Analysis of Storage Regarding Different Electronic Mail Client Application

Ying Zhang and Feng Gao

***Abstract*—With the development of internet technology, electronic mail (abbreviate as email) has been used widely as a communication tool and thus aroused great attentions of forensic investigators. In this situation email client, which is in the category of groupware environments, is popular because it helps users access and manage mails conveniently. Nevertheless most users utilize special email tools directly instead of focusing on the data structure and principle of email clients. In this paper we are going to analyze the storage of electronic mail regarding various popular client applications, including Outlook Express, Outlook, Foxmail, Windows live mail and Mozilla Thunderbird. It aims to discuss the default location of data file, the concrete approach that how emails are stored, data structure of storage files, as well as related configured files. In addition storage files of email clients with different versions or under various operation system would be discussed in this paper.**

***Index Terms*—E-mail clients, storage, pst, .dbx.**

I. Introduction

Email is the message transmitted through computer networks, which focus on internet primarily nowadays. As a means of information exchange, it is widely accepted by governments, enterprise, schools, hospitals, institutes and non-governmental organizations. According to the survey [1], the number of email subscribers in China has reached to 258,470,000 in 2015, with a utilization rate of 37.6%. On account of its universal usage, forensic investigators turn their attention to analysis of email itself.

For most users emails could be sent and received via web browser or client applications, while email clients are becoming more and more popular due to their easy access and management. The overview of primarily-used clients will be demonstrated in the follow sections, so as to introduce the basic concepts and explain corresponding technical background.

II. Overview

Email client is a computer program in the category of groupware environments used to access and manage a user's email [2]. Popular clients will be discussed as follow.

Manuscript received April 13, 2018; revised June 18, 2018. This paper was supported by the Special Basic Research, Ministry of Science and Technology of the People's Republic of China, project number: 2016GABJC24.

Ying Zhang and Feng Gao are with the Third Research Institute of Ministry of Public Security, Shanghai 201204 China (e-mail: zhangying@stars.org.cn, gaofeng@stars.org.cn).

A. Outlook Express

Outlook Express is a Usenet client based on NNTP Protocol, which is published by Microsoft. This client is included with Internet Explorer versions 3.0 through to 6.0. As such, it was bundled with several versions of Microsoft Windows, from Windows 98 to Windows Server 2003, and was available for Windows 3.x, Windows NT 3.51, Windows 95, Mac System 7, Mac OS 8, and Mac OS 9.

B. Outlook

Outlook is a client working on personal information management, while outlook express focus on internet mail and news mainly. It is part of Microsoft office suite and expand the capacity of Outlook express, including calendar, task manager, note taking and so on.

C. Foxmail

Foxmail is a freeware developed by Tencent. It is compatible with Chinese emails, as well as supporting multi-user management and multiple addresses. In addition to that, the client is able to download emails from different mail servers.

D. Windows Live Mail

Windows live mail is another freeware developed by Microsoft. It is the successor to Windows Mail on Windows Vista, which was the successor to Outlook Express on Windows XP. As a member of Microsoft live, it combines other live service to its interface, which enable users to activate Windows live message directly.

E. Mozilla Thunderbird

Mozilla Thunderbird is an open-source and free client developed by Mozilla foundation, which was designed for Mozilla Firefox browser users specifically. It is configured and customized easily. The latest official version is 52, with early version of 2.0.0.24 and 3.1.17. The early version is suitable for PC with low configuration while the latest one is more secure but with high resource occupation.

III. Analysis

Depending on clients, different measures would be taken to store data. Storage policy and related data structure will be discussed in this section.

TABLE I: Location of Folder File

Operation System	Location
Windows XP	C:\Documents and Settings\<user>\Local Settings\ Application Data\Identities\<entity>\Microsoft\ Outlook Express

DOI: 10.7763/IJCTE.2018.V10.1205

A. Outlook Express

In Outlook Express the backup copies switch to .dbx files from version 5.0. Default location of this file is described in Table I.

Nevertheless the backup file could be moved to any desired location. Choose 'Options' under 'Tools', and then select 'Maintenance'. After clicking 'Store Folder' the folder could be set as shown in Fig. 1 [3].

The account backup files are composed of Folders.dbx, Inbox.dbx, Sent Items.dbx, Deleted Items.dbx, Drafts.dbx, Pop3uidl.dbx, Offline.dbx and so on.

Folders.dbx is index of all files in Outlook Express folders, which records the number of folders and news groups, setting of folder synchronization, information of Hotmail and so on. Outlook Express will scan all .dbx files so as to build a new folders.dbx file if the original one was lost.

Inbox.dbx, Sent Items.dbx, Deleted Items.dbx and Drafts.dbx are default of system itself, which means Outlook express will regenerate them if they were deleted accidentally. We use forensic tool X-Ways Forensics to view above files and the content of Inbox.dbx is demonstrated in Fig. 2. After that we open Outlook Express to view the source code of corresponding mail and compare it to Inbox.dbx as shown in Fig. 2. By comparison it is known that mails in inbox are listed in Inbox.dbx in order without encryption.

Pop3uidl.dbx is used to store POP3 receiving records. Information such as whether the mails have been received by server or kept in server is stored in this one.

Offline.dbx is primarily for IMAP protocol, which retrieves email messages from a mail server over a TCP/IP connection [4].

B. Outlook

In Outlook the storage file is in format of.pst. Default location of this file is described in Table II. Nevertheless the backup file could be moved to any desired location. Choose 'Options' under 'Tools', and then select 'Mail Maintenance'. After clicking 'Data File' the folder could be modified.

TABLE II: LOCATION OF FOLDER FILE

Operation System	Location
Windows XP	C:\Documents and Settings\<user>\Local Settings\Application Data\Microsoft\Outlook
Windows 7	C:\Users\<user>\AppData\Local\Microsoft\Outlook

Previous research [5] has proved that this file contains information of messages, folders and attachments, which relies on the Exchange data stores. And this file is quite different from traditional database such as SQL and owns its unique data structure.

According to the technical documents published by Microsoft [6], the pst file structures are logically arranged in three layers: the NDB (Node Database) layer, the LTP (Lists, Tables and Properties) layer, and the Messaging layer. Nevertheless it could not be view directly without the help of specialized tools.

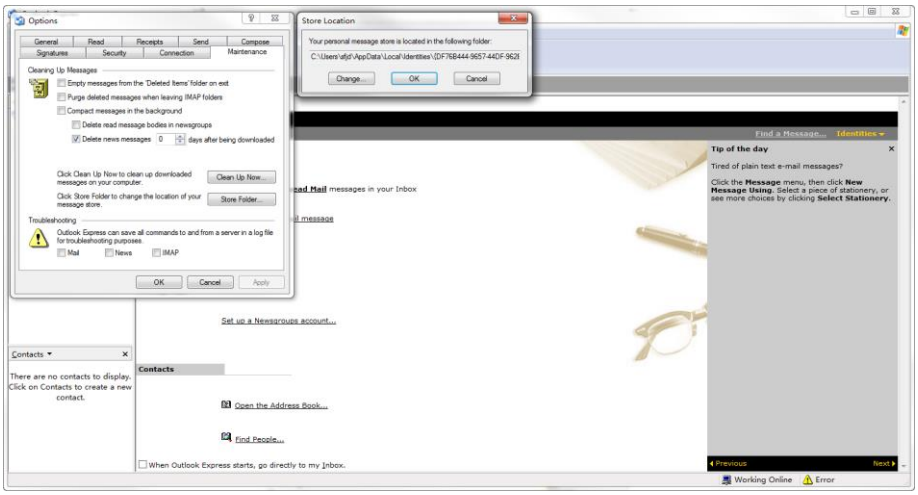


Fig. 1. Setting of backup files.

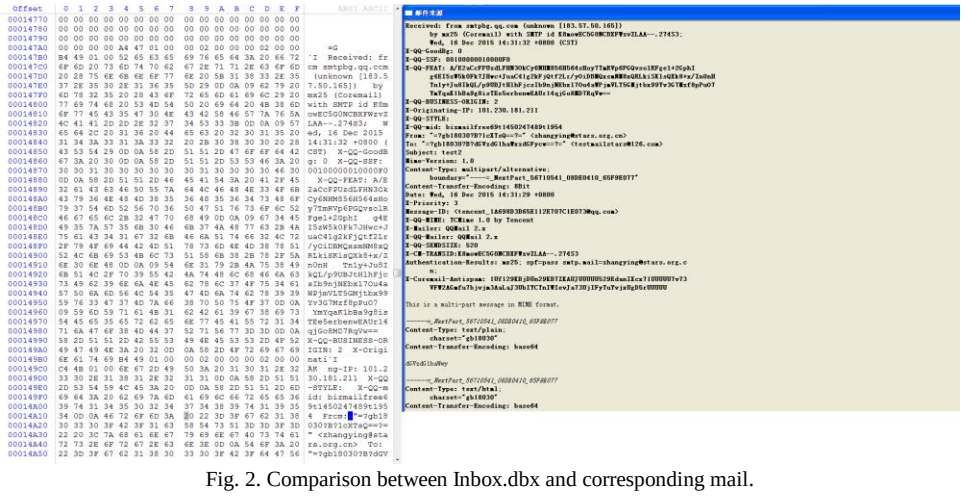


Fig. 2. Comparison between Inbox.dbx and corresponding mail.

C. Foxmail

The backup copies are in different formats depending on the version of client itself.

1) Foxmail 6.5

In this version backup copies are generally composed of in.BOX, in.INDX, out.BOX, out.INDX, sent.BOX, sent.INDX, spam.BOX, trash.BOX, trash.INDX and so on. There are two files corresponding to each folder, in the format of .BOX and .INDX.

Default location of above files is <install directory>\Foxmail\mail\<account>. Take in.BOX and in.INDX for example, in.INDX is the index of folder 'Inbox'. It starts with 48 bit hexadecimal digits, and the last bit of the third line is supposed to identify the number of mails. The next 16 bits are the beginning of index for the first mail, and index ID locates in the first bit of next 16 bits. Mails in 'Inbox' are stored in file in.BOX.

X-Ways Forensics is used again to view the content of in.BOX as shown in Fig. 3. After that we open Foxmail to view the source code of corresponding mail and compare it to in.BOX.

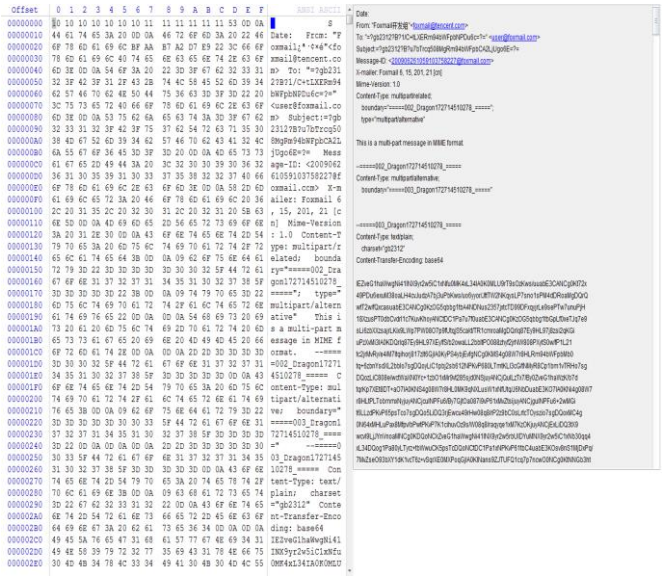


Fig. 3. Comparison between in.BOX and corresponding mail.

By comparison it is known that in.BOX is not encrypted as well. In addition to that, two mails are separated by 16 bytes of '1010101010101011 1111111111530D0A' as separator.

2) Foxmail 7.0

Default location of backup files is <install directory>\Foxmail 7\Data\Mails. A brand new storage policy has been taken in this version. Mails are not stored by account, and each mail is stored in the format of single folder instead of in in.BOX uniformly.

3) Foxmail 7.2

Default location of backup files is <install directory>\Foxmail 7.2\Storage\<account>\Mails. The mode that mails are stored by account has been used again. Each account refers to a folder with the same name, while each mail owns its single folder named by random numbers.

The file 'FMStorage.list' lies in the root directory, aiming to record the account list. And there is a file 'Index' in the folder corresponding to the account, which is used to record the sender, recipient and subject of mails as listed in Fig. 4.

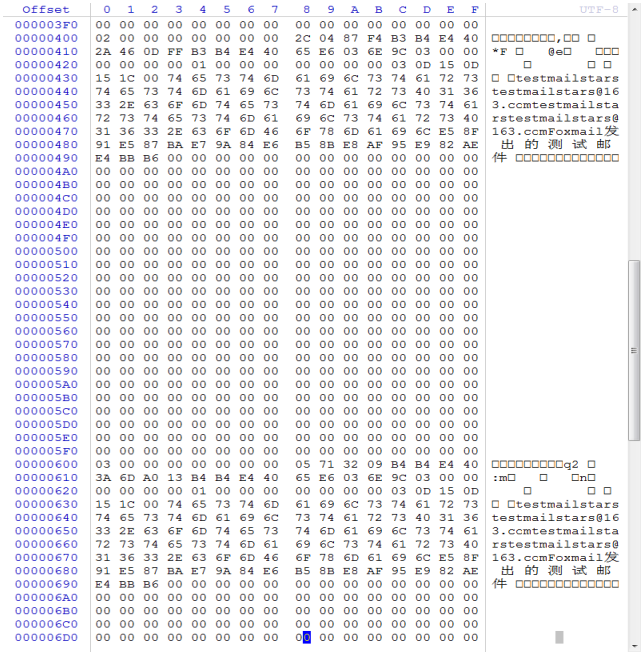


Fig. 4. Content of FMStorage.list.

D. Windows Live Mail

Default location of backup files is described in Table III. Also it could be moved to any desired location.

TABLE III: LOCATION OF FOLDER FILE

Operation System	Location
Windows XP	C:\Documents and Settings\<user>\Local Settings\Application Data\Microsoft\Windows Live Mail\<account>
Windows 7	C:\Users\<user>\AppData\Local\Microsoft\ Windows Live Mail\<account >

Choose 'Options' under 'Tools', and then select 'Advance' and 'Maintenance'. After clicking 'Store Folder' the folder could be found.

The root directory contains subfolder named by abbreviation of mailbox name, Backup, Contacts, Storage Folders, Outbox, Mail.MSMessageStore and so on.

Backup consists of Deleted Items, Drafts, Inbox, Junk E-mail, Sent Item, Account{…}.OceanAccount and so on, while the last one refers to information of account. Unlike other clients, each mail are stored in the corresponding folders with format of .eml directly. That means there is not any special folder file for storage in this client and users could access the mails without any tools.

E. Mozilla Thunderbird

The backup copies are in different formats depending on the version of client itself.

1) Thunderbird 52.5.0

TABLE IV: STRUCTURE OF FOLDER FILE

	Storage file	Storage file 2
Inbox	Inbox	Inbox.msf
Sent	Sent	Sent.msf
Trash	Trash	Trash.msf
Drafts	Drafts	Drafts.msf

msgFilterRules.dat
Popstate.dat

In this version the default location of backup copies is C:\Users\<user>\AppData\Roaming\Thunderbird\Pro-files\<random name>.default\Mail\<email retrieve server>. And the structure of backup is described in Table IV.

According to the name of storage file, each part of email accounts owns two files individually. One is without extension while another is with extension of msf. X-Ways Forensics is utilized to read the source code of file Inbox. Besides that source code of the first mail in inbox is compared to the content of file Inbox as shown in Fig. 5. By comparison it is found that all mails are listed in Inbox in order, separated by 2 bytes of '0D0A'. That means the file Inbox is used to stored inbox mails without encryption.

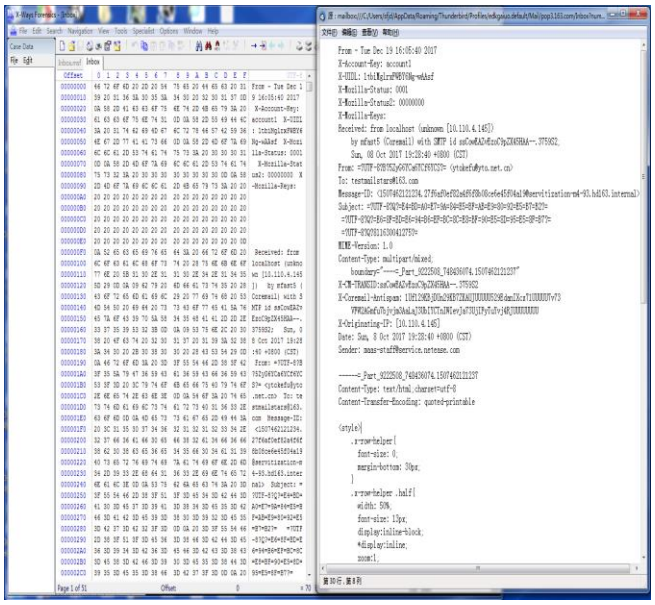


Fig. 5. Comparison between Inbox.dbx and corresponding mail.

Another file Inbox.msf is treated as index of corresponding storage file, informing users which storages files should be selected.

There are two configuration files except of storage files listed above. msgFilterRules.dat records the filter rules set by users. Popstate.dat is generated due to the utilization of pop server. It records users' settings for truncated messages and the approach of leaving messages on server.

2) Thunderbird 45.1.1

In this version the data structure is quite closed to the version 52.5.0. There are storage file and configured files with the same name as well.

Via analyzing the content of storage and configured files, it is found that this version utilizes the same principles used in the latest one.

3) Thunderbird 38.5.0

It is known that this version has the same structure with the version 45.1.1.

IV. Conclusion

With the popularity of email, email clients have been widely utilized by people, following the security issue. However, little attention has been payed to storage file of email clients in the previous research, while most researchers concentrate on the information of email itself such as email header. Therefore in this paper we demonstrate the storing method of different email clients, including format of backup copies, default location, components of backup files, content of components and so on.

Concretely speaking, the storage approach of emails has been analyzed firstly so as to confirm whether they are stored in individual files or integrated into one file. Then tools are used to examine storage files and related configured files. And content contained in above file has been analyzed carefully to find if there is any rule between them. Finally the relevance between storage and configured files was discussed aiming to reveal the storage mechanisms.

It is found that different clients own their unique approach to store, varying from different versions. For example, Outlook Express stores information in different folders, depending on the file structure of mails, while Outlook integrates all information into one file. Unfortunately we could not find the appropriate way to analyze .pst files. Besides that, meaning of certain files is unknown.

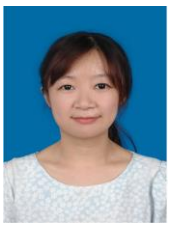
Considering above factors, it is expected that more research would be concentrate on analysis of .pst file. In addition, research on storage mechanism under other operating systems including Mac OS and Linux is supposed to be conducted in the future. And further research on email storage of multiple platforms including mobile devices will be summarized in our next study.

Acknowledgment

Ying Zhang thanks Hong Guo, who dedicates herself to the research of digital forensic and gives great patience and consistent encouragement to the author.

References

[1] [Online]. Available: http://www.chyxx.com/industry/201703/507891.html
[2] [Online]. Available: https://en.wikipedia.org/wiki/Email,
[3] [Online]. Available: https://msdn.microsoft.com/en-us/library/ff385210(v=office.12).aspx
[4] [Online]. Available: https://en.wikipedia.org/wiki/Email_client
[5] K. Dheepa and G. Annapoorani, "Knowledge discovery on extracted outlook Pst files and emails," in *Proc. 2015 Online International Conference on Green Engineering and Technologies*, 2016, pp.1-5.
[6] [Online]. Available: http://kb.mozillazine.org/Knowledge_Base



Ying Zhang was born in Jiangsu in 1985. Ying received the master's degree in global computing and multimedia from University of Bristol in United Kingdom in 2007, and the bachelor's degree in computer science from the university of Nanjing University of Aeronautics and Astronautics in China in 2006.

She worked as a C++ developer in Mode 7 Limited Company in Oxford in 2008. From 2009 she has worked as a digital forensic expert in the Third Research Institute of Ministry of Public Security in Shanghai. She worked as the technical investigation officer in Shanghai intellectual property court in 2016. Her current research concentrate on research on digital forensics, including data recovery, authenticity of email, blind digital image forensics and so on. Also she composed the industry standard of the People's Republic of China: Technical methods for E-mail examination.

Ms. Zhang is the member of China Computer Forensic Conference.