1-10-2015

A Comparative Study of Email Forensic Tools

Vamshee Krishna Devendran

Hossain Shahriar
Kennesaw State University, hsharia@kennesaw.edu

Victor Clincy
Kennesaw State University, vclincy@kennesaw.edu

Journal of Information Security, 2015, 6, 111-117
Published Online April 2015 in SciRes. http://www.scirp.org/journal/jis
http://dx.doi.org/10.4236/jis.2015.62012

Scientific Research Publishing

A Comparative Study of Email Forensic Tools

Vamshee Krishna Devendran, Hossain Shahriar, Victor Clincy

Department of Compute Science, Kennesaw State University, Kennesaw, GA, USA
Email: dvamsheekrishna@gmail.com, hshahria@Kennesaw.edu, vclincy@kennesaw.edu

Received 10 February 2015; accepted 9 April 2015; published 10 April 2015

Abstract

Over the last decades, email has been the major carrier for transporting spam and malicious contents over the network. Email is also the primary source of numerous criminal activities on the Internet. Computer Forensics is a systematic process to retain and analyze saved emails for the purpose of legal proceedings and other civil matters. Email analysis is challenging due to not only various fields that can be forged by hackers or malicious users, but also the flexibility of composing, editing, deleting of emails using offline (e.g., MS Outlook) or online (e.g., Web mail) email applications. Towards this direction, a number of open source forensics tools have been widely used by the practitioners. However, these tools have been developed in an isolated manner rather than a collaborative approach. Given that email forensic tool users need to understand to what extent a tool would be useful for his/her circumstances and conducting forensic analysis accordingly. In this paper, we examine a set of common features to compare and contrast five popular open source email forensic tools. The study finds that all email forensic tools are not similar, offer diverse types of facility. By combining analysis tools, it may be possible to gain detailed information in the area of email forensic.

Keywords

Email Forensic, Header and Content Analysis, Data Recovery, Search Option, Visualization

1. Introduction

Email is a common method of communication among two parties. It is a file transfer among two servers on a specific port number [1]. An email is usually composed using a client side application (Web Client, MS Outlook, Lotus notes) with a Sender's identity, then stored as a file, subsequently delivered to a destination user address through one or more number of servers. Although Email communication has been designed to make things simple, efficient, and powerful [2], the composing and communication have been under the focus of malicious intruders over the last few decades. We find it very common that email messages are the sources of transporting

How to cite this paper: Devendran, V.K., Shahriar, H. and Clincy, V. (2015) A Comparative Study of Email Forensic Tools. *Journal of Information Security*, **6**, 111-117. http://dx.doi.org/10.4236/jis.2015.62012

annoying, malicious, phishing, and spam contents.

In recent years, many technologies have been developed for examining and protecting emails that include spam detection, phishing email detection, content and attachment filtering (anti-virus engines). One key aspect of designing and developing these technologies is to conduct forensics investigation on sample emails to correctly identify important information such as the recipient name or identity, the path between the sender and the recipient used for transporting the email, the client-side application used to compose the email, the timestamp when a message was generated, a unique message ID, etc. In the literature, the examination and revealing of key information from an email is known as *Email Forensics* [3] [4]. A common example usage of forensics is to understand the key facts and rely on them for legal procedures.

Over the last two decades, many email forensic tools have been developed. According to Garfinkel [5], most of the tools (e.g., [6]-[10]) are diverse in nature, not necessarily built upon prior work. Rather, most tools have been developed in an isolated manner. Further, most of the forensics tools are not intended to solve any specific cyber or computer crime related problem. Rather, they are intended to discover or recover information. Given that a question naturally arises among stakeholders: to what extent existing tools are suitable for employing or conducting digital forensic investigation? This paper attempts to answer this question by comparing and contrasting a number of popular email forensic tools. In particular, our focus is on email header analysis phase offered by the tools. We examine the capability of a number of popular email forensic tools including MainXaminer [6], Add4Mail [7], Digital Forensic Framework [8], eMailTrackerPro [9], and Paraben Email Examiner [10]. Our work is complementary to previous research work that attempts to understand the capability of other types of forensic tools such as network forensics [4] and disk/memory forensic tools [11].

This paper is organized as follows: Section 2 provides a brief overview of email header elements and email forensics procedures in general. Section 3 provides an analysis of some open source tools based on a set of attributes. Finally, Section 4 concludes the paper.

2. Background and Related Work

2.1. What an Email Header Contains?

An email has header and body. The header part includes many important information such as sender's IP Address, mail user agents, servers in transit, message id field, and signatures field. We show an example of email header in **Figure 1**.

```
Received: (qmail 20564 invoked from network); 5 Jan 2006 16:11:57 −0000
From: foo<foo@foo.com>
To: bar@bar.com
Subject: Test
User-Agent: KMail/1.9
MIME-Version: 1.0
Content-Disposition: inline
Date: Thu, 5 Jan 2006 16:41:30 +0100
Content-Type: text/plain; charset = "iso-8859-1"
X-Originating-IP: [216.119.20.3]
Message-Id: <200601051641.31830.foo@foo.com>
X-HE-Spam-Score: 0.0
X-HE-Virus-Scanned: yes
Status: OR
Content-Length: 124
Lines: 26
```

Figure 1. Example of email header.

The *Received* field indicates the date and time when the email was received at the server. The *From* and *To* represent the sender and recipient, respectively. The *user-agent* field shows the client side application used to compose the email. The *Multipurpose Internet Mail Extensions* (*MIME*) version shows as 1.0. The *Content-Disposition* header indicates the presentation style. Here, the *inline* content-disposition means that the message should be automatically displayed (as text also indicated by the *content-type*) and no attachment is present. The *X-Originating-IP* indicates the source IP address where the email was originated from. The *Date* is the time when the email was generated. The *message id* is an automatically generated id where the timestamp information is present along with sender account information. The *X-HE-Spam-Level* contains the spam score computed at the client side (to prevent spamming). The *Status: OR* indicates that a message is downloaded but not deleted (please see http://www.faqs.org/rfcs/rfc2076.html for details; here, *R* means message is read or downloaded, and *O* means message is old but not deleted). The *Content-Length* indicates the length of the message (in bytes).

Cyber forensic e-mail analysis is employed to collect credible evidence to bring criminals to justice, in particular the header analysis [12]. A detailed header analysis can be used to map the networks traversed by messages. If multiple received field information is present, then the origin needs to be tracked from bottom to top *Received* field, where the bottom received address indicates the original sender's IP, and the top received address indicates the actual receiver's IP address.

2.2. Email Forensics Analysis Steps

A forensic investigation of e-mail can examine both email header and body. This paper will look at header examination. According to Marwan [12] an investigation should have the following:

- Examining sender's e-mail address
- Examining message initiation protocol (HTTP, SMTP)
- Examining Message ID
- Examining sender's IP address

Some other aspects that controls forensics step include the following properties:

1) Storage format of email: Server side storage format may include maildir (each email is kept separate in a file, for each user), mbox format (all email files are in a single text file). Server-side stores email in SQL Server databases. Reading different types of formats can be done for forensics analysis by using notepad editor and applying regular expression-based searches [1]. At the client-side, an email is stored as *mbox* format (Thunderbird) [1]. Client side may also store emails as *.PST* (MSOutlook), and *NSF* (Lotus Notes) files.

2) Availability of backup copy of email: When checking from the serve side, all copies are transferred to the client. This requires seizing the client computer. For webmail, copies are always saved at the server side [1].

3) Protocol used to transport email: Email can be initiated and transported based on SMTP or HTTP [12] depending on the email server applications.

3. Comparison Criteria of Email Forensic Tools

We compare email forensic tools based on a set of desired attributes required by forensic tool as specified by the seminal article of Garfinkel [2] that focuses on issues to support forensics investigation where most cases cannot be generalized. At the same time, expected criteria should be relevant to the reverse engineering capabilities, as well as scaling up data processing, considering the changes of technologies for storage, data transportation, and in between environment between devices processing and storing inputs.

We identify nine criteria that may be useful and present in forensic tools listed below:

1) requirement of input file in the hard disk,
2) search option
3) information extracted or provided by the tool
4) recovery capability
5) email file format supported
6) visualization support
7) operating system (OS) supported
8) extended device supported
9) and export format supported.

We exhaustively search the literature and various web page resources to identify email forensic tools available

in the real-world [4] [12]-[14]. We choose the following five open source email tools which are popular and widely used. Sections 3.1-3.5 discuss the tools in details with respect to the nine criteria.

3.1. MailXaminer [6]

1) *Input file in disk required*: This indicates the presence of email file at the local disk. MailXaminer [6] requires input file to be present in the disk.

2) *Search option*: This feature indicates how to perform search of interesting words in the content of an email. MailXaminer can perform plain text-based search.

3) *Information provided*: This feature indicates the information extracted and shown as part of forensic analysis. The MailXaminer tool shows the message, date and time details of an email.

4) *Recovery capability*: A forensic tools should have the capability to recover corrupted email or deleted email to be useful for investigation. The MailXaminer [6] can recover corrupted email and supports NSF, EDB, EML, PST, OST, OLM, IMM, TBB, MBX and MBOX type emails. It also has the capability to import corrupted contacts, calendar.

5) *Email format supported*: This feature indicates the file type supported by a tool. The MailXaminer [6] supports Gmail, yahoo, Hotmail, IMAP, Mozilla Thunderbird, Lotus Notes, Outlook, Exchange, Mac Outlook email format.

6) *Visualization format supported*: A forensic tool should allow investigator different types of display of the extracted information to enable more intelligence gathering. MailXaminer [6] supports different view options Hexa-decimal content inspection, Normal inspection, Property inspection, Email Header, MIME inspection, Email Hop View, viewing in HTML and RTF format.

7) *OS Supported*: Ideally, a forensic tool should support different types of operating systems to make it useful for email applications running on different platforms. The MailXaminer [6] can run on Windows (both 32 and 64 bit version).

8) *Export format*: A forensic tools should have friendly format for saving the examination results for compatible analysis with other forensic tools. The MailXaminer [6] provides export options in EML, PST, TIFF, PDF, MSG and HTML formats.

9) *Extended device support*: This feature indicates if a tool can act on plug-ins devices such as added hard disk or USB memory stick, etc. Currently, MailXaminer does not support such feature.

We show a snapshot of email forensic analysis that we obtained as part of our trial analysis with known and non-malicious emails in **Figure 2**. Note that for the rest other four tools, we will not re-introduce the features of forensic tools. We will rather indicate the presence, and supported feature. If a feature is not present, then we indicate N/A.

3.2. Add4Mail [7]

1) *Input file in disk required*: The tool can analyze emails stored in hard disk. Further, it supports email analysis directly from webmail services that use IMAP access (e.g. Gmail, Yahoo! Mail, AOL Mail, FastMail, GMX Mail, Outlook.com, Outlook 356). Thus, this tools supports both online and offline email analysis.

2) *Search option*: Email search can be performed and search activities can be performed in the exported pdf. The tool can filter the emails based on text, time, date, keywords, logical operators, and regular expressions. It can search mail by date, header content, and by message body content

3) *Information provided*: The tool shows the message, date and time details of an email.

4) *Recovery capability*: It can filter duplicate emails and recover emails from the trash folder. It can process deleted e-mail from *mbox* files and restore unpurged emails.

5) *Email format supported*: It can read import and export various email formats (most of the known formats available in the market). It can process large email files.

6) *Visualization format supported*: N/A.

7) *OS Supported*: Add4Mail [7] runs on Windows (both 32 and 64 bit version), Linux, and Mac.

h) *Export format*: Files can be exported into various formats and can be viewed by any other tool that can display the emails such as EML, PST, TIFF, PDF, MSG and HTML. It can save emails in plain text.

8) *Extended device support*: Mail folders and files can be processed even when disconnected (unmounted) from their email client including those stored on CD, DVD, and USB drives.

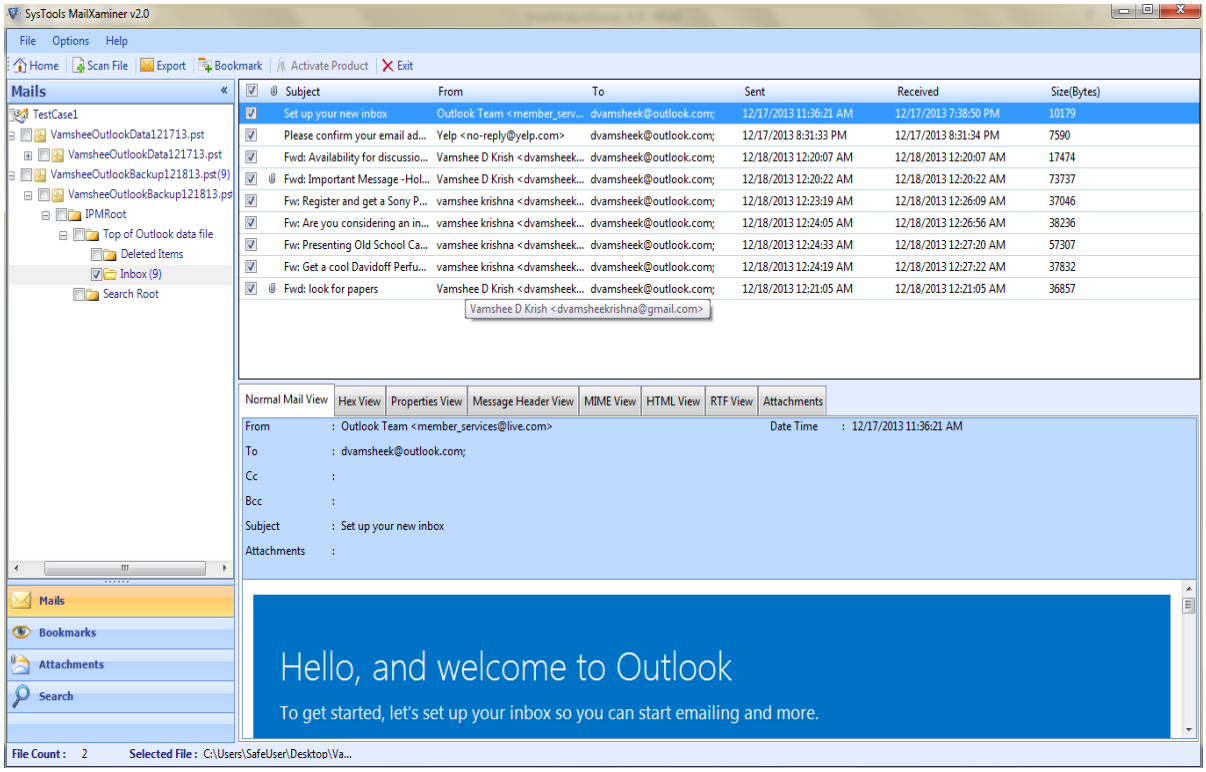

Figure 2. A snapshot of email analysis using MailXaminer for outlook mail [6].

Figure 3 shows a snapshot of email analysis (an example of Yahoo email loaded in the tool) using Aid4Email tool.

3.3. Digital Forensic Framework [8]

1) *Input file in disk required*: The tool can analyze emails stored in hard disk.
2) *Search option*: Search can be performed as regular expressions, based on dictionaries. Further, search can be performed based on the email content, tags, and time-line.
3) *Information provided*: The tool shows or retrieves the message, date and time details of email. It can also perform some features like virtual machine disk reconstruction.
4) *Recovery capability*: It basically can recover data from hard disk, flash drive etc.
5) *Email format supported*: It can also read 3 file formats, those are: Raw, Encase EWF, AFF.
6) *Visualization format supported*: N/A.
7) *OS Supported*: Supports only 32 bit version of windows. It runs on Windows and Linux OS.
8) *Export format*: Files can be exported into various formats and be viewed by any other tool that can display the emails such as EML, PST, TIFF, PDF, MSG and HTML. It can save the email in plain text.
9) *Extended device support*: It can recover data from USB or flash drive.

3.4. eMailTrackerPro [9]

1) *Input file in disk required*: It can analyze email files stored in local disk.
2) *Search option*: N/A. It rather focuses on analyzing email for possible spamming contents automatically.
3) *Information provided*: It provides the IP address that sends the message along with geographical location (city) of the IP address to determine the threat level or validity of an e-mail message. It can find the network service provider (*ISP*) of the sender. A routing table is provided to identify the path between the sender and receiver of an email. It also can check a suspected email against Domain Name Server blacklists to safeguard against spam and malicious emails. It also displays whether any port is open in any of the HTTP or FTP server in the tracked IP addresses.

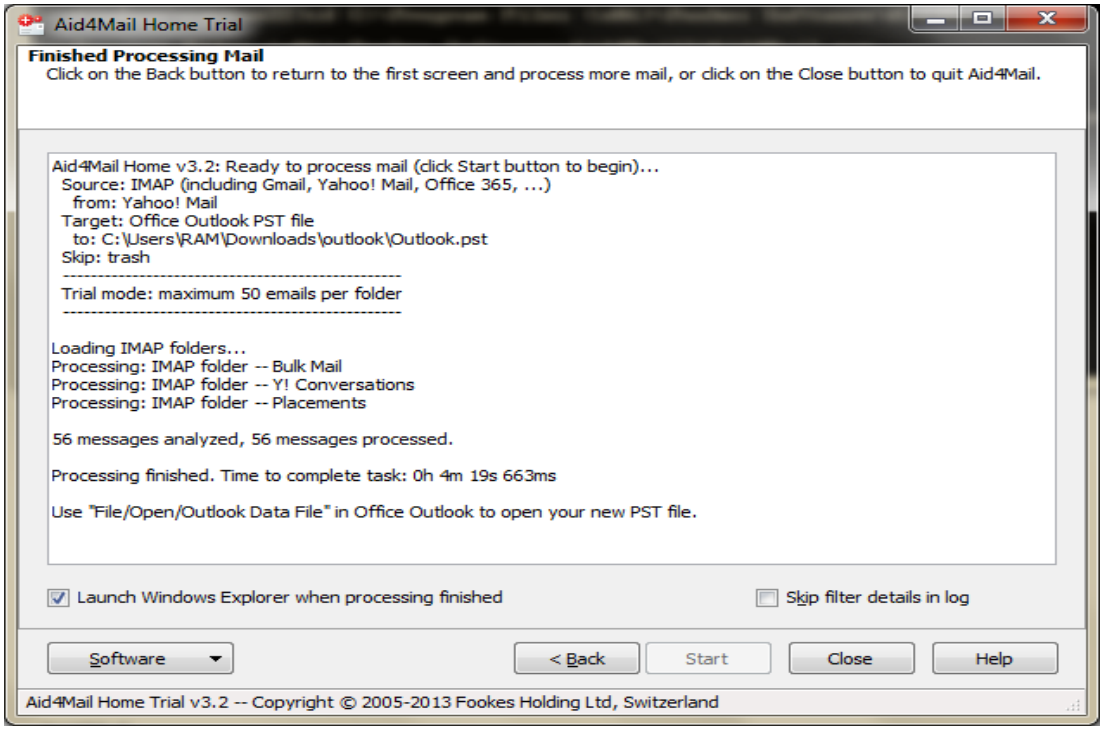

Figure 3. A snapshot of email analysis using Add4Mail.

4) *Recovery capability*: N/A.

5) *Email format supported*: It supports some of the common email file formats for AOL, AOL Web Mail, Eudora 7.1, Gmail, and Hotmail.

6) *Visualization format supported*: It offers visualization support similar to Outlook message display.

7) *OS Supported*: Windows.

8) *Export format*: It exports the list of IP addresses and domain names as a new tab, Excel, or HTML file.

9) *Extended device support*: N/A.

3.5. Paraben E-Mail Examiner [10]

1) *Input file in disk required*: The tool requires email to be present in the local hard disk.

2) *Search option*: It performs comprehensive analysis features, bookmarking, advanced boolean searching, and searching within attachments. The search is supported for various languages include UNICODE.

3) *Information provided*: The tool can examine email headers and bodies, provides information based on the search (including contents from attachments).

4) *Recovery capability*: It can recover deleted email from Exchange (EDB), Lotus Notes (NSF), and GroupWise email even though they may be deleted from the deleted items folder.

5) *Email format supported*: It supports various file formats including *Microsoft Exchange* (*EDB*), *Lotus Notes* (*NSF*), and *GroupWise* e-mail stores. *America On-line* (*AOL*), *Microsoft Outlook* (*PST*, *OST*), *Thunderbird*, *Outlook Express*, *Eudora*, *E-mail file* (*EML*), *Windows* mail databases. It supports more than 750 *MIME* Types.

6) *Visualization format supported*: N/A.

7) *OS Supported*: Windows.

8) *Export format*: The tool allows exporting email or mailboxes to PST or other formats. The output can be filtered (search terms, dates, or other criteria) further before exporting.

9) *Extended device support*: N/A.

4. Conclusions

This work comparatively analyzes five open source email forensic tools namely MailXaminer, Aid4Mail, Digi-

tal Forensic Framework, eMailTrackerPro, and Paraben Email Examiner. We compare the tools based on nine criteria: input file in disk, search option, information provided, recovery capability, format supported, visualization format supported, operating system supported, export format, and extended device support. Our analysis shows that among the five tools, Add4Mail can analyze emails stored both in hard disk (offline analysis) and on remote email servers (online analysis). In terms of the search option, Add4Mail has the highest amount of capability to gather information than other tools. Among all tools, the information provided by Paraben E-mail Examiner covers not only email header and body, but also the attached file contents. The recovery capability of Add4Mail and Paraben E-mail Examiner look better than the other three tools since they can recover emails from delete folder. In terms of Email format support, Paraben E-mail Examiner supports most of the known email formats and 750 MIME content types. However, for visualization support, MailXaminer provides various options to end users. Most of the tools support Windows as the preferred OS, while only a few support Linux. The Digital Forensic Framework supports a rich set of output file format. Finally, very few tools (Add4Mail, Digital Forensic Framework) support extended devices such as USB memory stick. The more coverage a tool has, the better it can be suitable to address various types of forensics activities and legal procedures.

Our future work includes measuring their runtime performance (CPU, memory) using a set of benchmark emails. We plan to check the compatibility issue based on our proposed criteria that may arise while using multiple tools in forensic investigation on the same source emails. Future study also includes how email forensic tools can be applied in conjunction with other complementary network and memory forensic tools.

References

[1] Conan Albrecht, Email Analysis. http://www.gsaig.gov/assets/File/other-documents/Forensics-EmailAnalysis.pptx.pdf

[2] McAfee SaaS Email Protection. http://www.mcafee.com/us/resources/solution-briefs/sb-saas-email-protection-solution-guide.pdf

[3] Banday, M. (2011) Analyzing Email Headers for Forensic Investigation. *Journal of Digital Forensics, Security, and Law*, **6**, 50-64.

[4] Meghanathan, N., Allam, S.R. and Moore, L.A. (2009) Tools and Techniques for Network Forensics. *International Journal of Network Security and its Applications*, **1**, 14-25. http://airccse.org/journal/nsa/0409s2.pdf

[5] Garfinkel, S.L. (2010) Digital Forensics Research: The Next 10 Years. *Digital Investigation*, **7**, S64-S73. http://dx.doi.org/10.1016/j.diin.2010.05.009

[6] MailXaminer. http://www.mailxaminer.com/

[7] Aid4Mail Forensic. http://www.aid4mail.com/

[8] Digital Forensics Framework. http://www.digital-forensic.org/

[9] EMailTrackerPro. http://www.emailtrackerpro.com/

[10] Paraben (Network) E-mail Examiner. http://www.paraben.com/email-examiner.html

[11] Garfinkel, S. (2006) Forensic Feature Extraction and Cross-Drive Analysis. *Digital Investigation*, **3**, 71-81.

[12] 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.

[13] Free Computer Forensic Tool. https://forensiccontrol.com/resources/free-software/

[14] Digital Intelligence Forensic Software. http://www.digitalintelligence.com/forensicsoftware.php