear y desarrollar contratos inteligentes que se ejecuten en la Máquina Virtual Ethereum (EVM de sus siglas en inglés). Los 'smart contracts' permiten que muchas de las partes de un negocio funcionen perfectamente por sí solas y además se lleve un registro de las mismas.

Solidity está diseñada específicamente para las aplicaciones Ethereum y se ejecuta sólo en la cadena de bloques Ethereum.

4 Entornos de desarrollo para Solidity

Remix, anteriormente conocido como Browser Solidity, proporciona un entorno de desarrollo integrado que permite escribir contratos inteligentes basados en Solidity.

Ethereum Studio es otro IDE que se caracteriza por tener especialización proporcionando un acceso completo a la red Ethereum, esto lo consigue a través de un intérprete de comando 'shell'. El plugin de Solidity para IntelliJ IDEA (y el resto de IDEs de la plataforma JetBrains), (https://plugins.jetbrains.com/plugin/9475-intellij-solidity).

A modo de ejemplo podemos mostrar un par de líneas de código de un contrato modelo:

```
1 contract TokenContractFragment {
2 // Balances for each account
3 mapping(address => uint256) balances;
4 // Owner of account approves the transfer of an amount to another account
5 mapping(address => mapping (address => uint256)) allowed;
6 // Get the token balance for account `tokenOwner`
7 function balanceOf(address tokenOwner) public constant returns (uint balance) {
8 return balances[tokenOwner];
9 }
10 // Transfer the balance from owner's account to another account
11 function transfer(address to, uint tokens) public returns (bool success) {
12 balances[msg.sender] = balances[msg.sender].sub(tokens);
13 balances[to] = balances[to].add(tokens);
14 Transfer(msg.sender, to, tokens);
15 return true;
16 }
17 ……
xx }
```

El plugin para Visual Studio está diseñado para permitir el desarrollo de contratos inteligentes de Solidity en este IDE de Microsoft. Para ello primero hay que adquirir Visual Studio y luego pasar a instalar la extensión. Visual Studio es un entorno rico de programación, integrado para la creación de aplicaciones para Windows, Android y iOS, así como aplicaciones web modernas y servicios en la nube. Su integración con Solidity propone a los desarrolles ir un paso más allá, facilitándoles el poder crear contratos inteligentes.

5 Red Pública vs Red Privada o Permisionada

Las Blockchains públicas, le permiten a cualquiera participar en ellas. Este tipo de red depende del número de usuarios para su correcto funcionamiento, por lo tanto motiva a la participación a través de un sistema de incentivo otorgando una recompensa que se les paga a los mineros que participan en la red. Los mineros son aquellos que ponen capacidad de cómputo a disposición de la red. Como cada transacción requiere ser procesada, la capacidad de procesamiento es fundamental. En la red pública de Ethereum se deberá pagar un fee (gas) por cada transacción que debe ser validada por lo mineros. Transacción es toda aquella acción que se haga sobre la red, por ejemplo:

- montar el contrato en la red
- modificarlo
- rescindirlo

La participación en una blockchain privada, requiere de una invitación, que a su vez debe ser validada por la red o a través de parámetros que se den a lugar. Dicha red se

conoce como una red autorizadora y pone una restricción a quién puede unirse. En la red privada no existe el costo de transacción pero sí conlleva un costo de construcción y mantenimiento que será afrontado por todas las instituciones que la conforman. Por ende, la decisión de utilizar una red privada frente a una pública depende de la cantidad de transacciones que se deban realizar por proyecto. En una red pública cada una de ellas tendrán un costo y dependiendo del tipo de contrato, este puede ser muy elevado, haciendo más rentable emprender la construcción de una red privada. Una analogía para describir la decisión de optar por construir una red privada o participar en una pública puede ser como la decisión de una empresa a la que le conviene construir un edificio de oficinas cuando el alquiler en un edificio no propio implique un costo mayor que el de edificar y mantener.

Un ejemplo de red privada puede ser Quorum, que ha sido desarrollada por JP Morgan y es open source. El punto diferenciador de Quorum es el hecho de que permite realizar transacciones privadas entre las partes.

Otro aspecto importante es respecto al algoritmo de consenso (Raft), que al igual que Proof of Stake, no exige el cálculo de un hash y valida bloques/transacciones en menos de 0.5 segundos. Pero la contracara es que no tiene tolerancia a fallos bizantinos, ´por lo que la seguridad tienen que darla las propias instituciones que conforman la red de forma que sus nodos no sean accesibles por un atacante. Luego, para el desarrollo se pueden usar Dapps, las cuales realizan peticiones a la red Blockchain desde un cliente. A continuación se agrega un cuadro en donde se han tomado 4 parámetros de comparación.

	Blockchain Públicas	Blockchain Privadas
Nivel de Acceso	• Sin Restricción	• Una sola organización
Participación	• Sin permisos • Anónimo	• Permisivo • Son conocidas los usuarios
Seguridad	• Mecanismo de consenso • Prueba de Trabajo/Prueba de Participación	• Pre-aprobación de participantes • Votaciones/Consensos múltiples
Rendimiento	• Baja velocidad de transaccionalidad	• Blockchain Liviana • Mayor velocidad de transacciones

Tabla 1: Cuadro comparativo entre BC Pública y Privada (Fte: https://nemespanol.io/blockchain-privada-vs-publica-cual-es-la-mayor-diferencia/)

6 Conclusiones

El mundo está girando alrededor de una revolución digital y por primera vez en la historia, Argentina avanza a la par de las mayores potencias. Blockchain es la disrupción por excelencia desde la invención de internet y sus funcionalidades han venido para quedarse. Es nuestra oportunidad de aprovechar nuestro capital humano calificado, para implementar soluciones que simplifiquen los vínculos entre personas y eficienticen el uso de recursos. Estamos convencidos que la idea de proponer un abordaje de la formalización contractual en el ámbito del marketing digital, permitirá traspolarlo al resto de los actos jurídicos existentes. Sabemos que el esfuerzo es grande pero también que estamos sobre el camino correcto.

Cada proyecto conllevará la necesidad de decidir si utilizar una red pública o privada dependiendo de las ventajas que apliquen a su situación particular. Como expertos en seguridad informática contamos con experiencia en ambas alternativas y el punto en cuestión es la robustez de la tecnología Blockchain, independientemente de ser privada o pública.

Conocimos la base teórica de esta arquitectura que a través de programas informáticos permite articular la revolución en la transparencia de las transacciones digitales, que se extiende transversalmente cualquiera sea el campo o disciplina sobre los que se implemente.

Hoy más que nunca, estamos preparados para liderar el cambio de paradigma en la formalidad contractual que tiene lugar en esta era, la era digital.

Referencias

1. Michael Mainelli, Alistair Milne; The impact and potential of blockchain on the securities transaction lifecycle May 2016.
2. Arvind Narayanan, Joseph Bonneau, Edward Felten, Andrew Miller, Steven Goldfeder; Bitcoin and Cryptocurrency Technologies. Oct 2015.
3. Melanie Swan.; Blockchain, blueprint for a new economy, O'Reilly Media, Inc., 2015.
4. Andreas M. Antonopoulos; Mastering Bitcoin. by O'Reilly Media, Inc., 2010.
5. UK Government Chief Scientific Adviser.; Distributed Ledger Technology: beyond block chain, Dec 2015.
6. IBM Institute for Business Value 'Device democracy: Saving the future of the Internet of Things' 2015. Available at http://www-01.ibm.com/common/ssi/cgi-bin/ssialias?subtype=XB&infotype=PM&appname=GBSE_GB_TI_USEN&htmlfid=GBE03620USEN&attachment=GBE03620USEN.PDF. Accedido 03/2018
7. Nakamoto S 'Bitcoin P2P e-cash paper' 2008. Available at http://satoshi.nakamotoinstitute.org/ emails/cryptography/1/. 03/2017
8. https://github.com/ethereum/. Accedido 02/2018
9. https://www.blintech.io/ Accedido 09/2018
10. https://www.bitcoinargentina.org/documentos-y-papersproyecto-impuesto-monedas-digitales/ Accedido 09/2018
11. https://miethereum.com/smart-contracts/solidity/ Accedido 04/2017
12. https://plugins.jetbrains.com/plugin/9475-intellij-solidity Accedido 05/2018

Deep Learning para la detección de ataques DDoS Smurf

Laura Fontanesi[1], Cecilia Viglianco[1], Eduardo Casanovas[1]

[1] Universidad de la Defensa Nacional – Facultad de Ingeniería CRUC – IUA
Av Fuerza Aérea 6500-Córdoba
lfontanesi7@gmail.com, ceciliaviglianco@gmail.com,
ecasanovas@iua.edu.ar
http://www.iua.edu.ar

Resumen. La principal característica de todas las Redes Neuronales Artificiales RNA, es su *capacidad de aprender*. Actualmente, el mundo de las redes neuronales está en auge mediante algoritmos de *Machine learning ML y Deep learning DL (Aprendizaje Profundo)*. Los sistemas de ML trabajan sobre grandes volúmenes de datos, identifican patrones de comportamiento y basándose en ellos, son capaces de predecir comportamientos futuros. Son muchos los sectores que están implementado Deep Learning siendo, el de la Seguridad Informática, el que está poniendo más énfasis en adaptar su tecnología a estas nuevas técnicas de Inteligencia Artificial, la que promete ser el futuro de la Ciberseguridad, disponiendo de máquinas que *aprenden* y *ajustan* sus algoritmos en tiempo real para que la cantidad de amenazas detectadas sea cada vez más exacta y con capacidad para predecir futuras. En el presente trabajo se verificará la aplicabilidad del Aprendizaje Profundo en la detección de Ataques DDOS (Distributed Denial of Service) Smurf.

1 Introducción

En general, las corporaciones, para proteger redes y nodos utilizan los conocidos IDS (Intrusion Detection System, IDS) o *Sistemas de Detección de Intrusos.* A pesar de los beneficios que los IDS proveen para mejorar la seguridad en sistemas, funcionan bajo la configuración de un conjunto de reglas estáticas que son determinadas por los administradores de los sistemas a priori. Preconfiguradas y mantenidas por los mismos, dejan total espacio a los ataques *zero day* y requieren de mayor uso de recursos si aumenta el nivel de sofisticación del ataque para el que se preparan.

Como consecuencia de lo anterior, los esfuerzos están dirigidos hacia la implementación de un IDS independiente de la intervención humana. Así, las técnicas ML, específicamente algoritmos Deep Learning, surgen como una solución prometedora: el desarrollo de Sistemas de Detección de Intrusos en una red **NIDS** (Network Intrusion Detection System) analizando el tráfico en tiempo real. La escasez de conjuntos de datos de calidad necesarios para llevar adelante lo expresado, conforma un problema en sí mismo y es la razón por la cual no se encuentra un producto operativo. Por el contrario, hay numerosos trabajos de investigación en análisis y ciencia de datos, como así también, implementando diferentes técnicas machine learning. El pre-

sente trabajo de investigación se centra en la hipótesis de la clasificación de ataques informáticos DDoS Smurf mediante la naturaleza de sus conexiones de red y atributos. Es nuestro objetivo esencial establecer un modelo óptimo para realizar la correcta clasificación de los mismos siendo congruente con la bibliografía consultada.

Este estudio comienza con el análisis del dataset KDD99-10 (variante reducida del KDD Cup 99). Luego del preprocesamiento de datos, se diseña e implementa una red neuronal del tipo Perceptrón Multicapa, clasificadora y con potencial para inferir en posibles perfiles de ataques.

Es de destacar que no se contempla realizar una comparación entre las diferentes variantes de preprocesamiento ni tampoco entre los distintos algoritmos que brinda el machine learning.

2 Desarrollo

2.1 Metodología

La investigación sigue una secuencia lógica y ordenada, comienzando con una exploración y examen del sistema en su totalidad, continúa con el preprocesamiento de los datos (data input) y concluye con el tratamiento del algoritmo DL como resultado y solución. Se enumeran las etapas de la misma:

-Análisis del sistema.
-Preparación de los datos.
-Diseño del modelo de red neuronal.
-Entrenamiento.
-Resultados. Evaluación y Prueba
-Implementación

2.2 Hardware y Software utilizados

Como para tener una idea respecto de los requerimientos de hardware necesarios para llevar adelante el trabajo, se detalla a continuación el equipo utilizado, una laptop con las siguientes características más relevantes:

-Procesador Core i5 4ta generación Ivy Bridge
-RAM 8gb
-Sistema operativo Fedora Linux

Asimismo, para el entrenamiento de la red, se trabajó en un servidor del laboratorio de Seguridad Informática de la Facultad de Ingeniería del CRUC-IUA de la UNDEF, el cual fue accedido mediante una conexión ssh, que consta de:

-Máquina virtual, procesador de 4 núcleos y 8gb de RAM,
-Sistema operativo Ubuntu Linux.

Referente al software, se listarán los frameworks y librerías a continuación:

-Preprocesamiento de datos: Weka 3.81,

-Backend: Tensorflow.
-Frameworks y librerías: Keras, Pandas, Scikit learn, Numpy
-Lenguaje: Python

2.3 Análisis del Sistema

Basándose en el estudio de investigaciones previas referidas a técnicas de ML y ataques informáticos se realiza una selección y adaptación de los componentes principales de nuestro sistema.

2.3.1 Selección del Dataset

En un primer momento, se había experimentado con la generación de un dataset propio montando una honeynet y capturando su tráfico. Posteriormente, se examinó la posibilidad de generar ese conjunto de datos a través de un script python. Se desistió de ambas por razones de tiempo, cantidad y calidad de los datos requerido.

Finalmente, mediante la exploración del dataset KDD Cup 99 y sus variantes NSL-KDD y KDD-10, se optó por hacer uso de este último por poseer más ejemplos de ataques que de conexiones normales.

Con un total de **42 atributos** y un **total de 494020 instancias**, su distribución es la siguiente:

Tabla 1. Distribución KDD99-10

Conjunto de Datos	DDos	Probe	U2R	R2L	Normal
KDD99-10	391458	4107	52	1126	97277

2.3.2 Selección del modelo de red neuronal

Como en Sabhnani & Serpen, el modelo a seguir fue una Red Neuronal Perceptrón Multicapa (MLP), la cual es una de las más utilizadas para clasificación. La mencionada es una red feed-forwad de tres capas: una de entrada, una oculta y una de salida.

2.4 Preparación de los datos

Del conjunto de datos KDD99-10 original, con un número total de 494020 instancias, se han eliminado todas aquellas diferentes a la que contiene normal y smurf como 'etiqueta' o label . Por tanto, el set de datos se redujo en un primer momento a:

-Número total de instancias: 378087
-Normal: 97277
-Smurf: 280790

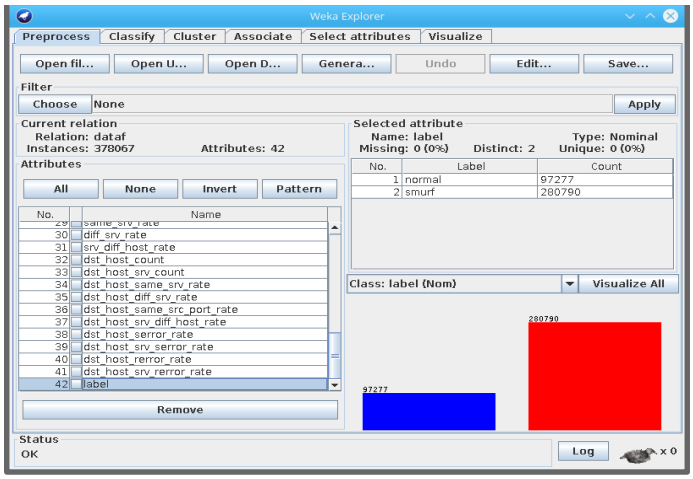


Fig. 1. Captura Weka Explorer dataset KDD99-10 primer filtrado

2.4.1 Selección de Atributos

Con el fin de obtener el subconjunto de datos más relevante para nuestro modelo, a partir de Weka *Explorer, Select Attributes*, y los siguientes métodos:

Sobre el conjunto de datos:

Cfs Evaluator

Consistency Subset Evaluator

Método de Búsqueda: BestFirst.

b) Sobre los atributos:

- Chi Square Attribute Evaluator:
- Information Gain Attribute Evaluator
- Método de búsqueda: Ranker

Cada uno de los métodos mencionados, ha establecido como resultado un "Ranking de atributos". Seleccionando aquéllos comunes se ha obtenido:

Atributos Seleccionados

Tabla 2. Nombre y descripción de atributos seleccionados

Nombre	Descripción	Tipo
protocol_type	Tipo de protocolo,ej: tcp, icmp	Nominal
service	Servicio de red destino	Nominal
src_bytes	Número de bytes de datos desde el	Nominal

	origen hasta el destino	
dst_bytes	Número de bytes de datos desde el destino hasta el origen	Numérico
count	Número de conexiones al mismo host en los últimos dos segundos	Numérico
srv_count	Número de conexiones al mismo servicio en los últimos dos segundos	Numérico
dst_host_same_src_port_rate	Porcentaje de conexiones mismo destino hacia el origen	Numérico
label	Etiqueta de clase	Nominal

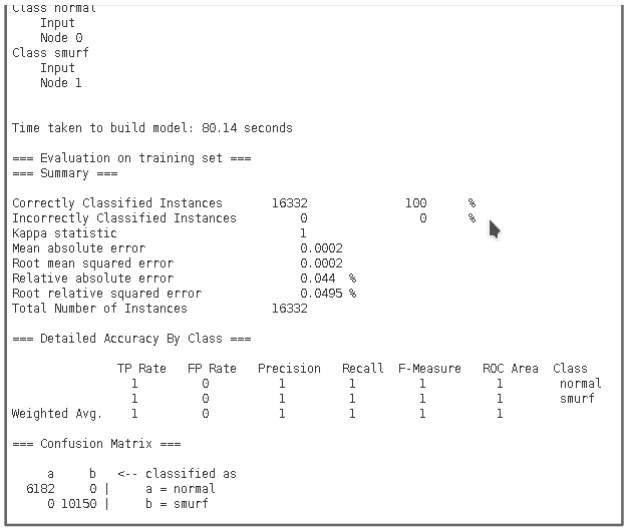
Solo por fines prácticos, se ha aplicado un nuevo filtro *(RemovePercentage*) para reducir nuestro dataset a un conjunto menor de instancias.

Por tanto, en el Dataset definitivo, podemos ver el resultado:

Total instancias: 16332

Total Atributos: 7 + 1 etiqueta

Con este set de datos definitivo, en Weka *Explorer Classify,* realizamos la clasificación de nuestro modelo. En la siguiente figura se puede ver el resultado de la aplicación del mencionado filtro



```
Class normal
    Input
    Node 0
Class smurf
    Input
    Node 1


Time taken to build model: 80.14 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances       16332               100      %
Incorrectly Classified Instances         0                 0      %
Kappa statistic                          1
Mean absolute error                      0.0002
Root mean squared error                  0.0002
Relative absolute error                  0.044  %
Root relative squared error              0.0495 %
Total Number of Instances            16332

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 1         0          1         1         1          1        normal
                 1         0          1         1         1          1        smurf
Weighted Avg.    1         0          1         1         1          1

=== Confusion Matrix ===

     a     b   <-- classified as
  6182     0 |     a = normal
     0 10150 |     b = smurf
```

Fig. 2. Captura Weka Classify dataset KDD99-10 segundo filtrado

2.4.2 Transformación de los datos

Por medio de la librería Scikit-learn de Python, particularmente con *sklearn.preprocessing.LabelEncoder,* se convirtieron los atributos categóricos a numéricos: *protocol_type, service y label.*

2.4.3 Normalización de los datos

También con el uso de Scikit-learn, *sklearn.preprocessing.MinMaxScaler,* se aplicó Escalado de Variables. Éste comprime los datos de entrada entre unos límites empíricos, máximo y mínimo de la variable, por lo que si existe un ruido en la entrada, el mismo será ampliado.

2.5 Diseño del modelo de la Red Neuronal

Nuestra red tipo Perceptrón Multicapa consta de tres capas: *una de entrada, una oculta y una de salida.* Se comenzó a diseñar el modelo sin la librería Keras. Luego optamos por ésta debido a su simplicidad y eficacia.

Los modelos en Keras son definidos como una secuencia de capas. Podemos crear un modelo secuencial (*Sequencial*) y agregar capas una a una hasta que cumplan nuestros requerimientos. Las capas completamente conectadas son definidas mediante la clase *Dense*. En esta clase se define el número de neuronas como primer argumento, como segundo argumento se define el método de inicialización, y la *función de activación*. Veamos ahora nuestra definición:

-Capa de entrada con 2 neuronas.
-Capa oculta con 16 neuronas, función de activación *relu*
-Capa de salida con 1 neurona y función de activación *sigmoid*

Tanto en las redes neuronales artificiales como biológicas, una neurona no sólo transmite la entrada que recibe, sino que también existe un paso adicional, una *función de activación*, que es análoga a la tasa de potencial de acción disparado en el cerebro. La función de activación utiliza la suma ponderada como entrada y la transforma una vez más como salida.

$$z = b + \sum wixi \qquad (1)$$

dónde:

xi es la i-ésima entrada,
wi, i-ésimo peso y
b, sesgo, umbral o bias.

La función ReLU (*Rectified Linear Unit)* o Rectificadora

Transforma los valores introducidos anulando los valores negativos (les asigna valor 0) y deja pasar sin modificaciones los positivos. Activation Sparse, sólo se activa si son positivos.

$$R(z) = max(0, z) \qquad (2)$$

La función Sigmoid

A la función sigmoid la podemos ver como una función aplastadora que transforma los valores introducidos a una escala (0,1), donde los valores altos tienen de manera asintótica a 1 y los valores muy bajos tienden de manera asintótica a 0.

$$R(z) = \frac{1}{1+e^{-z}} \qquad (3)$$

2.6 Entrenamiento

El aprendizaje supervisado es una técnica para deducir una función, que a partir de datos de entrada, les asigne la *etiqueta* de salida adecuada. Por tanto, actúa generalizando

Por la naturaleza de nuestro estudio, se ha seleccionado el tipo de aprendizaje superisado para la clasificación de ataques. Para el mismo, el método *model.compile()* y luego *model.fit()* provisto por *Keras.* Los parámetros establecidos fueron las siguentes:

Binary_crossentropy: se utiliza para evaluar el grado de error entre salidas calculadas y las salidas deseadas de los datos de entrenamiento. Como el nombre lo indica, utiliza la entropía cruzada para valores binarios en el cálculo del error.

Adam: es una combinación entre el algoritmo descenso del gradiente estocástico clásico y RMSprop. Adam es un método de tasa de aprendizaje adaptativo, lo que significa que calcula las tasas de aprendizaje individuales para diferentes parámetros.

Accuaracy: es una función que se utiliza para juzgar el rendimiento de su modelo.

Accuaracy calcula con qué frecuencia las predicciones coinciden con las etiquetas del set de datos.

Epoch: cantidad de ciclos completos por los que pasará el conjunto de datos de entrenamiento para lograr la convergencia.

Batch size define el número de muestras que tomará el algoritmo desde nuestros datos de entrenamiento.

2.7 Resultados. Evaluación y Prueba

Luego del entrenamiento, se evaluó el modelo y se muestra el resultado obtenido:

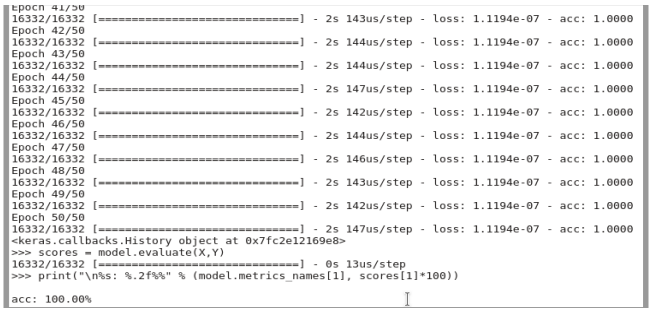


Fig. 2. Captura Weka Classify dataset KDD99-10 segundo filtrado

Una precisión del 100% (acc) lo que indica que la fase de entrenamiento ha sido exitosa. Esta es la demostración de que el algoritmo ha ajustado correctamente los pesos de la red en sus sucesivas iteraciones. En cuanto a la prueba, se utilizó un set de datos diferente incluído en el paquete KDD-10 mostrando que el porcentaje de acierto obtenido en el aprendizaje, efectivamente, fue correcto.

2.8 Implementación

Nos encontramos en esta etapa, analizando y experimentando para lograr un desarrollo o producto de fácil aplicación para demostración.

3 Conclusiones

El potencial de los algoritmos DL empleado en grandes volúmenes de datos, es indiscutible. Sin embargo, resulta fundamental para la eficiencia de estas soluciones, el diseño y construcción de su propio set de datos con sus convenientes 'etiquetas' y períodos de actualización, determinado el tipo de red neuronal a emplear así también como la adecuada parametrización de la misma. En esta dirección, surge la necesidad de integrar otra herramienta, junto a Weka, que incorpore clasificadores online para el

análisis en tiempo real como MOA (Massive Online Analysis) basada en flujos de datos continuos, que facilita la comparación con otros algoritmos conocidos y la generación de bases de datos sintéticas que simulan cambios de conceptos.

Por último y como principal conclusión hemos podido demostrar la aplicabilidad directa que tiene la utilización de estas técnicas en el marco de la ciberseguridad, dado que se ha podido verificar la deteccion de un ataque de DDOS (Distributed Denial of Service) Smurf.

Esta es la primer etapa en el armado de mecanismos predictivos para otros tipos de ataques los que serán caracterizados y parametrizados.

Referencias

1. Morate, D.: Manual de WEKA, Universitat Oberta de Catalunya, 2000.
2. García Gutiérrez, J.: Comenzando con Weka: Filtrado y selección de subconjuntos de atributos basada en su relevancia descriptiva para la clase, Universidad de Colorado, 2016.
3. Sabhnani, M., Serpent, G.: Application of Machine Learning Algorithms to KDD Intrusion Detection Dataset within Misuse Detection Context, in Proc the International Conf. on Machine Learning: Models, Technologies, and Applications, Las Vegas, vol. 1, pp. 209-215, 2003
4. Munivara Prasad, G., Dr. Rama Mohan Reddy, A., Dr Venugopal Rao, K.: DoS and DDOS Attacks, Defense, Detection and Traceback Mechanisms, Global Journal of Computer Science and Technology, E Network, Web & Security Volume 14 Issue 7 Version 1.0 , 2014
5. Rivero Pérez, J., Ribero, B., Ortiz, K.: Comparación de algoritmos para detección de intrusos en entornos estacionarios y de flujo de datos, Universidad y Sociedad pp 31-41. http://rus.ucf.edu.cu/, 2016
6. Rodriguez M.: Aplicación de Técnicas de Machine Learning a la Detección de Ataques, MISTIC, Rama, 2018

Comprensión del comportamiento del Ransomware a través del Análisis Forense

Santiago Moran Labat[1], Hernán Colmenarez[1], Eduardo Casanovas[1] y Carlos Ignacio Tapia[1]

[1] Universidad de la Defensa Nacional – Facultad de Ingeniería CRUC – IUA
Av Fuerza Aérea 6500-Córdoba
{santiagomoranlabat, hernan.colmenarez, carlosignaciotapia}@gmail.com, ecasanovas@iua.edu.ar,
http://www.iua.edu.ar

Resumen: Con la aparición de la era digital, los datos se han convertido en un bien muy valioso para todo tipo de usuarios, ya sean los dueños de los datos, empresas o entidades gubernamentales. La mayoría de estos usuarios o empresas no conocen o entienden el riesgo que esta información trae consigo en materia de seguridad informática. En este sentido, los ataques informáticos impactan la operación de cualquier empresa y los daños que genera, repercuten tanto en la productividad como en los costos legales, pérdidas de propiedad intelectual y daños a la reputación de la empresa. En muchas ocasiones el valor real de la información es mucho mayor a lo que las organizaciones habitualmente identifican. Debido a esto, durante los últimos años el termino Ransomware (un tipo de malware que prohíbe a usuarios acceder a los datos y pide algún tipo de pago a cambio de recuperar la información.) toma auge mientras afecta a miles de compañías y millones de usuarios a nivel mundial. Diferentes variantes de Ransomware aparecen cada año, pero todas con el mismo objetivo, atacar computadoras, bloquear el acceso a los archivos, solicitar un importe para desbloquear el equipo. Cuando un dispositivo es infectado con un Ransomware, es muy importante, para prevenir futuros ataques, utilizar técnicas, programas y herramientas forenses para la identificación de sus características y modos de infección con el propósito de definir estrategias de mitigación. En el presente trabajo se mostrarán las simulaciones de 3 variantes de ataques Ransomware, 2 en Windows, 1 en Linux, en ambientes aislados y seguros para posteriormente analizar su comportamiento y propagación; e indicar algunas conclusiones y controles preventivos según el tipo de Ransomware.

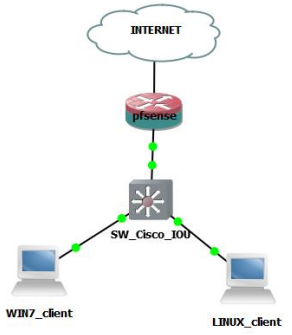
1. Introducción

En el presente trabajo se analiza diferentes muestras de Ransomware, una potencial amenaza para la información de las computadoras que utilizan Windows y Linux desde el año 2016 hasta actualmente (2019). El principal objetivo un Ransomware es infiltrarse en un sistema y bloquear el acceso a la información del sistema para pedir un

rescate monetario por alguna plataforma anónima como, por ejemplo: Bitcoin para su recuperación.

Para poder realizar este vector de ataque el atacante debe crear un programa capaz de cumplir con estas condiciones y además ser lo suficientemente robusto como para resistir todos los contrataques que pueda generar la victima (sistema y configuraciones del usuario) para defenderse desde el minuto 0 hasta la finalización del ataque, algo que asumimos que no es posible, por la entropía de situaciones, versiones de sistemas, etc. Entonces nuestra principal hipótesis es que el Ransomware va a realizar un flujo de comportamientos en cierto orden que en estos tres sistemas podemos de alguna forma interrumpir para detener el ataque y seguir teniendo acceso a los datos.

Para demostrar esto generamos un entorno seguro de prueba donde las muestras no se puedan propagar a otros sistemas externos, dentro de una máquina virtual de tipo Hyper-v con la siguiente topología de red:



Network	192.168.1.0/24
pfSense	192.168.1.1
WIN7_client	192.168.1.100
LINUX_client	192.168.1.101

Fig. 1. Topología

Comenzamos con el primer análisis, WannaCry en Windows:
Primero en la red WIN7_client (192.168.1.100) colocamos una máquina virtual de Hiper-V con Windows 7, a la cual le introducimos mediante un archivo cifrado con contraseña una muestra de WannaCry. Luego descomprimimos la muestra y comenzamos los análisis forenses.

Name	Date modified	Type	Size
WannaCry.EXE	5/15/2017 12:29 AM	Application	3,432 KB
WannaCry.rar	10/3/2018 3:27 PM	RAR File	3,403 KB

Fig. 2. Detalle de los virus utilizados

2. Análisis automatizados estáticos por comparación de hash para WannaCry.

Introduciremos nuestro archivo "WannaCry.EXE" cuyo hash es "ed01ebfbc9eb5bbea545af4d01bf5f1071661840480439c6e5babe8e080e41aa" en 3 páginas comerciales para ver qué resultados arrojan, en función de averiguar si la totalidad de soluciones de antivirus comerciales son capaces de detectar el malware. VirusTotal, 9 de 68 motores antivirus no lo detectaron como un malware, Metadefender, 6

de 35 motores antivirus no lo detectaron como un malware. Hibryd análisis lo detecta como un Ransomware, que utiliza protocolos de red en puertos inusuales, lee el GUID de la maquina criptográfica escribe datos en varios procesos remotos, lee el nombre del equipo, se comunica con 10 host, tiene capacidad para realizar consultas sobre información del kernel y los procesos, intenta eliminar fallas durante el inicio para ocultos cambios en el sistema, valores inusuales de entropía de 7.9998679751, posee habilidad para copiar y descargar archivos desde internet y para controlar servicios, posiblemente así instala los archivos necesarios para mantener persistencia y cifrar los datos, puede modificar las listas de control de acceso a los archivos, marca archivos para ser borrados, en su mayoría en C:\Users\%USERNAME%\. Está compilada con visual C++ 5.0 y trata de hacerse pasar por diskpart.

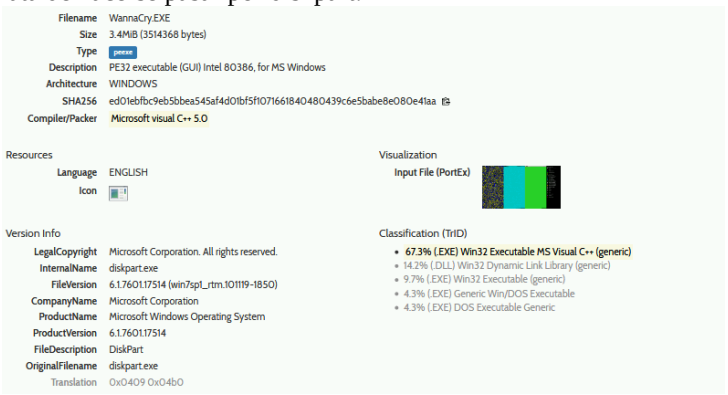


Fig. 3. Resultados del análisis a WannaCry.EXE

3. Resumen de los análisis automatizados para WannaCry

Los resultados de Metadefender, VirusTotal e Hybrid-Analysis nos confirman que actualmente WannaCry será detectado algunos de los actuales motores de antivirus, y además Hybrid nos arroja información extra sobre los comportamientos de red, directorios de interés, tipo de compilación, y afectación de procesos que genera WannaCry. Pero esta información es general, por lo que ahora utilizaremos herramientas más precisas para intentar entender el comportamiento más a nivel de código.

4. Análisis estáticos manuales para WannaCry

Con analizador de código hexadecimal (WinVi) confirmamos que la mayor parte del código está cifrado, pero algunas secciones de texto plano hacen mención a cifrado RSA, AES, código de creación de directorios, entradas de registro y soporte para diferentes versiones de Windows.

```
93 c1 00 0e  f1 a9 25 c8  f6 e8 8b c7  4d 69 63 72  |"Á° ñ©%Èöè‹ÇMicr|
6f 73 6f 66  74 20 45 6e  68 61 6e 63  65 64 20 52  |osoft Enhanced R|
53 41 20 61  6e 64 20 41  45 53 20 43  72 79 70 74  |SA and AES Crypt|
6f 67 72 61  70 68 69 63  20 50 72 6f  76 69 64 65  |ographic Provide|
72 00 00 00  43 72 79 70  74 47 65 6e  4b 65 79 00  |r°°°CryptGenKey°|
```

Fig. 4. Vista del estado del código

PEiD nos arroja que el compilador es Microsoft Visual C++6.0, el entry Point se encuentra en .text y la entropía es de 6.5, un valor alto. Por último, Ollydbg nos arroja los executable modules dentro del directorio C:\Windows\syswow64* y varias entradas para el registro de Windows principalmente en HKEY_LOCAL_MACHINE

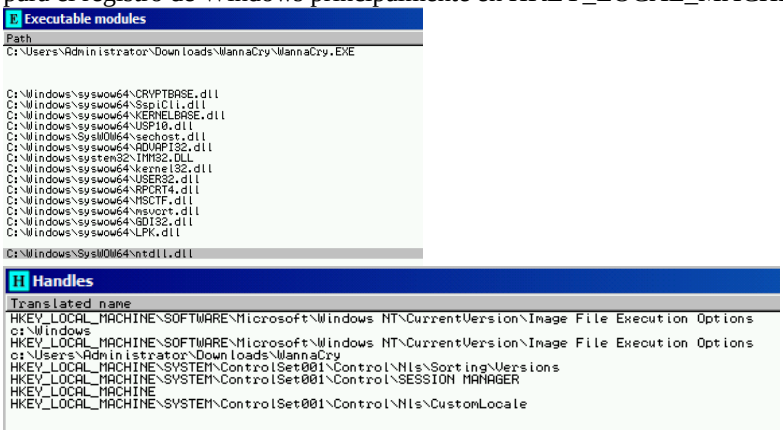


Fig. 5. Estado de los registros

5. Análisis Dinámico para WannaCry

Creamos un snapshot antes de iniciar la infección y otra luego de que el sistema está infectado para tener un parámetro a comparar.

Una vez que la víctima ejecuta el archivo WannaCry.EXE el procesador comienza a trabajar en niveles altos, por un tiempo indeterminado hasta que finalmente vuelve a un estado similar al que tenía previa infección y en la pantalla podremos apreciar que los archivos ya están cifrados con WannaCry que nos estará pidiendo que ingresemos un monto de dinero para pagar para no perder nuestros datos. Comparamos el registro de sistema con Regshot, antes y después del análisis, en total se cuentan 71 cambios en el registro entre los cuales vemos a HKLM\SOFTWARE\Microsoft\Windows NT\CurrentVersion\Schedule\TaskCache\ donde apreciamos que WannaCry atacó al antivirus de Microsoft eliminando tareas programadas.

Fig. 6. Ventana principal luego del ataque

Algo extraño ya que esta versión de Windows no tenía antivirus instalado. Esto nos da indicios sobre que el programador no puede prever la entropía de sistemas. Además de esto se modifica el Volume Shadow Copy, se cambia el fondo de pantalla y se modifica una entrada de registro de arranque de Windows HKLM\SOFTWARE\Wow6432Node\Microsoft\Windows\CurrentVersion\Run donde se invoca un archivo de nombre "buqvwehgtuzyug850" enlazado a otro de nombre taskche.exe con el que posiblemente agregue persistencia después de cada inicio de sistema. Autoruns nos confirma esta información sobre "buqvwehgtuzyug850".

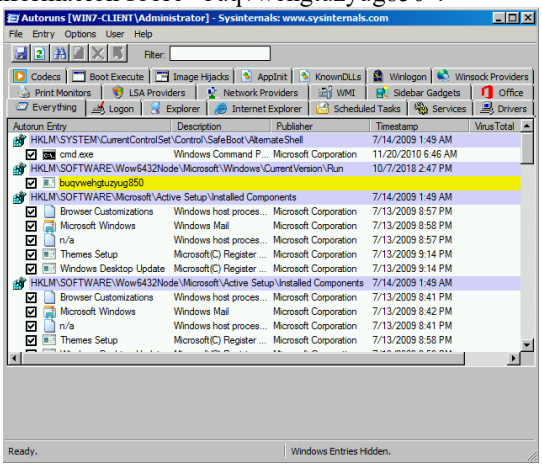


Fig. 7. Estado del Autoruns

Realizamos un análisis de red con Wireshark, para identificar si este WannaCry se intenta comunicar con alguien y se descubre que hay conexiones iniciadas por el proceso taskhsvc.exe y @WannaDecryptor@.exe, estos procesos son parte de los archivos instalados por la muestra. El tráfico descubierto es tráfico tcp por los puertos 443, 9050,

9001. Los tres son puertos usualmente utilizados por las redes TOR. También se observa como el tráfico por el puerto 9001 va cifrado con TLSv1.2

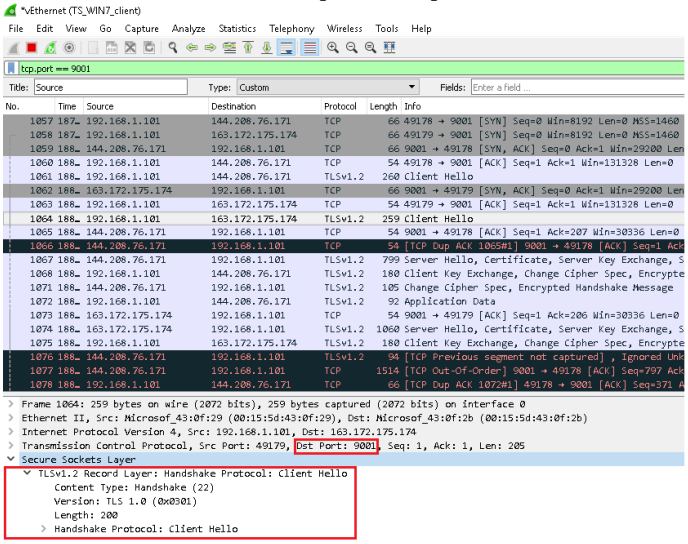


Fig. 8. Salida del wireshark

Por ultimo las direcciones ip dominos contactados por esta versión de WannaCry son Dominios: netimanmu.giannoug.gr ,maatuska.471.se ,tor01.zencurity.dk ,cotoaxi.tor.cool ,stoneghost.cridyn.com ,195-154-171-24.rev.poneytelecom.eu ,srv.hueskeedv.de ,tobi.gsgd.net ,parisdevspot11.webcannon.com ,Faraahar.redteam.net ,tor.dizum.com ,pakiplow.eu ,eos.fscore.de ,dvpn ,tor.noreply.org
IPs: 144.208.76.171 ,163.172.175.174 ,94.242.58.103 ,159.65.21.174 ,51.15.89.36 ,51.15.86.119 ,191.234.181.76 159.65.21.174

6. Contramedidas propuestas para WannaCry

- Se puede aprovechar la herramienta de Windows AppLocker para aplicar un bloqueo de aplicaciones por hash mediante una GPO. Por esta razón se recomienda el uso de versiones de Windows Enterprise o Ultimate.
- Bloquear, desviar o filtrar las IPs y dominios detectados durante este análisis.

7. Análisis automatizados estáticos por comparación de hash, de Obamas' blackMail

Esta muestra cuenta con el siguiente hash "0CD7440CA94D31212E21867439F38F0828823B76C94D566E81F5DFAF71574EBC", el cual VirusTotal y Metadefender detectan como amenaza 53/68 y 21/35. Hibryd Analisys detecta comportamientos muy similares a WannaCry sobre subprocesos, red, entropía, con la diferencia del uso de taskkill para los archivos kavasvc.exe, KVXP.kxp,

Rav.exe, Ravmon.exe, Mcshield.exe, VsTskMgr.exe, estos son ejecutables de antivirus comerciales, por lo que estamos viendo un intento de abarcar la entropía. Además, cuenta con cuentas de email dentro del código binario.

8. Análisis estático de Obamas' blackMail

El análisis con WinVi confirma la información relatada por los análisis automáticos, sobre taskkill y muestra un poco del código de la ventana principal.

OllyDog nos releva comportamientos similares a WannaCry sobre registros de sistema, y el mismo interés sobre las librerías del directorio C:\Windows\Syswow64\.

9. Análisis dinámco de Obamas' blackMail

En este análisis RegShot, nos revela que hubo una key eliminada, 23 keys nuevas, 5 valores eliminados, 66 valores agregados y 38 valores modificados. La mayoría corresponden a HKLM\SOFTWARE\Classes y a HKLM\SOFTWARE\Wow6432Node. Sumando en total 133 alteraciones. Autoruns nos aportó un solo cambio, y wireshark no detectó trafico de red. Un análisis con proces monitor nos permite comprender que el comportamiento fue el siguiente, modificar las llaves de registro de Windows, llamar al comando taskkill para detener posibles antivirus, luego intenta modificar 90 archivos *.exe, un valor significativo.

El comportamiento de este Ransomware es muy similar al de WannaCry, con la principal diferencia de que luego de bloquear el acceso a la información no intenta conectarse via red a algún servidor atacante, sino que proporciona números de teléfono chinos e emails.

10. Contramedidas propuestas

- AppLocker con la opción de hash, pero hay que personalizar un poco.
- Hosts, redirigir a localhost la direccion ftp.free3v.net
- Crear un archivo bs.ini en la raíz C:\ sin permisos para ningún usuario. Esto impide que el malware copie su configuración.

- Crear una carpeta ME dentro de C:\Windows sin permisos para ningún usuario. Aunque prácticamente con la creación del archivo bs.ini se evita la infección, todas estas modificaciones siguientes evitan que se copien otros archivos y entradas de registros basuras
- Lo tercero será crear algunas llaves en el registro de Windows, también sin ningún permiso porque son creadas por la muestra al ejecutarse. HKLM\SOFTWARE\Wow6432Node\Microsoft\Tracing\RASAPI32, HKLM\SOFTWARE\Wow6432Node\Microsoft\Tracing\RASMANCS, HKLM\SOFTWARE\Wow6432Node\360Safe, HKLM\SOFTWARE\Classes\.ec, HKLM\SOFTWARE\Classes\E.Document, HKLM\SOFTWARE\Classes\exe.

11. Análisis automatizado estático de EREBUS (Linux)

Iniciamos los análisis con el hash "0B7996BCA486575BE15E68DBA7CBD802B1E5F90436BA23F802DA66292C8A055F", donde lo detectan como positivo 41 por parte de VirusTotal y 24 por parte de Metadefender. Hybris Analisis, lo detecta como un malware que crea tareas en cron 96anacron, no hay tráfico de red, existen emails escritos dentro del binario ftp@example.com, intentos de anti virtualización, cambios de archivos y permisos,

12. Análisis estático de EREBUS

Trid y xdd nos revelan que es un archivo ELF, y los metadataos nos indican arquitectura x64. Con strings relevamos que intenta crear persistencia, se puede apreciar dos opciones, una por medio de un servicio bluetooth y la segunda por una tarea con cron que se ejecuta cada una hora.

```
main
/etc/rc.d/rc2.d/S25bluetooth
/etc/rc.d/rc3.d/S25bluetooth
/etc/rc.d/rc4.d/S25bluetooth
/etc/rc.d/rc5.d/S25bluetooth
/etc/rc.d/init.d/bluetooth
/etc/cron.hourly/96anacron
/etc/rc.d/init.d/
INSTALL_INIT_OK
../src/platform.nix.c
../init.d/bluetooth
INSTALL_INIT_ERR
/etc/cron.hourly/
INSTALL_CRON_OK
INSTALL_CRON_ERR
platform_pre_check
#!/bin/sh
# bluetooth:    Start/stop bluetooth services
# chkconfig:    2345 25 90
# description:  Bluetooth services for service discovery, authentication, \n#
terface Devices, etc.
bin="%s"
$bin
#!/bin/sh
# anacron's cron script
# This script updates anacron time stamps. It is called through run-parts
# either by anacron itself or by cron.
# The script is called "0anacron" to assure that it will be executed
# _before_ all other scripts.
# Don't run anacron if this script is called by anacron.
```

Se destaca que luego de cifrar deja dos archivos informando lo sucedido, también detectamos los siguientes dominions onion: 216.126.224.128, 7fv4vg4n26cxleel.onion.to, 7fv4vg4n26cxleel.onion.nu, 7fv4vg4n26cxleel.hiddenservice.net, 7fv4vg4n26cxleel.gbe0.top, qzjordhlw5mqhcn7.onion.to, qzjordhlw5mqhcn7.onion.nu, qzjordhlw5mqhcn7.hiddenservice.net, qzjordhlw5mqhcn7.gbe0.top, 7fv4vg4n26cxleel.onion, qzjordhlw5mqhcn7.onion. por

último identificamos los tipos de archivos y directorios de interés para EREBUS, principalmente /var/www/.

13. Análisis Dinámico de EREBUS

Obtenemos el mismo comportamiento que las muestras anteriores, un incremento del micro procesador por un tiempo, y luego cuando termina los archivos están cifrados. Se destaca que el demonio httpd PID3367 fue el que mantuvo ejecutándose como root. Con volatility y LiME determinamos que el proceso 3335, perteneciente a la muestra, tiene de hijo al 3367 quien se hace pasar como el demonio legitimo httpd de apache. Un linux_lsof nos revela que los proceso están ejecutándose dentro de /var/tmp/*

```
0xffff880035cb8000 muestra3     3335        0 /var/tmp/.651D8ED3E99B67B1A799D95BA1C36FA4.pid
0xffff880035cb8000 muestra3     3335        1 /var/tmp/.DCE774E95AC3F8ED11B79C067A18029E.pid
0xffff8800b6d98000 muestra3     3367        0 /var/tmp/.651D8ED3E99B67B1A799D95BA1C36FA4.pid
0xffff8800b6d98000 muestra3     3367        1 /var/tmp/.DCE774E95AC3F8ED11B79C067A18029E.pid
0xffff8800b6d98000 muestra3     3367        2 /var/tmp/.7453FBB38DCB8ED2F73735FA8C87B4BF.pid
0xffff8800b6d98000 muestra3     3367        3 /var/tmp/.2E2DC82D31210EFA4853C6E5540D3B15.res
0xffff8800b6d98000 muestra3     3367        4 socket:[20970]
0xffff8800b6d98000 muestra3     3367        5 []
0xffff8800b6d98000 muestra3     3367        6 /home/remnux/.mozilla/firefox/5p4qfubg.default
0xffff8800b6d98000 muestra3     3367        7 socket:[20968]
0xffff8800b6d98000 muestra3     3367        8 []
0xffff8800b6d98000 muestra3     3367        9 []
0xffff8800b6d98000 muestra3     3367       10 []
0xffff8800b6d98000 muestra3     3367       11 socket:[20976]
0xffff8800b6d98000 muestra3     3367       12 socket:[20965]
0xffff8800b6d98000 muestra3     3367       13 []
0xffff8800b6d98000 muestra3     3367       14 []
```

Wireshak revela que hubo contactos a las direcciones 216.126.224.[0,1,2,3,4,5,6,7]. Básicamente descargó las instrucciones de rescate, _DECRYPT_FILE.html y _DECRYPT_FILE.txt, también hace una copia de _DECRYPT_FILE.html y la renombre como index.html y como se puede observar también se elimina así mismo del directorio.

14. Contramedidas para EREBUS

- Se crea cada uno de los siguientes archivos con el comando touch dentro del directorio /var/tmp/
    - .651D8ED3E99B67B1A799D95BA1C36FA4.pid
    - .DCE774E95AC3F8ED11B79C067A18029E.pid
    - .2E2DC82D31210EFA4853C6E5540D3B15.res
    - .7453FBB38DCB8ED2F73735FA8C87B4BF.pid
    - .F216331543F45425AAA122BEA00B4CAF.conf
- Cambiar los atributos de los archivos con chttr +i

- Crear las siguientes entradas para iptables
    - iptables -I OUTPUT -d 216.126.224.0 -j DROP
    - iptables -I OUTPUT -d 216.126.224.1 -j DROP
    - iptables -I OUTPUT -d 216.126.224.2 -j DROP
    - iptables -I OUTPUT -d 216.126.224.3 -j DROP
    - iptables -I OUTPUT -d 216.126.224.4 -j DROP
    - iptables -I OUTPUT -d 216.126.224.5 -j DROP
    - iptables -I OUTPUT -d 216.126.224.6 -j DROP
    - iptables -I OUTPUT -d 216.126.224.7 -j DROP

6. Conclusión

Dados los datos que recolectamos, y aplicando un poco de experiencia del rubro, podemos concluir en que los atacantes no lograron vencer la entropía y existen una serie de contramedidas simples y bastante conocidas que nos permiten repeler los ataques, detalladas anteriormente.

Referencias

1. N. Milošević, Cornell Univ. Library, https://arxiv.org/ftp/arxiv/papers/1302/1302.5392.pdf, fecha de captura: 24/09/18
2. Wikipedia, OneHalf, https://en.wikipedia.org/wiki/OneHalf, fecha de captura: 25/09/18
3. Wikipedia, ILOVEYOU, https://en.wikipedia.org/wiki/ILOVEYOU, fecha de captura: 25/09/18
4. Wikipedia, Commwarrior (computer virus), https://en.wikipedia.org/wiki/Commwarrior_(computer_virus), fecha de captura: 25/09/2018
5. Wikipedia, Conficker, https://en.wikipedia.org/wiki/Conficker, fecha de captura: 25/09/2018
6. Juan A. Devincenzi, Técnicas y herramientas forenses para la detección de Botnets, Buenos Aires - Universidad de Buenos Aires, 2011, Introducción.
7. Mario Ávila, Detección de Malware Avanzado En Redes Organizacionales y Corporativas, Buenos Aires - Universidad de Buenos Aires, 2012, Cap. 4, (RESEÑA SOBRE STUXNET)
8. Eset, ACAD/Medre.A, https://www.welivesecurity.com/media_files/white-papers/ESET_ACAD_Medre_A_whitepaper.pdf, fecha de captura: 25/09/2018
9. Panda Security, CryptoLocker: Qué es y cómo evitarlo, https://www.pandasecurity.com/spain/mediacenter/malware/cryptolocker/, fecha de captura: 25/09/2018
10. Kaspersky, 10 síntomas de una infección maliciosa, https://www.kaspersky.es/blog/10-sintomas-de-una-infeccion-maliciosa/1348/, fecha de captura: 25/09/2018

Securización de una aplicación WEBRTC – autenticación y mitigación de ataques DoS, Password Cracking usando un SIP proxy

Andrés Felipe Macías Díaz[1], Fernando Boiero[1], Eduardo Casanovas[1]

[1] Universidad de la Defensa Nacional – Facultad de Ingeniería CRUC – IUA
Av Fuerza Aérea 6500-Córdoba
felipem1210@gmail.com, fboiero@gmail.com, ecasanovas@iua.edu.ar,
http://www.iua.edu.ar

Resumen. La VoIP (Voz sobre IP) es una tecnología en constante evolución debido al auge que tuvo y la gran cantidad de posibilidades que hay de desarrollar herramientas que usen los protocolos necesarios para llevar la data (voz, video, mensajes) de un extremo a otro.
Con el paso de los años, los extremos se fueron expandiendo y nuevas herramientas de software fueron naciendo, por ejemplo: softphones y aplicaciones que usan tecnologías web. Fue así como fueron apareciendo en paralelo protocolos de la capa de aplicación con los cuales pudieran entenderse dos aplicaciones hechas para cosas totalmente distintas. (ej, Un código javascript que emula un teléfono comunicándose con un SIP proxy). El ejemplo más claro de esto es WebRTC (Web Real Time Communications)[1], el cual es un estándar de comunicaciones implementado como un proyecto de software de la W3C y como un protocolo de comunicaciones establecido en el RFC 7478 que permite realizar comunicaciones en tiempo real sin plugins a través de una API implementada en cualquier tecnología web. El hecho de utilizar una tecnología web como Javascript para crear un teléfono implica una puerta abierta para ataques informáticos, a su vez, utilizar esta plataforma en la nube conlleva a considerar las amenazas a nivel del protocolo SIP.
En el presente trabajo se expondrá el desarrollo de las siguientes soluciones: diseñar un proceso de autenticación, con el cual se garantice la confidencialidad de los datos e implementar un mecanismo de defensa que garantice la disponibilidad del servidor de telefonía ante los ataques DoS.

1. Desarrollo

Vulnerar el servidor de telefonía es el objetivo de un atacante, pues así puede realizar llamadas a su antojo, sin pagar por ello y sin tener el permiso de hacerlo. Puede usar la conexión hacia la PSTN que tenga este servidor para revender esta conexión a un cliente que va a realizar llamadas a los destinos que quiera. También puede realizar ataques de denegación de servicio para inhabilitar el servidor. Ambos casos son críti-

cos para una empresa, pues afecta directamente el costo y disponibilidad de sus comunicaciones.

Blindar un servidor de telefonía con un SIP Proxy es añadir, en términos de seguridad, una capa mas de protección, ya que este componente puede encargarse del proceso de autenticación de los endpoints que quieren conectarse al servidor de telefonía[3] y realizar un filtro para las conexiones SIP. Teniendo esta premisa, en este trabajo se tomó como campo de acción una aplicación web que cumple la función de ser un webphone. Usa de backend el framework Django, codificado en python, y de frontend JavaScript, utilizando la librería JsSIP[2]; esta sería la aplicación basada en WebRTC, cumpliendo los parámetros que exige el RFC 7478. El servidor de telefonía está basado en Asterisk y como SIP Proxy se utiliza el software Kamailio usando su base de datos en PostgreSQL. La arquitectura de este sistema se detalla en la figura 1.

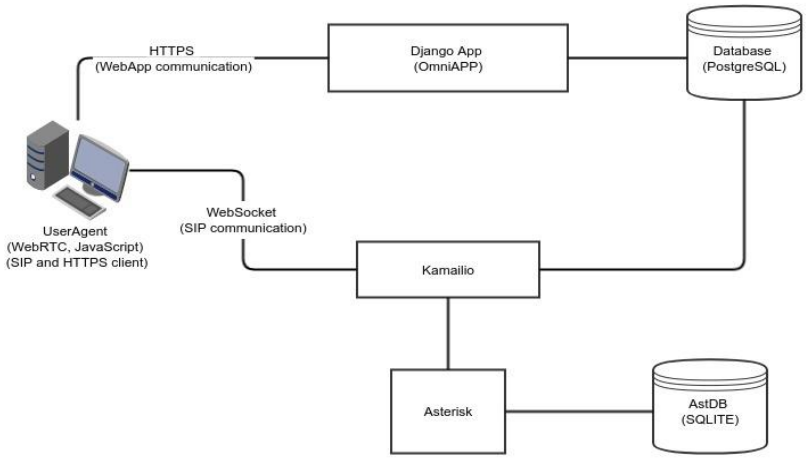


Fig. 1. Arquitectura del sistema usado como campo de acción

Kamailio es un software basado en módulos que pueden ser instalados, cargados y configurados a partir de un archivo de configuración que tiene un lenguaje de scripting nativo. A través de dicho archivo se programa el comportamiento del SIP proxy para lo que se requiere. A continuación se describen los módulos usado para el diseño del método de autenticación de usuarios SIP y mitigación de ataques DoS y password cracking:

- **auth_ephemeral[5]:** utiliza el concepto de credenciales efímeras para realizar autenticación. Dichas credenciales pueden ser generadas por un backend para luego ser enviadas en el mensaje SIP REGISTER[4]. El módulo especifica que las credenciales tienen que ser de esta forma:

username: combinación, separada por dos puntos, de un UNIX timestamp de expiración y el username SIP. Ejemplo: 15324352523:1002.
password: computada con el siguiente cálculo *base64(hmac-sha1(secret key, username)).* La secret_key debe ser única y compartida entre el generador de las credenciales y Kamailio.
ttl(opcional): tiempo de validez de las credenciales, es el cálculo entre el actual UNIX Timestamp - Timestamp en el username.

- **auth:** este módulo se encarga de enviar al endpoint la confirmación o rechazo de la autenticación solicitada.
- **pike[6]:** trackea el número de mensajes SIP realizados por una IP en un periodo de tiempo. Cuando ve que hay mucho tráfico entrante de dicha IP se loguea el evento y la IP será bloqueada por un tiempo.
- **htable[7]:** añade una tabla hash a la base de datos usada por kamailio. Así se guardan IP's que han sido bloqueadas.

Proceso de autenticación

Este proceso consta de dos grandes protagonistas: por un lado Django-python se encarga de generar las credenciales efímeras y escribirlas en un template HTML, para que el webphone las utilice con el fin de generar la conversación SIP con Kamailio. Por su lado Kamailio va a validar las credenciales recibidas y va a generar la respuesta al pedido de autenticación. En la figura 2 se grafica este proceso:

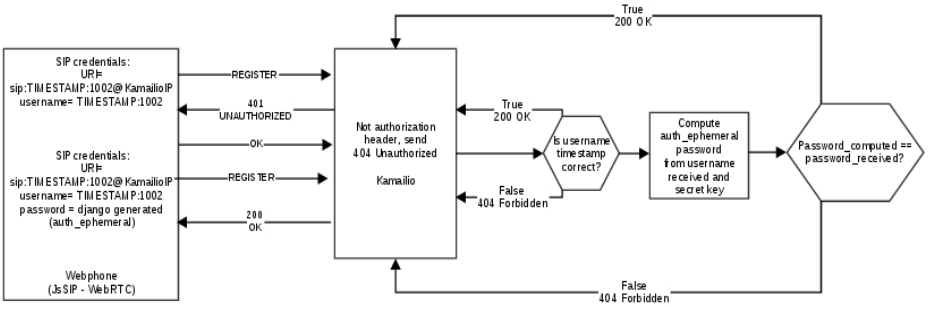


Fig. 2. Proceso de autenticación propuesto

En kamailio se configura el módulo auth_ephemeral escribiendo la secret_key que se va a usar para generar la contraseña. *modparam("auth_ephemeral", "secret", "34gggg$#$J").* Esta secret_key se puede consultar a kamailio usando un comando del módulo: *kamctl mi autheph.dump_secrets.*
Luego se añade el bloque de autenticación en la ruta AUTH.

```
route[AUTH] {
    if (is_method("REGISTER") || from_uri==myself) {
        # authenticate requests
        if (!autheph_check("$fd")) {
            auth_challenge("$fd", "0");
            $avp(uri) = $(fu{s.select,2,:});
            uac_replace_from("", "$avp(uri)");
            save("location","0x04","sip:$avp(uri)");
            exit;
    } }
```

A continuación se muestra el código de django que genera las credenciales efímeras: La primera función es la generar_usuario() que recibe como parámetro el usuario SIP y le concatena el timestamp generado a partir de la suma del timestamp de ejecución de la función mas 28800 segundos. Es decir por defecto se le da una validez de 8 horas a las credenciales

```
import time
 def generar_usuario(sip_extension):
   ttl = 28800
   date = time.time()
   timestamp = date + ttl
   user_ephemeral = str(self.timestamp).split('.')[0] + ":" +     str(sip_extension)
   return user_ephemeral
```

La segunda función es generar_contrasena() la cual va a generar la password haciendo el cálculo requerido por auth_ephemeral. Cabe anotar que la secret_key se obtiene llamando a un management command[8] creado para consultar a kamailio dicha key, utilizando el comando del módulo *autheph.dump_secrets*.

```
def generar_contrasena(self, sip_extension):
    out = StringIO()
    call_command('service_secretkey', 'consultar', stdout=out)
    secret_key = out.getvalue()[:-1]
    password_hashed = hmac.new(secret_key, sip_extension, sha1)
    password_ephemeral=password_hashed.digest().encode("base64").rstrip('\n')
    return password_ephemeral
```

Mitigación de ataques DoS y password cracking

Se configura el módulo pike añadiendo los parámetros necesarios. Con estos parámetros se van a tomar máximo 16 paquetes enviados por IP cada dos segundos y dicha IP se guardará en memoria durante 4 segundos:

```
# ----- pike params -----
modparam("pike", "sampling_time_unit", 2)
```

```
modparam("pike", "reqs_density_per_unit", 16)
modparam("pike", "remove_latency", 4)
```

Luego se configura los parámetros del modulo htable. Se define entonces la tabla ipban, la cual tendrá un tamaño de 8 columnas y tendrá un valor de autoexpiración de 600 segundos. Es decir, se ingresarán en la tabla un máximo de ocho IP's baneadas y la cantidad de tiempo que están baneadas será de 600 segundos

```
# ----- htable params -----
# ip ban htable with autoexpire after 5 minutes
modparam("htable", "htable", "ipban=>size=8;autoexpire=600;")
```

La lógica de baneo estaría asentada en la siguiente ruta llamada REQINIT, la cual revisa si la IP que manda el request está o no bloqueada. En caso de estarlo loguea el evento y envía un 403 Forbidden. En caso de no estar bloqueado pike revisa si se han enviado mas de los 16 paquetes permitidos en el límite de tiempo establecido. En caso de que si loguea que la IP se va a banear y guarda el bloqueo en la tabla ipban.

```
route[REQINIT] {
  if(src_ip!=myself) {
    if($sht(ipban=>$si)!=$null) {
     # ip is already blocked
     xlog("L_ALERT","request from blocked IP - $rm from $fu (IP:$si:$sp)\n");
     sl_send_reply("403", "Forbidden");
     exit;
     }
    if (!pike_check_req()) {
       xlog("L_ALERT","ALERT: pike blocking $rm from $fu (IP:$si:$sp)\n");
       $sht(ipban=>$si)=1;
       exit;
    } }}
```

2. Resultados

A continuación se detallará un cuadro donde se describe en resumidas cuentas el logro realizado con este trabajo, mostrando: ataque, tool de ataque, utilización de la tool, efecto sin protección, protección propuesta, efecto con protección.

Ataque	Tool de ataque
SIP impersonation: el atacante se registra a la central telefónica utilizando una cuenta SIP que no le pertenece, fingiendo ser un usuario auténtico del sistema.	No se usa una tool específica en este ataque. Si bien se puede obtener las credenciales SIP a través de fuerza bruta, en aplicaciones WebRTC hechas en JavaScript se puede obtener las credenciales SIP leyendo el HTML que renderiza las credenciales. Esto convierte al inspector del navegador en la tool de ataque.
SIP Flooding: esta ténica de ataque consiste en un ataque DoS con el fin de saturar un servidor SIP, a través de un envío masivo de peticiones OPTION o INVITE.	Se usó La herramienta SIPSAK https://github.com/nils-ohlmeier/sipsak
Utilización de Tool	**Efecto sin protección**
Se abre el inspector del navegador, se realiza el login de un agente en la plataforma Omnileads y se revisa en el código HTML las variables sipExt y sipSec. Verificar la sección Campo de Acción para conocer sobre la plataforma y la sección Pruebas realizadas para observar donde se obtienen las credenciales. Luego se ingresan los datos en un webphone de prueba ofrecido por la librería JsSIP https://tryit.jssip.net/ para simular una impersonación exitosa.	El atacante puede registrarse por tiempo ilimitado a la central telefónica conectada a la aplicación WebRTC sin ser un usuario auténtico del sistema. Esto le permite lograr el primer paso para poder utilizar la central para fines maliciosos.
Se ejecuta la herramienta con las siguientes opciones: *sipsak -F -vv -s sip:nobody@freetech.com.ar:5060* Donde: **-F:** activa el modo flood **-s SIPURI:** asigna el SIP URI al cual se le va a enviar el ataque **-v:** activa verbosidad	El SIP flooding hace que la central telefónica deje de responder a peticiones SIP reales, haciendo que se vuelva inoperativa
Protección propuesta	**Efecto con protección**
Cambiar el modo de generación de credenciales y autenticación de los usuarios SIP utilizando el módulo auth_ephemeral del SIP proxy Kamailio. Este módulo permite genrar credenciales efímeras cuyo tiempo de vida está estipulado en segundos.	El atacante, si bien aun podrá leer las credenciales SIP, no le van a servir para siempre, pues al ser efímeras tienen un periodo de validez. Luego de este periodo se vuelven inservibles.
Utilizar los módulos htable y pike del SIP Proxy Kamailio con el fin de realizar un baneo de IP cuando se detecte un envio masivo de mensajes de parte de una IP determinada	El atacante ya no podrá realizar un ataque simple de SIP Flooding, tiene que recurrir a técnicas mas avanzadas para poder hacer su ataque DoS.

Tabla 1. Resumen del trabajo realizado.

3. Conclusiones

El proceso de autenticación creado no permite que se mantenga la confidencialidad de las credenciales de autenticación SIP, sin embargo ofrece la capacidad de generar credenciales nuevas dejando las anteriores inválidas, por lo que disminuye drásticamente el riesgo de autenticidad de un usuario del sistema de telefonía.

El proceso implementado para evitar ataques de SIP Flooding aumenta drásticamente la disponibilidad del sistema telefónico pues evita ataques SIP simples realizando el bloqueo de IP.

Kamailio es un software ideal para ser utilizado como un SIP proxy donde se implemente la seguridad pertinente para aumentar la operatividad una central telefónica. A su vez, por ser software libre, su implementación es de bajo costo.

En la mayoría de los casos, las aplicaciones WebRTC no pueden asegurar un mínimo de confidencialidad en los datos que utilizan para funcionar, por ello es necesario remitirse a un SIP Proxy para evitar riesgos críticos que afecten la telefonía de una empresa.

Referencias

1. Manson, R.: Getting Started with WebRTC, Packt Publishing Ltd, p.7-9, 2013
2. Java Script SIP Library https://jssip.net/documentation/3.2.x/overview/ fecha de captura: 20/09/18
3. De Groef, W., Subramanian, D., Johns, M., Piessens, F., Desmet, L.: Ensuring endpoint authenticity in WebRTC peer-to-peer communication.
4. Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M.: RFC 3261. SIP: Session Initiation Protocol. AT&T. 2002
5. Dunkley, P., Kamailio auth_ephemeral module. Crocodile RCS Ltd.
6. Kamailio Pike module. Bogdan-Andrei Iancu. Voice Sistem SRL. 2003. http://kamailio.org/docs/modules/stable/modules/pike.html
7. Kamailio Htable module. Elena-Ramona Modroiu. Asipto. 2008-2011. http://kamailio.org/docs/modules/stable/modules/htable.html
8. The Web framework for perfectiponist with deadlines, https://docs.djangoproject.com/en/1.9/howto/custom-management-commands/ fecha de captura: 10/05/18

Cazadores de ciberamenazas como parte fundamental de un CyberSOC moderno

Santiago Nahuel Sarchetti[1], Carlos Ignacio Tapia[1]

[1] Universidad de la Defensa Nacional – Facultad de Ingeniería CRUC – IUA
Av. Fuerza Aérea Argentina 6500 – Córdoba
santiagosarchetti@gmail.com, carlosignaciotapia@gmail.com,
http://www.iua.edu.ar

Resumen. Actualmente la ciberseguridad se ha vuelto mucho más compleja que la mera instalación de firewalls, antimalwares o cualquier otra solución de Seguridad Informática disponible y efectuar el monitoreo de sus eventos. Es necesario, además, la incorporación de un equipo de Seguridad dedicado, específicamente para realizar la "caza" o "*hunting*" de amenazas en la red de la organización, definiéndose esta actividad como "el proceso proactivo e iterativo de búsqueda en la red en cuestión para detectar y aislar amenazas avanzadas que evadan las soluciones de seguridad existentes". Ésta se ha convertido en una actividad sumamente importante que permite a las organizaciones actuar antes que un ataque logre sus objetivos. En el presente trabajo se demostrará con dos ataques, uno vía PowerShell y otro a Active Directory, con el objetivo de recopilar información, evidencias y unir lo recabado para la creación de reglas que detecten más fácilmente ataques similares a futuro, permitiendo también compartir esta ciberinteligencia con otras organizaciones colaborativamente.

1 Introducción

Actualmente la ciberseguridad se ha vuelto mucho más compleja que la mera instalación de firewalls, antimalwares o cualquier otra solución de Seguridad Informática disponible y efectuar el monitoreo de sus eventos. Es necesario, además, la incorporación de un equipo de Seguridad dedicado, específicamente para realizar la "caza" o "hunting" de amenazas en la red de la organización, definiéndose esta actividad como "el proceso proactivo e iterativo de búsqueda en la red en cuestión para detectar y aislar amenazas avanzadas que evadan las soluciones de seguridad existentes". Ésta se ha convertido en una actividad sumamente importante que permite a las organizaciones actuar antes que un ataque logre sus objetivos.

Muchas organizaciones y empresas suponen que son seguras por el solo hecho de tener muchas de estas soluciones de Seguridad Informática implementada y con sus respectivas certificación ISO, pero la realidad nos demuestra que no es así. Independientemente de todo esto, las organizaciones y empresas siguen siendo víctimas de hackeos y filtraciones de información.

En muchas oportunidades estas herramientas y tecnologías, estando bien configuradas o no, tienen un enfoque de seguridad informática de tipo tradicional y muchas veces poco realistas para el tamaño o la envergadura de la empresa u organización, los responsable de IT en conjunto con los responsable de Seguridad Informática, crean y disponen de distintas redes para cada función (DMZ, Red de Finanzas, Administración, etc) y se comunican entre ellas a través de VLAN, Access List, Protocolos de Enrutamiento, etc, lo cual no significa que este incorrecta esta percepción, pero en este enfoque al llevarla a la práctica muchísimas veces no se aplica por muchos motivos, llevan consigo mucha configuración de muchos dispositivos distintos y de distintas tecnologías y/o marcas, errores humanos en la configuración, una infraestructura muy compleja, etc. Al final una estación de trabajo cualquiera termina logrando acceder a un servidor o servicio no autorizado sin ni siquiera notarlo. Una mala configuración de enrutamientos, o un SIEM que no reporte correctamente algunos eventos o la simple decisión de ignorar o eliminar alguna regla de correlación, nos puede llevar a resultados catastróficos para la empresa.

Los SIEMs hoy en día ahorran demasiado el trabajo de Seguridad Informática pero lo que sucede es que por más potente que sea nuestro SIEM implementado solo protegen lo que conocen o configuramos y no nos advierten o reportan de, por ejemplo, configuraciones erráticas, de cambios en las firmas digitales, de vulnerabilidades no conocidas y muchos otros.

La verdad de todo esto es que cuando el daño ya está hecho, uno como responsable piensa “esto lo pudimos haber evitado” pero la realidad es que cuando lo hacemos ya es demasiado tarde. Además, estas herramientas también tienen sus defectos y errores, como antivirus comunes que trabajan en base a firmas, vulnerabilidades de acceso en firewall en puertos de conocidos, en páginas web nuevas que no han sido categorizadas y no son filtradas, por robo de credenciales porque alguien lo dejo escrito en su escritorio o lo compartió con alguien más, y así muchos otros ejemplos.

Entonces ¿por qué no solo permitimos lo que conocemos y lo desconocemos lo bloqueamos? Porque en una empresa en constante expansión necesita de nuevos accesos, nuevos servicios todo el tiempo y si bloqueamos alguna transacción internacional o impedimos la entrada a algún sitio para realizar algún negocio los responsables de las pérdidas ocasionadas son los responsables de la Seguridad de la Información.

Otro problema muy frecuente, es que la mayoría de los ataques son por falta de actualización en sistemas operativos y servicios, y una persona con el objetivo de perjudicar a una empresa siempre buscan los activos más fáciles de vulnerar. ¿Y por qué estos activos no son correctamente actualizados? Porque muchos de los sistemas de las empresas, tienen sistemas de gestión a medida, que dejan de funcionar o no tienen soporte si actualizamos a otras distribuciones de sistemas operativos más seguras. Un ejemplo fue cuando Microsoft anuncio que ya no daría más soporte a su sistema operativo Microsoft Windows XP, rápidamente todos los bancos salieron actualizar todos sus Cajeros Automáticos porque las perdidas por alguna vulnerabilidad en algo tan cotidiano como un cajero ascenderían a miles de millones de dólares.

Entonces ¿qué podemos hacer ante todo esto? Primero aceptar el hecho que los hackeos van a suceder y van a seguir sucediendo en todo el mundo tarde o temprano. Y

en segundo lugar ser muy PROACTIVOS en la ciberseguridad y no esperar a que un evento menor se transforme en un incidente, y sobre esta base se hace Hunting donde las personas de este equipo se los llaman Hunters.

2 Cazadores de Amenazas o Hunters

Para poder desarrollar este tema primero hay que definir qué es Hunting, según SQRRL hunting es:

"[…] el proceso proactivo e iterativo de búsqueda en la red para detectar y aislar amenazas avanzadas que evadan las soluciones de seguridad existentes."

En este documento habla sobre la iteración que debe realizar el cazador como tarea fundamental para sus funciones (Figura 1).

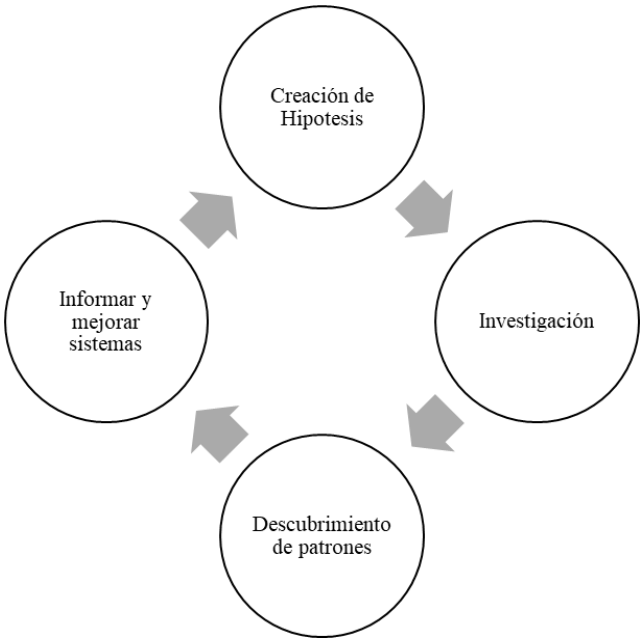


Fig. 1. Iteración de Cazador de Ciberamenazas (Fte: Sqrrl Data Inc., 2016 "A Framework for Cyber Threat Hunting" http://sqrrl.com/media/Framework-for-Threat-Hunting-Whitepaper.pdf)

Este bucle tiene el objetico buscar amenazas en su red a través de estos pasos:

1. Crear hipótesis: pensar que información importante maneja su empresa y que buscaría un Hacker, analizar los activos y procesos críticos que poseemos, como puede filtrarse esta información, y asumir siempre que el invasor ya paso todas las barreras de seguridad y se encuentra en la red interna de la empresa,

por eso es importante analizar siempre todas las anomalías, y no olvidar nunca de mantenerse bien informados de las últimas amenazas y vulnerabilidades.

2. Investigar las hipótesis que asumimos, recopilar información, registros, detectarlos y analizarlos si son eventos legítimos de la empresa o son causados por algún otro motivo extraño.
3. Descubrir nuevos patrones de comportamiento, corroborar las hipótesis declaradas y si existe o no una amenaza y en el caso de que existiese informar.
4. Y, por último, mejorar los sistemas de protección, reportar hallazgo recolectar estas evidencias y compartirlas como fuente de inteligencia para otras empresas que puedan beneficiarse de este conocimiento en alguna comunidad.

Según el autor Anton Chuvakin, las diferencias entre detección de amenazas y cazador de amenazas radica en su iteración (Figura 2).

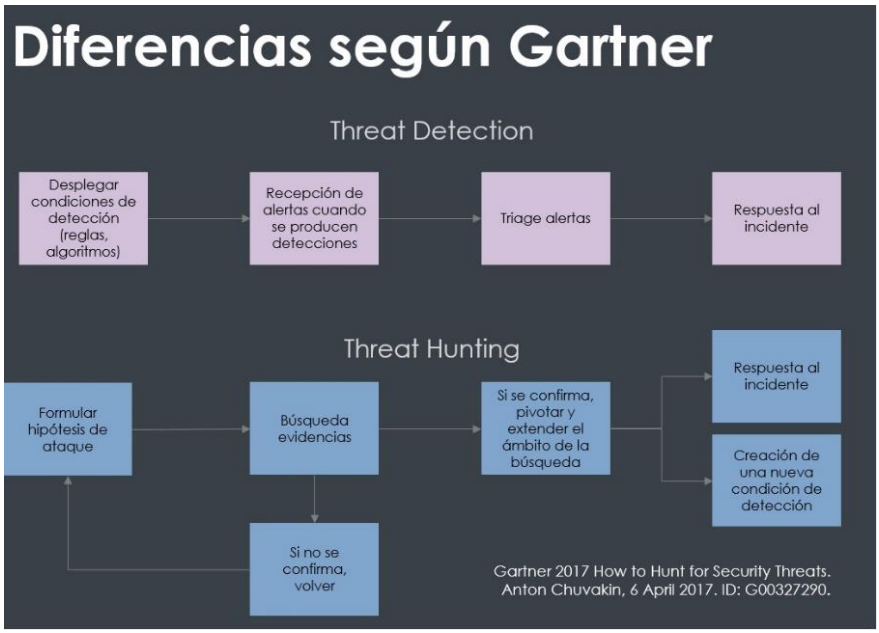


Fig. 2. Diferencia entre detección de amenazas y cazador de amenazas (Fte: *Gartner 2017 "How to Hunt for Security Threat" Anton Chuvakin, 6 de Abril 2017. https://www.pandasecurity.com/spain/mediacenter/adaptive-defense/threat-hunting-por-que-necesario/)*

Este ciclo es sumamente importante porque además de recopilar evidencias e información, gracias a los SIEMs actuales, podemos utilizar estos patrones e incorporarlos como reglas para futuras intrusiones que tengan el mismo patrón. Como Hunters tenemos que intentar descubrir comportamientos y recopilar la mayor cantidad de información, movimientos, horarios, que realizo primero y que después, en pocas pala-

bras realizar “Una Línea de Tiempo” porque mientras más precisa sea nuestras reglas mayores posibilidades de frenar futuras intrusiones o ex filtraciones tenemos.

Muchos intrusos prefieren u optan por métodos de larga duración, es decir, escanean o pruebas distintas herramientas y vulnerabilidades durante un tiempo largo para permanecer en la red sin ser detectados hasta que encuentra y logra sus objetivos. Dentro de las empresas puede pueden permanecer mucho tiempo y pueden hacer casi cualquier cosa como extraer información en poca cantidad entonces seria fácilmente camuflado dentro del flujo de datos normal de la empresa tratando pasar desapercibido. Estos ataques se llaman A.P.T. (Advanced Persistent Threat).

Ahora como Hunters para podemos enfocar nuestro esfuerzo en las distintas etapas de un ataque: en la intrusión inicial, durante los movimientos laterales del intruso, o cuando la información ya se está extrayendo. Si uno piensa en qué fase enfocar sus esfuerzos, en la etapa inicial de un ataque son muchísimos ataques diarios que una empresa recibe y la gran mayoría son bloqueados por las distintas soluciones informáticas (IDS, IPS, firewall, Antivirus, etc); en otra fase podría ser durante la filtración de información, pero sería ilógico intentarlo cuando la información ya fue fugada. Entonces la etapa más importante para realizar esta tarea sería durante los Movimientos Laterales del intruso.

Los Movimientos Laterales utilizan herramientas administrativas del sistema que son legítimos del sistema como por ejemplo “PowerShell” o “Active Directory” con el objetivo de escalar privilegios dentro de la red, realizar reconocimientos, y propagarse a otros sistemas, como cualquier Malware o agregando puntos de accesos adicionales.

Este extracto tiene como objetivo introducirnos en este concepto nuevo llamado Hunting o Threat Hunting sus alcances y objetivos pero el presente trabajo a desarrollar tiene como propósito final, siendo Hunters, 2 casos de estudios que dimos de ejemplo (ataques por PowerShell y Active Directory) y estudiar sus comportamientos, huellas, registros generados por los sistemas, como tomar todas las evidencias posibles para generar inteligencia y por último la generación de reglas según estos patrones para implementarlos en las distintas soluciones informáticas.

3 Integración de herramientas de investigación y automatización

A medida que la tecnología que utilizamos se expande, que las empresas proliferan cada vez más el BYOD (Bring your own device), de la expansión de internet, del trabajo en casa, cada vez es más difícil contener las amenazas, además si bien existen una amplia gama de soluciones de seguridad informática no todas se integran con otras herramientas o sistemas de forma sencilla, ni comparten información en formatos compatibles o estandarizadas y esto hace difícil la tarea de optimizar los recursos.

Por este motivo la automatización se está volviendo algo vital a la hora de plantear una estrategia de Seguridad Informática porque permite tiempo de respuestas ante incidente más rápidos, grandes ahorros de dinero y tiempo, y una mayor precisión en todas las operaciones de seguridad. Si omitimos la integración de todas las herramien-

tas, agregamos costos de esfuerzos humanos que deberán realizar las distintas tareas de forma manual.

En el presente trabajo final integrador he decidido por motivos meramente aplicativos utilizar todos sistemas y herramientas OPEN SOURCE en lo que sea herramientas de seguridad y muestras gratuitas de los sistemas operativos de Microsoft para realizar los ataques.

Para el presente trabajo he montado una infraestructura básica de una empresa con:

- Windows Server 2012 R2: con el servicio de Active Directory (dominio "tesis.local") y carpeta compartida. He aplicado, directivas y políticas de seguridad por defecto por motivos de tiempo en la implementación, y he actualizado a la última versión y aplicado todos los parches de seguridad hasta la fecha. Todos los registros generados por el SO son enviados al servidor de Graylog utilizando NXlog.
- Windows 10 PRO: en el dominio del Active Directory y conectado con la unidad compartida de la empresa donde se transfieren los archivos necesarios para sus funciones. Para enriquecer los eventos de seguridad, se instaló SYSMON. Todos los registros generados por el SO y SYSMON son enviados al servidor de Graylog utilizando NXlog.
- MISP v2.4.106: El cual tiene la capacidad de exportar reglas para distintos IDS y estudiar comportamientos extraños añadiendo distintos eventos e información.
- Graylog: Un sistema de gestión de eventos y registros de distintos sistemas operativos y aplicaciones que va a funcionar para tener toda la evidencia de los ataques que vamos a realizar. Además, va a estar integrado al Active Directory para el sistema de logueo.

En este trabajo al adoptar toda la infraestructura se denota la importancia de la integración de los distintos servicios, como cada uno interactuar con el otro y como con la falla de uno produce algún error en los otros. Por ese motivo va a ser importante tener bien documentado todos los registros y eventos para luego poder trabajar en la correcta correlación de los mismos y generación de reglas.

El Graylog va a tener la capacidad de poder recolectar información de las distintas fuentes y correlaciónalos, además se va a poder adoptar una dashboard con distintas informaciones de distintas fuentes para que como en cual empresa se pueda ver la información más importante de las distintas fuentes en varias pantallas.

4 Ejecución de pruebas y recolección de evidencias

Con la finalidad de poner en práctica el proceso de Threat Hunting, se seleccionó como casuística aquellas relacionadas con los ataques file-less vía Powershell, que

pueden ser ejecutados mediante Mimikatz para poder tomar control de un Dominio Windows. Esto implica que los anti-malwares tradicionales tendrán dificultades para detectar la amenaza basándose en firmas de archivos en file system ya que al trabajar con procesos ejecutados en memoria directamente no tendrá firma digital.

Para ejecutar las pruebas se siguieron los pasos que se muestran a continuación:

Se tomó el equipo Windows 10 (miembro del dominio "tesis.local") y mediante Poweshell, se realizó el siguiente Invoke-Expression y se procedió a ejecutar:

"IEX (New-Object Net.WebClient).DownloadString('https://raw.githubusercontent.com/PowerShellMafia/PowerSploit/master/Exfiltration/Invoke-Mimikatz.ps1'); Invoke-Mimikatz –DumpCreds"

Con esta parametrización, se volcaron las credenciales locales de Windows manejadas por el proceso LSASS.EXE. Igualmente, se pudo haber utilizado cualquier otro parámetro de Mimikatz distinto a "-DumpCreds" como por ejemplo "*lsadump::tickets*" o "*sekurlsa::logonpasswords*".

Es importante recalcar que la invocación Powershell o Comand funciona sólo con privilegios de Administrador o de SystenNT por esa razón es importante recalcar la importancia de la escalada de privilegios que se tiene durante un ataque, durante el presente trabajo utilizamos Akagi en su versión de x64 que proporciona multiples formas de escalada de privilegios de forma muy sencilla.

Monitoreando en el SIEM Graylog, se observó la siguiente secuencia de eventos, que define en conjunto los indicadores de compromiso que habilita a investigar en mayor profundidad, con un grado de certeza aceptable según la figura 3 y la figura 4.

Fig. 3. Se muestra los logs enviados por el IDS Snort.

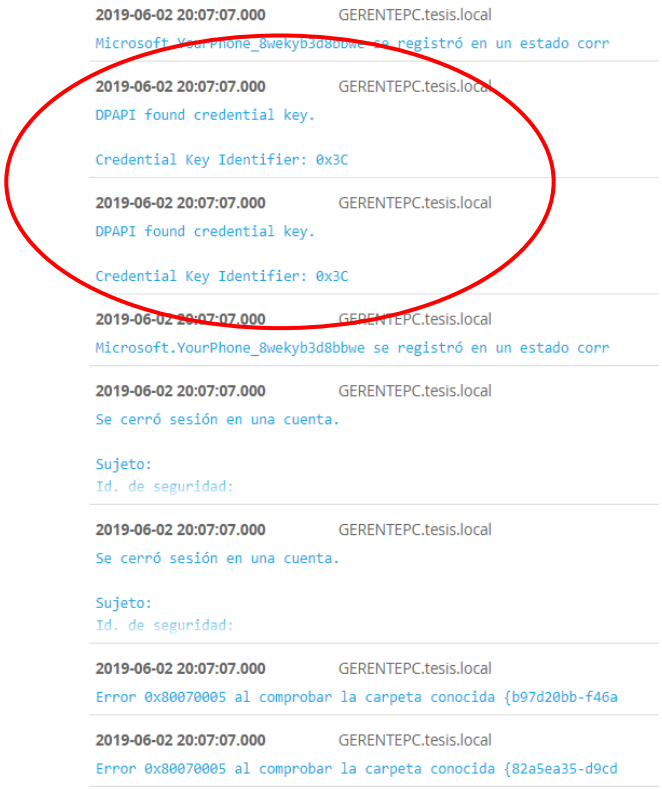


Fig. 4. Se muestra los logs enviado por la computadora atacada

Toda esta secuencia de eventos transcurre a lo sumo en 6 minutos, con lo que ese también será un parámetro para la configuración de alarmas. Es válido aclarar que es humanamente imposible dar seguimiento manual a esta cantidad de eventos multiplicados por el número de endpoints que se estén monitoreando, por lo cual es imperativo contar con un SIEM que permita almacenar los logs, correlacionar eventos y alarmar cuando se combinen las condiciones deseadas. Es importante destacar que el SIEM Graylog que utilizamos es un software Open Source y permite generar alarmas según las condiciones que nosotros configuremos y se encuentra disponible toda su documentación en su Página Oficial.

En otra variante de Mimikatz del software PowerShell Empire, la secuencia de detección es válida excepto por el uso de la máscara de acceso “0x1410” en lugar de la “0x143A”, y considerando que las detecciones pueden tener alta cantidad de falsos positivos, se debe combinar con el parámetro de nombre de proceso para que contenga el string “shell” (de forma que filtre por aquellos ejecutables con la cadena “shell” en su nombre).

Estas casuísticas en particular son de gran alcance en prácticamente cualquier organización de cualquier tamaño y de cualquier rubro de industria gracias a la amplia adopción del servicio de directorio de Microsoft.

Desde luego, en todos estos casos se habla de ataques o herramientas conocidas que tienen suficiente estudio en la comunidad para poder ser detectadas y monitoreadas. Lo importante de esta demostración de funcionamiento de Mimikatz con sus variantes es que el profesional de Threat Hunting debe mantenerse permanentemente en estado de alerta, conociendo los últimos ataques a las plataformas e infraestructuras para poder testearlas de primera mano, desde luego, tomando los recaudos pertinentes para que sus pruebas no conduzcan a afectación de servicio de la organización.

5 Conclusiones

Independientemente de las casuísticas que analice el responsable de Threat Hunting en una organización, el proceso en sí de mantener una mirada que imite la de un atacante real, será la que agregue valor y ayude en gran medida a estar con los controles de seguridad en correcto funcionamiento para prevenir amenazas y solucionar problemas o errores que a futuro causen una mayor repercusión para la empresa.

Sumado a la necesidad de dedicación horaria y procedimental del Threat Hunting, se destaca como esencial el soporte de un SIEM suficientemente flexible como para conectar e integrar diversas fuentes de información, configurar reglas de correlación que agreguen valor y permita alarmar dando inicio al proceso de gestión de incidentes.

Además, sumar a la necesidad de tener un CyberSOC lo suficientemente capacitado para poder determinar y detectar ataques o malfuncionamiento en las redes y no lamentar perdidas de ningún tipo para las empresas, y contar con laboratorios de forencia y pruebas para trabajar in situ en simulaciones y puesta en funcionamiento de distintas soluciones.

Referencias

1. Sqrrl Data Inc., 2016, “A Framework for Cyber Threat Hunting” http://sqrrl.com/media/Framework-for-Threat-Hunting-Whitepaper.pdf
2. Gartner 2017 ”How to Hunt for Security Threat” Anton Chuvakin, 6 de Abril 2017. https://www.pandasecurity.com/spain/mediacenter/adaptive-defense/threat-hunting-por-que-necesario/
3. Michael Mainelli, Alistair Milne; The impact and potential of blockchain on the securities transaction lifecycle May 2016.
4. Arvind Narayanan, Joseph Bonneau, Edward Felten, Andrew Miller, Steven Goldfeder; Bitcoin and Cryptocurrency Technologies. Oct 2015.
5. Melanie Swan.; Blockchain, blueprint for a new economy, O’Reilly Media, Inc., 2015.
6. Andreas M. Antonopoulos; Mastering Bitcoin. by O’Reilly Media, Inc., 2010.

Prototipo de aplicación para extracción de información de dispositivos móviles Android para uso forense

Gómez Palacios Pedro Nicolás[1], Colazo Danilo José[1], Mag. Ing. Solinas Miguel Angel[1]

[1]Laboratorio de Redes y Ciberseguridad (LARYC) FCEFyN
Universidad Nacional de Córdoba, Velez Sarsfield 1611 5000 Córdoba
{nicogp2, daniloc57}@gmail.com,
{miguel.solinas@unc.edu.ar}

Resumen. En este trabajo presentamos la investigación desarrollada en el Laboratorio de Redes y Ciberseguridad (LARYC) de la Facultad de Ciencias Exactas Físicas y Naturales de la Universidad Nacional de Córdoba cuyo objetivo es construir un prototipo de aplicación, con fines forenses, que permita la extracción de información de dispositivos móviles con sistema operativo Android.
Atacamos el problema de extracción de datos de un dispositivo móvil sin contar con que este se encuentre rooteado. Para lograr nuestro objetivo utilizamos librerías y protocolos específicos para ciertas marcas de celulares. Esto nos da la posibilidad de realizar con éxito la extracción.

Luego diseñamos y construimos una aplicación que incorpora como requerimiento funcional el uso de la librería y el protocolo para facilitar y automatizar la tarea. Si bien el trabajo es un prototipo, se continúa con su desarrollo para ampliar su uso a otras marcas de celulares y versiones de Android.

1 Introducción

La extracción de datos que forman parte de la evidencia y el examen forense de dispositivos móviles puede diferir de un caso a otro. Sin embargo, seguir un proceso de examen consistente ayuda al perito forense a garantizar que la evidencia extraída de cada teléfono esté correctamente documentada, sea segura y pueda ser defendida. Debido a la gran diversidad de *smartphones* no existe un proceso estándar establecido para el análisis forense de móviles. No obstante, en la Figura 1 se proporciona una vista general del proceso para la extracción de evidencia de smartphones [1] [2]. Este trabajo describe una propuesta para abordar la etapa de "procesamiento" con una aplicación propia. Se trata de un prototipo, del cual se ha diseñado su arquitectura y se se validado con una marca y algunos modelos. Este trabajo se organiza del siguiente modo: en 1 describimos los pasos del proceso de extracción de evidencia de un

dispositivo móvil; en 2 describimos la arquitectura de la aplicación; en 3 los primeros resultados obtenidos y en 4 algunas conclusiones.
Todos los métodos utilizados para la extracción de datos se deben probar, validar y documentar:

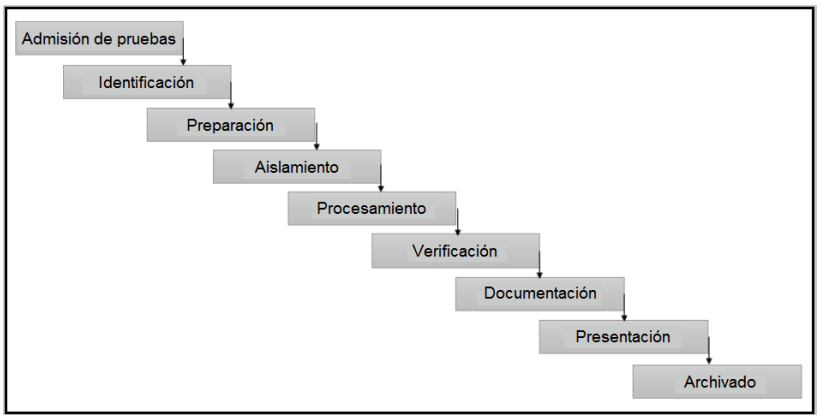


Figura 1. Proceso de extracción de evidencia en teléfonos móviles

El proceso se inicia con la fase de admisión de pruebas. Ella implica la confección de formularios de solicitud y documentos para registrar la información personal y el tipo de incidente en el que estuvo involucrado el dispositivo móvil. Además, se describe el tipo de dato que el solicitante está buscando que está en consonancia con los objetivos del perito forense.

La segunda fase es la de **identificación**. Aquí el perito debe identificar los siguientes detalles para cada examen de un dispositivo móvil: autoridad legal, objetivos del examen, marca, modelo e información de identificación del dispositivo, almacenamiento de datos extraíbles y externos, entre otras fuentes de evidencia potencial.

Por su parte, la fase de **preparación** implica una investigación sobre el *smartphone*. En particular, se analiza qué se examinará y los métodos y herramientas apropiados que se utilizarán para la adquisición de datos y la realización de dicho examen. Esto se hace generalmente según el modelo de dispositivo, el sistema operativo subyacente y su versión.

La fase de **aislamiento** es la cuarta etapa de este proceso. Dado que la mayoría de los teléfonos móviles intercambian datos de forma permanente gracias a su conectividad, en esta fase se corta todo medio de conectividad para evitar que se altere la evidencia. El aislamiento de la red se puede hacer colocando el teléfono en un paño de protección de radiofrecuencia y luego poniendo el teléfono en modo avión.

Una vez que esto se logró, comienza el **procesamiento** real del teléfono móvil. Los datos del teléfono deben ser extraídos utilizando un método probado (funciona con éxito), replicable (se puede realizar nuevamente) y legal (es acorde a la ley). La adquisición física es el método preferido ya que extrae los datos de memoria sin

procesar y el dispositivo generalmente se apaga durante el proceso de adquisición. Es en esta instancia donde se ubica la aplicación descrita en este trabajo.

Después de procesar el teléfono, sigue la fase de **verificación**. El perito debe constatar la precisión de los datos extraídos para asegurarse de que no se hayan modificado. Esto se puede lograr comparando los datos extraídos con los datos del teléfono o utilizando valores hash.

En la fase de **documentación** el perito forense debe registrar, durante todo el proceso de examen, en forma de notas actualizadas y relacionadas lo que se hizo durante la adquisición y el examen. Una vez que el perito completa la investigación, los resultados deben pasar por algún tipo de revisión por pares para asegurarse de que los datos se verifiquen y la investigación se complete.

Por otra parte, la fase de **presentación** permite que la información extraída y documentada desde un dispositivo móvil pueda presentarse claramente a cualquier otro perito o a un tribunal.

Finalmente, la fase de **archivado** permite preservar los datos extraídos del teléfono móvil. Es importante que los datos se conserven en un formato utilizable para el proceso judicial en curso, para futuras referencias (en caso de que el archivo de evidencia actual se corrompa) y para los requisitos de mantenimiento de registros.

1.1 Extración manual, lógica y física

Dado que los datos que residen en un dispositivo Android pueden ser parte de investigaciones civiles, penales o internas realizadas por parte de una empresa corporativa, al tratar con investigaciones que los involucran, el perito debe ser consciente de los problemas que deben ser atendidos durante el proceso forense. Esto incluye determinar si se permite el acceso con privilegios de root (a través del consentimiento o la autoridad legal) y qué datos se pueden extraer y analizar durante la investigación. Por ejemplo, en un caso que involucra acoso, el tribunal solo puede permitir extracción y análisis de mensajería SMS, registros de llamadas y fotos. En ese caso, tiene mas sentido capturar lógicamente solo esos elementos específicos. Sin embargo, se también se puede obtener una extracción física completa del dispositivo y luego examinar las áreas admisibles por el tribunal [1].

Las técnicas de extracción de datos en un dispositivo Android se pueden clasificar en tres tipos:

- Extracción manual de datos
- Extracción lógica de datos
- Extracción física de datos

La extracción manual datos implica navegar por el dispositivo de la forma en que lo haría un usuario y capturar cualquier información valiosa. La extracción lógica se refiere al acceso al sistema de archivos y la extracción física consiste en extraer una imagen bit a bit de todo el almacenamiento del dispositivo móvil. Los métodos de extracción utilizados para cada uno de estos tipos se describen en detalle en la siguiente sección.

Algunos métodos pueden requerir el acceso con privilegios de root sobre el dispositivo para acceder a la totalidad de los datos. Cada método tiene implicaciones diferentes y sus tasas de éxito van a depender de la herramienta, el método utilizado, la marca y modelo del dispositivo.

1.2 Extracción manual de datos

Este método de extracción directa, implica que el perito utilice la interfaz de usuario (IU) del dispositivo móvil para acceder al contenido que se encuentra en la memoria. Podrá ver contenidos tales como registros de llamadas, mensajes de texto y chats de Mensajería Instantánea (MI), por ejemplo Whatsapp. Luego el contenido de cada pantalla se captura tomando fotos con herramientas adaptadas para tal fin y se puede presentar como evidencia. Es recomendable documentar detalladamente el proceso utilizado. El principal inconveniente de este tipo de examen es que solo se pueden investigar los archivos a los que se accede a través del sistema operativo y su IU. Por otro lado es siempre limitado dada la cantidad de información que puede contener un celular.

Se debe tener cuidado al examinar manualmente el dispositivo ya que es fácil presionar el botón incorrecto y borrar o agregar datos. La extracción manual se debe usar como último recurso para verificar los hallazgos extraídos utilizando uno de los otros métodos. Ciertas circunstancias pueden justificar que el perito realice un examen manual como primer paso.

1.3 Extracción lógica de datos

Las técnicas de extracción lógica permiten recuperar los datos presentes en el dispositivo accediendo al sistema de archivos e interactuando con el sistema operativo. Estas técnicas son importantes porque proporcionan datos valiosos, funcionan en la mayoría de los dispositivos y son fáciles de usar.

Una vez más, el concepto de "*rooting*" entra en escena al extraer los datos. Las técnicas lógicas, en realidad, no requieren acceso con privilegios de *root* para la extracción de datos. Sin embargo, tener permiso de administrador en un dispositivo permite acceder a todos sus archivos. Por lo tanto, en móviles no rooteados sólo podrán extraerse algunos datos, mientras que en móviles con acceso *root* se podrá acceder a todos sus archivos presentes. En consecuencia, tener este tipo de acceso influiría en la cantidad y calidad de datos que se podrán extraer mediante técnicas de extracción lógica.

1.4 Extracción física de datos

La extracción física se refiere al proceso de obtención de una imagen exacta, bit a bit, del dispositivo [3] [4]. Es importante comprender que una imagen bit a bit no es lo mismo que copiar y pegar, donde sólo se copian los archivos disponibles, como los visibles, ocultos y aquellos relacionados con el sistema. Copiar y pegar se considera como una imagen lógica, en el cual los archivos eliminados y aquellos que no son accesibles no son copiados por el comando *copy*.

La extracción física es una copia exacta de la memoria del dispositivo e incluye información adicional, como espacio libre, espacio no asignado, totalidad de particiones, etc.

2 Desarrollo de la aplicación de extracción de datos

Se describe a continuación el diseño e implementación de una aplicación cuya finalidad es la extracción física de datos en dispositivos Android, diseñada y desarrollada en el contexto de un Proyecto Integrador de Ingeniería en Computación.

Para validar la funcionalidad de la aplicación se tuvo acceso a dispositivos de la marca LG. Por ello se investigó y utilizó LG Advanced Flash (LAF) [5]: un protocolo implementado en lenguaje Python para comunicación con dispositivos LG en "*Download Mode*" que permite la ejecución de "*shell comands"* arbitrarios con privilegios de root.

El patrón de arquitectura de software elegido para la contrucción de la aplicación es Model View Controller (MVC) [6]. Permite seprar los datos, la lógica de una aplicación y el módulo encargado de gestionar tanto eventos como comunicaciones. Para ello *MVC* propone la construcción de tres componentes distintos los cuales se denominan modelo, vista y controlador. Por un lado define componentes para la representación de información y por otro, la interacción del usuario.

A continuación, en la Figura 2, se presenta el Diagrama de Casos de Uso. Se puede observar dos escenarios. El primero es "Conectarse". Allí el técnico debe visualizar las instrucciones que debe ejecutar con el dispositivo para lograr la conexión; podrá visualizar el estado de la conexión y verificar que el procedimiento fue exitoso. El segundo es "Extracción mediante app" que extiende en el caso de uso "Conectarse", lo que implica que primero se debe realizar la conexión, luego seleccionar el modo de extracción y, finalmente, corroborar que el proceso finalizó correctamente mediante la visualización de un reporte de extracción.

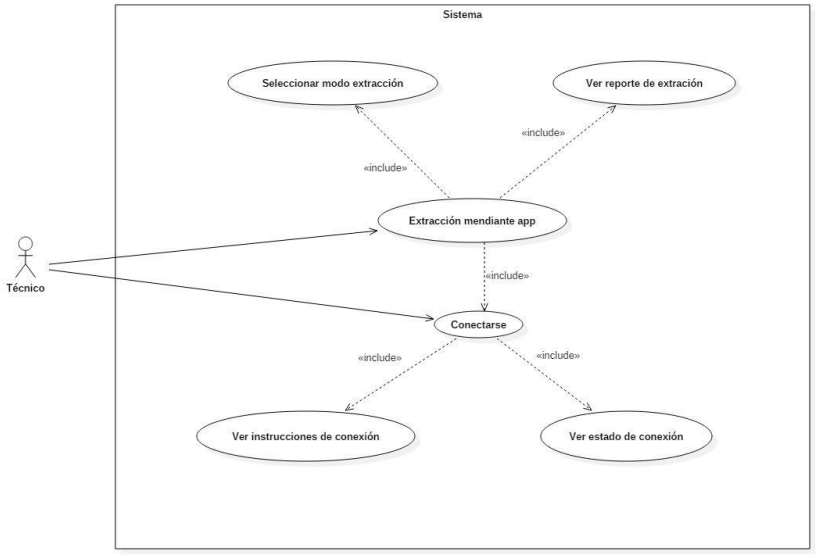


Figura 2. Caso de Uso del escenario extracción mediante App

En la Figura 3 se muestra el Diagrama de Secuencias del escenario “Extracción mediante app” que muestra el comportamiento de los componentes de software. Posteriormente, en la Figura 4, se expone el modelo estático como un Diagrama de Clases, de la aplicación desarrollada. Se muestra que como parte del Modelo se implementó el componente ControladorLAF que implementa el protocolo LG Advanced Flash, el cual permite la extracción física de dispositivos LG. Así diseñada, esta arquitectura habilita implementar y sumar componentes, como parte del Modelo, que permitan la extracción física de dispositivos de otras marcas.

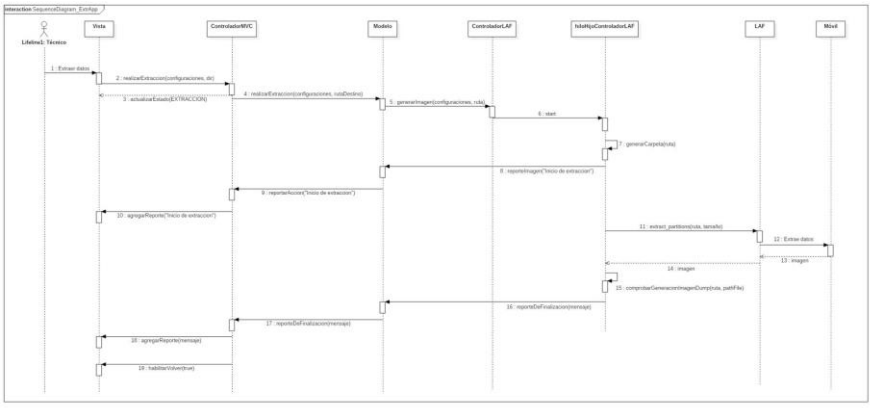


Figura 3. Diagrama de secuencias del escenario extracción mediante App

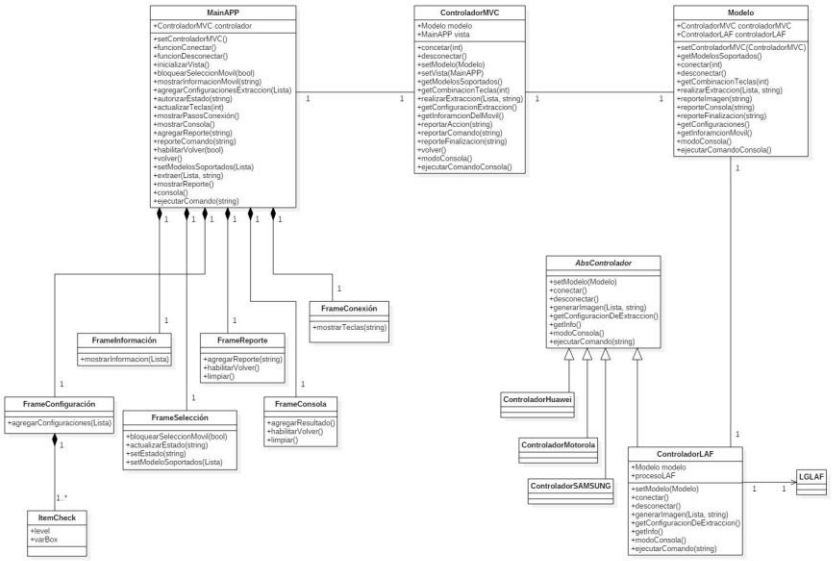


Figura 4. Diagrama de clases

3 Resultados obtenidos

Se corrió la aplicación en una netbook con cpu Intel Atom 1.6 GHz, 2 Gb de RAM, puerto USB 2.0 y espacio suficiente para almacenar la imagen del dispositivo. El sistema operativo utilizado fue la distribución de Linux denominada Forget Windows Use Linux (FWUL) basada en Arch Linux. Se trata de una distribución de Linux que facilita la conexión entre un dispositivo móvil Android y una PC ya que posee, entre otras cosas, todos los driver necesarios.

Para probar la funcionalidad de la aplicación, se realizó una primera extracción a un dispositivo LG modelo G2 mini con Android v4.2. Los resultados fueron positivos ya que se logró extraer el bloque de memoria interna de 7.3 Gb en un tiempo aproximado a 75 minutos. Luego el contenido pudo visualizarse mediante la aplicación open source Autopsy.

Posteriormente se realizaron extracciones de tres dispositivos mas : LG K4 con Android v5.1.1y LG Spirit con Android v5.0.1 (dos ejemplares) facilitados por el Ministerio Público Fiscal al LARYC para realizar estas pruebas. Los tamaños de memoria extraida fueron también de 7.3 Gb y el tiempo de extracción se mantuvo en valores superiores a la hora.

Para corroborar la integridad de los datos de las imágenes extraídas se llevó a cabo la comparación de valores hash. Por un lado, LG K4 tiene implementado y accesible el algoritmo MD5 para generar un hash de la memoria de almacenamiento previo realizar la extracción. Luego, se hizo un cálculo del mismo hash, sobre la imagen extraida y se compararon. Fueron iguales. Luego se verificó la integridad del proceso de extracción.

4 Conclusiones

La investigación llevada a cabo permitió encontrar una librería *open source* que hace uso de un protocolo denominado LG LAF. Este fue probado en diferentes dispositivos de marca LG con la intención de lograr realizar una extracción física de datos. Cabe destacar que, aunque para algunos modelos de dispositivos dicha librería funcionó, en otros modelos de la misma marca, no. La explicacion que encontramos fueron las siguientes:

- Desde Android v4.1 hacia atrás, el protocolo LG LAF no es soportado.
- Desde Android v6.0 en adelante, la extracción se puede realizar pero no visualizar debido a que los datos se encuentran encriptados con Full Data Encryption (FDE).
- En los casos que no se corresponden con las versiones de Android mencionadas, la extracción puede fallar por problemas con los drivers del sistema operativo de la máquina anfitrión. FWUL solucionó este problema.

Se continúan realizando pruebas sobre otros modelos de dispositivos LG y versiones de Android. Por otro lado, se está estudianto e implementando componentes para extracción de otras marcas, que implementan diferentes estrategias.

Referencias

1. Tamma, R.; Skulkin O.; Mahalik, H.; Bommisetty, S.: Practical Mobile Forensics Third Edition. Birmingham, UK, (2018).-
2. Practical Android Phone Forensics; consultado en abril de 2019; https://resources.infosecinstitute.com/practical-android-phone-forensics/#gref
3. Forensics – imágenes de disco paso a paso; consultado en abril de 2019; http://highsec.es/2013/08/forensics-imagenes-de-disco-paso-a-paso/
4. Imaging Android with ADB, Root, Netcat and DD | Digital Forensic Science; consultado en abril de 2019; https://dfir.science/2017/04/Imaging-Android-with-root-netcat-and-dd.html.-
5. GitHub - Lekensteyn/lglaf: LG Download Mode utility and documentation; consultado en abril de 2019 de, https://github.com/Lekensteyn/lglaf.-
6. Modelo–Vista–Controlador, Wikipedia, la enciclopedia libre; consultada en Abril del 2019; https://es.wikipedia.org/wiki/Modelo-vista-controlador.-

Presentación de proyecto para el análisis de incidentes de ciberseguridad o ciberataques durante las acciones de ciberdefensa de las infraestructuras críticas de la defensa nacional – InFoscopia–

Julio C. Liporace[1], *Adrián Buscaglia*[1], *Pablo Croci*[1], *Nicolás Díaz Pais*[1], *Darío Fernández*[1], *Verónica Ferreyra*[3], *Ignacio Martín Gallardo*[1,2] , *Elvira Quiroga*[12], *Fernando Vera Batista*[1] , *César D. Cicerchia*[1]

[1] Facultad de Ingeniería del Ejército (FIE), Ejército Argentino – Universidad de la Defensa Nacional;
[2] Centro de Investigación y Desarrollo de Sistemas Operacionales (CIDESO), Dirección General de Investigación y Desarrollo (DIGID) - Ejército Argentino

[3] Comando Conjunto de Ciberdefensa (CCCD), Estado Mayor Conjunto de las Fuerzas Armadas

{jcliporace,abuscaglia,pcroci,ndiaspais,dfernandez,igallardo,
equiroga,verabatista,cdcichercia}@est.iue.edu.ar
vferreyra@fuerzas-armadas.mil.ar

Resumen. El desarrollo actual de la Ciberseguridad y la Ciberdefensa está apoyado en la aplicación de procedimientos reactivos de detección, mitigación y remediación, que se aplican cuando el efecto de la agresión sobre una Infraestructura Crítica del Sistema de Defensa Nacional (ICSDN) ya ocurrió. Se ha adoptado un enfoque derivado de las normativas de Seguridad de la Información, tal como las formuladas por el estándar internacional ISO/IEC 27000. El aporte original de esta línea de investigación se basa en proponer una nueva metodología de Ciberdefensa, desarrollando un abordaje proactivo hacia las amenazas antes que éstas comprometan la infraestructura crítica, a fin de sorprender al agresor mediante una defensa dinámica y reducir sus posibilidades de éxito.
Las líneas de investigación de InFoscopia tienen como alcance desarrollar una metodología proactiva de análisis de eventos que ocurren antes de que el objetivo cibernético haya sido comprometido. Se enfocará en las ICSDN estratégicas (servicios esenciales de energía, transporte, financieros, comunicaciones e informática, alimentos, agua, químicos, nuclear o espacial) y de capacidades militares. Contribuirá al desarrollo de procedimientos proactivos por parte de los grupos de respuesta del tipo Computer Security Incident Response Team (CSIRT) o Centro de Operaciones de Ciberdefensa, encargados de la protección de las ICSDN.

Palabras Clave: Informática Forense. Informatoscopia. Ciber-seguridad. Ciberdefensa. Ciberataque. Infraestructura Crítica.

1. Contexto

La reciente Directiva de Política de Defensa Nacional[1] destaca que, en la atención del riesgo de "Ataques externos a objetivos estratégicos", el Sistema de Defensa Nacional debe focalizarse en "aquellas infraestructuras cuyo funcionamiento resulte crítico para el cumplimiento de las funciones vitales del Estado Nacional, su Defensa Nacional, el ejercicio de la soberanía y el resguardo de la vida y la libertad de sus habitantes."
El Ejército Argentino, desde hace un par de décadas patrocina la construcción de su propio sistema de comando y control para las Grandes Unidades de nivel táctico (Brigadas). Los sistemas de comando y control (C2) son de naturaleza socio – técnica y complejos en su concepción y diseño (Clay, 2007). Por su naturaleza y su finalidad de empleo, los sistemas de C2 constituyen una ICSDN (Dean, 2013). Por otra parte, el Ejército Argentino ha proporcionado los recursos humanos formados como Ingenieros Militares o en Sistemas de Computación para la creación y constitución del Comando Conjunto de Ciberdefensa[2].

En base a estos antecedentes, el grupo de trabajo del proyecto InFoscopia es responsable de abordar esta línea de investigación. Para su organización ha recibido el aporte de conocimiento de docentes y graduados del posgrado de Especialización en Criptografía y Seguridad Teleinformática y de la carrera Ingeniería en Informática, ambas de la misma unidad académica. Asimismo, los docentes son investigadores categorizados con trayectoria en otros proyectos de investigación de la FIE o docentes con experiencia en investigación aplicada al desarrollo tecnológico precompetitivo de sistemas militares.

InFoscopia es un proyecto aprobado y acreditado por el Programa de acreditación y financiamiento de Proyectos UNDEFI, convocado mediante Resolución Rectoral 154/18 de la Universidad de la Defensa Nacional (UNDEF). El financiamiento del proyecto está sustentado por parte de la FIE, mediante la asignación de cargos y horas de investigación de los docentes integrantes del grupo de trabajo, y de la UNDEF, por medio de un subsidio UNDEFI que se renueva anualmente.

Durante las actividades de investigación se intercambiarán conocimientos y experiencias con los siguientes grupos de investigación: Grupo de Investigación del CriptoLAB de la FIE / UNDEF; Grupo de Investigación Sistemas Operativos e Informática Forense de la Facultad de Ingeniería de la Universidad FASTA (sede Mar del Plata).

Por su parte, La Facultad de Ingeniería del Ejército, así como la Facultad de Ingeniería de la Universidad FASTA son miembros fundadores de la Red Universitaria

[1] Decreto 703/2018. DECTO-2018-703-APN-PTE - Directiva de Política de Defensa Nacional. Aprobación.
https://www.boletinoficial.gob.ar/pdf/pdfAnexoPrimera/5568234A01.pdf/20180731/0

[2] http://www.fuerzas-armadas.mil.ar/Dependencias-CIBDEF.aspx

de Informática Forense. Esta red constituye un apoyo fundamental para el Grupo de Investigación.

2. Introducción

Las Infraestructuras Críticas del Sistema de Defensa Nacional (ICSDN) constituyen recursos diversos y complejos, aunque en la actualidad todas tienen uno o más componentes de TIC (Edwards, 2014). Desde la perspectiva del Estado, se debe considerar una infraestructura como aquel conjunto de medios técnicos, servicios e instalaciones necesarios para el desarrollo de las actividades básicas de la sociedad. En este sentido, la mayoría de esas actividades proveen servicios esenciales de carácter estratégico a la sociedad, al gobierno y a los habitantes en su conjunto, tanto sean prestados por organizaciones de gestión pública como privada (Baggett & Simpkins, 2018).

La interrupción o perturbación severa de su funcionamiento, ocasionaría graves efectos sobre el normal desarrollo de las actividades básicas de la sociedad; por tal motivo deben ser consideradas infraestructuras críticas y su defensa, aún en tiempo de paz, es un deber fundamental del Estado. La componente TIC de cualquier ICSDN puede ser afectada desde el Ciberespacio, con un elevado y creciente riesgo que debe ser analizado y prevenido frente a un número importante de amenazas cibernéticas, cuyo comportamiento es dinámico, cambiante y de difícil predicción (Johnson, 2015).

Las amenazas cibernéticas progresan en su acción ofensiva, evolucionando entre las siguientes etapas (The Mitre Corporation, 2018):

- Exploración o Reconocimiento inicial.
- Adquisición intrusiva de servicios o procesos computacionales de la infraestructura / Desarrollo de herramientas acordes a la debilidad a ser explotada.
- Entrega o distribución
- Compromiso Inicial.
- Uso indebido / Escalamiento de Privilegios.
- Reconocimiento interno.
- Movimiento lateral.
- Establecimiento de la persistencia (consolidación).
- Ejecución de la Misión o cumplimientos de Objetivos.
- Exfiltración.

El conocimiento tecnológico que proporciona la Informática Forense, la Informatoscopia y su apoyo en las Ciencias de la Computación tiene su campo de aplicación en la Justicia (Consejo General del Poder Judicial (CGPJ) et al., 1996). Esta capacidad se sustenta en el empleo de técnicas científicas y analíticas especializadas sobre la infraestructura tecnológica y se desarrolla con la finalidad de identificar, preservar, analizar y presentar datos que sean válidos dentro de un proceso legal, en el marco de un delito informático que es objeto de investigación judicial. Es decir que estas disciplinas actúan ex post, luego de ocurrido un delito (Domínguez, 2013).

En la Ciberdefensa tienen prioridad la mitigación, contención o respuesta inmediata para detener los efectos de una ciberagresión o ciberataque, aunque se pueda mantener

el objetivo de la recolección de evidencias de la vulneración de los sistemas y preservar las pruebas, pero dejándolo en segundo orden de prioridad (Intelligence and National Security Alliance, 2018). Sin embargo, los métodos y herramientas de la Informática Forense pueden resultar de utilidad para apoyar el proceso de gestión de incidentes de Ciberseguridad y Ciberdefensa, pero bajo otra metodología enfocada en una acción proactiva, antes que el efecto de la agresión llegue a su punto culminante (Colbaugh & Glass, 2011).

La explotación del análisis forense digital y su capacidad de extraer muestras o pruebas de las computadoras, equipos móviles y otros dispositivos es fundamental para descubrir e interpretar datos electrónicos.

Un componente de soporte al desarrollo de la metodología es la "línea de tiempo" que permiten mostrar quién hizo, qué y cuándo, de manera de poder afirmar de forma concluyente que la Acción A causó el Resultado B (Amusategui López, 2016).

3. Líneas de investigación y desarrollo

Se identifican dos líneas principales de investigación de interés para InFoscopia; a saber:

- Análisis de amenazas cibernéticas y de tendencias de ciberagresiones a infraestructuras críticas estratégicas, para modelar el comportamiento de agresores cibernéticos que puedan afectar, especialmente, la disponibilidad3 de las infraestructuras críticas.
- Métodos de análisis forense digital en memoria, de ingeniería inversa de software y de análisis de malware, que puedan extenderse o aplicarse para estudiar procesos "vivos" en dispositivos de datos, red, seguridad informática o soporte. Se consideran de interés para elaborar un conjunto de Indicadores de Amenaza (*Indicators of Threat* - IoTh), por analogía a los Indicadores de Compromiso (*Indicators of Compromise* - IoC).

Para este proyecto, la primera línea de investigación resulta de interés para el modelado del comportamiento del agresor cibernético, especialmente en las etapas previas del ciberataque (exploración o reconocimiento inicial, adquisición intrusiva de servicios o procesos computacionales de la infraestructura, desarrollo de herramientas ofensivas acordes a la debilidad a ser explotada) y en momento inicial del asalto a la infraestructura, cuando empieza la entrega o distribución de una ciberarma (malware, tramas anómalas de datos, etc).

La segunda línea de investigación tiene interés, entre otros, sobre los métodos de recolección de información disponibles para archivos dependientes de la memoria RAM, como ser la RAM propiamente dicha y el archivo de paginación de los sistemas

[3] *Availability*, adoptada por el estándar internacional ISO/IEC 27000. Se refiere a la "propiedad de la información [o de un sistema] de estar accesible y utilizable cuando lo requiera una entidad autorizada".

operativos. También, el análisis de artefactos o piezas de software en tiempo de ejecución, los procesos desatados por el malware o software dañino y el análisis de tráfico en la red, desde y hacia el activo comprometido. Se pretende analizar y comprender eventos en tiempo de ejecución para clasificarlos como normales o anómalos con un grado de confianza aceptable.

El proyecto tiene previsto tres etapas:

- Formulación metodológica para activos bajo entornos Windows en redes Ethernet proponiendo indicadores de amenazas a la disponibilidad. Se validará con pruebas de concepto.
- Ampliación de la metodología para ambientes heterogéneos con dispositivos Linux y medios de redes y seguridad particulares de infraestructuras críticas. Se desarrollará un prototipo experimental.

Se completará con la experimentación y selección de parámetros definitivos de los indicadores de amenazas definidos.

4. Resultados obtenidos/esperados

El principal resultado esperado es desarrollar una metodología propia de análisis de eventos o incidentes aplicable a las fases iniciales de una ciberagresión o ciberataque que pudiera ocurrir en uno o más activos esenciales de una ICSDN, con la finalidad de dar respuesta a los siguientes propósitos:

- Detectar cómo y cuándo ocurrió una violación de la protección de una ICSDN.
- Identificar los activos esenciales o sistemas comprometidos y afectados.
- Determinar qué atacantes tomaron o cambiaron procesos.
- Contener y remediar los incidentes que se configuren.
- Desarrollar indicadores y fuentes clave de inteligencia de amenazas.
- Buscar violaciones adicionales usando el conocimiento de las tácticas, técnicas y procedimientos del agresor.

Con respecto a la primera línea, se ha comenzado por adoptar como referencia el marco de trabajo MITRE's Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK), desarrollada y mantenida por la organización estadounidense The Mitre Corporation. Se trata de una base de conocimiento accesible a nivel mundial, sobre tácticas, técnicas y procedimientos que utilizan los agresores cibernéticos.

Entre los primeros resultados de la segunda línea de investigación, por medio del empleo de varias herramientas forenses de análisis de memoria de computadoras con arquitectura Von Neumann, se logró validar la técnica de obtención de memoria para el análisis de cadenas de caracteres presentes en tiempo de ejecución. Como continuación de este trabajo, mediante la utilización de la herramienta Sysmon (de la suite Sysinternals de Microsoft) se pretende estudiar el comportamiento de los activos bajo condiciones normales de operación y compararlos con los efectos de eventos asociados a la fase previa de los ciberataques. La finalidad es obtener información detallada sobre las

creaciones de procesos, conexiones de red y cambios en el tiempo de creación de archivos. Al recopilar los eventos que se generan, se podrá realizar el análisis en vivo de los Logs para identificar actividades anómalas y comprender cómo operan los intrusos y el malware en la red de una infraestructura crítica, de manera de encontrar posibles Indicadores de Amenazas (IoTh) en un momento previo al compromiso de la disponibilidad del activo (weaponization en el framework PRE-ATTACK de Mitre).

5. Formación de recursos humanos

El Grupo de Investigación tiene conocimientos y experiencia sobre las técnicas y tecnologías de Informática Forense, Ciberseguridad y Ciberdefensa, Redes de Información, Sistemas Operativos de Computadoras, Sistemas de Control, Sistemas Electrónicos, Ingeniería de Software y Técnicas de Programación.

La estructura del equipo de trabajo está conformada por dos docentes e investigadores del posgrado de Especialización en Criptografía y Seguridad Teleinformática, tres docentes e investigadores de la carrera de Ingeniería en Informática, un tecnólogo posgraduado en la especialización mencionada con tesis vinculada al proyecto y estudiante de Doctorado en Ciencias Informáticas, dos tecnólogos con perfiles diferentes asociados a sistemas operativos y redes de computadoras, y un tecnólogo con experiencia en incidentes de ciberseguridad. Adicionalmente, se han incorporado dos alumnos de la carrera de Ingeniería en Informática y uno de Ingeniería Electrónica que desarrollarán tesinas de grado vinculadas al proyecto.

La formación de recursos humanos es un objetivo del proyecto y está reflejado en la constitución del grupo, donde se integran los conocimientos de docentes muy experimentados y jóvenes, personal técnico y tecnólogos de tres laboratorios asociados al proyecto (CriptoLab y CIDESO / Ejército; InFo-Lab / Universidad FASTA). Por último, se pretende becar desde un doctorando hasta alumnos de grado y de postgrado, e incorporar pasantes en el marco de la Red UNIF.

Referencias

1. Amusategui López, J. M. (22 de 12 de 2016). Universitat Oberta de Catalunya. Obtenido de http://openaccess.uoc.edu/webapps/o2/bitstream/10609/60765/6/jamusateguiTFM1216memoria.pdf
2. Baggett, R. K., & Simpkins, B. K. (2018). Homeland Security and Critical Infrastructure Protection, 2nd Edition. Santa Barbara, California. USA: Praeger Security International.
3. Clay, W. (2007). Network Centric Operations: Background and Oversight Issues for Congress. CRS Report for Congress.
4. Colbaugh, R., & Glass, K. (15 de 08 de 2011). IEEE Xplore Digital Library. (P. o. Informatics, Ed.) doi:10.1109/ISI.2011.5984062
5. Consejo General del Poder Judicial (CGPJ) et al. (31 de 12 de 1996). CITA.ES. Obtenido de http://cita.es/apedanica/INFORMAT.HTM

6. Dean, S. E. (2013). Cyber Defense: Securing Military Systems and Critical Civilian Infrastructure from an Electronic 9/11. HRISQ stands for Hampton Roads International Security Quarterly.
7. Domínguez, F. L. (2013). Introducción a la Informática Forense. Madrid. España: Ra-Ma.
8. Edwards, M. (2014). Critical Infraestructurre Protection. Ankara, Turquía: IOS Pres BV.
9. Intelligence and National Security Alliance. (31 de 08 de 2018). INTELLIGENCE AND NATIONAL SECURITY ALLIANCE (INSA). Obtenido de https://www.insaonline.org/wp-content/uploads/2018/08/INSA-Managing-Cyber-Attack-Critical-Infrastructure.pdf
10. Johnson, T. A. (2015). Cybersecurity: Protecting Critical Infrastructures from Cyber Attack and Cyber Warfare. St.Louis, Missouri. USA: CRC Press.
11. The Mitre Corporation. (22 de Nov de 2018). MITRE ATT&CK. Obtenido de https://attack.mitre.org/

INFOCONF 2019

Casos de Pericias Forenses Digitales (WCAPFD)

Informática Forense y Pericial (WIFP)

Aspectos Legales de la Actuación Forense (WALAF)

Presentación de casos relevantes de pericias forenses digitales/ Ciber Investigaciones / Análisis Forense en General / Análisis de Datos para Identificación de Actividades Maliciosas.

Identificación y comparación de imágenes en ambientes forenses

Bruno Constanzo, Martín Castellote, Santiago Trigo, Ana Haydée Di Iorio

Laboratorio de Investigación y Desarrollo de Tecnología en Informática Forense InFo-Lab
Universidad FASTA, Ministerio Público de la Provincia de Buenos Aires,
Municipio de General Pueyrredon
{bconstanzo, mcastellote, santiagotrigo, diana}@ufasta.edu.ar

Resumen. Las investigaciones digitales de la actualidad se enfrentan a una gran cantidad de contenido multimedia, que pone a los peritos e investigadores en la necesidad de analizar, buscar y detectar contenido específico en conjuntos de cientos de miles o millones de imágenes y videos. En particular, en las investigaciones de abuso sexual de menores esta problemática se ve ampliada por la escala de los casos, y por el fuerte impacto psicológico que tiene el mismo sobre los peritos e investigadores. Es de vital importancia implementar nuevas soluciones que tengan la capacidad de procesar el contenido en base a sus características visuales, y a partir de ellos poder hacer análisis de similitud, detección, y filtrado automático. En este trabajo se presentan 3 hashes perceptuales implementados como parte de una librería orientada al análisis forense de imágenes y video. Se explican los algoritmos, su implementación, se analiza su rendimiento, y se presentan casos de uso. Finalmente, se evalúa cómo se pueden utilizar en ámbitos de la justicia (tanto pericial como investigación), y se proponen alternativas para continuar el desarrollo.

Keywords: hashes perceptuales - procesamiento de imágenes - informática forense - investigaciones digitales

1 Introducción

La sociedad moderna ha experimentado avances tecnológicos vertiginosos y de gran impacto social en los últimos 10 años. La aparición del *smartphone* puso en los bolsillos de gran parte de la población un instrumento multimedia de enormes capacidades, tanto para generar contenido multimedia, como para compartirlo rápidamente con otras personas. Esta explosión multimedia impacta directamente sobre la tarea de los peritos informáticos y los investigadores judiciales, que en su actuación diaria deben enfrentarse ya no a miles, sino a cientos de miles, o incluso millones de imágenes y videos[1, 2].

En los casos de abuso sexual infantil, la problemática se ve agravada por el impacto psicológico que tienen las imágenes sobre los expertos que deben trabajar en estos casos[3]. Ya no es sólo un problema de escala o rapidez con la que se puedan analizar los

adfa, p. 1, 2011.

casos, sino que se convierte en una cuestión de salud psicológica y emocional de los investigadores.

Una de las herramientas que se puede plantear para enfrentar este problema es la utilización de hashes perceptuales, especializados para el caso de imágenes y fotografías digitales. Estos tipos de hashes tienen la ventaja frente a los hashes tradicionales basados en algoritmos que se computan sobre los bytes de cada archivo: al estar basados en los patrones visuales de la imagen, son resistentes a cambios en el archivo que no modifique la misma. Es decir, si una misma imagen se guarda como JPG, PNG, BMP o TIF, los hashes perceptuales mantendrán el mismo valor (o un suficiente nivel de similitud que posibilite detectar el contenido), independientemente del formato de archivo en que se guarde, a diferencia de los hashes MD5 y SHA-1 que usualmente se utilizan en las pericias informáticas.

En este trabajo se presentan los conceptos teóricos básicos de hashes e imágenes digitales que son necesarios para comprender estas técnicas, y luego se exponen tres algoritmos de hashes perceptuales que fueron implementados por los autores. Luego se muestran resultados de las pruebas realizadas utilizando estos hashes, y ejemplos de situaciones que se pueden resolver por medio de ellos. Finalmente, se llega a conclusiones sobre su utilización, y se plantea trabajo y aplicaciones futuras que podrían aportar a mejorar esta técnica, y el trabajo de los expertos en informática forense del país.

2 Marco teórico

2.1 Hashes: funciones y digestos

Las funciones de hash son funciones matemáticas o procedimientos que transforman un mensaje de entrada m en un digesto (de tamaño fijo) d, tal que $d = f(m)$. Dependiendo de la función de hash utilizada, y su objetivo, variará la complejidad en el cálculo del digesto, y se tendrán distintas garantías sobre el resultado.

Se las puede clasificar en base a su objetivo de la siguiente manera:

- **Funciones de índice** para tablas de hashes, que permiten ordenar información dentro de una estructura de datos de acceso aleatorio y, en base a una clave o identificador, rápidamente acceder a un contenido. Este tipo de funciones se utilizan para bases de datos, estructuras de dato de acceso rápido, o caches, entre otros.
- **Funciones de *checksum*** que tienen la capacidad de detectar cambios en un archivo que ha viajado por un medio de transmisión, de manera que se pueda verificar su integridad. Dado que están orientadas a detectar errores de transmisión, su enfoque no es tan riguroso como los hashes criptográficos.
- **Hashes criptográficos** orientados a utilización en criptografía, para asegurar comunicaciones, firmas digitales, verificación de passwords, entre otras. Las funciones de hash criptográficas idealmente cumplen con las siguientes propiedades:
  - Determinística: el cálculo repetido de d en base al mismo mensaje m siempre debe dar el mismo resultado.
  - Irreversible: no se puede inferir el mensaje m en base a d.
  - Avalancha: un cambio pequeño en m ocasiona grandes cambios en d.

  - Baja probabilidad de colisión, entendiendo como colisión al hecho que, para dos mensajes m_1 y m_2 distintos, $f(m_1) = f(m_2)$.
  - Debe ser de rápido cálculo.
- Hashes perceptuales orientados a identificar y detectar similitudes entre elementos multimedia de fuertes características sensoriales (visuales o auditivas).

Para la informática forense, son de particular interés los hashes criptográficos. Las funciones MD5 y SHA-1[4, 5] se utilizan para complementar la cadena de custodia y garantizar la validez e integridad[1] de las fuentes de evidencia digital[6 Cap. 5].

Pese a sus características para identificar y verificar archivos, tienen una problemática al enfrentarse a cambios en su estructura. Si se cambia el tipo de archivo, por ejemplo, de JPG a PNG, la imagen visualmente seguirá siendo la misma, pero los hashes criptográficos correspondientes no coincidirían entre el archivo JPG y el PNG. Lo mismo sucedería si el archivo contiene metadatos y estos fueran eliminados o modificados.

Antes de analizar en mayor detalle los hashes perceptuales que son objeto de este trabajo, es necesario introducir algunos conceptos de imágenes digitales necesarios para su presentación y análisis.

2.2 Imágenes digitales

Una imagen digital I se puede definir como una función $f(x, y)$, siendo x e y las coordenadas espaciales, y la amplitud de f en cualquier punto es la intensidad o nivel de gris de la imagen. Cuando f, x e y son todas cantidades discretas y finitas, entonces se trata de una imagen digital[7]. Es posible armar una matriz o arreglo de intensidades, de manera que:

$$f(x,y) = \begin{bmatrix} f(0,0) & f(1,0) & \dots \\ f(0,1) & f(1,1) & \dots \\ \dots & \dots & \dots \end{bmatrix}$$

A cada elemento de esta matriz se lo conoce como pixel, y se suelen limitar los valores que almacenan a un entero sin signo de 8 bits, que da un rango de posibles valores de entre 0 y 255[2]. Los valores menores que 0, o mayores que 255, son "saturados" y no se puede representar información por debajo o por encima de ellos, un fenómeno conocido como *clipping*.

Una imagen en escala de grises tendrá una única matriz de intensidades, que indicarán el nivel de intensidad de la imagen en cada punto. Si se trata de una imagen a color, la representación más usual consiste en utilizar tres matrices de intensidad, correspondientes a los colores rojo, verde y azul, o RGB (del inglés *red green blue*). A cada una de estas matrices de intensidad se las suele llamar "canal". Otros espacios de color

[1] Los posibles ataques criptográficos a estas funciones son conocidos por los autores. La discusión sobre utilizar otros tipos de hashes, cómo enfrentar la incertidumbre que plantean, y la probabilidad de encontrar un ataque real en entornos forenses, son objeto de otro trabajo.

[2] Otras representaciones utilizan enteros de más de 8 bits, o flotantes de 32 o 64 bits. Para la discusión de este trabajo, es indistinto.

representan la imagen digital en función de otras características, por ejemplo, HSL tiene tres canales dedicados al tono (*hue*), saturación (*saturation*) y nivel de brillo (*levels*).

De aquí en adelante en el trabajo se asumirá que las imágenes son matrices RGB con 8 bits de precisión por canal, ya que es una de las representaciones más utilizadas por los formatos de imagen, entre ellos JPG y PNG[8, 9]. Se deja de lado la cuestión de la codificación y compresión, porque esta, si bien puede traer ligeras modificaciones a la imagen en el caso de compresión con pérdida, no altera el hecho de representar la imagen como una matriz de intensidades.

Operaciones sobre imágenes.

Las representaciones como arreglo de valores o como matriz de una imagen permite aplicar operaciones matemáticas sobre la misma. Estas operaciones buscan realizar algún cambio sobre en la imagen que permita realizar un análisis o interpretación de la información que representa.

Las operaciones de imagen que se utilizarán en este trabajo son:

- Conversión RGB a escala de grises: cada pixel RGB de la imagen I se reemplaza por un único valor, $y = f(r, g, b)$, aplicando alguna función de transformación. Una de las más comunes es:

$$y = 0.299r + 0.587g + 0.114b \text{ [10]} \quad (1)$$

  Esta operación reduce la cantidad de canales de color de 3 a 1.
- Recorte (crop): el recorte de una imagen I en (x_0, y_0) con ancho w y alto h, $crop(I, x0, y0, w, h)$ genera una submatriz de imagen I' tal que su esquina superior izquierda es $I(x_0, y_0)$ y su esquina inferior derecha es $I(x_0 + w, y_0 + h)$.
- Curva de niveles: se procesa cada pixel p_{xy} de I de acuerdo a una curva que describe para cada valor [0..255] un valor correspondiente [0..255] que transforma al pixel. Esta operación permite cambiar el contraste de la imagen. Si la curva aplicada es biyectiva, la operación puede ser invertida, en caso contrario, será una operación que causará pérdida de información.
- Redimensionado (*downsample*): se transforma la matriz I de dimensiones (w, h) siendo w el ancho en pixels y h el alto en pixels, en una matriz I' de dimensiones (w', h') donde $(w' < w, h' < h)$. Cada pixel p'_{xy} tiene un valor interpolado en base a varios pixels p_{xy} originales. Esta operación reduce la resolución espacial de la imagen.
- Transformaciones geométricas: el producto matricial entre la matriz de intensidades A y una matriz M de 3x3 permite aplicar transformaciones afines u homografías a una imagen. Dependiendo de cómo se forme la matriz M, ésta puede aplicar múltiples transformaciones sobre la imagen en una sola operación, incluyendo traslaciones, rotaciones, sesgos, o cambios de perspectiva.
- Transformaciones de dominio: las operaciones que se mencionaron hasta ahora operan todas en el dominio espacial. Algunas operaciones de imagen pueden representarse mejor en un dominio transformado. En este trabajo es de interés la

Transformación Discreta del Coseno, o DCT por sus siglas en inglés, que permite obtener coeficientes que describen la estructura general de la imagen. Su fórmula es:

$$X_k = \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}\left(n + \frac{1}{2}\right)k\right] \quad k = 0, \dots, N-1$$

2.3 Hashes perceptuales

Un hash perceptual es aquel que se basa en características sensoriales de un archivo para realizar el cálculo. Para el caso de archivos de imagen, el objeto de interés de este trabajo, los algoritmos de hashes perceptuales se enfocan en los patrones visuales que se generan al renderizar una imagen.

En su estudio de hashes perceptuales, Zauner define los siguientes conceptos de interés[11]:

- **Modificación:** es una operación sobre el objeto de medios que **no afecta** sus características esenciales.
- **Manipulación:** es una operación sobre el objeto de medios que **sí afecta** sus características esenciales.

Los algoritmos de hashes perceptuales deben ser resistentes a las modificaciones, y no confundir las manipulaciones sobre un contenido. Es decir, si se compara la similitud de una imagen con otras versiones de la misma, que han sufrido modificaciones, debe identificarlas como la misma imagen, ya que no han cambiado sus características esenciales. Al contrario, si ha habido manipulación de la imagen, el hash perceptual debe identificarlas como imágenes distintas.

El resultado de los hashes perceptuales que se presentan en este trabajo es una secuencia de 64 bits, organizados en memoria como un arreglo de 8x8. La comparación entre dos hashes se realiza aplicando la distancia de Hamming[12] entre ellos, y definiendo un umbral de similitud. Este umbral es dependiente del tipo de hash que se haya aplicado. Un umbral de 0 exige al algoritmo que el hash de las imágenes debe ser idéntico. Un umbral más alto, permitirá detectar modificaciones sobre una imagen, pero descartar manipulaciones.

Como parte de este trabajo, se han implementado tres algoritmos de hashes perceptuales de imagen: Average Hash (*a_hash*), Difference Hash (*d_hash*) y DCT-Hash (*p_hash*)[3]. Se presentarán a continuación, con una breve discusión sobre sus ventajas y desventajas.

Average Hash.

[3] Para los nombres de los hashes se ha adoptado la convención de Zauner, pero también se brinda entre paréntesis el nombre popular con el que algunas librerías implementan estos hashes (ver [13-15])

Este hash perceptual es conceptualmente muy simple y fácil de calcular. Como modelo conceptual para entender otros hashes es interesante, pero no es tan robusto como las otras alternativas. El algoritmo para calcularlo es el siguiente:

- Convertir la imagen a escala de grises
- Redimensionar la imagen a 8x8. Se llamará a esta imagen, I'.
- Calcular el valor promedio de los 64 pixels, avg.
- Para cada pixel p'_{xy} de I', el valor de hash H en la posición (x, y) será $h_{xy} = 1$ si $p_{xy} > avg$, o $h_{xy} = 0$ en caso contrario.

De esta manera, el hash resultante describe la estructura macro de la imagen. Cualquier alteración sutil, o que afecte detalles finos, no generará cambios lo suficientemente grandes en el pixel p_{xy} afectado para impactar en el valor h_{xy} final. Si en cambio fueran alteraciones mayores, entonces se modificará tanto el valor de p_{xy}, como el promedio de la imagen. Un cambio de esa magnitud afectaría globalmente al hash H, modificando varios de sus elementos h_{xy}.

La mayor problemática de este algoritmo es que, al basarse en el valor promedio de los pixels, las modificaciones sobre el contraste de la imagen afectan al promedio de intensidad y generan grandes alteraciones en el hash, sin que haya cambiado la esencia de la imagen. Esto es particularmente notable si se realizan cambios sobre el canal verde, que como indica la ecuación (1), tiene una fuerte influencia sobre el nivel de gris calculado. Por estas razones, se concluye que este no es un hash robusto.

Difference Hash.

Este hash se basa en los gradientes horizontales de la imagen. Su cálculo es relativamente simple, pero el concepto en el que se basa es un mejor descriptor de la estructura macroscópica de una imagen que el utilizado por el Average Hash. El algoritmo para calcular este hash es el siguiente:

- Convertir la imagen a escala de grises.
- Redimensionar la imagen a 9x8. Se llamará a esta imagen, I'.
- Se toman dos subregiones de I' de tamaño 8x8, α y β, siendo α la subregión de más a la izquierda de I', y β la subregión de más a la derecha.
- El hash final H se construye como un arreglo de 8x8, donde valor de hash en la posición (x, y) se construye de la siguiente manera: $h_{xy} = 1$ si $\alpha_{xy} < \beta_{xy}$, $h_{xy} = 0$ en caso contrario.

Esta construcción aproxima el gradiente horizontal de la imagen escalada I'. Esto permite que el hash se construya en base a los bordes de imagen, que son más resistentes a las alteraciones que resultan problemáticas para el Average Hash. Otra ventaja de este algoritmo es que es muy simple de implementar, y su velocidad de cálculo es elevada ya que sólo hay que hacer comparaciones entre valores adyacentes en memoria.

No tiene ninguna desventaja en sí, pero tiene debilidades comunes a los tres algoritmos que se mencionan más adelante.

DCT-Hash.

Este hash se basa en coeficientes DCT calculados sobre la imagen. Si bien el cálculo es más exigente que las alternativas vistas hasta ahora, la implementación del algoritmo lograda no presenta grandes diferencias de velocidad con respecto a los anteriores. El algoritmo para calcular este hash es el siguiente:

- Convertir la imagen a escala de grises.
- Redimensionar la imagen a 32x32. Se llamará a esta imagen I'.
- Aplicar la transformada DCT sobre I'.
- De $DCT(I')$, quedarse con el segmento superior izquierdo de tamaño 8x8. Se llamará a esta imagen I''.
- Calcular la mediana de los valores en I'', avg.
- Para cada pixel ${p''}_{xy}$ de I'', el valor de hash final H en la posición (x, y) se construye de la siguiente manera: $h_{xy} = 1$ si ${p''}_{xy} > avg$, $h_{xy} = 0$ en caso contrario.

La DCT es una transformación relacionada con la transformada de Fourier que permite expresar señales de manera compacta. Al igual que la transformada de Fourier, transforma la imagen del dominio espacial al dominio de la frecuencia. La DCT tiene una gran capacidad de almacenar información en relación al bajo costo computacional que exige su cálculo (comparado con otras transformaciones al dominio de la frecuencia). Por esta razón, se utiliza para la compresión con pérdida en JPG. Esta misma capacidad se aprovecha en el cálculo de este hash para representar la estructura general de la imagen, con pocos coeficientes que en última instancia son transformados a los bits del hash final.

Al igual que el Difference Hash, no tiene ninguna desventaja propia, más allá de las debilidades comunes que tienen estos tres algoritmos. Durante algún tiempo se discutía que el cálculo de la DCT resultaba en un impacto notable en la velocidad de cálculo, sin embargo, se verá que en la implementación realizada por los autores todos los algoritmos son equivalentes en velocidad de ejecución.

3 Implementación y resultados

La implementación realizada de los algoritmos se encuentra disponible en un repositorio de GitHub de los autores[16]. Para la misma, se utilizó el lenguaje de programación Python, utilizando las librerías Numpy[17], SciPy[18] y OpenCV[19]. A modo ilustrativo, se incluye el código final de la implementación del Difference Hash:

```
def d_hash(source):
    y = cv2.cvtColor(source, cv2.COLOR_BGR2GRAY)
    y = cv2.resize(y, (9, 8))
    hash_ = y[:, 0:8] < y[:, 1: 9]
    return hash_.astype(np.uint8)
```

Con la forma en que se ha elegido representar los hashes, el cálculo de la distancia de Hamming se simplifica a una función que calcula el bitwise-XOR entre dos hashes

perceptuales, y luego la sumatoria de elementos sobre el arreglo resultante. Dicha función puede expresarse en Python de la siguiente manera:

```
def hamming(h1, h2):
    return np.sum(h1 ^ h2)
```

De esta forma, si los elementos de hash coinciden, contribuyen con 0 a la sumatoria, y si son distintos contribuyen con 1, siendo esta la definición de la distancia de Hamming.

Como experimento, se realizó una búsqueda sobre un volcado de WhatsApp compuesto por 8628 imágenes. La búsqueda estaba orientada a encontrar imágenes visualmente similares a alguna de las presentes en una base de 40 imágenes de interés. En repetidas pruebas, se determinó de manera experimental los umbrales adecuados que permiten detectar modificaciones sobre las imágenes base para cada algoritmo de hash perceptual utilizado.

Hash	Umbral	Dectadas (Falsos positivos)	Sin detectar
Average	3	6 (2)	7
Difference	6	10 (1)	3
DCT	12	7 (0)	4

Tabla 1. Resultados de búsqueda con hashes perceptuales en el conjunto de prueba.

Algunas consideraciones para tener en cuenta:

- Un total de 11 imágenes que deberían ser encontradas. Estas imágenes son modificaciones de 6 imágenes base del conjunto de interés.
- Tanto Average Hash como Difference Hash generaron falsos positivos. De haber bajado más su umbral de detección, se habrían disminuído los falsos positivos generados, pero se habrían dejado de detectar imágenes de interés.
- En el caso de DCT-Hash el umbral parece relativamente alto (en comparación con los otros), también está en un delicado equilibrio. De aumentarlo, se comienzan a generar falsos positivos sin lograr encontrar las 3 imágenes faltantes del conjunto de interés. En caso de disminuirlo, se dejarían de detectar algunas de las imágenes que generan un match.
- Una de las imágenes de interés no es detectada a menos que los umbrales de detección se eleven por encima del nivel que en generan una alta tasa de falsos positivos.

Con respecto a la velocidad de cálculo de las funciones, no se encontraron mayores diferencias entre una y otra. Tomando como ejemplos una imagen pequeña de 160x160, una imagen mediana de 2400x1400 y una imagen grande de 4608x3456, del cálculo repetido de cada función de hash se obtuvieron los siguientes tiempos:

Hash	Pequeña	Mediana	Grande
Average	72.8 µs ± 4.41 µs	4.68 ms ± 106 µs	25.9 ms ± 2.35 ms
Difference	103 µs ± 2.8 µs	11.7 ms ± 274 µs	25.6 ms ± 3.22 ms
DCT	214 µs ± 42.9 µs	12.9 ms ± 1.89 ms	24 ms ± 2.08 ms

Tabla 2. Tiempos de cálculo de las funciones de hash perceptual sobre una imagen pequeña, mediana y grande. Obtenido con el comando %timeit de IPython.

Las mediciones se realizaron en Python 3.6.3 de 64 bits, bajo Windows 10, y las librerías Numpy, SciPy y OpenCV compiladas con la Intel Math Kernel Library, en un equipo de pruebas con un procesador Intel Core i5 7200U y 8GB de memoria RAM, utilizando el comando %timeit de IPython[20].

Analizando el rendimiento en cada caso, se concluye que la operación que más tiempo lleva en cada una de las funciones de hash es el redimensionado.

Es notable que, pese a lo que indica la literatura previa[13, 14], no se ha observado una diferencia notable en el rendimiento del DCT-Hash respecto al Average Hash o Difference Hash una vez superado un cierto umbral de tamaño (ver columna "Grande" en la Tabla 2). Para imágenes de tamaño medio, el Average Hash presenta una cierta ventaja con respecto a los otros dos, pero su rendimiento es pobre para detectar variaciones de una imagen base.

4 Conclusiones y trabajo futuro

La experiencia de implementación de los algoritmos y de las pruebas realizadas se ha llegado a las siguientes conclusiones:

- La utilización de hashes perceptuales permite realizar un análisis de alta velocidad en busca de imágenes específicas. Esta búsqueda es resistente a modificaciones de imagen, con distinto nivel de robustez dependiendo tanto del algoritmo utilizado, como del umbral de distancia que se decida utilizar.
- Los umbrales de detección deben elegirse teniendo en cuenta el delicado equilibrio que se debe alcanzar si se pretende buscar variantes de las imágenes que hayan sufrido modificaciones. Al contrario, si solamente se buscan matches totales, puede utilizarse un umbral de 0.
- Este trabajo se enfocó en las imágenes en RGB con 8 bits de color por canal. Sería útil ampliar la problemática para considerar las imágenes almacenadas en RGBA, y analizar las posibilidades y problemáticas que se abren en este caso.
- ImageHash tiene una variante del Difference Hash que calcula los gradientes verticales. Es posible implementar esta variante en el desarrollo realizado, para agregar una herramienta más que permita identificar de manera unívoca una imagen en base a sus hashes perceptuales.
- Los hashes se han planteado con un tamaño fijo en las dimensiones de imagen que, en todos los casos, resulta en 64 bits, que pueden almacenarse de manera compacta en un entero largo. Es posible extender los hashes para que las imágenes tengan otros tamaños, que resulten en hashes de más bits, por ejemplo, imágenes intermedias de 16x16 resultarían en hashes de 256 bits.
- La utilización de hashes perceptuales abre la puerta a generar bases de datos propias de las instituciones nacionales, que permitan reconocer contenido multimedia ya identificado en investigaciones realizadas por las fuerzas de seguridad y los investigadores judiciales. Para esto, es necesario continuar el trabajo de estudio de estas

técnicas, definir los alcances y las posibilidades que brindan, y también las estrategias de mitigación para los falsos positivos. Ese trabajo permitirá implementar bases de datos de contenido sensible basadas en los hashes perceptuales, complementaria de las bases de datos de hashes criptográficos que hoy se utilizan.

Analizando el rendimiento en cada caso, se concluye que la operación que más tiempo lleva en cada una de las funciones de hash es el redimensionado.

Es notable que, pese a lo que indica la literatura previa[13, 14], no se ha observado una diferencia notable en el rendimiento del DCT-Hash respecto al Average Hash o Difference Hash una vez superado un cierto umbral de tamaño (ver columna "Grande" en la Tabla 2). Para imágenes de tamaño medio, el Average Hash presenta una cierta ventaja con respecto a los otros dos, pero su rendimiento es pobre para detectar variaciones de una imagen base.

5 Agradecimientos

Los autores desean agradecer a la Universidad FASTA, al Ministerio Público de la Provincia de Buenos Aires y al Municipio de General Pueyrredon por brindar un espacio único de trabajo como es el InFo-Lab.

En especial, desean agradecer a Ayrton Betti y Axel Ziegler, estudiantes que realizaron sus práctica profesionales en el Laboratorio participando de este proyecto, a Mónica Pascual, secretaria de investigación de la Facultad de Ingeniería de la Universidad FASTA, y a Roberto Giordano Lerena, decano de la Facultad de Ingeniería.

Referencias

1. Ministerio Público Fiscal de la Ciudad Autónoma de Buenos Aires: Operación Luz de Infancia: golpe internacional contra la pornografía infantil. (Noticia). Disponible online en: https://www.fiscalias.gob.ar/project/operacion-luz-de-infancia-golpe-internacional-contra-la-pornografia-infantil/.
2. Diario Clarín: Mega operativo contra la pornografía infantil: secuestran casi un millón de archivos con videos y fotos. (Noticia). Disponible online en: https://www.clarin.com/sociedad/operativo-luz-infancia-secuestran-millon-archivos-pornografia-infantil_0_b2QxGiDz_.html.
3. Whelpton, J.: The Psychological Effects Experienced by Computer Forensic Experts Working with Child Pornography. Masters Thesis. University of South Africa. Febrero 2012. Disponible online en: https://core.ac.uk/download/pdf/43169033.pdf.
4. Rivest, R.: The MD5 Message-Digest Algorithm. RFC 1321. Abril 1992.
5. Federal Information Processing Standards Publication: Secure Hash Standard. FIPS PUB 180-1. Abril 1995.
6. Di Iorio, A. H. et al: El Rastro Digital del Delito. Universidad FASTA Ediciones. Marzo 2017.

7. Gonzalez, R. C., Woods, R. E: Digital Image Processing. 3ra Edición. Pearson Prentice Hall. 2008.
8. CCITT International Telegraph and Telephone Consultative Committee: INFORMATION TECHNOLOGY – DIGITAL COMPRESSION AND CODING OF COTINUOUS-TONE STILL IMAGES – REQUIREMENTS AND GUIDELINES. 1993. Disponible online en: https://www.w3.org/Graphics/JPEG/itu-t81.pdf.
9. Boutell, T., Lane, T., et al: Portable Network Graphics (PNG) Specification and Extensions. Disponible online en: http://www.libpng.org/pub/png/spec/.
10. Documentación de OpenCV 4.1: Color conversions. Disponible online en: https://docs.opencv.org/4.1.0/de/d25/imgproc_color_conversions.html.
11. Zauner, C.: Implementation and Benchmarking of Perceptual Image Hash Functions. Master's thesis, Upper Austria University of Applied Sciences, Hagenberg Campus. 2010.
12. Hamming, R.: Error detecting and error correcting codes. The Bell System Technical Journal. Volúmen 29, Número 2. Abril 1950.
13. Krawetz, N.: Looks like it. (Blog personal). Disponible online en: http://www.hackerfactor.com/blog/index.php?/archives/432-Looks-Like-It.html.
14. Krawetz, N.: Kind of Like That. (Blog personal). Disponible online en: http://www.hackerfactor.com/blog/index.php?/archives/529-Kind-of-Like-That.html.
15. Buchner, J.: ImageHash. Software libre, disponible online en: https://github.com/JohannesBuchner/imagehash.
16. Constanzo, B.: phantom. Software libre, disponible online en: https://github.com/bconstanzo/phantom.
17. Oliphant, T.: Python for Scientific Computing. Computing Science and Engineering. Volúmen 9, Número 3. Junio 2007.
18. Jones, E., Oliphant, T., Peterson, P.: SciPy: Open source scientific tools for Python. http://www.scipy.org/. 2001.
19. Bradski, G.: The OpenCV Library. Dr. Dobb's Journal of Software Tools. Noviembre 2000.
20. Pérez, F., Granger, B. E.: IPython: A System for Interactive Scientific Computing. Computing in Science and Engineering. Volúmen 9, Número 3. Mayo-Junio 2007.

Análisis del Marco Normativo de la Protección de los Datos Personales a Aplicarse en la Argentina en Proyectos de Cloud Computing Implementados en el Exterior del País

Juan González Allonca[1], Cintia Gioia[1], Jorge Eterovic[1], Mario Krajnik[1], Walter Ureta[1], Sergio Conde[1], Sergio Bonavento[1] and Santiago Igarza[1]

[1] Universidad Nacional de La Matanza (UNLaM), Departamento de Ingeniería e Investigaciones Tecnológicas, Florencio Varela 1903, B1754JEC San Justo, Buenos Aires, Argentina
{gonzalezallonca, cgioia, eterovic, mkrajnik, wureta, sconde, sbonavento, asigarza}@unlam.edu.ar

Resumen. Al momento de iniciar un proyecto de cómputo en la nube (cloud computing) es determinante adecuarse a la normativa local y a su vez, analizar las cláusulas relativas a la seguridad de la información, especialmente a la protección de los datos personales. Existe legislación aplicable que determina la extensión de responsabilidad de usuario y proveedor. Este estudio se propone presentar un proceso de análisis que permita describir y evaluar las regulaciones aplicables en la Argentina relacionadas con servicios de cómputo en la nube en el exterior del país, como la transferencia internacional de datos personales y la prestación por cuenta de terceros de servicios de tratamiento de datos personales. El proceso de análisis propuesto logra identificar y valorar el grado de cumplimiento con la normativa local, lo que facilitará la toma de decisiones informadas, basadas no sólo en criterios técnicos o económicos, sino también regulatorios. A su vez, identifica los principales ejes donde se deben abordar pericias en contextos de cómputo en la nube.

1 Introducción

En los últimos años, gran cantidad de empresas se han visto atraídas por las ventajas técnicas y los bajos costos de mantenimiento que ofrece el esquema de cómputo en la nube [1]. Flexibilidad, accesibilidad, autoservicio bajo demanda, escalabilidad, gestión de grandes volúmenes de datos, son algunos beneficios que ofrece este esquema. Sin embargo, estas ventajas muchas veces no contemplan cuestiones críticas como la seguridad de la información, cumplimiento normativo y privacidad de los datos [2].

Actualmente, la información es el activo más importante de las organizaciones [3], por lo que asegurar la privacidad de la información durante su ciclo de vida es crucial a la hora de utilizar estos servicios.

El desconocimiento o la no aplicación de la normativa vigente pueden transformarse tanto en pérdida de confianza o daño en la imagen de una empresa o perjuicio económico como en responsabilidades jurídicas [4][5]. Las preocupaciones por estos inconvenientes, por lo general, son lo suficientemente importantes para algunas empresas y organizaciones, tanto que las llevan a evitar implementar sus sistemas en arquitecturas de cómputo en la nube.

En el mundo conviven múltiples legislaciones relacionadas con la transferencia internacional [6], lo que dificulta establecer una estrategia global en términos de la utilización de servicios de cómputo en la nube.

Como señala Etro [7], en un informe realizado por el Foro Económico Mundial en 2010, en el que se consultaba al sector industrial, gobiernos y académicos respecto de los principales obstáculos para la adopción de servicios cloud, sus respuestas se concentraban en tres cuestiones de localización de los datos: privacidad, confidencialidad y las relacionadas con la propiedad y los derechos de los datos en la nube.

Por este motivo, a partir del presente estudio se define un proceso de análisis que permite a las empresas u organismos locales describir y evaluar la reglamentación vigente referida a la protección de datos personales en proyectos de cómputo en la nube en el exterior del país. Este proceso de análisis posibilita verificar el grado de cumplimiento con la normativa vigente, sumando el aspecto regulatorio a los análisis de viabilidad de un proyecto de cómputo en la nube.

2 El Modelo de Cloud Computing

Hablar de Cloud Computing es presentar un concepto de servicios de cómputo por demanda [8]. Se trata de un nuevo esquema en el uso de los recursos de tecnológicos y de sus modelos de consumo y distribución [9]. Este modelo presenta un cambio importante en el paradigma computacional actual, la transformación de la infraestructura y las aplicaciones, de un mundo claramente dominado y administrado por las organizaciones, a otro donde un tercero confiable y conocido le brinda servicios de infraestructura y uso de aplicaciones [9].

La Cloud Security Alliance (CSA) es la Guía para la Seguridad en áreas críticas de atención en cloud computing y describe cinco características esenciales que evidencian similitudes y diferencias con las estrategias de computación tradicionales:

- Autoservicio por demanda. Un consumidor puede abastecerse unilateralmente de tiempo de servidor y almacenamiento en red, según sus necesidades, de forma automática sin requerir la interacción humana con cada proveedor de servicios.
- Amplio acceso a la red. Las capacidades están disponibles en la red y se accede a ellas a través de dispositivos estándar (p.ej., PC, teléfonos móviles y tablets).

- Reservas de recursos en común. Los recursos, como por ejemplo el almacenamiento, el procesamiento o la memoria del proveedor, son compartidos y pueden ser utilizados por múltiples clientes. Estos recursos son asignados dinámicamente y reasignados en función de la demanda de los consumidores. El cliente, por lo general, no tiene control o conocimiento exacto sobre la ubicación de los recursos. Usualmente, el proveedor no revela el lugar, aunque se puede especificar una ubicación genérica, como región o país.
- Rapidez y elasticidad. Las capacidades pueden suministrarse de manera rápida y elástica, en algunos casos, de manera automática, para poder realizar el redimensionado correspondiente rápidamente. Para el consumidor, las capacidades disponibles para abastecerse a menudo aparecen como ilimitadas y pueden adquirirse en cualquier cantidad y en cualquier momento.
- Servicio supervisado. Los sistemas de nube controlan y optimizan el uso de los recursos de manera automática, utilizando una capacidad de evaluación en algún nivel de abstracción adecuado para el tipo de servicio (p.ej., almacenamiento, procesamiento, ancho de banda, y cuentas de usuario activas).

3 Descripción del Problema

Como quedó demostrado, la implementación de servicios de cómputo en la nube ofrece múltiples ventajas, tanto desde un enfoque técnico (flexibilidad, accesibilidad, autoservicio bajo demanda, escalabilidad, gestión de grandes volúmenes de datos, etc.) como económico (bajos costos de implementación y mantenimiento, facturación por demanda, entre otros). Sin embargo, una de las grandes dificultades que se presentan a la hora de implementar estos servicios es de índole legal, más precisamente, cuando se transfieren datos personales de un país a otro para luego aplicarles un proceso informático [10][11][12][13][14][15].

Por lo tanto, al momento de iniciar un proyecto de cómputo en la nube, es necesario adecuarse a la normativa local y analizar las cláusulas relativas a la seguridad de la información, especialmente a la protección de los datos personales. Existe legislación aplicable que determina la extensión de responsabilidad, tanto del cliente, como del proveedor de servicios de cómputo en la nube. Aunque, aún no existe como práctica generalizada realizar un análisis previo en este sentido, donde se le permita al usuario conocer su nivel de riesgo y de cumplimiento normativo. Según Enrique Larrieu-Let (2013), presidente del Instituto de Auditores Internos de Argentina, “actualmente se carece de un marco de trabajo específico estructurado y completo para la identificación y evaluación de riesgos en Cloud Computing, es decir, la panacea aún no existe”.

A su vez, Mather [16] describe los contradictorios puntos de vista y nociones existentes en distintos países sobre los derechos a la privacidad y la protección de los datos personales, lo que genera múltiples batallas legales, disputas políticas y regulaciones conflictivas. Algunos ejemplos de regulaciones en tensión que presentan los auto-

res son las Reglas Federales de Procedimiento Civil de EE. UU. (FRCP) y la Directiva de la Unión Europea sobre protección de datos personales.

Debido a que existen múltiples proveedores en distintos países, sumado a diferentes modelos y formas de despliegue de cómputo en la nube, las metodologías actuales no definen una serie de pasos a seguir para realizar un proceso de análisis que permita describir y evaluar la reglamentación local vigente referida a la protección de datos personales en proyectos de cómputo en la nube. A la falta de un proceso de análisis regulatorio se le suman riesgos propios de ese modelo de negocio, que podrían generar responsabilidades legales tanto en el país de origen como en el de destino.

En este contexto, el proceso de análisis propuesto logra identificar y valorar el grado de cumplimiento con la normativa local, lo que facilitará la toma de decisiones informadas, basadas no sólo en criterios técnicos o económicos, sino también regulatorios. A tales fines, resulta pertinente determinar las garantías y requisitos necesarios para proteger adecuadamente los datos personales que se transfieran a países sin legislación adecuada en los términos del artículo 12 del Anexo I al Decreto N° 1558/01.

El proceso permite efectuar un análisis de la normativa y de los riesgos en la implementación de un servicio de cómputo en la nube, que permite controlar la aplicación del derecho fundamental a la protección de datos de los titulares. Luego del análisis, es posible contar con un registro directo de los incumplimientos normativos identificados y promueve la adopción de las medidas necesarias para eliminarlos o mitigarlos.

Uno de los principales beneficios de esta metodología radica en que, a partir de su realización, en las etapas iniciales de la implementación de un servicio de cloud computing se logran identificar los posibles riesgos y corregirlos anticipadamente, evitando los costos y complicaciones derivados de descubrirlos a posteriori, cuando el servicio está en funcionamiento o, lo que es peor, cuando la lesión de los derechos se ha producido, lo que implica pérdidas económicas y también daños a la imagen para la organización, cuya reputación se ve afectada.

A su vez, la ejecución de este proceso otorga transparencia a la gestión de los datos personales, base de una relación de confianza entre su titular y el usuario de ellos. Esto permite planificar las medidas ante posibles impactos en la privacidad y gestionar los vínculos con terceras partes implicadas en la transferencia de datos, otorgando más garantías para ellas.

4 Marco Legal: Legislación y Jurisdicción Aplicable

¿Cuál es la importancia de la privacidad y por qué la legislación argentina la protege? Es decir, ¿de dónde surge la necesidad de tomar medidas técnicas para su protección? La privacidad es un derecho humano fundamental y se encuentra receptado en tratados internacionales, leyes, disposiciones y jurisprudencia. Es el derecho que pro-

tege la libertad individual y de expresión, la intimidad y la dignidad personal, e incluye el derecho a la protección de datos personales y la figura del Habeas Data.

Ahora bien, ¿cuál es la relación que existe entre privacidad, protección de datos y habeas data? "De manera general, se puede decir que la protección a la privacidad es el género y la protección de datos la especie. Y todavía en un sentido más estricto queda la figura de habeas data, la cual se opera como un derecho de acceso a la información personal dentro del régimen de datos personales" [17].

El derecho a la privacidad se sustenta en principios fundamentales como el honor y la dignidad personal. Como lo afirma la Secretaría de Asuntos Jurídicos, OEA (2012), "el derecho a la privacidad va más allá de la protección de datos, abarca el respeto de la vida familiar, preferencias religiosas, políticas y sexuales, la intervención de las comunicaciones, el uso de cámaras ocultas, los análisis genéticos, etc. La protección de la vida privada y la protección de la intimidad son necesarias para el orden jurídico y como garantía de respeto a la dignidad personal".

La protección de datos es un derecho a la intimidad personal que tienen las personas contra un tratamiento incorrecto, no autorizado o contrario a las normativas vigentes de sus datos personales por tratadores de datos. Al proteger los datos personales frente al riesgo de la recopilación y el mal uso de sus datos personales, se ampara, en consecuencia, la privacidad de las personas.

Dentro del derecho de protección de datos personales, como se muestra en la Figura 1, se encuentra la acción de Habeas Data. Se trata de un recurso legal mediante el cual las personas agraviadas pueden informarse sobre datos referidos a ellos y el propósito de su recolección. A su vez, permite exigir, dependiendo el caso, su rectificación, actualización o supresión de información personal alojada en bancos o registros de datos, públicos o privados.

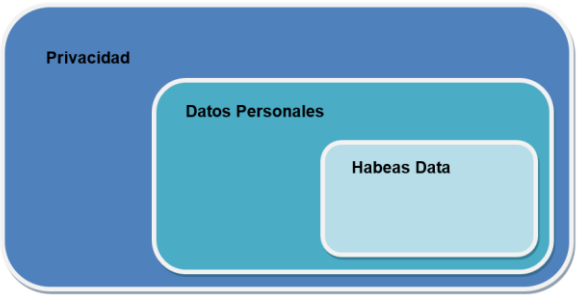


Fig. 1 Relación entre Privacidad, Protección de Datos y Habeas Data

Nuestro país cuenta con una amplia tradición en materia de protección de datos personales, que se manifiesta en tres niveles distintos. En el primer nivel, se encuentra la Constitución Nacional que, luego de su reforma en el año 1994, incluyó el artículo

43 que, en su párrafo tres, contempla el llamado habeas data, de la siguiente forma: Toda persona podrá interponer esta acción para tomar conocimiento de los datos a ella referidos y de su finalidad, que consten en registros o bancos de datos públicos, o los privados destinados a proveer informes, y en caso de falsedad o discriminación, para exigir la supresión, rectificación, confidencialidad o actualización de aquellos. No podrá afectarse el secreto de las fuentes de información periodística.

Como se advierte, esta reforma de la Constitución Nacional ha establecido un instituto que carecía de antecedentes en el derecho federal, aunque ya se encontraba en las constituciones provinciales: la acción de habeas data [6]. Se trata de un procedimiento especialmente necesario a partir del aumento del uso de las computadoras, que pueden compilar la información y datos personales afectando el honor y la privacidad de las personas [4]. La acción también está establecida para tomar conocimiento de estos datos y, en su caso, exigir la supresión, rectificación, confidencialidad o actualización.

El segundo nivel está representado por la Ley N° 25.326 de Protección de los Datos Personales, que tiene por objeto la protección integral de los datos personales asentados en archivos, registros, bancos de datos, u otros medios técnicos de tratamiento de datos, sean estos públicos, o privados destinados a dar informes, para garantizar el derecho al honor y a la intimidad de las personas, así como también el acceso a la información que sobre las mismas se registre. A su vez, el Poder Ejecutivo reglamentó dicha ley por medio del decreto N° 1558/01, en el que se crea la Dirección Nacional de Protección de Datos Personales, que es el órgano de control de la ley, primero en América Latina y el tercero del hemisferio sur.

En tercer nivel, está la interpretación y la aplicación que hacen los jueces de estas normas. A partir de este desarrollo legislativo, Argentina fue declarada país adecuado por la Unión Europea en materia de Protección de Datos Personales, de conformidad con la Directiva 95/46/CE [18].

5 Fases del Proceso de Análisis de Cumplimiento Normativo

El modelo de un procedimiento de análisis que permite describir y evaluar la reglamentación local vigente referida a la protección de datos personales en proyectos de cómputo en la nube, a través de proveedores en el exterior del país, que busca encontrar una solución al problema planteado, con el fin de identificar y valorar el grado de cumplimiento con la normativa local, lo que facilita la toma de decisiones informadas, basadas no sólo en criterios técnicos o económicos, sino también regulatorios.

Esta herramienta metodológica de evaluación de cumplimiento normativo permite evaluar el grado de impacto en la protección de datos personales para servicios de cloud computing, lo que permite tomar las medidas necesarias para evitar o minimizar los impactos negativos.

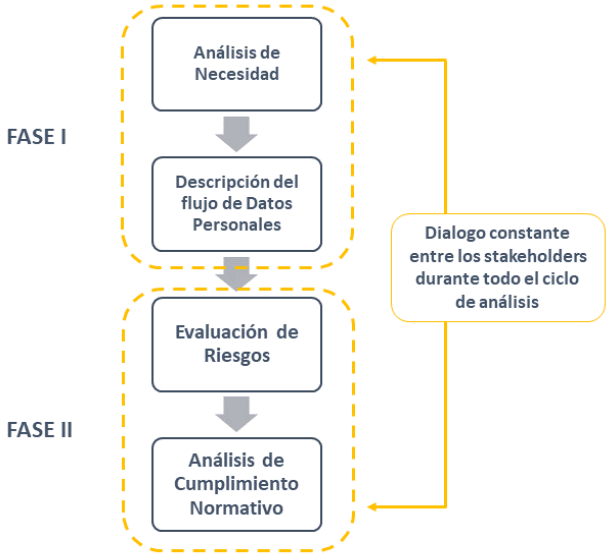


Fig. 2 Etapas del Proceso de Análisis de Cumplimiento Normativo.

El uso de servicios de cómputo en la nube permite una amplia variedad de beneficios para los usuarios en términos de agilidad, movilidad y reducción de costos vinculados a los recursos de procesamiento por lo que el procesamiento de datos sin duda crecerá [19].

Dado que es difícil imaginar una organización que no tenga una cierta cantidad de datos personales (relacionados con los empleados, por ejemplo), es probable que externalizarlos sea un potencial obstáculo para el procesamiento de datos en la nube. Por lo tanto, para asegurarse de que esta actividad no colisiona con las normas locales, una organización tendrá que determinar si un proveedor de servicios en la nube procesará sus datos personales ajustado a derecho.

Una solución a este problema es la aplicación de una metodología que permita describir y evaluar la reglamentación local vigente referida a la protección de datos personales en proyectos de cómputo en la nube y la evaluación de riesgos en materia de protección de datos personales para servicios de cloud computing. A través de esta metodología, que consiste en un enfoque analítico para mejorar la gestión del tratamiento de datos personales, será posible identificar la norma aplicable, los riesgos y establecer un nivel de criticidad sobre ellos. La aplicación de esta metodología se aplica de forma simultánea y complementaria con la legislación vigente que los responsables del tratamiento de datos deben cumplir.

Con este procedimiento se intenta brindar a los responsables del tratamiento de información de carácter personal y, a su vez, permitirles:

- Tener un punto de vista racional de los riesgos derivados del procesamiento de datos personales en el exterior del país;
- Determinar medidas de seguridad necesarias y suficientes con el fin de "adoptar las medidas técnicas y organizativas que resulten necesarias para garantizar la seguridad y confidencialidad de los datos personales, de modo de evitar su adulteración, pérdida, consulta o tratamiento no autorizado, y que permitan detectar desviaciones, intencionales o no, de información, ya sea que los riesgos provengan de la acción humana o del medio técnico utilizado. (artículo 9, Ley N° 25.326, 2000).

La propuesta de proceso aquí planteada facilita un marco común a organizaciones tanto del ámbito público como privado, que procesen datos personales en la nube fuera del país, independientemente de si se encuentran en un país con legislación adecuada o no.

Este proceso establece una serie de medidas que deben implementarse en conjunto con la normativa local en materia de protección de datos, por ello, quienes implementen servicios de cloud computing, no sólo deberán analizar si éste se adecúa a la norma, sino también, si por el carácter o tipo de información que se trata, no son necesarias medidas adicionales de acuerdo con la normativa argentina.

Fase I - Definición de Conveniencia y Lineamientos del Proyecto: En la primera fase del proceso se realizan las acciones necesarias para establecer las bases del proyecto, identificar las partes involucradas, evaluar el impacto de la implementación de un servicio de cómputo en la nube en la organización y analizar el flujo de datos personales.

Fase II - Evaluación de Riesgos y Cumplimiento Normativo: La evaluación de riesgos se utiliza en diversas áreas (seguridad aérea, finanzas, seguridad física, etc.) Esta metodología se centra en la seguridad de la información, más precisamente, en los riesgos vinculados con los datos personales.

En el tratamiento de datos es posible identificar como riesgos aquellos provenientes del tratamiento de datos de carácter personal. Dichos riesgos están compuestos por un incidente determinado y todas aquellas amenazas que lo pueden hacer posible; cómo estos incidentes pueden hacerse realidad.

Es posible estimar los riesgos en el tratamiento de protección de los datos personales en términos de severidad (la magnitud del riesgo) y probabilidad. En el caso de cómputo en la nube, esencialmente depende del nivel de identificación de los datos personales y el nivel de consecuencia de los potenciales impactos.

La probabilidad representa la factibilidad de que un hecho pueda ocurrir. En esencia, depende del nivel de vulnerabilidad de los factores de soporte frente al nivel de capacidad para explotarlo de las fuentes de riesgo.

6 Principales ejes a donde se deben abordar pericias en contextos de cómputo en la nube

La adquisición y el tratamiento de la evidencia digital son actividades estratégicamente decisorias, al momento de generar pruebas informáticas que permitirán dirimir situaciones dudosas y los respectivos autores o culpables. Es importante considerar las nuevas alternativas y problemática que se generan al abordar pericias en contextos de cómputo en la nube, lo cual se centra en el debido conocimiento y control de la gestión en la nube, el cumplimiento contractual entre el prestador y el cliente, la disponibilidad del servicio en la nube y la confiabilidad, seguridad y confidencialidad sobre los servicios [20].

Actualmente existe un vacío normativo en relación con el tratamiento de la evidencia digital en servidores externos que puede ser causa de que la evidencia no sea aceptada en una instancia judicial. La evidencia digital en la nube plantea nuevos ejes técnicos y legales que deben considerarse:

En términos legales:

- Conocer las responsabilidades legales tanto en el país de origen como en el de destino. Considerar la legislación y jurisdicción aplicable que determina la extensión de responsabilidad, tanto del cliente, como del proveedor de servicios de cómputo en la nube.
- Aplicar la metodología de Análisis de Cumplimiento Normativo (descripta en la sección anterior) de forma reactiva, de manera que permita describir y evaluar el grado de cumplimiento de la reglamentación local vigente referida a la protección de datos personales en proyectos de cómputo en la nube y así delimitar la investigación forense a realizar.

En términos técnicos:

- La evidencia digital está ubicada en un servidor externo de un tercero y sólo es accesible a través de este.
- El tercero puede ser conocido o no, estar en la misma jurisdicción o en otra diferente, aunque mayormente es en el extranjero.
- Tipos de nubes: públicas (recursos de la nube propiedad de terceros), privadas (recursos informáticos exclusivos de una organización) o híbridas.
- Existen múltiples proveedores en distintos países, con diferentes modelos y formas de despliegue de cómputo en la nube.
- La mayoría de los proveedores de servicios en la nube sólo entregan información mediante orden judicial, lo cual implica que las solicitudes de acceso a evidencia digital remota deben ser realizadas por el Juzgado.

- Los proveedores de servicios poseen sus propias políticas de entrega de información a las autoridades, junto con diversas condiciones unilaterales impuestas por los mismos, como ser los períodos de retención establecido por sus términos y condiciones y las formas en que deben ser realizados los pedidos de información, sean de conexión, tráfico o de contenido.

En Argentina, la recolección de evidencia en la nube suele realizarse de dos formas. La primera consiste en la solicitud de conservación de datos al proveedor de servicios. Dicha petición suele ser aceptada por los proveedores de servicios en la nube, si la misma es efectuada por un oficial de las fuerzas de la ley, desde una dirección de correo oficial de la fuerza a la que reporta. La segunda forma es la obtención de datos, la cual puede requerir una orden judicial o un exhorto diplomático. Todo este proceso está fuertemente influenciado por los términos y condiciones de cada empresa que presta el servicio, los cuales suelen establecer la forma en que deben ser hechos los pedidos [21].

Existen diferentes escenarios de discusión que con el crecimiento de la computación en la nube se hace esencial que sean tratados, como ser:

- Recolección remota de evidencia ubicada en el servidor externo, previa autorización del Juez.
- Obtención de evidencia digital recopilada desde perfiles públicos de redes sociales o diversas fuentes abiertas, tratadas como información pública.
- Obtención de evidencia digital mediante orden de allanamiento, en donde se verifica si se utiliza la facilidad de acceso a la nube para almacenar datos. En caso afirmativo, se deberá tratar de obtener las características de acceso a tal información, a los efectos que la misma sea "bajada¨ a un dispositivo, para luego ser secuestrado y posteriormente analizado [20], previa orden del Juez.

6 Conclusiones

Como conclusión del presente trabajo, se desprende que, al momento de iniciar un proyecto de Cloud Computing, no sólo deben evaluarse variables relativas a la rentabilidad, capacidad tecnológica y ventaja de negocios, sino también analizar el cumplimiento normativo y las cláusulas sobre seguridad de la información, especialmente las relativas a la protección de los datos personales. Aplicar la legislación local en materia de protección de datos personales le permite al usuario de servicios de cómputo en la nube conocer la extensión de su responsabilidad y la de su proveedor ante un eventual incidente. De este modo, el usuario podrá valorar qué delega en este modelo y qué cuestiones prefiere reservarse, pudiendo tomar una decisión basada en información concreta.

Por otro lado, este trabajo presenta las bases de una metodología de evaluación de riesgos en materia de protección de datos personales en servicios de cómputo en la

nube que permitirá cuantificar la magnitud de los riesgos existentes y, en consecuencia, jerarquizar racionalmente su prioridad de corrección.

En términos de pericias en contextos de cómputo en la nube, se plantean nuevos paradigmas, desafíos y escenarios que obliga a los peritos informáticos a capacitarse y prepararse para manejar nuevos aspectos no solo técnicos sino legales en el tratamiento de la evidencia digital almacenada en la nube, de manera de evitar errores en los procedimientos que vuelvan inválida la prueba recolectada.

Referencias

1. Gartner, Mindy Cancila, Douglas., Toombs, Alan D Waite, y Elias Khnaser. «2017 Planning Guide for Data and Analytics.» 2017, Recuperado 3 de marzo 2019.
2. Rao, R, y K Valemadhava & Selvamani. «Data Security Challenges and Its Solutions in Cloud Computing.» Procedia Computer Science, nº 48, 2015: 204-209.
3. World Economic Forum, WEF. «Unlocking the Value of Personal Data: From Collection to Usage. Industry Agenda.» Editado por Geneva. Word Economic Forum, 2013: 7-9.
4. Peyrano, G. Régimen legal de los datos personales y el habeas data. Buenos Aires, Argentina. ISBN: 9501418626., Buenos Aires, Argentina: De-palma., 2002.
5. Palazzi, P. La protección de los datos personales en la Argentina. Buenos Aires, Argentina. Errepar, 2004.
6. Palazzi, P. «La transmisión internacional de datos personales y la protección de la privacidad.» Ad-Hoc. ISBN: 950-894-318-1, Buenos Aires, Argentina, 2002.
7. Etro, F. «The Economic Consequences of the Diffusion of Cloud Computing» en Dutta, Soumitra; Mia, Irene. The Global Information Technology Report 2009 – 2010 ICT for Sustainability.Foro Económico Mundial - INSEAD. Londres. 2010.
8. Mell, P, y T Grance. «The Nist Definition of Cloud Computing.» NIST Special Publication 800-145, National Institute of Standards and Technology, Department of Commerce, U. S., 2011, Recuperado 15 de marzo 2019.
9. Catteddu , D, y Hogben , G. «Cloud Computing - Benefits, and risks recommendationes for information security.» Enisa, 2009.
10. Anuar, N, Gani, A, Hashem, I.A, Khan, S.U., Mokhtar, S., y Yaqoob, I. The rise of "big data" on cloud computing: Review and open research issues. Vol. 47. Infrmation System, 2015.
11. Bernardino, J., Cámara, J., Neves, P.C., y Schmerl, B.R. Big Data in Cloud Computing: features and issues. 2016, recuperado 10 de marzo 2019.
12. Azer, M., y El.Zoghby, A. «Cloud computing privacy issues, challenges and solutions.» 12th International Conference on Computer Engineering and Systems (ICCES). 2017. 154 - 160.
13. Brodkin, J. «Gartner: Seven cloud-computing security risks.» Infoworld, 2008, 1 - 3.
14. Pavolotsky, J. Top Five Legal Issues for The Cloud. Forbes. Forbes, 2010, Recuperado 16 Marzo de 2019.
15. Pearson, S, y A Benameur. «Privacy, security and trust issues arising from cloud computing. Cloud Computing Technology and Science (CloudCom).» IEEE Second International Conference on IEEE, 2010.

16. Mather T, Kumaraswamy S, Latif S. «Cloud Security and Privacy» O'Reilly Media, Inc., Sebastopol, CA. 2009
17. Organización de los Estados Americanos (OEA). «Interrelación entre protección a la privacidad, protección de datos y habeas data.» Secretaria de Asuntos Jurídicos, 2012, Recuperado 4 de abril de 2019.
18. European Parliament and of the Council. «Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data.» 1995.
19. Orban, S. «The Fast and the Furious: How the Evolution of Cloud Computing Is Accelerating Buil-der Velocity. AWS Cloud Enterprise Strategy Blog.» Recuperado 17 de marzo 2019.
20. Piccirilli, Mg.Lic. Darío A. Protocolos a aplicar en la forensia informática en el marco de las nuevas tecnologías (pericia – forensia y cibercrimen). La Plata, Buenos Aires: Tesis Doctoral en Ciencias Informáticas - UNLP - Facultad de Informática, 2015.
21. Asociación por los derechos civiles. La investigación forense informática en América Latina. Vol. 2. ADC por los Derechos Civiles, 2018, Recuperado 20 de abril de 2019.

Guía técnica para el diseño de laboratorios judiciales de informática forense

Ana Haydée Di Iorio[1], Sabrina Lamperti[1], Lucía Coppes[2], Bruno Constanzo[1]

[1] InFo-Lab, Laboratorio de Investigación y Desarrollo de Tecnología en Informática Forense Universidad FASTA, Ministerio Público de la Provincia de Buenos Aires, [2] Municipalidad de General Pueyrredon Mar del Plata, Buenos Aires
[1] {diana, slamperti, bconstanzo}@ufasta.edu.ar, [2] luciacoppes@gmail.com

Abstract. El desarrollo de las tecnologías de la información y las comunicaciones ha traído como consecuencia un incremento en la cantidad de información digital, y la necesidad de utilizarla como evidencia es un reto creciente. La Informática Forense constituye una disciplina que pretende dar respuesta a una demanda cada vez mayor de los organismos de justicia, y va adquiriendo un lugar cada vez más preponderante en los institutos forenses.
Instalar un laboratorio de informática forense requiere conocer y considerar diversos aspectos claves, tanto desde el punto de vista estrictamente técnico como desde el punto de vista normativo, institucional, organizacional, estratégico, edilicio, tecnológico y de recursos humanos.
El producto "Laboratorio de Informática Forense" puede ser excelente de manera aislada, disponer del software apropiado, contar con el personal capacitado, cumplir con todas las previsiones arquitectónicas y de infraestructura, y, sin embargo, si no se adapta al contexto en el que está inserto desde el punto de vista estratégico e institucional, puede no resultar útil. Es decir, puede ser un "buen producto" pero una "mala solución" a las necesidades de la organización, y, en definitiva, de la ciudadanía. Por esta razón, deben definirse específicamente qué servicios se brindarán, quiénes serán los demandantes de los servicios y quiénes recibirán sus resultados.
Se presenta en este trabajo un conjunto de aspectos a considerar para el diseño y gestión de laboratorios de informática forense judiciales.

Keywords: Laboratorios Forenses, Informática Forense, Pericias Informáticas, Investigación criminal

1 Introducción

La investigación criminal consiste en la realización de acciones sistemáticas integradas para llegar al conocimiento de una verdad relacionada con el fenómeno delictivo, a través de un conjunto de saberes interdisciplinarios [3]. Comprende el manejo de estrategias que contextualizan la relación entre la víctima, el delincuente y el delito como tal; el estudio de las técnicas orientadas a contrarrestar, controlar y prevenir la acción delictiva; el dominio de la investigación como proceso metodológico que se basa en los principios y teorías de las respectivas ciencias, en los procedimientos jurídicos y la reconstrucción del hecho mediante las circunstancias de tiempo, modo y/o lugar para sustentar, en forma técnico-científica, los resultados conducentes al esclarecimiento de un presunto delito y a la identificación de sus autores [4].

En la Provincia de Buenos Aires, donde rige el sistema acusatorio, la actividad de investigación es actualmente llevada a cabo, bajo la dirección del Agente Fiscal, por personal de distintas fuerzas policiales [1], así como por Auxiliares Letrados, Secretarios o Instructores Judiciales del Ministerio Público Fiscal [2], a lo que se suma la labor de los peritos.

En este contexto, el desarrollo de las TIC's ha traído como consecuencia un incremento en la cantidad de información digital que se transmite y se almacena, por lo que la necesidad de utilizarla como evidencia es un reto creciente. El modelo de perito tradicional y las estructuras de laboratorios e institutos forenses hasta hoy conocidas no se ajustan a lo que hoy se requiere, por lo que esta realidad exige una revisión de las estructuras establecidas, tanto desde la infraestructura técnica y edilicia como desde lo organizacional.

Además, para representar un aporte con validez legal, la labor informático forense no debe dejarse librada a la improvisación del profesional, así como tampoco a procedimientos rígidos y rutinarios que no se adecúen a las variantes impuestas por los cambios tecnológicos. Al contrario, la aplicación forense de la informática requiere disponer de una infraestructura flexible y suficiente, procesos de trabajo adecuados, formación y actualización profesional, todo ello en el marco de un escenario sumamente cambiante y frecuentemente imprevisible.

2 Diseño de laboratorios de informática forense judiciales

Existen varios caminos para diseñar una organización o una nueva área dentro de ésta, como es el caso de un laboratorio de informática forense. Uno de los posibles cursos de acción es abstracto, basado en modelos ideales y/o inspirados en laboratorios de otras organizaciones que manifiestan un buen desempeño. Otras opciones están atadas a lo coyuntural, a las modas y/o a la improvisación. Ninguna de estas vías contribuye a obtener resultados sostenibles en el tiempo.

El otro camino, más arduo, pero también más productivo, comienza con la observación y análisis de la realidad de la organización en la que se insertará este laboratorio, como por ejemplo: demandas y necesidades actuales, características del entorno,

características de la institución madre, previsión de los posibles escenarios futuros, entre otras.

Como resultado de esta práctica, que debería ser parte de la cultura de una organización, surgen algunas preguntas clave: ¿Cuáles son las principales demandas y necesidades insatisfechas, actuales y futuras, a las que se enfrenta nuestra organización?, ¿de qué modo podrán ser satisfechas? para, finalmente, estar en condiciones de discutir y proyectar eficazmente el diseño de nuevas estructuras informático forenses.

Cabe destacar que, en este aspecto, no existen guías previas que establezcan los lineamientos sobre los cuales diseñar, implementar y llevar adelante laboratorios forenses en esta disciplina. Se ha dado un primer paso consistente en hacer visible la necesidad de consolidar los laboratorios forenses como estrategia frente a la investigación criminal por parte del poder judicial, y en especial, cuando es llevada adelante por los Ministerios Públicos. Así, se ha señalado que "existe un antes y un después en la investigación criminal, cuyo punto de inflexión ha sido la concreción del proyecto de laboratorios de investigaciones forenses, está claro que la ciencia y la tecnología se han instalado, también, en —todo— el interior del país; los ministerios públicos fiscales han variado sus paradigmas investigativos, han cambiado, para siempre, la cultura en la investigación criminal, dando, ahora sí, posibilidad de que se concrete o de cumplir con la obligación estatal de brindar investigaciones serias, conducentes y eficaces" [7].

3 Aspectos a considerar en el diseño organizacional de laboratorios de informática forense

Delinear la misión, visión y objetivos de una organización, es visualizar su razón de ser, y hacerla explícita [9-10] Los integrantes de una institución deben conocer el espíritu de la organización que integran, y su función dentro de ella. Para el caso de los laboratorios de informática forense, que responden a una estructura marco que los contiene, la misión, visión y objetivos deben estar en concordancia con los de la institución a la que responden.

El ámbito estatal encuentra los principios organizadores de su labor y estructuras tanto en la Constitución Nacional y Provincial como en las regulaciones orgánicas que regulan el funcionamiento de los entes estatales.

Por ejemplo, en la Provincia de Buenos Aires, "*El Ministerio Público es el cuerpo de Fiscales, Defensores Oficiales y Asesores de Incapaces que, encabezados por el Procurador General, actúa con legitimación plena en defensa de los intereses de la sociedad y en resguardo de la vigencia equilibrada de los valores jurídicos consagrados en las disposiciones constitucionales y legales*"[5].

Más específicamente, una de las ramas del Ministerio Público, es el Ministerio Público Fiscal o de la Acusación, que actúa en el área penal. De acuerdo con el portal del Ministerio Público bonaerense, "*Una de las ramas del Ministerio Público es la que conforman los Fiscales. Estos son los encargados de la persecución de los delitos y de la defensa de los intereses generales de la sociedad. Entre otras funciones*

los fiscales reciben denuncias, dirigen la investigación de los hechos criminales y son los encargados de llevar a juicio a los acusados por la comisión de delitos de acción pública.- No solo eso, a lo largo del proceso penal asisten y acompañan a las víctimas de delitos, a la vez que promueven la solución pacífica a los conflictos que se generan entre particulares a través de medios alternativos como la mediación penal, la suspensión del proceso a prueba y el principio de oportunidad"[6].

Los objetivos, en cambio, se orientan decididamente a la práctica en un determinado período. Partiendo de una situación presente, establecen qué debe hacerse para llegar a la situación futura deseada, asignando los recursos y medios que se emplearán para ello. Son los caminos a transitar para hacer realidad la visión y cumplir con la misión institucional.

No sólo basta con definir la misión, visión y objetivos, es preciso comunicar a los integrantes de la institución lo esperado y los procesos de medición de los resultados. El esclarecimiento de la misión, la elaboración de la visión y la definición de objetivos de una organización son procesos estratégicos replicables en los subsistemas que la integran.

Es importante considerar, en todos estos temas, las siguientes cuestiones que tienen especial impacto en la misión, visión y objetivos del laboratorio:

1. *Estructura funcional dentro del Organigrama.*
2. *Infraestructura y equipamiento.*
3. *Tareas administrativas y de gestión.*

Sumado a esto, es relevante tener presente algunas cuestiones jurídicas inherentes a la actividad profesional que se desarrollará.

Los aspectos legales y reglamentarios tienen gran incidencia sobre la conformación y funcionamiento de los laboratorios de informática forense. En primer lugar, es necesario identificar el conjunto de normas que resultan aplicables al ámbito en el cual se desempeñará el laboratorio. Considerando que no todos los laboratorios comparten el mismo contexto, es posible que tengan diferencias en cuanto a su regulación jurídica.

Para facilitar la labor de identificación de normas y reglamentos aplicables, conviene establecer criterios clasificatorios acerca de las diversas disposiciones en juego. Por ejemplo, las *normas relativas al empleo de datos, las normas regulatorias de la profesión informática, las normas procesales, las normas relativas a los vínculos con aquellos terceros que inciden en el desempeño del laboratorio, las normas regulatorias de los laboratorios en sí mismos.*

3.1 Servicios de un laboratorio de informática forense

Los servicios que ofrece un laboratorio de informática forense son las tareas que, dentro de su ámbito de incumbencia, puede realizar éste a solicitud del agente fiscal, del defensor o del juez, dependiendo de la organización a la que brinde servicios. Estas tareas estarán determinadas por las solicitudes que requieran de conocimientos informáticos específicos en un proceso judicial. En resumen, estos servicios definen todo aquello que se le puede solicitar al laboratorio. Cada una de estas tareas de-

penderá de la función específica que se le asigne en virtud de lo detallado en el apartado anterior.

La Guía Integral de Empleo de la Informática Forense en el Proceso Penal de la Provincia de Buenos Aires, Res PG SCBA 483/16 [11], distingue tres roles básicos a desempeñar por los informáticos forenses: rol de asesoramiento, de investigación o pericial, cada uno de los cuales incluye un conjunto de servicios. Es importante, entonces, como primer paso, distinguir, acorde la misión y visión, si las actividades vinculadas al asesoramiento e investigación estarán incluidas o no dentro del laboratorio, cuya razón de ser es netamente pericial.

En la descripción de los servicios, es menester diferenciar entre genéricos y específicos. Los servicios genéricos tienen en común una técnica, un objetivo, un objeto de estudio, o una prestación, sin especificar una tecnología particular y concreta. Por otra parte, los servicios específicos se aplican a una tecnología precisa y limitada a un entorno o área de estudio. Por ejemplo, un servicio genérico de acuerdo a PURI [3] es "Adquisición de imagen de datos (copia forense)" y uno específico dentro de este género sería "Adquisición de Dispositivos Móviles", dado que para realizar una imagen de un dispositivo móvil se necesita determinada tecnología, que puede no ser empleada en otro tipo de adquisición.

Los servicios específicos se encuentran ligados a una tecnología en particular y un conocimiento determinado y concreto de la misma. Esto implica que, no solo se necesita hardware y software acorde sino también personal capacitado en la materia. Además, dado que el avance tecnológico es continuo, un servicio específico puede quedar obsoleto y entonces deviene necesario, para el laboratorio, incorporar otros que hayan surgido del desarrollo de nuevas tecnologías.

3.2 Recursos humanos

La descripción del puesto de Informático Forense consiste en la elaboración de un documento que recoja las competencias, su definición conceptual, las actividades, requerimientos y responsabilidades correspondientes a cada uno de los roles y funciones establecidos [8]. Para definir un puesto de trabajo no sólo se tiene en cuenta las funciones compuestas por las actividades, requerimientos y responsabilidades de cada uno de sus roles, sino que también se pone un especial énfasis en el perfil y en el tipo de competencias que debe reunir la persona para poder asumir el reto tanto laboral como profesional y el compromiso ético con los objetivos institucionales.

Si bien dentro del Ministerio Público bonaerense existe el denominado "cargo", que podrá ser ocupado por un profesional capacitado en informática forense (perito, instructor, analista, etc.), es necesario ir más allá de los requisitos formales e identificar y definir los perfiles necesarios de acuerdo a los roles a desempeñar.

La instancia de identificación del perfil de competencias de cada rol de trabajo, permite reflexionar respecto a su asignación y el grado en el que deben ser solicitadas.

La descripción de los puestos y la relación existente entre ellos, la explicación de los cargos, los grados de autoridad y responsabilidad, las funciones y las actividades

previstas por cada uno de los integrantes de un laboratorio de informática forense Judicial, constituyen la base para alcanzar los siguientes objetivos:

- Facilitar el proceso de selección de personal.
- Identificar las necesidades de capacitación y desarrollo del recurso humano.
- Precisar las funciones requeridas a cada puesto.
- Propiciar el establecimiento de estándares tecnológicos y laborales de los diferentes laboratorios.
- Permitir el ahorro de tiempo y esfuerzos en la ejecución del trabajo.
- Servir de medio de integración y orientación al personal, facilitando su incorporación a las distintas funciones y su labor interdisciplinaria.
- Proporcionar el mejor aprovechamiento de los recursos humanos.

Tanto la descripción de puestos, como el organigrama son esenciales en el diseño organizacional.

El organigrama se concibe como la representación gráfica de la estructura de una institución, de forma tal que se pueda observar en ella la relación de jerarquía que tiene entre sus funciones [8].

Los organigramas permiten:

1. Promover la comprensión de las funciones dentro del laboratorio.
2. Orientar a los nuevos integrantes ante las relaciones y complejidades estructurales.
3. Proporcionar una imagen gráfica del aspecto íntegro de actividades y funciones de la organización y de las actividades y personal vitales para las mismas.

4 Aspectos a considerar a nivel Infraestructura

El laboratorio de informática forense debe radicarse en un espacio físico determinado, el cual deberá ser diseñado en función del capital humano, la localización física de quienes solicitarán su labor, los servicios que brinde, la infraestructura tecnológica y el equipamiento que requiera.

En función de ello, se puede esbozar un conjunto de criterios básicos para esta planificación y los aspectos a considerar. En tal sentido, como criterios para la planificación de los espacios de un Laboratorio de Informática Forense, deben atender el cumplimiento de la normativa vigente, la flexibilidad en su implementación, y su factibilidad de acuerdo a las condiciones sociales, políticas y económicas.

En el diseño edilicio debe considerarse y determinarse de manera pormenorizada el equipamiento e insumos intervinientes, las personas y sus roles, las situaciones geográficas y los espacios requeridos. En este punto resulta imprescindible la interacción continua con miembros del equipo e invitados especiales, los cuales puedan describir pormenorizadamente la actividad, y exponer detalles ante el diseñador arquitecto.

Algunas consideraciones metodológicas respecto al conjunto de actividades a llevar a cabo para elaborar una propuesta de diseño que cumpla con los requerimientos

y expectativas del laboratorio, consisten en el diseño de una carta de intención, en la determinación de actores periciales, de la variedad y alcances del objeto pericial, en la evaluación del equipamiento general y específico para la ejecución de las pericias, la forma apropiada en que deberán resguardarse las evidencias, así como el conocimiento general de la situación geográfica y las especializaciones a que se orientará el laboratorio atendiendo a la realidad contextual en la que trabajará.

5 Sistemas de Gestión de la Calidad

Pensar en el diseño de un Sistema de Gestión de Calidad (SGC) para los Laboratorios de Informática Forense requiere, en primer lugar, de un análisis de su situación actual, que identifique sus falencias y puntos de mejora, para luego poder contemplar de forma integrada todas sus características y necesidades; y desarrollar el SGC que aborde dicha situación.

La implementación del SGC provocará que los Laboratorios de Informática Forense cuenten con métodos normalizados junto a su documentación, herramientas debidamente validadas, procesos y pautas que contribuyan a garantizar la eficacia de toda la actividad que se desarrolla cotidianamente.

Por otra parte, el personal que conforma el Laboratorio puede desarrollar competencias técnicas específicas con relación a las herramientas, los sistemas, los procesos, contando de esta forma con pautas que fidelicen su labor diaria, garantizando la validez de su actividad pericial, y la efectividad y confiabilidad de las evidencias digitales extraídas, evitando nulidades.

El siguiente paso de este trabajo, en el proyecto que le continuará, pretenderá avanzar sobre la necesidad de aplicar parámetros de calidad a los laboratorios informático-forenses mediante la obtención de medidas de gestión, como por ejemplo: frecuencia de solicitud de ciertos servicios, tiempos de resolución, carga de trabajo, utilización de los recursos y de sistemas de modelado y simulación para analizar cómo se comportaría un laboratorio particular ante distintas situaciones, o evaluar qué cambio tendría un mayor impacto en el funcionamiento del mismo.

Los productos y servicios son generados a través de procesos. Un proceso de creación de bienes y/o servicios es un conjunto de actividades interrelacionadas que transforman un estado de cosas inicial (entrada) en un estado final (salida). Para el Ministerio Público Fiscal, cada reclamo o controversia vinculados con deberes y/o derechos sería el estado inicial; y el estado final consistiría en la resolución judicial o extrajudicial que decide la cuestión, los cuales generan consecuencias y efectos concretos.

A su vez, cada una de las actividades o subprocesos que componen este proceso, implicaría también transformaciones internas, de un estado de cosas inicial (entrada) en un estado final (salida). Por otra parte, es imposible generar un producto o servicio de la nada: siempre se requiere contar con recursos o insumos. Los productos o servicios generados por los diferentes subprocesos no siempre son recibidos por los usuarios finales, sino que algunos de ellos son insumos para otra actividad subsiguiente. Por ejemplo, en el marco procesal penal, la labor de los expertos produce informa-

ción y conocimiento, como insumos necesarios para las tareas de investigación, negociación y/o litigación que llevan a cabo los fiscales para defender los intereses sociales que se encuentran en juego en cada caso concreto. Los fiscales e investigadores son, así, "clientes internos" de los peritos, es decir, destinatarios finales de los servicios brindados.

Puede verse, entonces, que además de definir e internalizar la misión, visión y objetivos generales de una institución o entidad, es recomendable hacer lo propio con las estructuras destinadas a cumplir con los distintos subprocesos de trabajo. Para que esta labor de definición sea productiva, estos subprocesos deben ser contextualizados, ya que sólo adquieren sentido si contribuyen a cumplir con la misión, visión y objetivos básicos de la organización.

El alineamiento e integración de cada subproceso de trabajo en función de la misión general de la institución puede dar lugar a discusiones al interior de la organización. Ahora bien, ello es no sólo inevitable, sino también necesario si es que deseamos que la entidad esté centrada en las necesidades de sus beneficiarios finales. La satisfacción de los clientes internos (operadores forenses en sentido amplio, autoridades jerárquicas, entre otros) y el cumplimiento de estándares de calidad autónomos son valores a tener en cuenta, pero si se los eleva al rango de valores absolutos, podría pervertirse el destino público de los subprocesos de trabajo y de las estructuras que los llevan a cabo.

6 Conclusiones

Este trabajo tiene como fin exponer una serie de cuestiones a considerar a la hora de diseñar un laboratorio de informática forense, el que, como toda dependencia destinada a brindar un servicio, está inserto en un contexto funcional, institucional y geográfico.

El producto "Laboratorio de Informática Forense" puede ser excelente de manera aislada, disponer del software apropiado, contar con personal capacitado, cumplir con todas las previsiones de infraestructura, y, sin embargo, si no se adapta al contexto en el que está inserto desde el punto de vista estratégico e institucional, puede no resultar útil. Es decir, puede ser un buen producto pero una mala solución, según las necesidades de la organización, y, en definitiva, de la sociedad.

Un buen diseño exige, en principio, definir su misión, visión objetivos, y a la par definir explícitamente qué servicios brindará el laboratorio, quiénes podrán demandar estos servicios y cómo recibirán los resultados.

A partir de estas definiciones, se podrá determinar los requisitos a fin de cumplir con dichos servicios de la manera esperada. Esto es: infraestructura edilicia y tecnológica, recursos humanos, procesos operativos, estratégicos y de soporte, entre otros.

Esperamos que este documento aporte las pautas mínimas y necesarias para la creación, puesta en funcionamiento, rediseño y gestión adecuada y óptima de los laboratorios de informática forense.

Sería deseable que, además, el empleo de este documento impulse y oriente la generación de un sistema de medición y evaluación de los procesos periciales. Ello

podrá sentar las bases para la definición de programas de calidad en este tipo de laboratorios. A su vez, este tipo de programas podrá ser extendido, con las necesarias adaptaciones, a los institutos forenses (de los cuales suelen depender algunos laboratorios) y a las áreas o redes de servicios periciales.

Bibliografía

1. Arts. 56, 267, 268, 293 a 297 del CPPBA
2. Arts. 39 y 80 de la Ley 14442 de la Provincia de Buenos Aires
3. Di Iorio, Ana et al, El rastro digital del delito. Aspectos técnicos, legales y estratégicos de la Informática Forense. Universidad FASTA. 2017
4. Lago, Montejo Vicente, La práctica de la investigación criminal: inspección técnica ocular. Disponible en: https://www.editorialreus.es/static/pdf/9788429019841_primeras_paginas_la-practica-de-la-investigacion.pdf
5. Ley 14.442. art. 1° de la Provincia de Buenos Aires. Disponible en: http://www.gob.gba.gov.ar/legislacion/legislacion/l-14442.html
6. Sitio web del MPBA disponible en https://www.mpba.gov.ar/comunidad
7. Laboratorios regionales de investigación forense. - 1a ed. - Ciudad Autónoma de Buenos Aires: Infojus, 2014 Disponible: http://www.saij.gob.ar/docs-f/ediciones/libros/Laboratorios_Regionales_de_Invest._Forense.pdf
8. CHIAVENATO, Adalberto. Introducción a la Teoría General de la Administración. Bogotá: McGraw-Hill, 1999.
9. Estrategia Magazine – Año 2, Edición Nº 41, Sección Administración. "La misión: comenzar con un fin en la mente"
10. Ortiz, Sergio (2003); "¿Cómo generar una visión?" en Visión y Gestión Empresarial. Capítulo2. Ed. Thomson Editores, España.
11. Di Iorio et al. "Guía Integral de Empleo de la Informática Forense en el Proceso Penal. (2015)". Ed. Universidad FASTA.

Infraestructura de clave pública Argentina; estado, normativa, casos de aplicación y dificultades

Jorge Javier [1], Solinas Miguel[1], Bosch Luis Antonio[1]

[1] Laboratorio de Redes y Ciberseguridad, Facultad de Ciencias Exactas Físicas y Naturales, Universidad Nacional De Córdoba

{javier.jorge, miguel.solinas, luis.bosch}@unc.edu.ar

Resumen. Este trabajo reúne el estado del arte de la firma digital en Argentina, a casi veinte años de la sanción de la ley que la creó. Consta de la construcción de un análisis de tipo documental y testimonial que muestra los avances más relevantes que se han logrado con respecto al conocimiento de la firma digital hasta el momento de su redacción. El trabajo aporta una mirada amplia sobre el estado de la tecnología, su grado de implementación y los problemas por resolver. Además, incluye una recopilación de documentos de referencia, ideas y conceptos.

1 Introducción

El 12 de marzo del 2019 se publicó en el Boletín Oficial de la República Argentina el Decreto 182/2019 que aprueba, en su Anexo I, la reglamentación de la Firma Digital, la cual regula la utilización del documento electrónico, la firma electrónica y la firma digital y su eficacia jurídica en el marco de la Infraestructura de Firma Digital establecida por la Ley N° 25.506 y su modificatoria y el Plan de Modernización del Estado y de la Simplificación y Desburocratización de la Administración Pública. Esta nueva reglamentación recoge la experiencia en la implementación de la firma digital a lo largo de los casi 20 años. Hay que tener en cuenta que en el 2001, año de la sanción de la ley, muy pocos hogares tenían acceso a internet, recién en el año 2010 [1] el número de hogares con acceso a internet ascendía a cerca de cinco millones, mientras que en el momento en que se escribe este trabajo, el informe de marzo 2019 [2], dice que existen cerca de cuarenta millones de accesos a internet entre fijos y móviles. Este nuevo contexto es más favorable para implementar estas tecnologías que están disponibles desde hace ya muchos años.

Montar toda la infraestructura necesaria para brindar un servicio adecuado de firma digital no es tarea sencilla y la legislación establece requisitos para ello. Respecto de las herramientas de implementación existen herramientas de software libre disponibles, que se pueden consultar en [3].

Tengamos presente que la firma digital significa aquella que cuenta con un certificado digital (el "Certificado Digital") a través del cual puede verificarse: (i) la autoría de la firma; y (ii) autenticidad y/o integridad del documento firmado digitalmente. Explicar cómo funciona el Certificado Digital y quien lo otorga excedería en extenso este

artículo, se puede consultar [3], pero a fin de ser claros vamos a resumir en que el Certificado Digital es parte fundamental del proceso informático de creación de la firma digital y es emitido por el Ministerio de Modernización o alguna otra dependencia que fuera autorizada por la Autoridad de Aplicación de la Ley de Firma Digital.
El trabajo está organizado de la siguiente manera: en 2 revisamos brevemente el marco legal; en 3 las autoridades licenciadas al mes de Mayo del 2019; en 4 presentamos algunas aplicaciones para las cuales la tecnología está madura en Argentina; en 5 recordamos las dificultades que puede presentar su implementación y cerramos en 6 con algunas conclusiones.

2 La Infraestructura de Firma Digital de la República Argentina (IFDRA)

2.1 Marco legal

La reglamentación de la firma digital en Argentina está dada por Ley N° 25.506 de Firma Digital que reconoce y establece las condiciones para el empleo de la firma electrónica y de la firma digital y su eficacia jurídica, y crea la Infraestructura de Firma Digital de la República Argentina. A ello se suma el Decreto N° 2628/2002, Decreto Reglamentario de la Ley N° 25.506 y la Resolución Nº 399e/2016 del Ministerio de Modernización que reemplaza la Decisión Administrativa Nº 927/2014 y la Disposición SSTG Nº 7/2015. Esta última establece los procedimientos y condiciones que se deberán cumplir para emitir certificados digitales en el ámbito de la Infraestructura de Firma Digital de la República Argentina [4].

Posteriormente a la sanción de la ley, la Firma Digital fue incorporada por el El Código Civil y Comercial de la Nación estableciendo: “En los instrumentos generados por medios electrónicos, el requisito de la firma de una persona queda satisfecho si se utiliza una firma digital, que asegure indubitablemente la autoría e integridad del instrumento” (art. 288). Adicionalmente, El Código Civil y Comercial de la Nación también previó los contratos celebrados a distancia, incluyendo a aquellos celebrados a través de medios electrónicos (art. 1105).

Recientemente el decreto 182/2019 reglamenta y revisa múltiples aspectos de la Ley N° 25506 y la actualiza en base a la experiencia adquirida y los avances tecnológicos. En esta se define la Infraestructura de Firma Digital, que está compuesta por:

1. La Autoridad Certificante Raíz de la REPÚBLICA ARGENTINA.
2. El Ente Licenciante conformado por la SECRETARA DE GOBIERNO DE MODERNIZACIÓN de la JEFATURA DE GABINETE DE MINISTROS y la SECRETARÍA DE MODERNIZACIÓN ADMINISTRATIVA dependiente de la SECRETARÍA DE GOBIERNO DE MODERNIZACIÓN de la JEFATURA DE GABINETE DE MINISTROS.
3. Los certificadores licenciados, incluyendo sus autoridades certificantes y sus autoridades de registro, según los servicios que presten.
4. Las autoridades de sello de tiempo.

5. Los suscriptores de los certificados.
6. Los terceros usuarios.
7. Los certificadores reconocidos por la Autoridad de Aplicación.
8. El Organismo Auditante establecido en el artículo 34 de la Ley N° 25.506 y su modificatoria.
9. Los prestadores de servicios de confianza.

2.2 Sellado de tiempo (time stamping)

¿Desde qué instante comienza la cobertura de la póliza de un seguro que ha contratado? ¿Presentó a tiempo su reclamo? ¿Cuándo se emitió una factura? Cuando se realizan operaciones en Internet el conocimiento del instante de tiempo en el que ocurrieron es importante y solicitar a un tercero de confianza que dé constancia de la fecha y hora es fundamental a la hora de aportar pruebas, garantizando la fuente de tiempo que fue empleada y asegurando la integridad de los datos así sellados. El sellado de tiempo es un servicio que pueden brindar los certificadores licenciados. Para profundizar en el tema de sellados de tiempo y su implementación se puede referir a [6].

El art. 17 de la ley de firma digital define al certificador licenciado, además de referirse a la expedición de los referidos certificados, indica que "... presta otros servicios en relación con la firma digital… " La Decisión Administrativa 927/2014 "Certificaciones Digitales. Política, Contenidos, Requisitos y Formularios. Aprobación". Incluye el concepto y lo describe como parte de la infraestructura de firma digital argentina. La Resolución 399 E/2016 del Ministerio de Modernización, también menciona los SELLOS DE TIEMPO. Recientemente el Decreto 182/2019, actualiza lo establecido en la decisión 927/2014. Existen actualmente múltiples prestadores que brindan este servicio.

2.3 Estándares aprobados

La Infraestructura de Firma Digital de la República Argentina (IFDRA) está conformada por un conjunto de componentes que interactúan entre sí, permitiendo la emisión de certificados digitales para verificar firmas en condiciones seguras, tanto desde el punto de vista técnico como legal. La IFDRA ha adoptado los siguientes estándares tecnológicos:

1. Formato de los certificados y de las listas de certificados revocados: ITU–T X509.
2. Generación de las claves: RSA, DSA o ECDSA.
3. Protección de las claves privadas de certificadores y suscriptores: FIPS 140.
4. Políticas de certificación: RFC 5280 y 3739.

El listado completo de los estándares aprobados para la IFDRA, así como las condiciones bajo las cuales deben ser utilizados, se encuentra descripto en la Decisión Administrativa Nº 6/2007 (Anexo 3).

2.4 Requisitos para el licenciamiento de certificadores

El Anexo I de la Resolución Nº 39e/2016 del Ministerio de Modernización establece los procedimientos y condiciones que se deberán cumplir para emitir certificados digitales en el ámbito de la Infraestructura de Firma Digital de la República Argentina. Para comenzar con la solicitud de licenciamiento es necesario presentar documentos técnicos donde se detalla:

a) Política Única de Certificación, con los datos del solicitante.
b) Acuerdo tipo con suscriptores.
c) Términos y condiciones tipo con Terceros Usuarios ("relying parties").
d) Política de Privacidad.
e) Contratos con los proveedores de la infraestructura tecnológica, de corresponder.
f) Manual de Procedimientos.
g) Plan de Cese de Actividades.
h) Plan de Seguridad (incluye política y procedimientos de seguridad).
i) Plan de Continuidad de las Operaciones.
j) Descripción de la plataforma tecnológica.
k) Descripción de los servicios que brinda.

La especificación detalla el contenido de los documentos y los requisitos asociados a ellos. Todos estos documentos, en caso de ser aceptados, serán posteriormente utilizados para las auditorías.

2.5 Seguridad física de una autoridad certificante

Dentro de los requisitos de los entes certificantes cabe destacar los requisitos de seguridad física. Las autoridades certificantes deberán implementar un sistema de seguridad física que cuente con CUATRO (4) niveles de acceso físico, por lo menos, para llegar desde las áreas de libre circulación al ambiente donde reside su equipamiento informático afectado a la firma de certificados y CRLs. Además cada certificador deberá disponer de DOS (2) niveles adicionales para la protección de los elementos críticos vinculados a la activación de la clave privada de cada autoridad certificante y otros elementos críticos. Estos DOS (2) niveles pueden consistir en cajas de seguridad, gabinetes reforzados o compartimentos, de uso exclusivo de cada certificador. Debe tenerse en cuenta que en el caso que varias autoridades certificantes pertenecientes a distintos certificadores licenciados utilicen la misma infraestructura tecnológica, se deberá contar con N+1 cajas de seguridad, gabinetes o compartimentos, siendo N la cantidad de certificadores licenciados, a los que se agrega un contenedor adicional para el resguardo de otros elementos de operación del dispositivo criptográfico "HSM" (Hardware Security Module), que deban ser compartidos.

3 Certificadores licenciados

El Ministerio de Modernización actúa como Ente Licenciante otorgando, denegando o revocando las licencias de los certificadores licenciados y supervisando su accionar. Los certificadores licenciados son entidades públicas o privadas que se encuentran habilitados por el Ente Licenciante para emitir certificados digitales, en el marco de la Ley 25.506 de Firma Digital. Al momento de escribir este trabajo están licenciadas las siguientes instituciones/empresas, con una breve descripción de la Política de Seguridad referida a quiénes podrán suscribir un CD y los usos que se les podrá dar [17]:

3.1 AFIP (Administración Federal de Ingresos Públicos)

Suscriptores
Podrán suscribir los certificados digitales emitidos por la Autoridad Certificante de la AFIP (AC de la AFIP) [8] las personas físicas que utilizan los servicios de la AFIP mediante Clave Fiscal, y siempre que cumplan todos los requisitos expuestos en la presente política. La AFIP es suscriptora de un certificado, para ser usado en relación con el servicio OCSP de consultas sobre el estado de los certificados.
Usos
Las claves correspondientes a los certificados digitales que se emitan bajo la presente Política Única de Certificación podrán ser utilizadas en forma interoperable en los procesos de firma digital de cualquier documento o transacción y para la autenticación o el cifrado.

3.2 ANSES (Administración Nacional de Seguridad Social)

Suscriptores
Los certificados digitales emitidos [9] bajo la presente Política Única de Certificación tienen como suscriptores a aquellas personas físicas que desempeñan funciones para la ANSES con independencia del tipo de relación pudiendo ser funcionarios, empleados, adscriptos, personal designado desde otros organismos, pasantes o contratados de cualquier naturaleza; sin perjuicio de su posible ampliación previa notificación al ente licenciante.
Usos
Las claves correspondientes a los certificados digitales que se emitan bajo la presente Política Única de Certificación podrán ser utilizadas en forma interoperable en los procesos de firma digital de cualquier documento o transacción y para la autenticación o el cifrado.

3.3 Box Custodia de Archivos SA

Suscriptores
Podrán ser suscriptores de los certificados emitidos por la Autoridad Certificante AC - BOX CUSTODIA FIRMA DIGITAL [10] las personas físicas o jurídicas sean éstas

públicas o privadas y aquellos que presten otros servicios relacionados con la firma digital, tales como los enunciados en el artículo 10 de la Decisión Administrativa Nº 927 de fecha 30 de octubre de 2014.

Usos

Las claves correspondientes a los certificados digitales que se emitan bajo la presente Política Única de Certificación podrán ser utilizadas en forma interoperable en los procesos de firma digital de cualquier documento o transacción y para la autenticación o el cifrado.

3.4 Digilogix SA

Suscriptores

Podrán ser suscriptores de los certificados digitales emitidos por la AC -DIGILOGIX ANEXO [11] a) Las personas físicas y/o jurídicas relacionadas con las funciones, entre otras, de clasificación y/o guarda de documentación pública o privada, procesos de despapelización y/o digitalización y/o desarrollo e implementación de sistemas o aplicativos que protejan la autoría e integridad de la documentación tratada. b) Las personas físicas y/o jurídicas relacionadas, entre otras, con la gestión administrativa y documental, como ser: recibos de sueldo, correos electrónicos, órdenes de compra, facturas comerciales, documentos laborales, documentos comerciales, contratos, entre otros documentos. e) Las personas físicas y/o jurídicas vinculadas, entre otras, actividades a las relacionadas con funciones de tramitación y administrativas aduaneras.

Usos

Las claves correspondientes a los certificados digitales que se emitan bajo la presente Política Única de Certificación podrán ser utilizadas en forma interoperable en los procesos de firma digital de cualquier documento o transacción y para la autenticación o el cifrado.

3.5 Encode SA

Suscriptores

Según los términos de la presente Política Única de Certificación [12], se define la Comunidad de Suscriptores de certificados digitales a todas las Personas Físicas o Jurídicas de naturaleza Pública o Privada o responsables autorizados de aplicaciones y sitios seguros, que suscriban certificados de firma digital o provisión de servicios
vinculados de firma digital con ENCODE S.A.

Usos

Las claves correspondientes a los certificados digitales que se emitan bajo la presente Política Única de Certificación podrán ser utilizadas en forma interoperable en los procesos de firma digital de cualquier documento o transacción y para la autenticación o el cifrado.

3.6 Lackaut SA

Suscriptores

Podrán ser suscriptores de los certificados emitidos por la Autoridad Certificante AC - LAKAUT [13] las personas f!sicas o jurídicas sean éstas públicas o privadas y aquellos que presten otros servicios relacionados con la firma digital, tales como los enunciados en el artículo 10 de la Decisión Administrativa Nº 927 de fecha 30 de octubre de 2014.
La AC -LAKAUT será además, suscriptora de un certificado para ser utilizado en relación con el servicio On Line Certificate Status Protocol (en adelante, OCSP) de consulta sobre el estado de los certificados digitales.

Usos

Las claves correspondientes a los certificados digitales que se emitan bajo la presente Política Única de Certificación podrán ser utilizadas en forma interoperable en los procesos de firma digital de cualquier documento o transacción y para la autenticación o el cifrado.

3.7 Oficina Nacional de Tecnologías de la Información (ONTI)

Suscriptores

Podrán ser suscriptores de los certificados emitidos por la AC ONTI [14] las personas humanas, que requieran un certificado digital para firmar digitalmente cualquier documento o transacción, pudiendo ser utilizados para cualquier uso o aplicación, como así también para autenticación o cifrado.
La AC ONTI emite también un certificado para ser usado en relación con el servicio On Line Certificate Status Protocol (en adelante, OCSP) de consulta sobre el estado de un certificado.
Asimismo, la AC ONTI emite certificados de aplicación, y presta el servicio de sello de tiempo, según lo dispuesto en el artículo 9° de la Resolución N° 399-E/2016° del 5 de octubre de 2016 del entonces MINISTERIO DE MODERNIZACIÓN.

Usos

Las claves correspondientes a los certificados digitales que se emitan bajo la presente Política Única de Certificación podrán ser utilizadas en forma interoperable en los procesos de firma digital de cualquier documento o transacción y para la autenticación o el cifrado.

3.8 Prisma Medios de Pago SA

Suscriptores

Los suscriptores de certificados [15] serán personas físicas o aquellas personas jurídicas que tengan por objetivo cualquiera de los usos de certificados digitales y emitidos a nombre de personas físicas, personas jurídicas, aplicaciones, sitio seguro y de autoridad de competencia y de sello de tiempo de acuerdo a lo definido, y en los términos expresados en la presente “Política única de Certificación”.

Uso

Las claves correspondientes a los certificados digitales que se emitan bajo la presente Política Única de Certificación podrán ser utilizadas en forma interoperable en los procesos de firma digital de cualquier documento o transacción y para la autenticación o el cifrado.

3.9 Tecnologías de Valores SA

Suspendió sus actividades.-

3.10 Ministerio de Modernización

Suscriptores
Podrán ser suscriptores de los certificados emitidos por la AC MODERNIZACIÓN-PFDR [16] las personas humanas que requieran un certificado digital para firmar digitalmente cualquier documento o transacción, pudiendo ser utilizados para cualquier uso o aplicación, como así también para autenticación o cifrado. La AC MODERNIZACIÓN-PFDR emite también un certificado para ser usado en relación con el servicio On Line Certificate Status Protocol (en adelante, OCSP) de consulta sobre el estado de un certificado. Los suscriptores de certificados de la AC MODERNIZACIÓN-PFDR generan sus claves en la Plataforma de Firma Digital Remota (en adelante, PFDR). En el caso de los Oficiales de Registro, las claves son generadas en dispositivos que cumplan con certificación "Overall" FIPS 140 (versión 2) Nivel 2 o superior. Estos certificados deberán ser emitidos por alguna de las AUTORIDADES CERTIFICANTES pertenecientes al MINISTERIO DE MODERNIZACIÓN.
Usos
Los certificados digitales que se emitan bajo la presente Política Única de Certificación podrán ser utilizados en forma interoperable en los procesos de firma digital de cualquier documento o transacción y para la autenticación o el cifrado.

3.11 Solicitantes de Licencia

Las siguientes organizaciones han presentado su solicitud de licenciamiento, en los términos de la normativa vigente, con el fin de convertirse en certificadores licenciados y como tales, componentes de la Infraestructura de Firma Digital de la República Argentina.

- AdeA Administradora de Archivos S.A.
- Alpha 2000 Soluciones Informáticas S.R.L.
- Suprema Corte de Justicia de la Provincia de Buenos Aires.
- Minder S.A
- Poder Ejecutivo de la Provincia Córdoba

4 Aplicaciones disponibles

Brindar los servicios de una autoridad de certificación de prueba o a nivel de prototipo, con herramientas open source es una tarea relativamente sencilla. En unas pocas horas podrá estar en condiciones de emitir su propio certificado de acuerdo al estándar X.509.

El problema surge cuando se pretende encontrar una aplicación de esta tecnología. Contar con una idea de implementación novedosa no es suficiente, es imprecindible contar con un plan de negocio que contemple una solución rentable, y en general se suelen subestimar los costos de puesta en marcha de este tipo de tecnologías.

Desde sus inicios, fue una tecnología orientada a aumentar la confianza del consumidor en Internet como un vehículo para realizar comercio electrónico y a la esperea de aplicaciones de la tecnología PKI. Esto en alguna medida se ha logrado, pero ahora queda por desplegar aplicaciones que faciliten a los ciudadanos el uso extendido de esta tecnología. Por ejemplo, es una interesante propuesta cuando se habla de sustentabilidad ya que pone en serio cuestionamiento el soporte en papel. Una transformación digital dentro de las áreas de administración sin PKI se hace impensable. Brindar seguridad a dispositivos de IoT es una alternativa interesante y también posible [7] . Facilitar la autenticación de personas para pago con dispositivos móviles; son alguna de las aplicaciones para las cuales la tecnología PKI está madura en Argentina.

5 Dificultades

Si bien la reglamentación y el marco normativo es extenso y complejo, se observa que desplegar una AC en el marco de la Ley de Firma Digital Argentina, no es tan sencillo como poner en un ambiente de laboratorio una herramienta de código abierto que realice la gestión de los certificados. Existen múltiples aspectos a tener en cuenta a la hora de implementar una AC para su licenciamiento, no solo a nivel de implementación tecnológica en términos de infraestructura necesaria. Los aspectos legales también presentan un desafío.

Por otro lado se observa que la firma digital comienza a utilizarse en actividades cotidianas, sin embargo también se observa en la sociedad un profundo desconocimiento del tema, este desconocimiento es peligroso, y puede llevar este sistema al fracaso. Se observa que es indispensable comenzar con planes de capacitación y concientización sobre el uso y buenas prácticas de la firma digital. El usuario es parte fundamental de la infraestructura y debe ser consciente de ello.

6 Conclusiones

El reciente decreto 182/2019 actualiza y ordena el marco legal de la infraestructura de firma digital argentina. Infraestructura que dadas las condicione materiales actuales comienza a ser de uso cotidiano para un porcentaje muy alto de la población. Del relevamiento de Autoridades de Certificación licenciadas se desprende que la mayoría han obtenido una licencia para suscribir certificados para los mismos perfiles y usos. AFIP y ANSES emiten certificados para uso interno o de personas físicas que interactúan directamente con ellos. Es interesante la propuesta del Ministerio de Modernización que a través de una red de Autoridades de Registro distribuidas territorialmente [5] permite obtener un CD gratuito la mayoría de las veces, llevando un token criptográfico o un smartphone el cual quedará vinculado al proceso de firma digital remota. Es interesante mencionar que entre las AR que brindan este servicio se encuentran empresas

privadas como Arcor, gobiernos provinciales y Registros de Automotor. Esto evidentemente ayudará a la difusión del uso de esta tecnología en los próximos años.

Por otro lado es interesante observar que después de casi veinte años de desarrollo de la tecnología, no hemos superado el obstáculo de poder tener al menos una Autoridad de Certificación que pueda emitir Certificados SSL para autenticación de dominios cuyo CD de AC Raíz pueda estar incluido en los principales navegadores de internet. Si bien existen soluciones como Let's Encrypt [6] que ofrecen certificados gratuitos, su principal limitación es la vigencia de seis meses. Por lo que muchos sitios siguen eligiendo pagar por sus certificados. Una primera estimación, nos lleva a pensar que en el país se están invirtiendo alrededor de 50 millones de u$s anuales en servicios de certificados, tecnología que es estándar, de dominio público y que localmente ha madurado al punto de poder brindar estos servicios a precio de moneda local.

Referencias

1. Nota de Prensa de Acceso a Internet, Cuarto Trimestre 2010, Instituto Nacional de Estadísticas y Censo (INDEC), consultado Abril 2019, https://www.indec.gob.ar/uploads/informesdeprensa/internet_03_11.pdf
2. Informes Técnicos. Vol. 3, nº 44, ISSN 2545-6636, Instituto Nacional de Estadísticas y Censo (INDEC), consultado Abril 2019,(https://www.indec.gob.ar/uploads/informesdeprensa/internet_03_19.pdf)
3. Solinas M., Castello R.J., Tula L., Gallo C., Jorge J., Bollo D.: Implementación de una infraestructura de clave pública con herramientas de software libre. 10mas Jornadas Argentinas de Software Libre, JSL 2013; 42 JAIIO; http://42jaiio.sadio.org.ar/proceedings/simposios/Trabajos/JSL/10.pdf
4. Normativa de Firma Digital AR, consultada en Abril 2019; https://www.argentina.gob.ar/firmadigital/normativa
5. Autoridades de Registro de la República Argentina, consultada en Abril 2019; https://firmar.gob.ar/docs/listaARs.pdf
6. Utilizando Software Libre para un servicio de Sellado Digital de Tiempo; http://sedici.unlp.edu.ar/bitstream/handle/10915/4025/Tesis_.pdf?sequence=3
7. Vučinić Mališa y otros, OSCAR: Object Security Architecture for the Internet of Things, Grenoble Alps University, CNRS Grenoble Informatics Laboratory UMR 5217, France STMicroelectronics, Crolles, France.-
8. https://acn.afip.gov.ar/afipacn/p_home.xhtml; consultado en Abril 2019.-
9. http://www.anses.gob.ar/firmadigital; consultado en Abril 2019.-
10. https://pki.boxcustodia.com; consultado en Abril 2019.-
11. http://www.digilogix.com.ar; consultado en Abril 2019.-
12. http://www.encodesa.com.ar; consultado en Abril 2019.-
13. https://lakautac.com.ar; consultado en Abril 2019.-
14. https://pki.jgm.gov.ar; consultado en Abril 2019.-
15. http://ac.banelco.com.ar; consultado en Abril 2019.-
16. https://firmar.gob.ar; consultado en Abril 2019.-
17. https://www.argentina.gob.ar/modernizacion/administrativa/firmadigital/acraiz; consultado en Abril 2019.-

Los derechos humanos, las garantías constitucionales y las investigaciones criminales en ambientes digitales. Tensiones, límites y desafíos

Pablo Casas[1] y Antonela Mandolesi[1]

[1]Juzgado Penal, Contravencional y de Faltas Nº 10, Poder Judicial CABA.
{pablocruzcasas, mariaantonelamandolesi}@gmail.com

Resumen. La propuesta del trabajo parte desde el convencimiento de una necesaria reconfiguración y adaptación a las nuevas dinámicas sociales que se generan en esta era de entrega masiva de datos, en un contexto donde tiene lugar el conocido y vertiginoso avance en el poder de las herramientas de explotación y procesamiento de esos datos. El debate que nos propusimos está guiado a través del análisis de casos concretos de investigaciones criminales donde se debatió acerca de los límites y formas en la obtención de evidencia digital, dando como resultado las interpretaciones que se plasmaron en recientes fallos jurisprudenciales locales, Europeos y de los Estados Unidos de Norteamérica. En ellos el objeto de decisión se centró en la búsqueda de la construcción de límites en la posibilidad de accionar autónomamente por parte de los investigadores, no con el objeto de entorpecerla, ni hacerla más lenta, sino por la necesidad de que ese tipo de actividades, que tienen una factibilidad de injerencia muy grande a la intimidad de las personas, tenga que ser informada dentro del proceso penal dado los altos riesgos de afectación a la vida de las personas humanas, que hoy lleva intrínseca la obtención de nuestros datos y a los cuales debemos proteger dentro del derecho de la autodeterminación informativa. En definitiva se trata de pensar y discutir estos desafíos con el objetivo de balancear la necesidad de que las investigaciones sean eficientes pero sin poner en riesgos aspectos que hacen a nuestro desarrollo personal como individuos a los que se nos reconocen y garantizan el respeto de los derechos humanos como parte fundamental de hacer efectiva la vida en libertad en un estado democrático.

1 Datos personales, privacidad, autodeterminación informativa y derechos humanos

En nuestro país luego de la reforma constitucional del año 1994 en la que se incorporó el instituto del hábeas data (art. 43, tercer párrafo, CN), se sancionó la ley 25.326 de Protección de los Datos Personales, hoy con estado parlamentario un proyecto de reforma, que reglamenta justamente dicha acción constitucional[1].

[1] https://www.argentina.gob.ar/sites/default/files/mensaje_ndeg_147-2018_datos_personales.pdf

El art. 1º establece que tiene por objeto la protección integral de los datos personales asentados en archivos, registros, bancos de datos, u otros medios técnicos de tratamiento de datos, sean éstos públicos, o privados destinados a dar informes, para garantizar el derecho al honor y a la intimidad de las personas, así como también el acceso a la información que sobre las mismas se registre, de conformidad a lo establecido en el artículo 43, párrafo tercero, de la Constitución Nacional.

Por su parte, en el art. 2º define como "datos personales" a la información de cualquier tipo referida a personas físicas o de existencia ideal; y señala expresamente que el "titular de los datos" es la persona cuyos datos sean objeto del tratamiento al que se refiere la ley, es decir, el usuario, con independencia de que esa información se encuentre en poder de un tercero.

Finalmente, el propio artículo define al "tratamiento de datos" como operaciones y procedimientos sistemáticos, electrónicos o no, que permitan la recolección, conservación, ordenación, almacenamiento, modificación, relacionamiento, evaluación, bloqueo, destrucción, y en general el procesamiento de datos personales.

A través de la regulación constitucional de la acción de hábeas data y de la regulación legal, se reconoció que la información personal almacenada en bancos de datos merece protección contra eventuales abusos de poder, que se inician muchas veces a partir del desconocimiento por parte de los titulares de los datos de las grandes acumulaciones de información en bancos de diferente naturaleza.

En este sentido, Pablo Palazzi apunta que "el derecho a la protección de los datos personales es la respuesta a la acumulación y el tratamiento de datos personales en forma automatizada en ordenadores. Consiste en otorgar a los individuos una facultad de control sobre los propios, a través de toda una serie de reglas y principios que incluyen la calidad de ciertos datos, el consentimiento para su tratamiento, acciones judiciales, limitaciones a los bancos de datos en su contenido, en el tiempo y en la forma de tratamiento, en las cesiones o transferencias a terceros y en la intervención de agencias especiales del Estado destinadas a tutelar esos derechos"[2]

La llamada era digital nos está haciendo transitar por realidades y desafíos muy prometedores, pero al mismo tiempo con necesidad de reflexionar, pero también de actuar en campo de los nuevos interrogantes que se abren y algunos que se profundizan en muchos aspectos de los que no escapa la investigación criminal.

Hoy resulta fácil advertir el corrimiento de los límites físicos de protección de la vida privada o del derecho a la intimidad, pero no por ello dejaron de ser parte de los derechos humanos que sostienen libertades democráticas fundamentales, por tanto constitucionalmente protegidas.

Bajo este prisma es necesario abordar la problemática de la protección de la privacidad en la era de la digitalidad, con el objetivo de aportar a la construcción de nuevas formas para controlar el modo en que el Estado ámbitos íntimos de nuestras vidas.

Si bien somos conscientes que la evolución de la tecnología también trae aparejada una correlativa complejización de la investigación de determinada categoría de delitos, y que no es posible perder de vista la necesidad de garantizar investigaciones criminales eficientes, de todas maneras no debemos perder de vista la necesidad imperiosa de

[2] PALAZZI, Pablo: La protección de los datos personales en la Argentina, Errepar, 2004, p. 3.

adaptar rápidamente los conceptos jurídicos tradicionales para que no se sigan erosionando los contornos del derecho a la vida privada, frente al vertiginosa evolución de las tecnologías y de las formas de comunicación, cuando la experiencia histórica en la materia demuestra que las leyes y la jurisprudencia tardan en reconocer esta realidad y que ese encorsetamiento tiene la potencialidad de impactar fuertemente sobre los derechos de las personas.

Estamos convencidos de que la protección de nuestros datos personales debe ser entendida a partir del concepto del derecho a la autodeterminación informativa como parte de los derechos humanos de las personas.[3]

En palabras de Johanna Caterina Faliero, "La protección de datos personales, como aquella disciplina jurídica tuitiva que se encarga de proteger los datos personales de los titulares, débil jurídico de la relación de tratamiento de datos frente al responsable de estos, con miras a la preservación de su derecho de autodeterminación informativa, representa en la actualidad, en nuestra era de datos, un punto regulatorio clave en las normativas regionales y locales de todo el mundo. La tutela de los datos personales, no sólo como un desprendimiento de un derecho personalísimo del individuo, sino también como una manifestación y extensión de la soberanía nacional, se sitúa como una de las preocupaciones cardinales en las agendas gubernamentales de los estados[4].

La primera sentencia en la que se reconoció la autodeterminación informativa como un derecho fundamental de las persona humana fue dictada por el Tribunal Constitucional Alemán de 15 de diciembre de 1983[5], en la que se interpretó que resultaba lícito el recopilamiento de gran parte de los datos del censo referidos a nombre, apellidos, dirección, estado, nacionalidad, utilización de la vivienda, fuente de los medios principales de subsistencia, datos académicos y profesionales, rama de actividad, pero se declaró que resultaban ilícitos, entre otros, los preceptos relativos al cotejo de datos para ser utilizados contra las personas obligadas a suministrar esa información.

Con la capacidad de recolección de aquella época, ya se reconoció que la proliferación de centros de datos y los avances tecnológicos han permitido producir "una imagen

---

[3] En este sentido, ver PALAZZI, Pablo A.: Delitos contra la intimidad informática, CDYT Colección Derecho y Tecnología, Buenos Aires, 2019, capítulo I Las nuevas tecnologías y la protección penal de la privacidad, pag. 17/40. El autor hace foco precisamente en esta idea, al señalar que los cambios legales son lentos frente a las nuevas tecnologías. Los cambios tecnológicos y los cambios sociales -indica el autor-, por su parte, se dan sin formalidades y se instalan progresivamente en la sociedad influyendo en los conceptos legales. Luego, cita como ejemplo la discusión que se generó en el sistema estadounidense de regulación de las interceptaciones telefónicas, que dio lugar a un debate judicial y legislativo que en ese país duró casi ochenta años. Concluye entonces que cada nueva década obliga a replantear las interpretaciones pasadas en función de las nuevas tecnologías, como sucedió en el debate por el uso de escáneres términos, los teléfonos inteligentes, los simuladores de torres de celulares o la captación masiva de datos personales en internet.

[4] FALIERO, Johanna Caterina, El Futuro de la regulación en protección de datos personales en la Argentina, La Ley, Thomson Reuters, publicado en: Sup. Esp. LegalTech 2018 (noviembre), 05/11/2018, 55, Cita Online: AR/DOC/2375/2018

[5] Sentencia de 15 de diciembre de 1983, Ref. 1 BvR 209/83, Fondo, Ley del Censo http://www.derecho-chile.cl/sentencia-de-15-de-diciembre-de-1983-del-tribunal-constitucional-federal-aleman-ley-del-censo/

total y pormenorizada de la persona" convirtiéndose así el ciudadano en "hombre de cristal".

En dicho precedente, con un lenguaje que este siglo ya nos acostumbró, se dijo que "(...) la autodeterminación del individuo presupone -también en las condiciones de las técnicas modernas de tratamiento de la información- que se conceda al individuo la libertad de decisión sobre las acciones que vaya a realizar o, en caso, incluyendo la posibilidad de obrar de hecho en forma consecuente con la decisión adoptada. El que no pueda percibir con seguridad suficiente que informaciones relativas a él son conocidas en determinados sectores de su entorno social y quién de alguna manera no sea capaz de aquilatar lo que puedan saber de él sus posibles comunicantes puede verse sustancialmente cohibido en su libertad de planificar o decidir por autodeterminación. No serían compatibles con el derecho a la autodeterminación informativa un orden social y un orden jurídico que hiciese posible al primero, en el que la persona ya no pudiera saber quién, qué, cuándo y con qué motivo sabe algo sobre él. Quien se siente inseguro de sí en todo momento se registran cualesquiera comportamientos divergentes y se catalogan, utilizan o transmiten permanentemente a título de información, procurará no llamar la atención con esa clase de comportamiento. Quien sepa de antemano que su participación, por ejemplo, en una reunión o en una iniciativa cívica va a ser registrada por las autoridades y que podrán derivarse riesgos para él por este motivo renunciara presumiblemente a lo que supone un ejercicio de los correspondientes derechos fundamentales [artículo 8° y 9° de la Ley Fundamental (17). Esto no sólo menoscabaría las oportunidades de desarrollo de la personalidad individual, sino también el bien público, porque la autodeterminación constituye una condición elemental de funcionamiento de toda comunidad fundada en la capacidad de obrar y de cooperación de sus ciudadanos".

En la misma línea la Corte Interamericana de Derechos Humanos interpretó que la protección a la vida privada, establecida por el texto de la Convención, no se agota exclusivamente en las referencias al domicilio y a la correspondencia consagradas en su texto, proponiendo entonces una interpretación dinámica de la cláusula examinada[6].

También el Tribunal Europeo de Derechos Humanos estableció que "vida privada" es un término amplio, no susceptible de una definición exhaustiva y que su protección no se limita a un "círculo íntimo" en el que el individuo desarrolle su vida personal, excluyendo de su conocimiento a cualquier otra persona externa, sino que también protege el derecho de establecer y desarrollar relaciones con otras personas y con el mundo exterior. Existe una zona de interacción de la persona con otras, incluso en un contexto público, susceptible de protección como "vida privada"[7].

Si partimos de la idea conceptual tradicional de que la "intimidad física" supone la libertad y una carta de protección contra cualquier injerencia arbitraria del estado en la familia, el domicilio, la correspondencia, la comunicaciones y los papeles privados, es

---

6 Corte IDH, caso "TRISTÁN DONOSO v. Panamá". Excepción Preliminar, Fondo, Reparaciones y Costas. Sentencia del 27 de enero de 2009. Serie C Nº 193; párr. 29.

7 TEDH, "PERRY c. Reino Unido", núm. 63737/00, párr. 36; "PECK c. Reino Unido", núm. 44647/98, párr. 57 y 59.

posible definir a la “intimidad informativa” como el derecho de cada individuo de definir cómo, quién y bajo cuáles circunstancias y condiciones se puede acceder a su información personal.

En este marco conceptual es inserta al derecho a la autodeterminación informativa, como derivación de la garantía de intimidad, una faz de un derecho humano a proteger. Que puede ser definido, de manera general, que cada persona tiene el derecho personal de conocer, decidir y disponer libremente sobre sus datos personales, lo que alcanza también el derecho de disponer sobre quiénes pueden acceder a ellos y para qué propósito.

La derivación básica del reconocimiento de la existencia de ese derecho está dada por la correlativa imposibilidad de que particulares, o incluso el propio Estado, accedan injustificadamente a esa información personal, salvo que se den determinadas hipótesis excepcionales -como podría serlo la existencia de una investigación criminal en la que se haya reunido cierto grado de mérito-, respetando ciertas formas y siempre que se garantice la instancia de control de fundamentación y razonabilidad, que por imperativo constitucional debe ser realizada por el órgano jurisdiccional.

2 Límites, formas y roles en el proceso penal

Resulta poco controvertible que cierta parte de la información de carácter personal que en el marco de investigaciones criminales se requiere a las empresas que las almacenan, se encuentra amparada por la garantía de la privacidad, y como tal el acceso a la misma por parte de los investigadores sin orden judicial podría redundar en una afectación a dicha garantía (art. 17 PDCyP, art. 11.2 CADH, art. 12 DUDH, arts. 18 y 19 CN y art. 12 inc. 3º y 13.8 CCABA).

En ese sentido, de acuerdo con una interpretación amplia y dinámica del derecho a la intimidad, consideramos que los datos de tráfico de todo usuario de un correo electrónico y de redes sociales (registro de direcciones de IP históricamente asignadas al usuario, registro de la información transaccional), registrados en las bases de datos de las empresas de telecomunicaciones, por las redes sociales o por cualquier otra plataforma digital, constituyen información personal almacenada y, en consecuencia, por imperativo constitucional, su relevamiento en el marco de una investigación penal sólo puede ser ordenado por el juez competente.

Este es el motivo por el cual consideramos que nos encontramos frente a una medida que no puede ser dispuesta unilateralmente por los fiscales que dirigen las investigaciones, sin intervención del único órgano constitucionalmente habilitado para permitir el acceso a ciertos ámbitos reservados.

En este punto, más allá de la propuesta de readaptación dinámica del alcance de la garantía a la que invitamos, no está de más recordar que el derecho a la privacidad se encuentra reconocido en nuestra Constitución Nacional desde el año 1853, a través de los arts. 18 y 19, que reconoce la inviolabilidad de ciertos espacios físicos, que alcanzan el domicilio y los papeles privados.

Si bien estas normas no requieren en rigor la orden de un juez para acceder a esos espacios, sino solo una reglamentación razonable, lo cierto es que el programa constitucional prevé que los jueces son custodios de las garantías constitucionales y corresponde a ellos decretar las injerencias excepcionales en dichas garantías, toda vez que son quienes están en mejores condiciones de objetividad y de serenidad para hacerlo. De allí que la mayoría de los ordenamientos procesales, cuando regulan esta clase de medidas probatorias, exigen la orden judicial de allanamiento como condición de su validez.

Por otra parte, el artículo 13.8 de la Constitución de la Ciudad Autónoma de Buenos Aires, reconoce que la privacidad alcanza también a las escuchas telefónicas y a la "información personal almacenada", y también prevé expresamente el requisito de la orden judicial para disponer medidas que signifiquen un avance estatal sobre esos espacios de reserva.

La creciente complejización de las investigaciones, y la necesidad cada vez más frecuente de requerir medidas de prueba digital incluso en la investigación de delitos comunes, obligan a redefinir hasta dónde puede hacerse extensivo el requisito de orden judicial, frente a la problemática planteada en la introducción de este artículo. Máxime, frente a la predominancia de modelos procesales que ubican a los fiscales como directores de la investigación, y supeditan la intervención de los jueces únicamente en aquellos casos en los cuales los fiscales así la requieran.

En este punto, resulta particularmente interesante traer a la cita el reciente caso "BENEDIK c. Slovenia" (rta. 24/04/2018), del Tribunal Europeo de Derechos Humanos, en el que se discutió justamente la facultad de los investigadores de acceder autónomamente a los datos personales de los usuarios almacenados por las empresas proveedoras de servicio de internet; concretamente, a qué usuario se había asignado una determinada IP, en el marco de una investigación por distribución de imágenes con contenido de abuso sexual de personas menores de edad.

El TEDH entendió que existió una violación al derecho a la privacidad, previsto en el art. 8 de la Convención Europea de Derechos Humanos (análogo al art. 11 de la Convención Americana sobre Derechos Humanos).

Destacó que el art. 8 CEDH protege el derecho de identidad y desarrollo personal, y el derecho de establecer y desarrollar relaciones con otros seres humanos y con el mundo exterior, así como el derecho a la autodeterminación informativa, al considerar que el concepto de "vida privada" es un término amplio, y que por ello incluye el derecho a la privacidad con respecto al procesamiento automático de "datos personales", entendiendo por tal concepto a cualquier información relativa a un individuo identificado o identificable.

Por otra parte, explicó que la información del imputado asociada con la IP dinámica, no era información que estuviera accesible y por lo tanto no podía ser comparada a la información encontrada tradicionalmente en un directorio público. Para poder identificar a una persona a través de una IP dinámica, la empresa prestadora del servicio debía acceder a la información almacenada concerniente a eventos de telecomunicaciones particulares, por lo que el uso de esa información, por sí sola, podía dar lugar a consideraciones sobre la vida privada.

Tal como lo hizo este último tribunal, cuando se trate de medidas referidas a investigaciones por delitos relacionados con imágenes de abuso sexual infantil, resulta importante tomar en cuenta las disposiciones de la Convención sobre Ciberdelito (Budapest), aprobada por nuestro país mediante Ley Nº 27.411 (BO del 15/12/2017), que obliga a los Estados a llevar a cabo medidas que permitan a las autoridades combatir esos crímenes.

Pero que también dispone que dichas medidas deben ser llevadas a cabo de conformidad con el art. 15 de la misma Convención, que establece que la instauración, ejecución y aplicación de los poderes y procedimientos previstos en esa sección se sometan a las condiciones y salvaguardias previstas en su derecho interno, así como que las "condiciones y salvaguardias incluirán una supervisión judicial u otra forma de supervisión independiente, los motivos que justifiquen su aplicación, así como la limitación del ámbito de aplicación y de la duración de dicho poder o procedimiento" (destacado agregado).

Desde la perspectiva propuesta, ante el vacío de la mayoría de las legislaciones procesales que aún no contienen una regulación específica que establezca las pautas para la obtención de determinada clase de información, siempre que se trate de medidas que impliquen acceder a información de carácter personal, resulta indispensable la exigencia de una orden judicial, la cual únicamente será legítima siempre que el caso reúna cierto mérito sustantivo, y en tanto la medida sea necesaria para el avance de la investigación e indispensable en una sociedad democrática, conforme el estándar que en este punto imponen los arts. 30 y 32.2 CADH.

3 Datos de Geoposicionamiento

Una medida que puede resultar problemática desde la perspectiva de la protección del derecho a la privacidad se vincula con el acceso por parte de los investigadores a los listados de celdas de conexión de los teléfonos móviles, con su correspondiente ubicación geográfica.

Los usuarios de los teléfonos celulares mantienen una razonable expectativa de privacidad respecto del registro de los sucesivos y constantes movimientos que van quedando capturados por las celdas de geolocalización de las antenas de las empresas prestatarias del servicio, y por lo tanto, para que puedan ser reveladas en el marco de una investigación penal, también se requiere la orden de un juez, por lo que resulta inadmisible que los investigadores accedan a esa información haciendo uso de sus facultades investigativas autónomas.

A los fines de comprender adecuadamente las características técnicas de este tipo de medidas, conviene señalar sucintamente que el análisis de las celdas de localización permite determinar, sobre la base de la información con que cuentan las operadoras, la ubicación de todas las terminales móviles que se activaron dentro de una "celda" (rango de cobertura geográfica de una antena) en un momento determinado.

Sobre la base de la información obtenida, se puede concluir que determinado usuario se encontraba en determinado horario, en las cercanías del lugar de los hechos; o

bien, se puede determinar el presunto lugar de comisión del hecho, cuando no está suficientemente circunscripto.

Es importante destacar que los registros del posicionamiento de los usuarios quedan almacenados en los archivos de las empresas prestatarias con independencia de si el titular utiliza o no el teléfono móvil. En efecto, se conocen ciertas técnicas de investigación llevadas a cabo en países europeos como Alemania, Holanda, Francia y España, que permiten acreditar la ubicación de personas mediante los denominados "SMS silenciosos" (Stille SMS), a través de los cuales las autoridades envían mensajes de texto a un destinatario cuya ubicación se pretende conocer, quien no se anoticia del mismo, no obstante lo cual el acto de comunicación queda registrado y archivado entre los datos de la operadora referidos a ese usuario[8].

Se trata entonces de una medida que tiene fuertes repercusiones desde el punto de vista de la protección de la privacidad y de la información personal que las empresas de telecomunicaciones tienen almacenada respecto de sus usuarios, únicos titulares de dichos datos, conforme los lineamientos constitucionales y normativos establecidos precedentemente.

La relevancia de la cuestión se hace más notoria si se contrasta con el hecho de que ninguno de nosotros nos desprendemos habitualmente de nuestros teléfonos celulares, y que en la actualidad se trata de accesorios que parecen formar parte de la anatomía de muchas personas.

Sin perjuicio de que este tipo de medida probatoria no se encuentra específicamente regulada en nuestro ordenamiento procesal, resulta necesario priorizar una interpretación constitucionalizada de las normas procesales y orgánicas que regulan las funciones propias de quienes tienen en sus manos el poder y deber de investigar.

Con relación a esta temática, el TEDH ha interpretado que la práctica estatal de conservar datos de localización de un individuo constituye una injerencia en la vida privada de las personas aún cuando se trate de datos recogidos en lugares públicos (TEDH, "AMANN c. Suiza", rta. 16/02/2000, párr. 65-67 y "ROTARU c. Rumania", rta. 4/05/2000, párr. 43-44; más recientemente "SHIMOVOLOS c. Rusia", rta. el 21/06/2011 en un caso referido al archivo de datos por parte del Estado de desplazamientos de un ciudadano a través de viajes en avión y tren).

Por otra parte, estas preocupaciones también fueron atendidas muy recientemente por la Corte Suprema de Estados Unidos, al resolver el caso "CARPENTER v. United States" (rta. el 22/06/2018), en el que se estableció por mayoría que para obtener las celdas de localización de un teléfono celular en una investigación criminal, se requiere una orden judicial de registro, que se adecue al estándar probatorio de la "expectativa razonable de privacidad", ya que se trata de una medida que impacta en la garantía consagrada por la Cuarta Enmienda.

De hecho, la Corte de Estados Unidos concluyó que para que procediera la medida en se requería el mismo estándar de convicción que para la emisión de una orden de

[8] Al respecto, ver https://www.p-lib.es/derechos-y-libertades/sms-invisibles; y en el mismo sentido, https://www.genbeta.com/activismo-online/la-policia-usa-mensajes-de-texto-invisibles-para-localizar-y-rastrear-moviles

intervención telefónica o un allanamiento. Sostuvo así que no alcanzaba con los "motivos suficientes" dijo que se requería una orden judicial de registro fundada en "causa probable"[9].

El Juez Roberts, actual Presidente de dicho tribunal, lideró el voto de la mayoría, reconociendo justamente que el desarrollo de la tecnología requiere que se encuentren formas de preservar la privacidad de los ciudadanos de las injerencias del estado, frente a herramientas de investigación que permiten a las autoridades acceder a áreas que normalmente estaban fuera de la vista de los investigadores.

En particular, enfatizó que el hecho de que la información en cuestión obrara en poder de las empresas prestatarias del servicio como consecuencia de una entrega voluntaria de parte de los usuarios, no alcanzaba para interpretar una pérdida de interés en la privacidad de esa información.

Máxime si se tiene en cuenta que la información que almacenan las empresas no es información que el usuario les "entrega", en el sentido estricto del término, ya que no existe ningún acto afirmativo del usuario que habilite a que las empresas de telecomunicaciones almacenen y suministren sus datos de localización, sino que los dispositivos de telefonía móvil van generando esos registros automáticamente para permitir el acto de comunicación -o incluso en total ausencia de un acto comunicacional-, siendo que es esta última, y no otra, la finalidad propia de dichos dispositivos.

A pesar de que la localización de los usuarios a través de las antenas de las empresas de telecomunicaciones puede constituir un elemento necesario para la operatividad del servicio, de todas maneras el servicio al que suscribe un usuario que utiliza un teléfono móvil es el de servicio de telecomunicaciones y no el de geolocalización.

En función de ello, creo que es posible sostener válidamente que la sociedad en su conjunto tiene una expectativa legítima de que las agencias estatales, a través de las entidades prestatarias del servicio de comunicaciones, no monitoreen sus movimientos.

Con relación a esto último, resulta especialmente interesante recordar que en sus anteriores precedentes vinculados a la interpretación del alcance de la garantía consagrada por la Cuarta Enmienda de la Constitución, la Corte estadounidense había establecido la denominada "third party doctrine" ("doctrina de terceros") según la cual los individuos no se encontraban protegidos por dicha cláusula constitucional respecto de los registros que eran de titularidad o que eran controlados únicamente por un tercero. sí, en el caso "MILLER" (425 U. S., at 437438) referido a la obtención de registros bancarios de un ciudadano sin orden judicial, y en "SMITH" (442 U. S., at 737) referido a los registros de las llamadas salientes de un teléfono fijo, obtenido también sin orden judicial, el tribunal había interpretado que las autoridades no habían registrado nada que efectivamente perteneciera a los sujetos investigados.

---

[9] Los investigadores habían accedido a los registros de las celdas en virtud de una orden judicial prevista por la "Stored Communications Act", que requería que se demostrara que existían "motivos fundados" para creer que los registros eran relevantes para la investigación en curso (18 U. S. C. §2703(d). Esa demostración exige un estándar probatorio muy inferior que el de "causa probable" requerido para una orden de registro ("United States v. Martínez-Fuerte", 428 U. S. 543, 560–561 (1976). Bajo el estándar previsto por la SCA, en cambio, sólo se debía acreditar que la prueba de las celdas de geoposicionamiento podía ser relevante para una investigación en curso.

El fundamento de la aplicación de esta doctrina en este último caso fue que cuando un individuo realiza una llamada, voluntariamente habilita a que la compañía prestadora del servicio accediera a esos registros que hacen a la corriente diaria del servicio que la empresa presta.

Sin embargo, al revisar la naturaleza de la medida vinculada con la solicitud de registros de celdas de geoposicionamiento, tuvo especialmente en cuenta que el nivel de especificidad y detalle de la información que permite acceder esta prueba es muy parecida a la del GPS de un vehículo, y que no requiere ningún esfuerzo producirla.

El propio Juez Roberts reconoció en su voto que cuando se gestó la "third party doctrine" (1979) nadie podía imaginar una sociedad en la cual las personas vamos acompañadas de un dispositivo móvil con las características actuales, suministrando a las compañías involucradas en el sector, entre ellos los de telefonía celular, como en el caso, no sólo los registros de las llamadas que realiza, sino también un detallado registro de muchísimos de los movimientos de las personas.

En definitiva, lo interesante pasa por admitir que el hecho de que cierta información personal de un usuario se encuentre almacenada por las empresas de telecomunicaciones, por sí solo no permite sortear el derecho del usuario de reclamar la protección de su privacidad y la obligación positiva del estado en protegerla con los debidos contralores para acceder a ella.

Esta conclusión se ve reforzada, por el hecho de que la geolocalización de un individuo ni siquiera forma parte del servicio específico que la empresa le presta a los ciudadanos.

Frente a la solicitud de determinación de las celdas correspondientes a un teléfono móvil, los desarrollos tecnológicos modernos cuentan con aptitud suficiente como para revelar tanto el lugar en cuyas cercanías se encuentra un dispositivo de telefonía móvil en un momento determinado, como las ubicaciones en las que ha estado en momentos anteriores, o bien el recorrido que estimativamente podría atravesar en el futuro.

Si bien este tipo de medida generalmente apunta a conocer los datos acontecidos en el pasado, lo cierto es que la tecnología actual permite un grado de especificidad tal que incluso resultaría posible requerir datos referidos a tiempo real, que obran en las bases almacenadas de la empresa prestataria del servicio.

Como ya señalamos, la generalizada utilización del teléfono móvil por parte de la mayoría de los miembros de las comunidades en la sociedad moderna, y la posibilidad de obtener una georreferenciación constante con independencia del conocimiento y de la voluntad de los usuarios, permite afirmar que la localización a través de la telefonía móvil constituye una de las herramientas más modernas para situar personas en el espacio en un momento determinado, desplazando otras clases de vigilancia[10].

[10] Así, por caso, en el sitio https://www.zeit.de/datenschutz/malte-spitz-data-retention puede visualizarse un ejemplo impactante de cómo la información aportada por las empresas de telefonía móvil permiten observar todos los movimientos de un diputado alemán del Partido Verde durante un período de tiempo prolongado. Ese monitoreo se realizó a partir del entrecruzamiento de datos de localización que poseen almacenadas las empresas prestatarias del servicio de telefonía móvil y otras fuentes de información disponibles públicamente

4 A modo de conclusión

No cabe dudas que frente a la aparición de nuevas formas de criminalidad que requieren de especiales técnicas de investigación y de nuevos medios de prueba, y frente a la irrupción de nuevas tecnologías de la información, es indispensable no sólo redimensionar el alcance de la garantía de la intimidad para extender el ámbito de protección desde el mundo físico hacia el entorno digital, sino también el alcance de la labor jurisdiccional.

El rol del juez de garantías en la etapa de investigación no puede quedar reducido a la simple emisión de órdenes de allanamientos, requisas o interceptaciones de comunicaciones, ya que estas medidas de prueba fueron ideadas exclusivamente para investigación de hechos acontecidos en el mundo físico y no en el mundo digital.

De lo contrario, frente al auge de nuevas técnicas de investigación y las nuevas formas de vigilancia, quedaría absolutamente desbalanceado el esquema de distribución de competencias entre el órgano de acusación encargado de impulsar la investigación y el órgano jurisdiccional encargado de decidir y de velar por las garantías del imputado con carácter previo a la consumación de la injerencia estatal.

En este último esquema, las posibilidades reales de ejercer el rol jurisdiccional de velar por el respeto por las garantías frente a una serie de medidas que tienen la potencialidad de impactar en un altísimo grado sobre la vida privada de las personas, quedaría reducido a una mínima expresión.

En síntesis, antes del desarrollo de internet y del auge de la sociedad de la información y del conocimiento, las fronteras de la privacidad eran tangibles, y estaban definidas espacio-temporalmente, por ciertas barreras de carácter físico. Sin embargo, en la sociedad moderna esas barreras tradicionales se fueron tornando progresivamente más difusas, por lo que es necesario dimensionar la relevancia e impacto que tiene esta cuestión para poder redefinir el alcance con el que tradicionalmente se interpretaron ciertas garantías, como así también el rol de cada uno de los operadores del proceso frente a los nuevos medios de prueba disponibles en la sociedad de la información y del conocimiento.

Es innegable la utilidad procesal que pueden tener ciertas medidas de prueba digital a las que nos referimos en este artículo, tales como el acceso al registro histórico de las direcciones de IP asignadas a determinado usuario, o el listado de celdas de conexión de un número investigado. Pero sería un gran error considerar que el argumento de la eficacia alcanza para legitimar a los investigadores a requerir esa información autónomamente, cualquiera sea el estado procesal de la causa y el grado de sospecha que haya logrado reunir respecto de la hipótesis delictiva que persigue.

Antes bien, la intervención jurisdiccional constituye una garantía necesaria en virtud del innegable impacto que esta clase de medidas generan desde la perspectiva de la intimidad.

El requisito referido a la exigencia de una orden judicial, como así también el estricto escrutinio al que se debe someter esta clase de medidas, en atención a su naturaleza y gravitación constitucional, no puede quedar relativizado por la gravedad o el

carácter aberrante de los delitos que se persiguen, ni mucho menos por criterios de pragmatismo o de celeridad.

En este sentido, siguen vigentes a pesar del tiempo transcurrido, las reflexiones del ex Ministro Petracchi de la CSJN al emitir su voto en el conocido precedente referido a la garantía de la inviolabilidad del domicilio, "FIORENTINO", rta. 27/11/1984, con remisión a un párrafo del Juez Frankfurter en el precedente "ESCOBEDO v. Illinois" (378, US, 478, p. 490), cuando señalaba: "Por medio de la declaración de Derechos, los fundadores de este país subordinaron la acción judicial a restricciones legales, no para conveniencia de los culpables sino para protección de los inocentes…" y "Podemos afirmar, con certeza, que el delito se combate con mayor eficacia cuando se cumplen rigurosamente los principios que han inspirado las restricciones constitucionales sobre la acción de los policías".

Siguiendo tales lineamientos, por la trascendencia actual y potencial de la medida desde la perspectiva de la intimidad, creemos que el pedido de datos de geoposicionamiento o el acceso a los logs de conexión de un usuario pueden proceder únicamente por orden judicial, luego de someterla a un escrutinio estricto, siempre que haya causa probable de la comisión de un delito y de la posible participación en el hecho de determinada persona, y en la medida que la prueba resulte necesaria y proporcional en atención a la gravedad del delito investigado

3° Conferencia Nacional de Informática Forense

INFOCONF 2019

FCEFyN

Facultad de Ciencias Exactas, Físicas y Naturales

Construyendo Educación Pública

2019 | Año de la Exportación

ISBN 978-950-33-1553-8