Recovering Evidentiary E-mail for Non-Repudiation Forensics

Lokendra Kumar Tiwari, Saurabh Mishra, Shefalika Ghosh Samaddar

Abstract*: Computer Forensic, the upcoming branch of forensic science where acquiring, preserving, retrieving and presenting content processed electronically and stored digitally, is used for legal evidence in computer related crimes or any other unethical practice involving manipulation of digital content. Such digital content can take many forms which are manifested by different file formats and digital artifacts". This paper concentrates on acquisition of deleted e-mail from mailbox of web servers satisfying two tier, three tier and n-tier technology. A detailed survey of several possibilities are included for non-repudiation forensic. A case study of a particular file type using suitable forensic tool is cited as a proof of concept towards this claimed inference to provide digital evidence in case of non-repudiation by sender and/or by receiver. This is simply conducted by using Encase a proprietary Digital forensic tools. The whole process is captured in step by step fashion to have a better understanding of the mechanism used. Recovery of files/emails have certain kinds of legal hurdles, the paper have addressed them as well. This paper contributes to the extend the recovered email can used as a ready digital evidence in any court of law.

***Keywords*: E-mail recovery, DBX file recovery, EnCase, Outlook Express mail recovery.**

I. INTRODUCTION

"Forensics is the process of using deductive scientific knowledge in the collection, analysis, and presentation of legally valid evidence to the court of law". "The goal of cyber technology in forensics is to explain the current state of various digital artifacts" [1]. Another definition is "computer forensics is considered to be the use of analytical and investigative techniques and tools to identify, collect, examine and preserve evidence/ information which is magnetically stored or encoded" [2,3].

Forensic deals primarily with the recovery and analysis and presentation of latent evidence in specified media. Latent evidence can be thought of as fingerprints left on a windowpane or DNA evidence recovered from a blood stains to a deleted e-mail recovery for the purpose of non-repudiation either by the sender or by the receiver. Considering that modern forensic technique has had almost six decades of experience in the U.S courts and existence of 800 year as a discipline, the specialization of computer forensics, in comparison, is still considered in its infancy and yet to build up the mechanism for development of technique [4].

Computer forensics as a methodology can be defined as the legally consistent collection, analysis and legally valid presentation of data from computer systems, networks, communication streams (wired and wireless) and storage media (online and offline) admissible in a court of law [5]. It is a merger of the discipline of computer science and forensic methodology in real time as well as the examination of latent data for evidence building.

The Block Diagram below demonstrate step by step process of computer forensic analysis

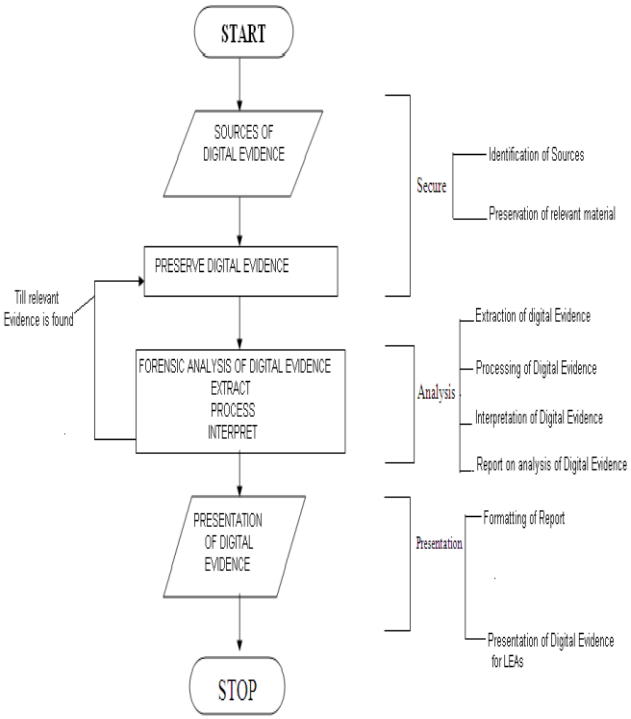

Figure 1 Computer Forensic Methodology

Any computer forensic process of file or email recovery must go through the above mentioned stages.

1.1 Role of Computer Forensic in Data Recovery

Computer forensic mechanism makes use of analytical and investigative techniques to identify, collect, examine and preserve evidence in the form of crime-related contents and/ or information. While collecting digital evidence, crypto graphical hash of evidence file is calculated and stored in safe storage for future reference. This paper uses in-built Hash mechanism of the selected tool Encase 6.0. The method described here is an attempt to

Revised Manuscript Received on July 22, 2019.

* Correspondence Author

Lokendra Kumar Tiwari, Department of Computer Science, Ewing Christian College, Allahabad, Uttar Pradesh India

Saurabh Mishra, School of Management Studies, Motilal Nehru National Institute of Technology, Allahabad, Uttar Pradesh India

Shefalika Ghosh Samaddar, Department of Computer Science and Engineering, NIT Sikkim, India

throw light on practical insight to deal with a situation where digital evidence in the form of a particular e-mail is deleted. The purposeful deletion is to show non-existence of the e-mail on demand. Such deletion of e-mail may lead to repudiation to validate the wrongful action of the cyber-criminal. Successful recovery of such legally validated email leads to non-repudiation. Growing use of forensic tools does contribute to timeliness and promptness of desired action to be taken in such cases. A forensic investigation can be initiated as typical criminal investigation or civil litigation using sophisticated digital forensic tools and techniques.

Forensic investigation provides digital evidence when data loss occurs in the following manner

- Abuse on Internet, denial of service, etc
- Industrial espionage resulting in marketing disadvantage
- System Damage in an accident resulting in financial loss as well as loss of goodwill.
- Cases of criminal fraud and deception.
- Computer usage by criminals for storage and communication during active and passive attacks
- Recovery of deleted (intentional or unintentional) emails.

The last point has been dealt with in this paper by suggesting a novel methodology on usage of tools so that it can be used in digital evidence.

The sections are devised in the following manner.
Section 1 introduces the subject of computer forensics and role of computer forensic in authenticated data recovery, particularly in case of non-repudiation. Section 2 is a survey of the state-of-art in e-mail recovery. This section also provides the tools presently available that can be used to recover deleted e-mail and prove its existence in a legally valid manner. The special purpose software can be used in wireless media by utilizing Access Data's Universal Forensic Extraction Device or UFED offered by a number of security service providers like Cellebrite, AccessData and Forensic Guru [6, 7]. This increases possibility of extending this methodology to wireless medium as well. Section 3 investigates the methods of e-mail recovery under the circumstances of emptying of inbox, intentional deletion and cleaning of trash folder, steganographic reformatting of e-mail, moving a particular e-mail to a hidden folder etc. Section 4 provides a practical method of deleted e-mail recovery using a computer forensic tool namely, Encase 6.0. Practical insight is the conclusion of the technique adopted to recover the e-mail to be used as latent evidence in the court of law. Section 5 traverses path towards future direction of research in deleted file recovery in general and deleted e-mail recovery in particular. The role of such digital evidence can be of paramount importance in case of modern warfare, terrorist nabbing and banking fraudulence.

The limitation of the proposed technique is also discussed in this section.
The document successfully addresses a method of e-mail recovery to provide assistance in case of verifiable non-repudiation. There can be a family of such methods to be adopted depending upon the suitability of various file formats commonly used in e-mail based communication over internet. Email recovery comes under the acquiring phase of forensic procedures involving collection of data or useful content from mailboxes in such a way that may be used for prosecution in computer related crime. This paper addresses a method to be adopted in this phase only.

1.1.1 Origin of the problem

When file is deleted from system including emails the forensics investigators collect deleted files from incident node and present details report using National Legal Standard. This will proved direct or indirect evidence[8]

1.1.2 Our contribution

Recovering of deleted file(e-mail) is achieved by using proprietary tool encase and whole mechanism is captured in step by step fashion to have a better understanding of the methodology. This recovery can also be used in any court of law for legal proceedings.

II. RELATED WORK

There are many forensic tools that are commercially accessible for playing varied forensic procedures. "A variety of open supply tools are accessible to recover the deleted emails [9]. Most of the tools are meant for casual users and so, cannot be employed by authorities for any legal purpose. Such tools don't seem to be thought-about as weapon within the arsenal of laptop rhetorical professionals. Computer forensic, by rule, avoids operating directly on the evidentiary material. This stems from the actual fact that physical proof to be made at the court of law, must always be command pristine". The necessity for glorious disk imaging method victimization relevant tools is preponderating. The National Institute of Standards and Technology (NIST) has developed many tools used for disc drive imaging tool analysis [10].

The tool should provide correct documentation of forensic processes involved and performed [10,11,12]. Acronis True image 6.0 takes an exact image of a hard disk drive or separate partitions, performs a complete backup image or a clone of it online in Windows and can take care of FAT 16/ 32, NTFS, as well as Linux Ext2, Ext3, ReiserFS file systems [11, 5, 13, 14, 15, 16]. SafeBack is used to create mirror-image (bit-stream) files of disks or disk partitions as the case may be [17]. EnCase can be used to mount images of hard drives or CDs as read only local drives. Together with VMware, the virtual machine infrastructure software, Encase enables the booting and examination of a computer under investigation to a state when the evidence was first captured [18]. The jobs performed by these software are:

- Imaging volatile memory: (dd, safeback and EnCase) [12, 18]
- Disk and file imaging: (Linux dd, SafeBack, SnapBack and DatArrest) [12, 17, 19]
- Write blockers: (FastBloc2 FE from Guidance Software, Firefly) [8]

- Integrity code generators and checker: (md5sum of Linux OS) [9]

2.1 Evidence collection methodology in * tier technology

The process of evidence collection methodology is defined in case of 2 tier, 3 tier (client server) and n tier (web server) technology.

2.1.1 Email evidentiary data in two tier technology (Client Server Model)

Most of the present systems run in this mode. Small enterprises rely on client-server model for their day to day business and would like to cope with the forensic requirement of the present day through deployment of appropriate tools. Network security of such systems depends on development of detecting and forewarning products against cyber-cheat and intrusion. There is no method which can be said absolutely effective against hacker intrusion. The expected intrusion can cause interference and damage. Utilization of forensic technology against hacker intrusion provides:

- Remote login
- Anonymous file access
- Mail service
- Non-authorized access
- Detection before attack
- Protocols decode
- System proxy attack detection
- Forewarn in case of exercising super user authority from an account others than 'root' account.
- Using source beyond authority
- Illegal operation detection
- Modification of information
- Exposing information related to authorized and unauthorized use of resources

2.1.2 Email evidentiary data in three tier technology (Client to server with a middle tier)

Collection of E-mail evidentiary data follow of similar technique as presented in 2.1.1 with the exception that in this case, forensic technique may reveal many meta-data information of middle-tier which are of forensic use.

2.1.3 Email evidentiary data in n-tier technology (web server model)

N-tier technology relates to web servers and web services. Securing a web service requires to protect the basic elements of the server and services. Their interactions with various kind of threat, vulnerability usually in case of an attack are recorded for analysis. The recorded attack pattern is analyzed. "An attack pattern is based on the analysis of observed attack exploitation [20]"

There are various vulnerability categorization and catalogues such as US-CERT vulnerability notes [21], "MITRE Common Vulnerabilities and Exposures (CVE)" [22], "MITRE common weakness Enumeration (CWE)", "National Vulnerability Database (NVD)" [23], "Security focus Vulnerability Database" [24], "Bug Traq List" [25] and "Open Source Vulnerability Database (OVSDB)" [26] etc. "The Web Application Security Threat Classification" [27], "Open Web Application Security Project", "Software Vendor's Security bulletins and Advisories and Enterprise Vulnerability Description Language" [28] etc. are web based information bases. Evidentiary data Collection in case of web services must consult these databases. It must follow some security standards such as Secure Socket Layer (SSL) and Transport Layer Security (TLS), XML Data Security, Security Association's Mark-UP Languages (SAML), SOAP Message Security, XML Key Management Standards (XKMS) [29], WS-Trust [30] etc. It should follow Access Control Policy Standards like extensible Access Control Mark-up Language (XACML), XML right Mark up Language (XrML) [31] etc. Applying the technique of remote forensic is complex and requires a virtual forensic audit system to be deployed with various kinds of services required for this purpose.

2.2. E-mail Recovery for Email investigation

Deleted e-mail recovery with legal binding in 2 tier and 3 tier technology follows a step by step procedure in case of e-mail recovery. Web services system based E-mail recovery in a n-tier architecture is out of the scope of the present paper and may be taken as a separate study in due course of time.

2.2.1 Email recovery in 2/3 tier Architecture

E-mails can be a simple communication through the protocols like SMTP and POP3 to secret code exchange using the protocols like TLS 1.0 or IPSec. In most instances POP3 based e-mail clients send passwords in clear text unencrypted across internet encouraging security breaches. It can even be presentation of confidential report to a series of Memorandum of Understanding (MoUs). The corporate world instead of using free email services like Yahoo or Hotmail uses a customized email system due to varied requirements. There is other trusted e-mail protocol like biometric based authentication but are of limited use [32]. In a Microsoft dominated world, Outlook Express or Microsoft Outlook is widely used for such purpose. They configure their account and use it as a vehicle to send and receive the emails in various modes. It may happen that a file in Outlook Express is corrupt while receiving it or deleted, intentionally or accidently by the receiver. Even the receiver may deny the receipt of the mail after a complete deletion to ensure non-recovery. However, in such cases, detection of the trace through signature of the file is important and in some cases reconstruction through the fragments is made possible by right usage of forensic software. Many information including sender, recipient, date of sent and received, attachments, whether it was read and marked as unread, flags set etc. can be discovered about the particular e-mail. E-mail can be extracted from sending computer, receiving computer, any number of servers in between the two, or from archive in a back-up system. E-mail forensics is primarily a investigative study of source and content as evidence. E-mail headers contain sender, receiver, subject, whether it was composed in a reply mode, time stamp etc.

Any ordinary e-mail reveals

- The host that added the received line
- The host/ IP address of the incoming SMTP connection
- The reverse-DNS look-up of that IP address
- The name of the sender used in SMTP 'Hello; command

III. PROPOSED TECHNIQUE FOR EMAIL RECOVERY USING PROPRIETARY TOOL ENCASE

Every file email or otherwise contains metadata related to it. Encase has number of inbuilt powerful tools that provides efficient analysis of different file formats related to email and related metadata

E-mail metadata collection is a pre-recovery process to get insight into the piece of evidence in the form of e-mail.

3.1 Type of metadata in case of E-mail

Types of metadata in case of E-mail are of the following with element of consideration:

▪Headers ▪ Attachment
-To -Encrypted / plaintext
-From -Text/multimedia document
-CC ▪ Secured protocol used
-BCC ▪ Server identification
-Exploring more ▪ Node identification
-Subject
-Message length

It may be noted here that it may not be possible to find the actual user (physical), through the mapping between the logical user and physical user. It is possible with a high success rate when an user ID is having a one to one relation with a particulars node ID in a network. However, forensic techniques are available to provide legal proof of user identification.

3.2 Digital Evidence Preparation

Preparation of Digital Evidence includes following different stages as listed in Block Diagram above which includes imaging, getting copy of work out analysis, analysis of result, analysis of metadata for building up evidence, presentation of result. As far as detailed legal process is concerned it is beyond the scope of this paper

3.3 Email Analysis tools and techniques

Similar algorithm as adopted here in the case study can be tried for different software of similar purpose like Cyber check 3.0 [33]. The software provides a separate module for training and tracking of email. The module is rightly named 'Email Tracer 2.0' [33]. "EnCase 6.0 has the ability to find, parse, analyze, display documents of various email formats, including Outlook PSTs/OSTs ('97-'03), Outlook® Express DBXs, Lotus Notes NFS, webmail such as Hotmail, Netscape and Yahoo; UNIX mbox files e.g. files used by Mac OS X; Netscape; Firefox; UNIX email applications; and also of the format AOL 6, 7, 8, 9."

IV. ENCASE 6.0 BASED E-MAIL RECOVERY

A simple case study has been considered taking the matching pair of outlook Express email and Encase 6.0. EnCase 6.0 can read Outlook Express .DBX files. The file structure is parsed first. The entries and the records tables on the table-pane list individual emails by leaving their subject line unaltered. The record table-pane lists the attachments with the emails, if any. The view-pane displays the contents of the selected email or attachment as and when required. The whole process is transacted in read only mode. Therefore, it can be run a number of times without any fear of losing or altering evidentiary data. Saving such displays for comparison of results is not difficult.

Deleted e-mails can be recovered forensically from users email client and e-mail servers. They may be recovered from the hard disk of the client and the server. Deleted emails and attachments can be retrieved from unallocated clusters of hard disk and may be stored for analysis. Even the fragments of a file are able to give its signature and detection of a file is made possible in most of the cases. Sometimes data encryption of e-mails makes it difficult for meaningful recovery of the file. Such cases can be dealt with timely help from metadata manipulation.

Web based e-mail services such as Gmail, Yahoo Mail or Hotmail use a browser to interface with the e-mail server. The browser sends the information to the disk drive in the system that is used to recover or generate the e-mail saving a copy to the disk. It is, again, possible to extract HTML based e-mail from disk drive.

4.1 Essential Steps to recover deleted e-mail

When any file or email is deleted from media say hard disk it is very difficult to recover it unless we use any tool. For the current case study also we have used a proprietary tool called encase which is recognized by US Courts. The fact are presented through a very common scenario in corporate world [34]. Which is given below.

- A leading manufacturing company say (A) has prepared a bid for certain working assignment
- The bid contained various financial details as required.
- The bid was confidential and only one or two people including top level management has the knowledge of it.
- The bid details were stored in the system of CEO having access of his PA only.
- The PA leaked the bid detail through an email and deleted the same (email was configured on outlook)
- Since the mail was deleted cyber forensic expert was called to do the investigation

4.2 A novel methodology of deleted e-mail recovery

Since the PA has send the mail configured in outlook of his own his activities are traced as follows

Step 1: Assuming he send the mail from his account configured on outlook as shown below.

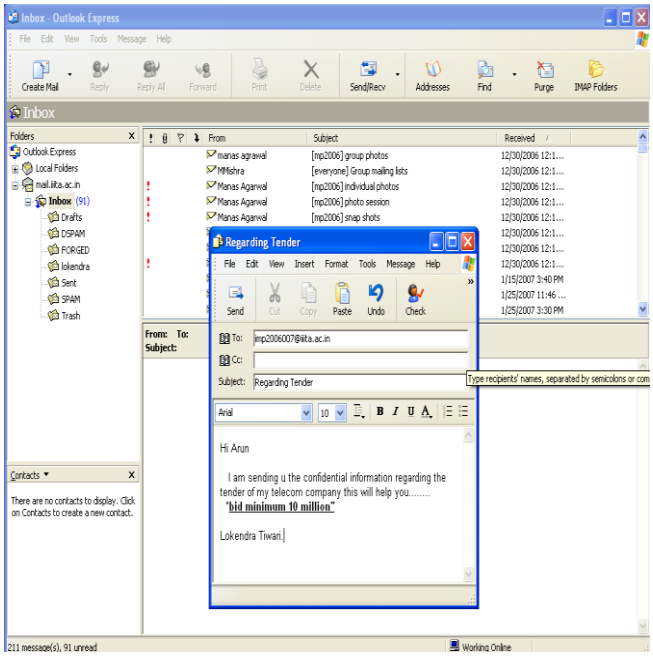

Figure 2: Window of composed email.

The screen shot shows such an email has been sent through an email.

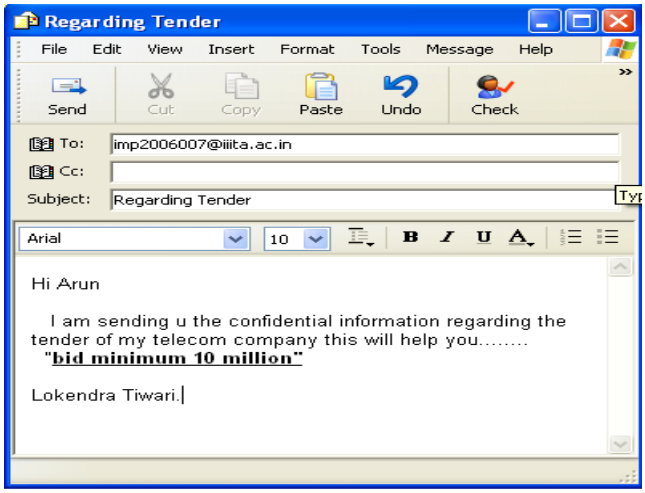

Figure 3: Content of composed email and address of recipient

Step 2: The copy of the mail which is present in the sent folder has also been intentionally deleted to remove all traces.. The file has been deleted as shown below. The deletion is made further untraceable by deleting it permanently (by emptying the trash bin).

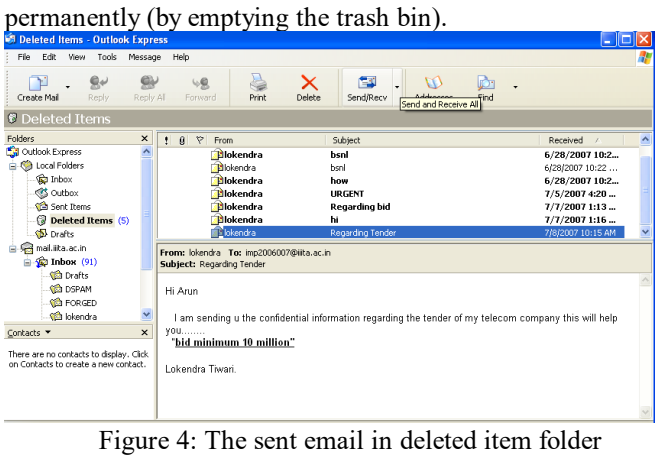

Figure 4: The sent email in deleted item folder

Step 3: The evidentiary email is now permanently removed from the 'Sent items' as well as from 'Deleted items' by emptying the folder 'Deleted Items'.

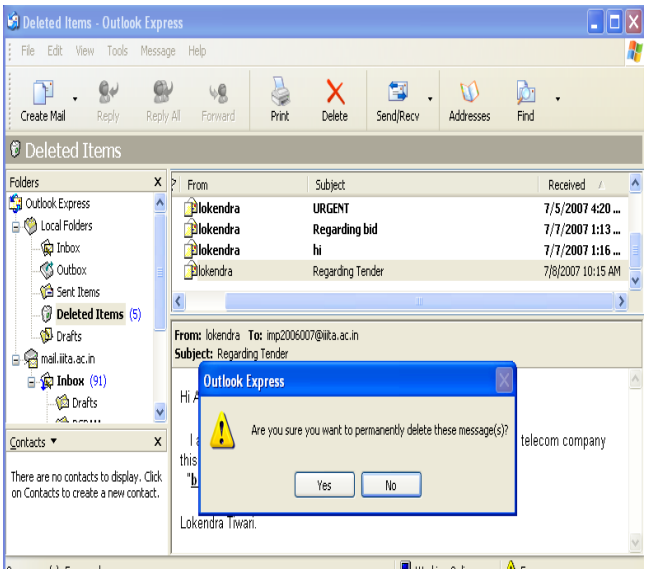

Figure 5: Window showing permanent deletion of sent email.

E-mails may contain *In-Reply-To* headers that allows to be reconstructed by using simple techniques and tools.

Step 4: The evidentiary e-mail may be recovered using search option of EnCase 6.0. Search window will appear and "*search for email option*" is selected as shown below.

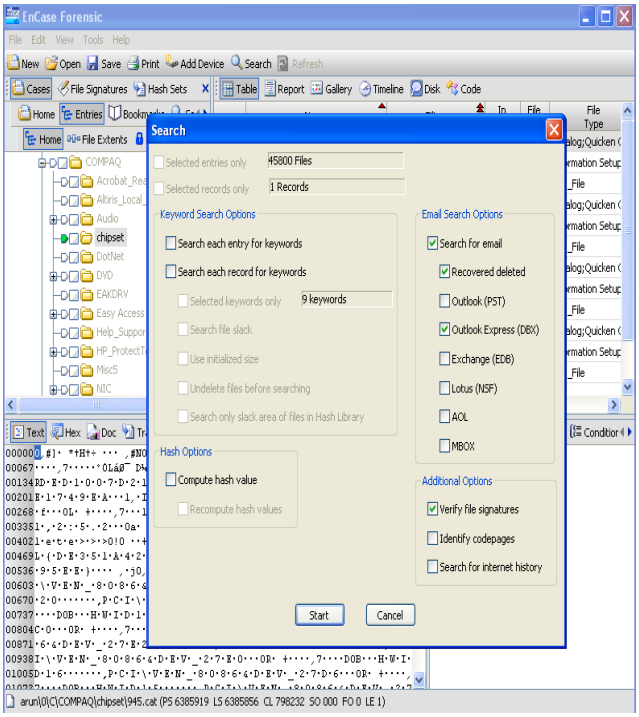

Figure 6: Encase 6.0 interfaces for searching email

On clicking on the start button searching will start scanning different files of different folders and would verify the file signature.

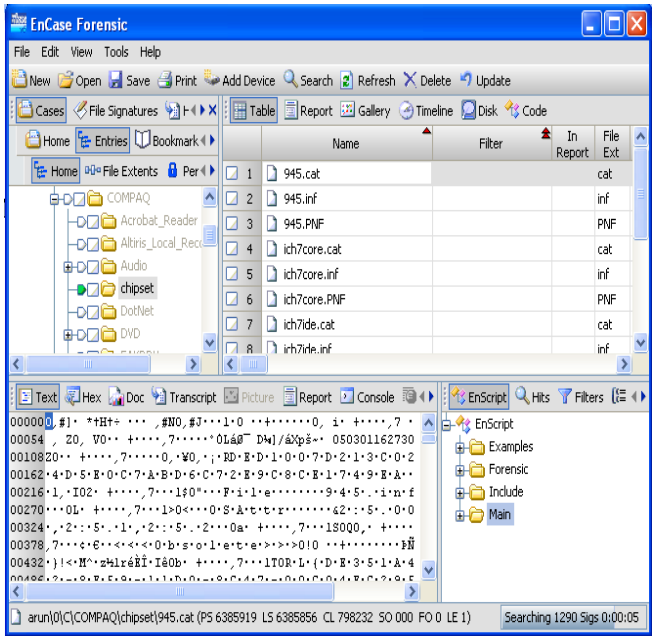

Figure 7: Window showing parsing of email at bottom right

Step 5: After the search is complete, the result will be displayed by confirmation of data in a window. Otherwise it shows the unavailability of data

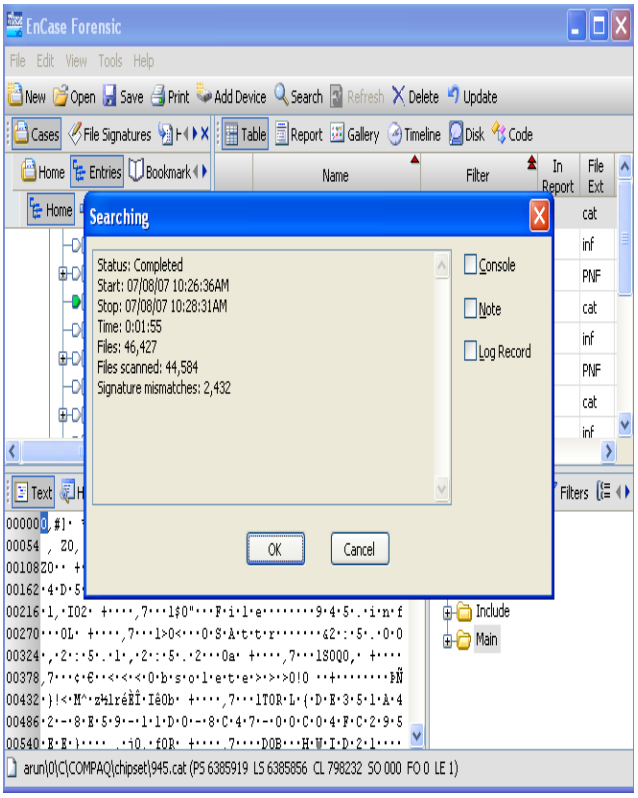

Figure 8: Window showing search is completed

Step 6: The records of e-mails of various categories can be obtained from recovered data by selecting appropriate file type, format and signature. Some of the attributes are available at the time of investigation; others are divulged in course of metadata analysis.

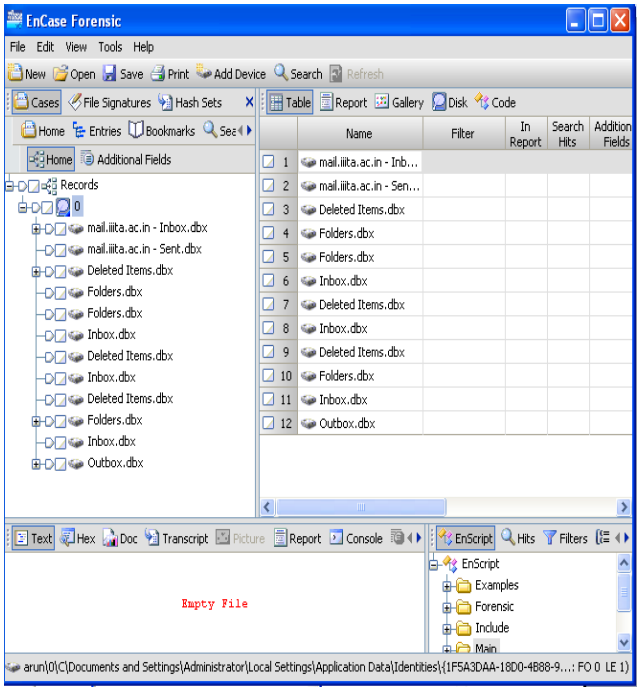

Figure 9: Window showing various recovered email format.

Step 7: Outbox.dbx is likely to provide traces of the desired email (if any) to be recovered.

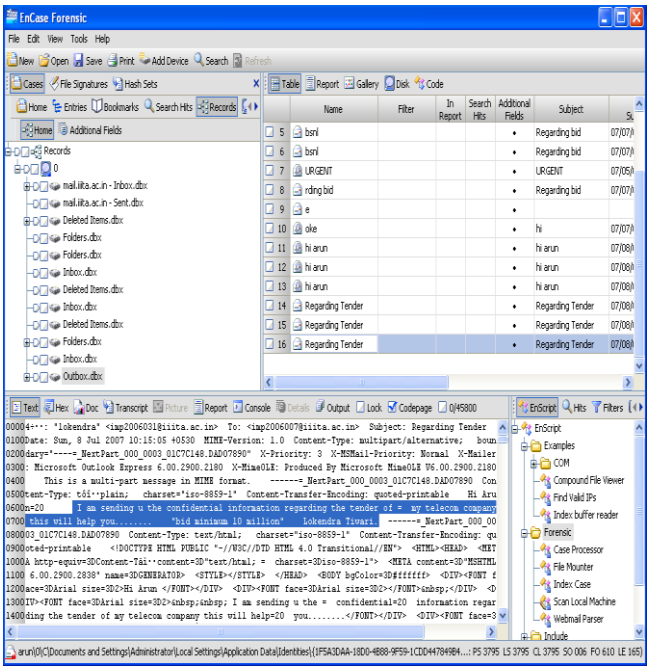

Figure 10: Window showing content of recovered email.

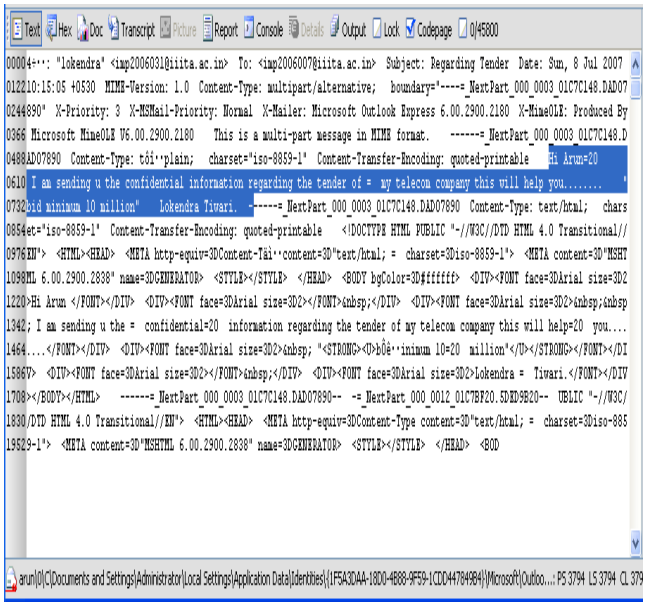

Figure 11: Detailed view of recovered email

V. CONCLUSION AND FUTURE DIRECTION OF WORK

This method can be extended for different file formats. "A well-defined strategy of suitability of different file formats and forensic software can be obtained as a result. Such formulation will serve the purpose of a guide book for the investigators of e-mail forensic. This ability to provide empirical evidence and identification of original author of e-mail misuse is of paramount importance in successful prosecution of an offender". In order to solve the problem of non-repudiation and computing forensic of such security issue with legal effect, a digital record is required to be maintained which can provide the accessory data to detect and obtain proof. This record should be maintained under self-security so that it cannot be destroyed or juggled by simple manipulation of computational data. The technical difficulty of digital forensic is that the forensic systems are mostly applied as tools. There are very few analytical commands at the operating system level to capture digital data access which may provide acceptable legal effect. The auditing devices of any computing system require enough hardware and software with certain programmable analysis capability. The auditing device must use a high level of self protection. The forensic auditing device may be designed on the platform of Network Forensic or Remote Forensic in future, enabling the user to activate or deactivate the system on the fly. There may be authorized body who are empowered to activate or deactivate this auditing device. The forensic auditing device may have the capabilities of the forensic tools and mobile Forensic Extract Device [35]. Building up evidentiary data in that case will be speed (hence time) and space effective. The other possibility is to generate digital forensic web services. These service can be activated through a publish subscribe architecture of service orientation. However due to a number of vendor specific or proprietary services, remote forensic deployment in Service Oriented Architecture (SOA) requires serious standardization effort.

VI. LIMITATIONS OF THE PROPOSED TECHNIQUE

This method may not work in case of counter forensics. For example, let us consider "MAFIA (Metaspoilt Anti-Forensic Investigation Arsenal)" [36]. Culprit who carries out any criminal activity tries to take all possible precautions he may uses various programs/scripts to wipe out all possible evidences of crime,

In the case of MAFIA, users can also change the time stamping in such cases investigators usually fails to obtained the desired evidences.

"Legally admissible data are required to be authentic. It is really easy to alter the real documents just by saving it another time. The computer forensic experts must ensure that no one else can break into their systems and change data that they are working on."

REFERENCES

1. https://en.wikipedia.org/wiki/Computer_forensics accessed on 01 August 2019
2. Asha Joseph , K. John Singh, "Review of Digital Forensic Models and A Proposal For Operating System Level Enhancements" International Journal of Computer Science and Information Security (IJCSIS), Vol. 14, No. 11, November 2016
3. https://articles.forensicfocus.com/, accessed on 11 September 2019
4. https://www.ontrack.com/uk/, accessed on 11 September 2009
5. https://www.cru-inc.com/industries/forensics/, accessed on 12 September 2019
6. Linux "dd" : www.redhat.com, accessed on 17 September 2019
7. http://www.x-ways.net/winhex/forensics.html, accessed on 11 September 2019
8. SafeBack :www.forensics-intl.com, accessed on 19 September 2019
9. http://www.digitalintelligence.com/software/guidancesoftware/encase/, accessed on 19 September 2019
10. SnapBack DatArrest :https://www.forensicmag.com/article/2006/08/software-imaginganalysis-tools-part-1, accessed on 19 September 2019
11. http://www.kb.cert.org/vuls/ accessed on 12-3-2019
12. http://cve.mitre.org/ accessed on 12-8-2019.
13. http://nvd.nist.gov/ accessed on 12-8-2019.
14. http://www.securityfocus.com/vulnerabilities accessed on 12-7-2019
15. http://www.webresourcesdepot.com/9-free-and-open-source-bug-tracking-softwares/ accessed on 12-7-2019.
16. https://blog.osvdb.org/ accessed on 12-8-2019.
17. http://projects.webappsec.org/Threat-Classification-Previous-Versions accessed on 12-7-2019.
18. http://xml.coverpages.org/appSecurity.html accessed on 12-7-2019.
19. http://www.w3.org/TR/xkms2/ accessed on 12-7-2019.
20. http://docs.oasis-open.org/ws-sx/ws-trust/200512/ws-trust-1.3-os.html accessed on 12-8-2019.
21. http://www.xrml.org/ accessed on 12-3-2010.
22. http://www.cdac.in/HTML/press/2q05/spot487.asp accessed on 12-7-2019
23. https://www.digitalforensics.com/ accessed on 12-7-2019