Estudio comparativo de desempeño de herramientas para el análisis forense de correos electrónicos

Esteban A. Rivetti[1], Beatriz P. de Gallo[2]
[1] *IEsIIng /Facultad de Ingeniería, Universidad Católica de Salta*
Campo Castañares S/N, Salta, Argentina
earivetti@ucasal.edu.ar
[2] *IEsIIng /Facultad de Ingeniería, Universidad Católica de Salta*
Campo Castañares S/N, Salta, Argentina
bgallo@ucasal.edu.ar

Abstract

Este trabajo intenta dar respuestas a los interrogantes que comúnmente se observan cuando se solicita el análisis forense de un correo electrónico. Para ello, se consideran distintas herramientas forenses, las cuales son evaluadas y analizadas en dos escenarios de trabajo, para poder determinar e identificar las más eficientes valorando principalmente los resultados obtenidos y la performance que presentan.

1. Introducción

El correo electrónico es un método de comunicación entre dos partes, emisor y receptor. El mensaje recorre un camino, desde un servidor a otro hasta llegar a su destino. En medio del recorrido existen varios actores que cumplen distintas funciones para lograr el objetivo final. Ahora bien, no es simple poder identificar y dar respuestas a diversos interrogantes que se solicitan en una pericia de correo electrónico.

Para simplificar esto, existen distintas herramientas de análisis forenses de correo electrónicos.

En este trabajo se intenta dar respuesta a estos interrogantes haciendo una comparación exhaustiva de las herramientas, enfocándose en la fase de análisis de encabezados y en la performance de cada una de ellas.

Se analizarán seis herramientas, las cuales son las más populares y utilizadas en el ámbito de las pericias informáticas, *Aid4mail* [1], *Email TrackerPro* [2], *Mail Navigator* [3], *Osforensics*[4], *E-mail Ex*aminer[5] y *MailXaminer*[6][1].

Este trabajo se organiza de la siguiente manera: el apartado 2 aborda una breve descripción de la ontología para el análisis forense de correos electrónicos. La sección 3 se presentan las herramientas que se consideraran en este trabajo, en la sección 4 se describe los 2 casos de estudio que se tomaron como ejemplo, en la primera parte se trabajará con un correo electrónico, y luego con un conjunto de correos de una bandeja en particular. En la sección 5 se detallan las conclusiones arribadas.

2. Ontología para el análisis forense de un correo electrónico

El análisis forense no debe presentarse como un reporte técnico sino como información sistemática y con sentido semántico en el marco de la causa judicial [7].

En el trabajo citado los autores plantean como objetivo el de contar con un marco de referencia común apoyándose en tecnologías semánticas, enfocándose en la trazabilidad de un correo electrónico, es decir, en todo el camino que recorre un correo electrónico hasta llegar a su destino final.

En particular, introducen una ontología que define los principales conceptos y relaciones que representan este camino el cual se describe sintéticamente en la Figura 1.

De estos componentes interesan aquellos referidos a los datos que las herramientas forenses identifican de manera automática, como ser: fechas, direcciones IP, nombres de cuentas.

3. Herramientas de análisis forenses de correos electrónicos

Antes de iniciar con el estudio, se presentan las herramientas con las que se trabajaran, las cuales son las más utilizadas en el ámbito forense.

Aid4Mail soporta más de 40 formatos de correo electrónico y programas de cliente de correo, así como muchos servicios populares de correo web y cuentas remotas a través de IMAP.

Como resultado, los desarrolladores de esta marca indican que es posible procesar prácticamente cualquier tipo de buzón que llegue a su destino.

Las carpetas y archivos de correo local se pueden procesar fácilmente cuando se desconectan de su cliente

[1] Si bien existen herramientas especializadas y de grandes capacidades para el análisis forense, el presente estudio se orienta solo a aquellas de uso libre y disponibles en internet.

de correo electrónico, incluidos los almacenados en discos duros externos y medios como DVD y dispositivos USB. *Aid4Mail* puede leer archivos mbox de sistemas Mac y Linux sin conversión previa.

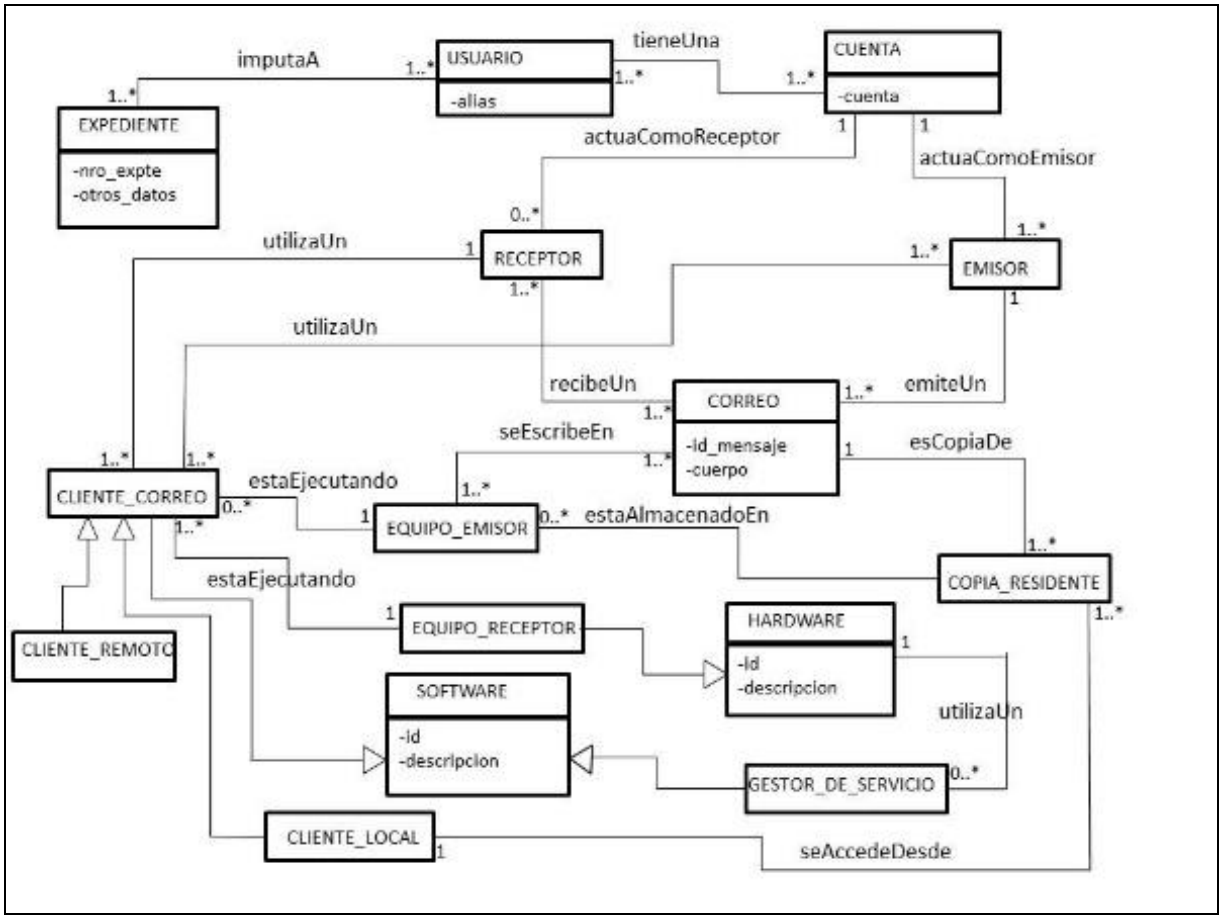

Figura 1: Vista Parcial de la ontología propuesta en [7]

EmailTrackerPro no sólo ofrece la capacidad de rastrear un correo electrónico usando el encabezado de correo electrónico, sino que también viene con un filtro de spam (edición avanzada), que escanea cada correo electrónico a medida que llega y advierte al usuario si se sospecha de spam. Un encabezado de correo electrónico contiene toda la información necesaria para rastrear su procedencia. Tiene la huella de cada servidor a través de los cuales pasó el correo electrónico, y puede brindar información sobre el lugar de origen del correo electrónico. Una vez identificada la dirección IP[2] del servidor, es posible encontrar al propietario o responsable de ese servidor, mediante la información Whois[3], quien provee la información de contacto para la organización que registró y es responsable de la dirección o sitio web que se está rastreando. La característica más valiosa de *EmailTrackerPro* es la capacidad de rastrear más de una dirección IP o nombre de dominio a la vez. Se puede trazar tantas direcciones IP y nombres de dominio como sea necesario y se envían los resultados a una nueva pestaña o un archivo Excel / HTML.

MailNavigator fue creado a partir de dos herramientas para la lectura de email y canales de noticias: FILTER, que es un poderoso sistema para la búsqueda de correos en ficheros de los distintos programas de e-mail; y NAVIGATOR, que es un lector de correos y noticias con funciones avanzadas.

OSForensics permite extraer pruebas forenses de computadoras rápidamente con búsquedas e indexación de archivos de alto rendimiento. Puede identificar archivos sospechosos y actividad con coincidencia hash, comparaciones de firmas de unidad, correos electrónicos, memoria y datos binarios. También permite administrar

[2] Una dirección IP es un número que identifica, de manera lógica y jerárquica, a una Interfaz en red (elemento de comunicación/conexión) de un dispositivo (computadora, tableta, portátil, smartphone) que utilice el protocolo IP (Internet Protocol)

[3] WHOIS es un protocolo TCP basado en petición/respuesta que se utiliza para efectuar consultas en una base de datos que permite determinar el propietario de un nombre de dominio o una dirección IP en Internet.

una investigación digital y crear informes de datos forenses recopilados.

La empresa creadora de *E-mail Examiner* ha sido una de las pioneras en soluciones para dispositivos móviles, smartphones y correo electrónico, y su enfoque de trabajo en la movilidad le permitió avanzar en muchas otras áreas de la innovación incluyendo la investigación y el desarrollo en el Internet de Cosas (IoT) con el Forensics of Everything TM (FoE). *E-mail Examiner* permite analizar los encabezados, los cuerpos y los archivos adjuntos de los correos electrónicos. Analiza el mensaje de principio a fin, incluyendo la clasificación y análisis detallado de archivos adjuntos. Soporta los principales tipos de correo electrónico que se almacenan en equipos locales para análisis, generación de informes y exportación/conversión de datos.

Systools Software, creador de *MailXaminer*, se dedica al negocio de proporcionar herramientas de alta tecnología con interfaz de usuario amigable. Ha contribuido con la recuperación de datos, soluciones de copia de seguridad, así como herramientas forenses de investigación y análisis de correo electrónico. El primer lanzamiento importante realizado en el campo de las aplicaciones de eDiscovery fue *MailXaminer*, es un conjunto completo de herramientas para la documentación, análisis, examen y notificación de evidencias de correo electrónico.

4. Casos de estudio

Para realizar el estudio propuesto se consideran dos casos de estudio: el primero con el análisis de un único correo electrónico y el segundo, con un conjunto de correos electrónicos de una cuenta en particular.

4.1. Análisis de un único correo electrónico

En primera instancia se evaluó un correo en particular, y se lo analizó con cada una de las herramientas. El correo fue descargado desde una cuenta con dominio gmail y fue convertido a extensión txt.

La Figura 2 muestra la cabecera del correo electrónico de ejemplo.

```
x-store-Info:sbevkl2QZR7OXo7WID5ZcdV2tiiWGqTn+TqXcEmOv5qA/2pYAZ9atdTRoF2b9
UsV5+ovG653QDIg/PBof6bNpNSdzThJNZwHB1bM5P4ejIltoXvCRVaY1REwin7oM4eZiPC
Authentication-Results: hotmail.com; spf=pass (sender IP is 209.85.218.45;
identity alignment result is pass and alignment mode is relaxed)
 smtp.mailfrom=beagallo@gmail.com; dkim=pass (identity alignment
result is pass and alignment mode is relaxed) header.d=gmail.com; x-hmca=pass
 header.id=beagallo@gmail.com
X-SID-PRA: beagallo@gmail.com
X-AUTH-Result: PASS
X-SID-Result: PASS
X-Message-Status: n:n
X-Message-Delivery: Vj0xLjE7dXM9MDtsPTA7YT0xO0Q9MTtHRD0xO1NDTD0w
X-Message-Info: Received: from mail-oi0-f45.google.com ([209.85.218.45]) by
COL004-MC1F49.hotmail.com over TLS secured channel with Microsoft SMTPSVC(7.5.7601.23143);
      Sun, 3 Apr 2016 02:15:56 -0700
Received: by mail-oi0-f45.google.com with SMTP id w85so232434oiw.0;
      Sun, 03 Apr 2016 02:15:56 -0700 (PDT)
DKIM-Signature: v=1; a=rsa-sha256; c=relaxed/relaxed;
      d=gmail.com; s=20120113;
      h=mime-version:reply-to:date:message-id:subject:from:to;
      bh=bgnN7sU9UwNlq/GoA2dI5v4r18+YnqvHqOhDPqxsISs=;
X-Google-DKIM-Signature: v=1; a=rsa-sha256; c=relaxed/relaxed=1e100.net; s=20130820;
      h=x-gm-message-state: mime-version:reply-to:date:message-id:subject
       :from:to;
      bh=bgnN7sU9UwNlq/GoA2dI5v4r18+YnqvHqOhDPqxsISs=;
      X-Gm-Message-State:
AD7BkJLjESAAOz7So1S0OgruZxoRF/gato0ZAbfeBU1NtdPzHShxP6dj4y1UziyRAiLtob+MPPlzw==
MIME-Version: 1.0
X-Received: by 10.202.169.212 with SMTP id s203mr5615498oie.35.1459674956500;
 Sun, 03 Apr 2016 02:15:56 -0700 (PDT)
Received: by 10.182.29.8 with HTTP; Sun, 3 Apr 2016 02:15:56 -0700 (PDT)
Reply-To: beagallo@gmail.com
Date: Sun, 3 Apr 2016 06:15:56 -0300
Message-ID:<CAH18OQWt2cqp955q_LekqxucpZLMj7BmwVTX-+hV4knBB-yNhA@mail.gmail.com>
Subject: =?UTF-8?Q?investigaci=C3=B3n?=
From: "Ing. H. Beatriz P. de Gallo" <beagallo@gmail.com>
To: josearaoz@hotmail.com, Esteban Rivetti <erivetti@hotmail.com>
Content-Type: multipart/alternative; boundary=001a113ce79e4547cd052f910fee
Return-Path: beagallo@gmail.com
X-OriginalArrivalTime: 03 Apr 2016 09:15:56.0824 (UTC) FILETIME=[6E226980:01D18D89]
```

Figura 2: Cabecera del correo electrónico

Los criterios seleccionados con los que se evaluaran las herramientas, se derivan del trabajo “Hacía una Ontología para el soporte de la trazabilidad del correo electrónico en la Forensia Digital” [7], los cuales buscan

responder los interrogantes de un punto de pericia que usualmente se solicitan en un análisis forense de un correo electrónico más otros datos relevantes para el proceso de comparación.

Los interrogantes a los que hacemos mención en el párrafo anterior, se pueden buscar en los puntos de pericia que usualmente se proponen al solicitar un análisis forense de un correo electrónico.

- De aquí surgen los criterios seleccionados para poder evaluar las herramientas propuestas.¿Cuáles son los componentes informáticos a través de los cuales se escribe y se lee un correo electrónico?
- ¿Cuál es la fecha, hora y dirección IP de emisión del correo electrónico?
- ¿Cuál es la fecha, hora y dirección IP de recepción del correo electrónico?
- ¿Cuál es el nombre de usuario y dirección de e-mail del Autor del mismo?
- ¿Cuál es el nombre de usuario y dirección de e-mail del Receptor del mismo?
- ¿Es posible establecer la trazabilidad del mensaje desde que se envía hasta que se recibe?
- ¿Cuáles son los diferentes actores/servicios que participaron de la transmisión?

Una investigación forense de correo electrónico puede incluir tanto la cabecera del correo electrónico como el cuerpo del mismo [8]. Este trabajo se enfocará en el examen de los datos de cabecera del correo electrónico.

Según Marwan [9] una investigación forense de correos electrónicos debe tener lo siguiente:

- El análisis del remitente dirección de correo electrónico.
- Análisis del mensaje de protocolo de inicio (HTTP, SMTP)
- El análisis e identificación de ID de mensaje
- Examinar la dirección IP del remitente

Tabla 1: Cuadro Comparativo de Resultados del Análisis del Correo Electrónico

	Aid4Mail	**eMailTrackerPro**	**Mail Navigator**	**OsForensics**	**E-mail Examiner**	**MailXaminer**
Determinación de IP	No	Si	No	No	No	Si
Determinación ID	No	No	No	No	Si	Si
Identificador del emisor	Si	Si	No	No	No	Si
Identificación del receptor	No	Si	No	No	No	Si
Determinación de Fecha y Hora	Si	Si	No	No	Si	Si
Visualización de los resultados	No	Si	No	Si	Si	Si
Opción de Búsqueda	Si	Si	Si	Si	Si	Si
Incluye Encabezados	No	Si	No	Si	Si	Si
Formatos de exportación	Si	Si	No	No	Si	Si
Dispositivos externos	Si	No	No	Si	No	Si
Admite bandejas de correos	Si	No	Si	Si	Si	Si
Capacidad de Recuperación	Si	No	No	Si	Si	Si
Usabilidad (Siendo 5 el valor más óptimo)	2	4	2	5	4	4

Tariq Banday [10] señala que una investigación forense es el estudio de la fuente y el contenido del mensaje de correo electrónico como evidencia para identificar el remitente real y el destinatario de un mensaje, los datos temporales de transmisión, el registro detallado de la transacción de correo electrónico, la intención del remitente, etc. El análisis forense implica la investigación de metadatos[4], búsqueda de palabras clave, exploración de puertos, etc. para la atribución de la autoría y la identificación de estafas o delitos realizados a través de correos electrónicos.

Al analizar las herramientas forenses de correos electrónicos, además de dar respuesta a los interrogantes que anteriormente mencionamos, es necesario evaluar otros criterios como la capacidad de procesamiento, la recuperación de correos eliminados, los formatos admitidos de entrada y salida, y la identificación de campos claves para el análisis de un correo electrónico como ser el ID del mensaje, ip intervinientes, fecha y hora del correo y la facilidad de uso para el usuario.

Para este caso de estudio se consideró el correo electrónico cuyo encabezado se muestra en la Figura 2, y se procedió a realizar el análisis forense con las seis herramientas citadas.

Los resultados obtenidos para cada herramienta se pueden organizar en base a las entidades señaladas en la ontología [7]:

- Determinación de IP
- Determinación ID
- Identificador del emisor
- Identificación del receptor
- Determinación de Fecha y Hora

[4] *Metadatos* no tiene una definición única. Según la definición más difundida de metadatos es que son datos sobre datos. El concepto de metadatos es análogo al uso de índices para localizar objetos en vez de datos. Por ejemplo, en una biblioteca se usan fichas que especifican autores, títulos, casas editoriales y lugares para buscar libros. Así, los metadatos ayudan a ubicar datos.

Otros elementos de interés para el análisis, son las características de las herramientas y su comportamiento. De esto surgen otras variables de análisis:

- Visualización de los resultados
- Opción de Búsqueda
- Incluye Encabezados
- Formatos de exportación
- Dispositivos externos
- Admite bandejas de correos
- Capacidad de Recuperación
- Usabilidad (Escala 1 al 5)[5]

Los resultados arribados se muestran en la Tabla 1.

Del análisis realizado y los resultados que se muestran en la Tabla 1, surgen distintas apreciaciones de cada una de las herramientas, las cuales se detallan a continuación:

- ***Aid4Mail***

La fortaleza de esta herramienta está en la capacidad para exportar en distintos formatos, aunque es muy complicada para utilizar y visualmente nada amigable con el usuario. No permite realizar análisis del encabezado de un correo en particular. Permite realizar análisis de correo electrónicos almacenados en el equipo como también en los servidores de correos remotos. La opción de búsqueda es muy completa.

- ***eMailTrackerPro***

La herramienta es completa, al realizar un análisis del encabezado construye una tabla con todas las ip por la que pasó el correo electrónico. Además, realiza de forma automática la identificación del propietario del dominio o IP y permite visualizar el encabezado de forma clara. Se puede generar un reporte con el resumen de los resultados. También puede comprobar si un correo electrónico sospechoso existe en las listas negras para protección contra el spam y correos electrónicos maliciosos. Asimismo, muestra si cualquier puerto está abierto en cualquiera de los protocolos HTTP o FTP en las direcciones IP de seguimiento [8].

- ***Mail Navigator***

Esta herramienta es un cliente de correo, que tiene como aspecto más notorio el sistema de búsqueda de palabras claves en un correo electrónico o en bandejas de entradas. Soporta varios formatos, aunque es una herramienta muy poco amigable al usuario.

- ***OsForensics***

Al analizar un correo electrónico con este software, se puede visualizar el cuerpo del correo y se puede seleccionar si se desea ver el encabezado. No brinda otra capacidad para la obtención de los datos necesarios para el análisis forense (dirección IP, cuentas, etc.).

- ***E-mail Examiner***

Al realizar el análisis de una bandeja de correos, se puede visualizar todos los correos. Al seleccionar uno, muestra información importante como ser el ID del mensaje, fechas y horas de los envíos. Es compatible con gran variedad de formatos. La información proporcionada por la herramienta abarca no sólo la cabecera del correo electrónico y el cuerpo, sino también el contenido del archivo adjunto [8].

- ***MailXaminer***

La herramienta consta de distintos filtros de búsquedas, palabras claves o frases. Se configura para conectarse a correos electrónicos de Gmail, Hotmail, Live Exchange Server. Tiene una función Email Hop, que permite analizar el recorrido de comunicación de un correo electrónico en particular, indicando los enrutadores y servidores identificándolos con la IP. Adicionalmente, gráfica en un mapa la ubicación física de cada servidor que intervino, e indica el recorrido que tuvo el correo.

Como conclusión del cuadro comparativo, se presentan las tres herramientas más completas y que obtienen los datos más relevantes del encabezado en un correo electrónico:

- *MailXaminer*
- *EMailTrackerPro*
- *E-mail Examiner*

Cada una de estas herramientas, son bastante completas, pero en el puesto número 1 y 2, están las herramientas que arman una tabla de ruteo e indican la IP de todos los servidores por lo que ha pasado un correo electrónico, además identificando el emisor y el receptor. La herramienta *MailXaminer*, tiene la ventaja de admitir dispositivos externos, recuperar los correos que fueron eliminados, e identificación del ID del mensaje, lo que hace que se encuentre en el primer lugar.

4.2. Análisis de múltiples correos electrónicos de una cuenta

Las herramientas que se van a analizar para este tipo de prueba, son las dos que mejor se comportaron en la primera parte y admiten análisis de correos masivos. Al ser ambas una versión trial, se debe tener en cuenta que tienen limitaciones funcionales.

En este caso, los criterios a evaluar, es la cantidad de correos identificados y el tiempo de procesamiento, para

5 La Usabilidad es la medida de la calidad de la experiencia que tiene un usuario cuando interactúa con una herramienta o sistema. Los criterios a evaluar serán la eficiencia en el uso de los diferentes elementos ofrecidos en las pantallas y la efectividad en el cumplimiento de las tareas que se pueden llevar a cabo a través de ellas.

complementar el cuadro comparativo del punto anterior. Se considerarán las herramientas *E-mail Examiner* y *MailXaminer*.

La bandeja de entrada a analizar tiene 1162 correos, es un archivo con extensión .pst y tiene un volumen de datos de 34.489 Kb. Ambas herramientas admiten este formato, por lo cual no es necesario ningún proceso ETL[6]

Luego de ejecutar el análisis forense de la cuenta ejemplo con ambos programas, los resultados obtenidos son los siguientes:

Criterio	*E-mail Examiner*	*MailXaminer*
Cantidad de Correos Identificados	1158	1162
Tiempo de procesamiento (en segundos)	5	68

5. Conclusiones

El análisis y estudio de las herramientas forenses de correos electrónicos, se llevó a cabo con el fin de evaluar cada una de ellas y poder identificar las beneficios y propiedades que presentan, enfocando el estudio en los campos e información relevantes que contiene el cuerpo o código fuente de un correo electrónico.

Cada una de ella tiene una potencialidad en alguna característica, lo cual significa que, según el tipo de análisis forense a realizar se pueden complementar para llegar a una conclusión más detallada y con mayor sustento.

Según las pruebas realizadas y evaluadas, se llega a la conclusión que la herramienta más completa es *Mailxaminer*, la cual brinda diversas funcionalidades para llevar a cabo el análisis de un correo electrónico o bandeja de correos, y se resalta la usabilidad de la misma.

Una característica a remarcar y de gran utilidad, es la conformación de una tabla de ruteo donde identifica todos los servidores por lo cual tuvo que pasar en su recorrido un correo electrónico y si uno desea lo grafica en un mapa geográfico marcando la ruta del mismo.

De este trabajo se derivan otras líneas de investigación como ser analizar la perfomance de estas herramientas cuando se trata de correos o cuentas corporativas donde intervienen servidores de correos privados, cuentas de webmail, cuentas en dispositivos móviles con distintos tipos de clientes.

El desafío es analizar el comportamiento del correo electrónico con distintas variables, y verificar con las herramientas si se logran los mismos resultados.

[6] Extract, Transform and Load (extracción, transformación y carga). Es el proceso que permite mover datos, transformarlos y cargarlos en distintas fuentes para que puedan ser analizados.

6. Referencias

[1] *Aid4mail*, se puede obtener y consultar sus características en http://www.aid4mail.com/download-free-trial (página consultada el 30/06/2017)

[2] *Email Trackerpro*, se puede obtener y consultar sus características en http://www.emailtrackerpro.com/ download.html (página consultada el 30/06/2017)

[3] *Mail Navigator*, se puede obtener y consultar sus características en http://www.mailnavigator.com/download.html (página consultada el 30/06/2017)

[4] *OsForensics* se puede obtener y consultar sus características en http://www.osforensics.com/download.html (página consultada el 30/06/2017)

[5] *E-mail Examiner* se puede obtener y consultar sus características en https://www.paraben.com/products/e3-emx (página consultada el 30/06/2017)

[6] *MailXaminer* se puede obtener y consultar sus características en https://www.mailxaminer.com/download.html (página consultada el 30/06/2017)

[7] Gallo Beatriz P. de, Vegetti Marcela, Leone Horacio, *Hacía una Ontología para el soporte de la trazabilidad del correo electrónico en la Forensia Digital*, CIIDDI 2017, Cuba.

[8] Devendran, V. K., Shahriar, H., & Clincy, V. (2015). A Comparative Study of Email Forensic Tools. *Journal of Information Security*, *6*(2), 111.

[9] Marwan A.Z. (2004) *Tracing E-mail Headers*. Proceedings of Australian Computer, Network & Information Forensics Conference, November 2004, School of Computer and Information Science, Edith Cowan University Western Australia, 16-30

[10] Tariq Banday, *Analysing E-Mail Headers for Forensic Investigation*. Journal of Digital Forensics, Security and Law, Vol. 6(2)