Análisis Forense a correos electrónicos en Outlook

Jorge Andrés Ortiz Castro, Leidy Gisell Bautista Montilla
Universidad Piloto de Colombia
Bogotá, Colombia.
Jaoc21@gmail.com
Leidybautista4@hotmail.com

Resumen — La búsqueda, parametrización y reconstrucción de correos es fundamental ante casos como denuncias, insultos, amenazas, etc. en el que el medio utilizado es el e-mail o correo electrónico. El uso de correos electrónicos para envío y recepción de mensajes es la actividad más frecuente y masiva en la Gran Telaraña Mundial (INTERNET), no solo para enviar mensajes y anexos a nivel personal sino también a nivel empresarial. Pocas herramientas son capaces de reconocer los formatos más usuales que se utilizan en clientes y servidores de correo electrónico. Quizás la suite comercial más completa es FTK. Mientras que en el mundo Open, son muy pocas las que existen y dan cobertura.

Uno de los gestores de correos más empleados, especialmente por empresas públicas y privadas es el programa Microsoft Outlook, que al igual que otras aplicaciones de Microsoft Office también tiene sus diferentes versiones, en este caso se enfocará en la versión 2010.

Abstract — The search, configuration and post reconstruction is essential in cases such as complaints, insults, threats, etc. wherein the medium used is the e-mail or mail. The use of emails for sending and receiving messages is the most frequent and massive in the World Wide Web (INTERNET) activity, not only to send messages and attachments on a personal level but also at the business level. Few tools are able to recognize the most common formats used on clients and email servers. Perhaps the most complete business suite is FTK. While in the Open world, very few that exist and provide coverage.

One of the mailer managers post more used, especially public and private companies is Microsoft Outlook, which like other Microsoft Office applications also has different versions, in this case will focus on the 2010 version program.

Índice de Términos — Cadena de Custodia, Computación Forense, Datos Activos, Datos Borrados, Datos Recuperados, Digitalizar, Extensión de un archivo, Extraer, Firma Digital, Imagen Forense, Rastros digitales.

I. INTRODUCCIÓN

Debido al uso masivo que se dan a los correos electrónicos a través de internet, este se convirtió en un elemento esencial a la hora de enviar mensajes y diversos tipos de archivos, pero por desgracia también surgieron diversas amenazas en los diferentes ámbitos.

En el caso de los correos existen amenazas como interceptación y apropiación de información sensible o confidencial de la empresa, recepción de falsos correos electrónicos, correos Phishing, correos Spam, correos con virus adjuntos en imágenes (esteganografía), entre otras.
Hoy en día son muchos los correos electrónicos que se emplean como elementos probatorios que pueden incriminar a un sospechoso o resolver una determinada situación, pero que al final terminan cayéndose en un juicio por no salvaguardarlos o no tomar las medidas pertinentes.

Los objetivos que se deben alcanzar con el análisis forense de correos electrónicos en Outlook podemos considerar:

Analizar los diferentes riesgos y amenazas que se pueden presentar en el manejo y administración de los correos electrónicos Outlook y que se pueden materializar en un incidente.

Establecer los tipos de métodos y acciones que se emplearán en los correos electrónicos como elementos probatorios en un caso.

Comprender con gran detalle la composición y funcionamiento del gestor de correos Outlook.

II. DESARROLLO

Microsoft Outlook [1] es un programa de organización ofimática y cliente de correo electrónico de Microsoft, y forma parte de la suite Microsoft Office.
Puede ser utilizado como aplicación independiente para trabajar día y noche o con Microsoft Exchange Server para

[1] Tomado de Microsoft Outlook Página Web - www.**microsoft**.com/en-us/**outlook**-com

dar servicios a múltiples usuarios dentro de una organización tales como buzones compartidos, calendarios comunes, etc.

A. Características de Outlook.

Microsoft Outlook es una aplicación de gestión de correo, así como agenda personal, que nos permite la comunicación con miles de personas en todo el mundo a través de mensajes electrónicos. Entre las características más comunes tenemos:

Administrar varias cuentas de correo electrónico desde un único lugar.

Puede administrar fácilmente los mensajes de correo electrónico de varios buzones. Sincronizar varias cuentas de correo electrónico de servicios como Hotmail®, Gmail® o de prácticamente cualquier otro proveedor con Outlook 2010®.

Administrar fácilmente grandes volúmenes de correo electrónico y personalizar tareas comunes en comandos de un solo clic.

Búsquedas para encontrar fácilmente lo que requiere.
Crear mensajes de correo electrónico que llamen la atención. Por medio de herramientas de office.

B. Carpeta personal de archivos Outlook.

Microsoft Outlook emplea el folder personal (PFF (Personal Folder File)) para almacenar e-mails, citas, tareas, contactos, notas, entre otros.

Existen tres (3) diferentes tipos de PFF conocidos:

El **libro personal de direcciones (PAB),** el cual contiene el libro de direcciones de los contactos. Estos archivos tienen la extensión **.pab.**

La **tabla personal de almacenamiento (PST),** que contiene elementos como mensajes de correo electrónico, citas, tareas, notas, etc., y se usa como archivos de buzones actuales y archivados. Estos archivos tienen la extensión **.pst**. El formato PST también se conoce como el formato de archivo de carpetas personales (PFF).

La **tabla de almacenamiento sin conexión (OST)**, que contiene elementos como mensajes de correo electrónico, citas, tareas, notas, etc., y se utiliza como fuera de línea en los archivos del buzón junto con Microsoft Exchange. Estos archivos tienen la extensión **.Ost**. El formato OST también se conoce como el formato de archivo de carpetas sin conexión (OFF).

El formato de archivos subyacente de estos archivos es el mismo que el nombre real es desconocida, pero se ha conocido como el formato de archivo de carpetas personales (PFF), debido a su uso más común.

Nota: El archivo de almacenamiento de Outlook Express es .dbx, pero debido a que fue descontinuado en el año 2005, y dejó de ser operativo en el año 2009.

III. OUTLOOK PST

Nombre. outlook.pst — formato de Microsoft Outlook .pst file.

Arquitectura lógica de un archivo PST. Las estructuras de archivos PST están dispuestos lógicamente en tres capas: la capa NDB (Base de datos del nodo), el (listas, tablas y propiedades) capa de LTP y la capa de mensajería.

El siguiente diagrama muestra la jerarquía lógica de estas capas, y qué abstracciones son manejados por cada capa.

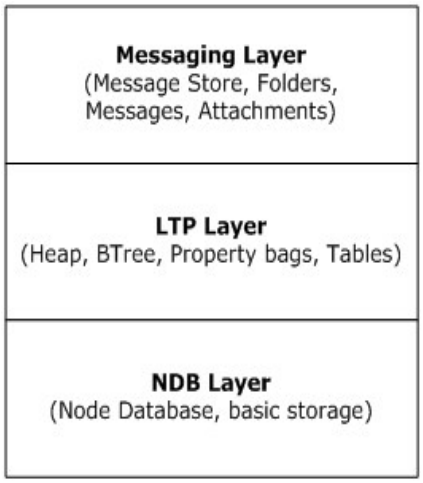

Figura 1. Arquitectura lógica de un archivo pst.

A. Organización Física de un formato de archivo PST.

Esta sección proporciona una visión general de la distribución física de un archivo pst. El siguiente diagrama muestra la organización de archivos de alto nivel de un PST.

Este formato de archivo se organiza con un elemento de encabezado seguido de asignación de páginas de información en intervalos regulares que se entremezclan con los bloques de datos extensibles. La sección de encabezado incluye metadatos acerca de la PST y la información que señala a las secciones de datos que contienen el almacén de mensajes y su contenido. Las siguientes secciones cubren cada uno de estos elementos con más detalle.

Encabezado (Header). La cabecera reside en el principio del archivo, y contiene tres grupos principales de información: Metadata, registro raíz, y el mapa libre inicial (FMap) y el mapa de la página libre (FPMap). (Ver Figura 2).

Metadatos y estado del archivo PST. Los metadatos incluyen información tal como números de versión, sumas de comprobación, contadores persistentes, y las tablas del espacio de nombres. Usando esta información, una aplicación puede determinar la versión y el formato del archivo PST, que determina la disposición de los datos posteriores en el archivo.

Registro Root. El disco raíz contiene información sobre los datos reales que se almacenan en el archivo PST. Esto incluye la raíz de la NBT y BBT, tamaño y la información de asignación necesaria para gestionar el espacio libre y el crecimiento de los archivos, así como información de integridad de los archivos.

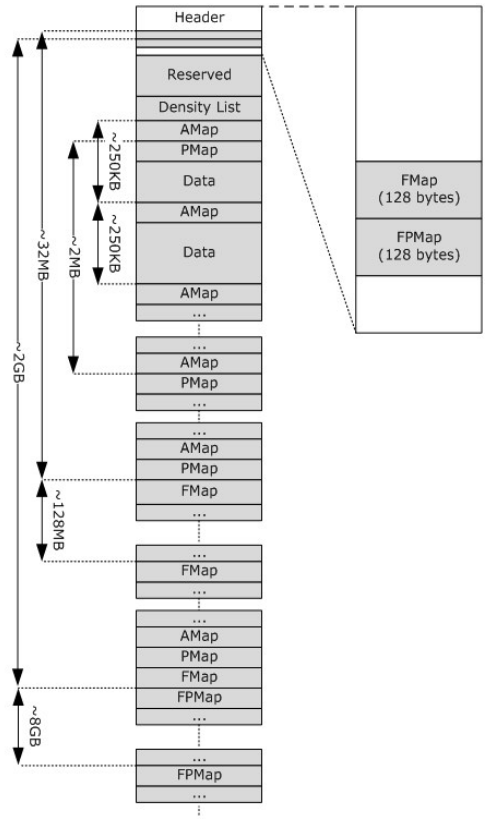

Figura 2. Diagrama de un Archivo PST.

Libre mapa inicial (FMap) y la página de mapas gratuita (FPMap). Mapas gratis (FMAP) y mapas gratis Página (FPMaps) se utilizan para buscar el espacio libre contiguo dentro de un archivo PST. <1> FMAP y FPMaps.

Datos reservados. Un número de octetos se han reservado entre el final de la cabecera y el inicio de la Lista de Densidad (DList). Una parte de este espacio se reserva para la futura expansión de la estructura de encabezado del archivo PST, mientras que el resto se reserva para la persistencia, los datos específicos de la implementación transitorios.

Lista Densidad (DList). La Lista de densidad consiste en una lista ordenada de referencias a la asignación Mapa páginas (AMAP).Se ordena en orden ascendente de densidad (es decir, al descender la cantidad de espacio libre disponible). Su función es optimizar la asignación de espacio de modo que el espacio mencionado por páginas con la más abundante espacio libre (es decir, la densidad más baja) se asigna primero. Sólo hay una DList en el PST, que siempre se encuentra en una posición desplazada en el archivo PST fija. Para obtener más detalles acerca de los detalles técnicos de la DList.

Mapa de asignación (AMAP). Una página del Mapa de asignación de una página de tamaño fijo que se utiliza para realizar el seguimiento del estado de asignación de la sección de datos que sigue inmediatamente a la página AMAP en el archivo. La página entera AMAP puede ser visto como una matriz de bits, donde cada bit corresponde al estado de la asignación de 64 bytes de datos. Una página AMAP aparece más o menos cada 250 kilobytes en el PST.

Página Mapa (PMAP). Un mapa de páginas es un bloque de datos que es de 512 bytes de tamaño (incluida la parte superior), que se utiliza para el almacenamiento de casi todos los metadatos en el PST (es decir, el BBT y NBT). El pmap se crea para optimizar la búsqueda de páginas disponibles. El pmap es casi idéntica a la AMAP, excepto que cada bit de los mapas pmap el estado de asignación de 512 bytes en lugar de en lugar de 64, ya que cada bit de la pmap cubre ocho veces los datos de un AMAP, una página pmap aparece aproximadamente cada 2 megabytes (o una pmap por cada ocho AMaps).

Sección de Datos. Secciones de datos son conjuntos de datos alrededor de 250 kilobytes de tamaño que contienen asignaciones. Cada asignación individual está alineada con un límite de 64 bytes, y es en tamaños que son múltiplos de 64 bytes. Todos los elementos mencionados por el BBT se asignan de estas secciones de datos. Secciones de datos están representadas por los bloques marcados con "datos".

Libre mapa (FMap). Una página FMap proporciona un mecanismo para localizar rápidamente el espacio libre contiguo. Cada byte en el FMap corresponde a una página AMAP. El valor de cada byte indica el número de bits más largas y libres que se encuentran en la página AMAP correspondiente. Debido a que cada bit en los mapas AMAP a 64 bytes, el FMap contiene la máxima cantidad de espacio libre contiguo en el que AMAP, hasta cerca de

16 kilobytes. En general, debido a que cada AMAP cubre alrededor de 250 kilobytes de datos, cada página FMap (496 bytes) cubre alrededor de 125 megabytes de datos.
Sin embargo, existe un caso especial para el FMap inicial, el encabezado contiene una FMap inicial, que está a sólo 128 bytes, y que cubre los primeros 32 megabytes de datos.

Mapas Página Free (FPMap). Un FPMap es similar a la FMap excepto que se utiliza para encontrar rápidamente páginas libres. Cada bit de la FPMap corresponde a una página pmap, y el valor del bit indica si hay alguna página libres dentro de esa página PMAP. Con cada pmap abarca alrededor de 2 megabytes, y una página FPMap a 496 bytes, se deduce que una página FPMap abarca aproximadamente el 8 gigabytes de espacio.
Sin embargo, existe un caso especial para el FPMap inicial, el encabezado contiene una FPMap inicial, que está a sólo 128 bytes, que cubre los primeros 2 gigabytes de datos.
Archivos PST ANSI sólo contienen la FPMap inicial en la cabecera y hay páginas FPMap adicionales. Esto limita el tamaño de un archivo PST ANSI a alrededor de 2 gigabytes.

Relación con los protocolos y otras estructuras. Este formato de archivo utiliza estructuras descritas en [MS-OXCDATA] y etiquetas propiedades descritas en [MS-OXPROPS].

Declaración de Aplicabilidad. Este formato de archivo permite a los implementadores a leer y escribir archivos PST que son compatibles con otras implementaciones de este protocolo.

Visión de conjunto. Nivel bajo o elementos primitivos en un archivo pst. Son identificados por un valor I_ID. Mayor nivel o elementos compuestos en un archivo pst. Son identificados por un valor d_id. Hay dos árboles B separados indexados por estos valores I_ID y d_id. A partir de Outlook 2003, el formato de archivo ha cambiado desde que tiene 32 bits de punteros, a una con 64 bits de punteros.

Para llevar a cabo la visualización del encabezado del archivo pst, simplemente lo abrimos en cualquier editor hexadecimal.

Archivo del encabezado de 32 bits. El archivo del encabezado de 32 bits está ubicado en el offset cero (0) en el archivo pst. Al abrir el archivo pst en el editor hexadecimal obtenemos lo siguiente:

```
0000  21 42 44 4e 49 f8 64 d9  53 4d 0e 00 13 00 01 01
0010  00 00 00 00 00 00 00 00  50 d6 03 00 bd 1e 02 00
0020  08 4c 00 00 00 04 00 00  00 04 00 00 0f 04 00 00
0030  0d 40 00 00 99 0a 01 00  18 04 00 00 0d 40 00 00
0040  0d 40 00 00 11 80 00 00  02 04 00 00 0a 04 00 00
0050  00 04 00 00 00 04 00 00  0f 04 00 00 0f 04 00 00
0060  0f 04 00 00 0d 40 00 00  00 04 00 00 00 04 00 00
0070  04 40 00 00 00 04 00 00  00 04 00 00 00 04 00 00
0080  00 04 00 00 00 04 00 00  00 04 00 00 00 04 00 00
```

```
0090  00 04 00 00 00 04 00 00  00 04 00 00 00 04 00 00
00a0  0c 09 00 00 00 00 00 00  00 04 27 00 00 24 23 00
00b0  c0 09 0a 00 00 c8 00 00  bc 1e 02 00 00 7e 0c 00
00c0  b4 1e 02 00 00 54 00 00  01 00 00 00 23 55 44 d1
00d0  5a 4f ce 6b 80 ff ff ff  00 00 00 00 00 00 00 00
00e0  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
00f0  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
0100  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
0110  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
0120  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
0130  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
0140  00 00 00 00 00 00 00 00  00 00 00 00 3f ff ff ff
0150  ff ff ff ff ff ff ff ff  ff ff ff ff ff ff ff ff
0160  ff ff ff ff ff ff ff ff  ff ff ff ff ff ff ff ff
0170  ff ff ff ff ff ff ff ff  ff ff ff ff ff ff ff ff
0180  ff ff ff ff ff ff ff ff  ff ff ff ff ff ff ff ff
0190  ff ff ff ff ff ff ff ff  ff ff ff ff ff ff ff ff
01a0  ff ff ff ff ff ff ff ff  ff ff ff ff ff ff ff ff
01b0  ff ff ff ff ff ff ff ff  ff ff ff ff ff ff ff ff
01c0  ff ff ff ff ff ff ff ff  ff ff ff ff 80 01 00 00
01d0  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
01e0  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
01f0  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
```

```
0000  firma(signature)                                  [4 bytes]
0x4e444221 constant
000a  Tipo de indice (indexType)                        [1 byte]
0x0e constant
01cd  Tipo de encripción (encryptionType)               [1 byte]
0x01 in this case
00a8  Tamaño total del archivo(total file size)         [4 bytes]
0x270400 in this case
00c0  backPointer1                                      [4 bytes]
0x021eb4 in this case
00c4  offsetIndex1                                      [4 bytes]
0x005400 in this case
00b8  backPointer2                                      [4 bytes]
0x021ebc in this case
00bc  offsetIndex2    [4 bytes] 0x0c7e00   in this case
```

Figura 3. Encabezado de un archive PST.

Sólo se admiten los tipos de índices 0x0E, 0x0F, 0x15 y 0x17, 0x00 y tipos de cifrado, 0x01 y 0x02. Índice de tipo 0x0e es el formato de Outlook de 32 bits mayor. Índice de tipo 0x0f parece ser rara, y hasta el momento los datos parece ser idéntica a la de los archivos tipo 0x0E. Índice de tipo 0x17 es el formato de Outlook de 64 bits más reciente. Índice de tipo 0x15 parece ser rara, y de acuerdo con el proyecto libpff debe tener el mismo formato que los archivos de tipo 0x17. Se encontró en un archivo pst de 64 bits creado por recuperación visual. Puede ser que los tipos de índice inferior a 0x10 son de 32 bits, y los tipos de índice mayor o igual a 0x10 son 64 bits, y el orden bajo cuatro bits del tipo de índice es cierto subtipo o número de versión secundaria.

En el tipo de cifrado 0x00 hay un cifrado tipo 0x01 que es la encriptación "compresible" y es un simple código de sustitución, y el tipo 0x02 es el cifrado "fuerte" que es un simple de tres rotor cifrado Enigma de la Segunda Guerra Mundial.

El OffsetIndex1 es el archivo de desplazamiento de la raíz del árbol b index1, que contiene (I_ID, desplazamiento, tamaño desconocido) tuplas de cada elemento en el archivo. backPointer1 es el valor que debe aparecer en el indicador principal de ese nodo raíz.

El OffsetIndex2 es el archivo de desplazamiento de la raíz del árbol b index2, que contiene (d_id, DESC-I_ID, TREE-I_ID, PADRES Y d_id) tuplas de cada elemento en el archivo. backPointer2 es el valor que debe aparecer en el indicador principal de ese nodo raíz.

IV. CONCLUSIONES

Este análisis permitirá saber qué medidas tomar para salvaguardar correos electrónicos que se podrían tomar como futuras pruebas en un caso.

Muchas de las modalidades de estafa en un correo electrónico como Phishing (Pesca), se materializan debido al desconocimiento general por parte del usuario.

Muchos de los correos electrónicos son de origen desconocido y muchos de ellos son spams o correos basura y su procedencia no siempre es del usuario original.

El establecimiento de las políticas de seguridad en una empresa es indispensable para concientizar a los usuarios en el uso correcto del correo electrónico para la detección de correos sospechosos que evitarían un incidente de seguridad.

V. BIBLIOGRAFIA

[1] COMPUTACION FORENSE. Descubriendo los rastros informáticos. Jeimy J. Cano M. Editorial Alfaomega, 2009.

[2] WINDOWS FORENSIC ANALYSIS. Harlan Carvey. Editorial Syngress, 2007.

[3] REVISTA ENTER. Edición 140, Planeta digital, Noviembre de 2010.

[4] WIKIPEDIA. PáginaWeb. Microsoft Outlook. https://es.wikipedia.org/wiki/Microsoft_Outlook.

[5] INFORME BASE PRUEBA PERICIAL. Adalid Abogados Corp. Página Web http://adalidabogados.com/material/informe%20base%20prueba%20pericial.pdf

[6] ANÁLISIS FORENSE DE SISTEMAS INFORMÁTICOS. Helena Rifà Pous (coordinadora), Jordi Serra Ruiz (coordinador), José Luis Rivas López, 2009. http://webs.uvigo.es/jlrivas/downloads/publicaciones/Analisis%20forense%20de%20sistemas%20informaticos.pdf