Novel Approach for Email Forensics

Mrityunjay
HMR Institute of Technology and Management
Delhi, India

Utsav Chauhan
HMR Institute of Technology and Management
Delhi, India

Mrs. Shally Gupta
(Asst. Professor)
CSE Department
HMR Institute of Technology and Management, Delhi, India

***Abstract* - Communication system has so far developed a lot it was not till now that we need a system to track what people share with each other with the use of Artificial Intelligence. Due to increasing crime it has become our priority to analyze what people share with each other. As soon as cell phone tapping has come into existence, criminals have switched over to digital media to share information on their mission. So, a system is required to keep a track of all the emails and other digital media and point out a suspicious activities to prevent cybercrime.**

Index Terms— Digital Forensics, Email Forensics, Manual Method, Accessdata's FTK, EnCase, Sawmill, DBXtract [11]

I. INTRODUCTION

Computer Forensics is a branch of digital forensic science pertaining to evidence found in computers and digital storage media. The goal of computer forensics is to examine digital media in a forensically sound manner with the aim of identifying, preserving, recovering, analyzing and presenting facts and opinions about the digital information.

Although it is most often associated with the investigation of a wide variety of computer crime, computer forensics may also be used in civil proceedings. The discipline involves similar techniques and principles to data recovery, but with additional guidelines and practices designed to create a legal audit trail.

When we come up to track our communication activities on the internet, then we come through a sub branch of Computer Forensics namely **Network Forensics.**

Network Forensics

Network forensics is a sub-branch of digital forensics relating to the monitoring and analysis of computer network traffic for the purposes of information gathering, legal evidence, or intrusion detection. Unlike other areas of digital forensics, network investigations deal with volatile and dynamic information. Network traffic is transmitted and then lost, so network forensics is often a pro-active investigation.

Network forensics generally has two uses. The first, relating to security, involves monitoring a network for anomalous traffic and identifying intrusions. An attacker might be able to erase all log files on a compromised host; network-based evidence might therefore be the only evidence available for forensic analysis. The second form relates to law enforcement. In this case analysis of captured network traffic can include tasks such as reassembling transferred files, searching for keywords and parsing human communication such as emails or chat sessions.

II. EMAIL ANALYSIS

Email analysis is the task preformed in the network forensics. Email analysis is the process which involves analysis of emails sent and received at different ends. In current era, there are very less ways to analyze emails. Most widely accepted method is the **Manual Method of Email Analysis** [10]**.**

While preforming manual method for email analysis, we try to spot spoofed messages which are sent through SMTP (Simple Mail Transfer Protocol). By analyzing them we can decode the message being sent. After decoding, all IP addresses are analyzed and their location is traced. A timeline of all event is made (in universal standard time) and is checked further for suspicious behavior. Server logs are checked at the same time to ensure that all the activities are mentioned in the timeline so formed. If any suspicious activity is found, the mails are recovered and can be used as evidence against the sender. Email is extracted from the client server which keeps a copy of sent mails until a specific number.

A case study involving the use of Manual Method for Email Analysis

- An email attached to a $20 million dollar lawsuit purported to be from the CEO of "tech.com" to a venture capital broker. The message outlined guaranteed "warrants" on the next round of finding for the broker.
- "tech.com" filled counterclaim and claimed the email was forgery. Their law firm engaged a team to determine the validity of the message.
- The team imaged all of the CEO's computers at his office and his home. Email server backup tapes were recalled from the client servers.
- All hard drivers and email servers were searched for "questioned" message. There were no traces of any such mail on any of the hard drive or mail spool.
- When the time stamps and message id's were compared with the server logs then it was found that the "questioned" message have not gone through either "tech.com's" webmail or mail server at the time indicated by the date/time stamp on the message.

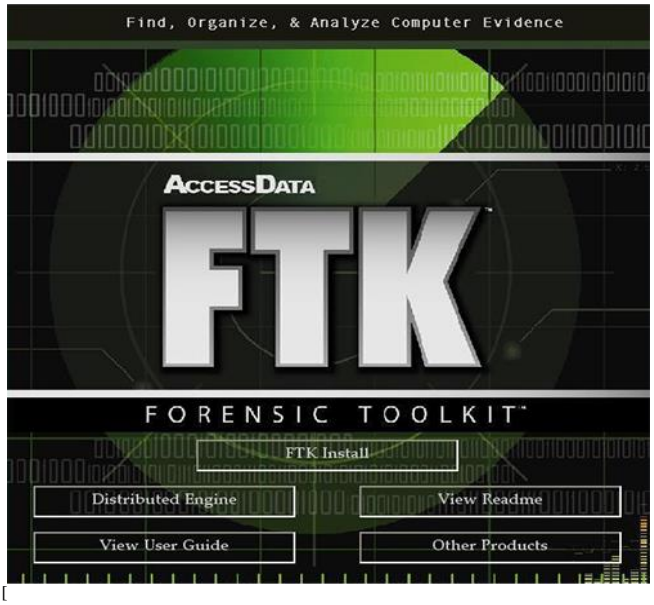
- Based on the analysis the defendants filed motion to image and examine broker's computers.
- Federal judge issued subpoena and the team arrived at the broker's business, he refused to allow his system to image.
- Broker's lawyer went into the state court, on a companion case, and got judge to issue an order for a new court appointed examiner.
- The examination revealed direct proof of the alteration of a valid message's header to create a "questioned" email.

Now, this is quite a bit long and tiring procedure which would involve too many mails to be analyzed which would be too much time consuming. Time being the most expensive entity, we need to save the time as much as we can. To save this time certain tools are present which helps to reduce the work burden. These tools are discussed in the next section.

III. TOOLS AVAILABLE FOR EMAIL ANALYSIS

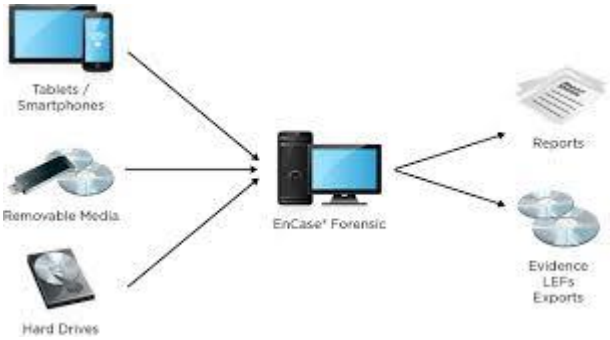
To save this time, we have designed certain toots which would help us in the faster analysis of the former. Some of the tools are as follows:

1. *Accessdata's Forensic Toolkit (FTK)*[1][12]
   A toolkit whose stable version was released back on June 3, 2013, Accessdata's Forensic tool kit is a court cited digital investigation platform built for speed, stability and ease of use. It provides comprehensive processing and indexing up front, so filtering and searching is faster than with any other product in use. This means we can "Zero-in" on the relevant evidence quickly, dramatically increasing our analysis speed. Furthermore, because of its architecture, FTK can be setup for distributed processing and incorporate web-based management and collaborative analysis.
   The only **disadvantage** of Accessdata's Forensic Toolkit is that it has to be manually run by the user to search for any suspicious activity.

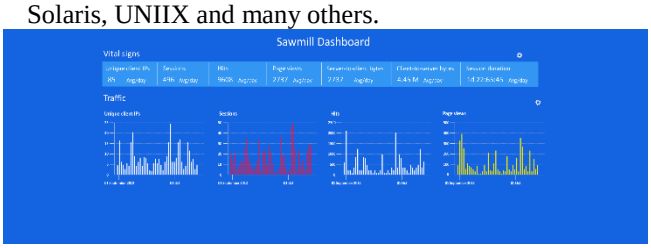


[

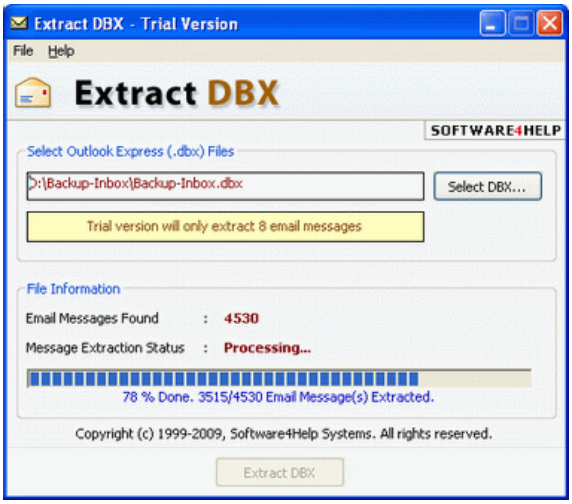
2. ***EnCase***[2][13]
   Released on March 24, 2016, EnCase is the shared technology within a suite of digital investigations products by Guidance Software. The software comes in several products designed for forensic, cyber security, security analytics, and e-discovery use. The company also offers EnCase training and certification. The basic purpose of this program is to recover the damaged data which could be useful as an evidence against the accused by the prosecution.
   Again the **disadvantage** of using this software is that it is able just to recover certain type of data, it cannot analyze anything about the source and the sender of the piece of text or media.



3. ***Sawmill-GroupWise log analyzer***[3][14]
   Initially released in the 1997, Sawmill-GroupWise log analyzer is a program that can process log files in Novell GroupWise Post Office format, and generate dynamic statistics from them. It can export all the data so formed to a MySQL, Microsoft SQL Server, Oracle Database, Aggregate them and generate a dynamically filtered reports, all through a web interface. It can perform the log files analysis on any platform, including windows, Linux, FreeBSD, OpenBSD, Mac OS, Solaris, UNIIX and many others.



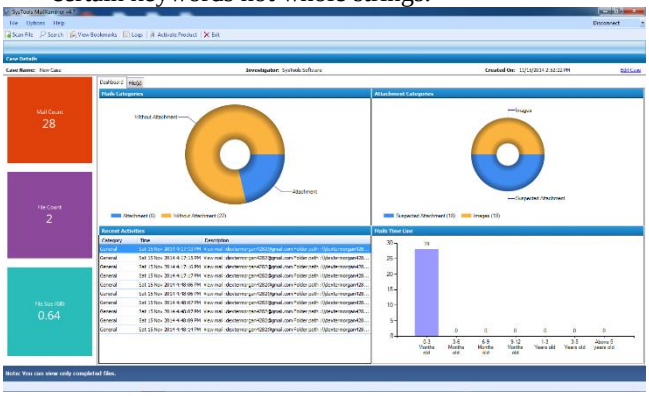
4. ***DBXtract***[4][15]
   DBXtract is a software program similar to EnCase. It is used for recovery of corrupted data which can be used as an important evidence against the accused by the prosecution.
   Again the **disadvantage** of using this software is that it is able just to recover certain type of data, it cannot analyze anything about the source and the sender of the piece of text or media.



5. *MailXaminer*[5][9][16]

MailXaminer is a digital forensic program released globally in December 2013, built to allow the examination of email messages from both web & application based email clients. The application is being developed by SysTools Inc., with the slogan 'Simplifying Email Forensics. MailXaminer first loads messages from the chosen email storage source and arranges them hierarchically for the purpose of evidence analysis and extraction. The product name derived from a combination of 'Mail' and 'Examiner', denoting it as a platform to examine emails. The programming of the application provides carving out of deleted evidence or evidence from damaged sources in cases of evidence spoliation. Post analysis, the software serves output generation in court admissible digital formats. It is able to search specific keywords on all the mails present on the computer system.

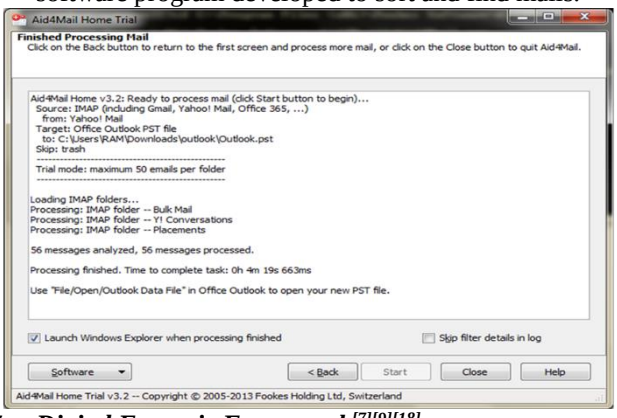
This software program has a disadvantage that it do not have a real-time working and moreover it can just find certain keywords not whole strings.



6. *Add4Mail*[6][9][17]

Another tool developed for helping in the mail sorting purpose only. This software can find emails which can be searched by any particular keyword. The output provided by this software program is the message written in the email along with the date, time and other information specific to the mail. This software program can also be used to fetch some deleted mails from their trash folder.

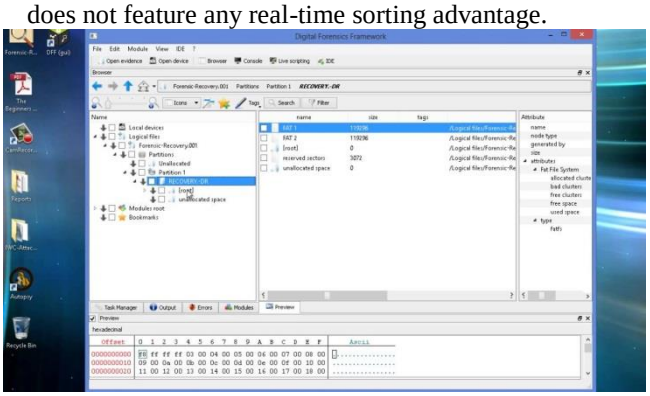
The major **disadvantage** of this software is that it can just find keywords which the user search for. It has no artificial intelligence and so is completely a manual software program developed to sort and find mails.



7. *Digital Forensic Framework*[7][9][18]

Released back on February 28, 2013, Digital Forensics Framework is computer forensics open-source software used by professionals and non-experts to collect, preserve and reveal digital evidence without compromising systems and data. It provides us the details like the message, date, time and other details of the mail. It is similar to Add4Mail in some ways.

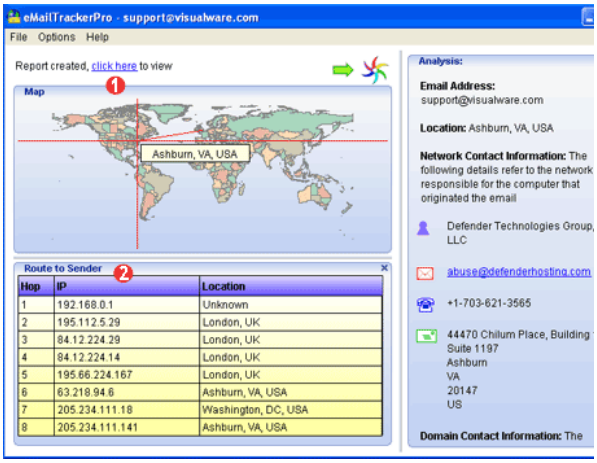
The major **disadvantage** of this software is that it can just find certain strings which the user search for. It has no artificial intelligence and so is completely a manual software program developed to sort and find mails. It does not feature any real-time sorting advantage.



8. *eMailTrackerPro*[8][9][19]

Yet another software made available to us by Visualware is a software program which can be used in checking for spamming contents automatically. It provides the IP address that sends the message along with geographical location of the former to determine the threat level or validity of an email. It check for its domains from its blacklisted server.

The disadvantage associated with this software is that it would be unable to find a spammer which is not blacklisted into its database.



S.No	Name	Release	Advantages	Disadvantages
1.	**Accessdata's Forensic Toolkit (FTK)**	June 3, 2013	Thoroughly filters all emails	Has to be run manually
2.	**EnCase**	March 24, 2016	It can recover damaged data	It can only recover damaged data and cannot analyze anything
3.	**Sawmill-GroupWise log analyzer**	1997	It can Analyze mails and can export them in the SQL format	It can only analyze mails for certain specific keywords
4.	**DBXtract**	--	It can recover damaged data	It can only recover damaged data and cannot analyze anything
5.	**MailXaminer**	December 2013	It can be used to analyze a mail for a complete string.	It do not have a real-time working and moreover it can just find certain keywords not whole strings
6.	**Add4Mail**	--	This software can find emails which can be searched by any particular keyword	It can just find keywords which the user search for
7.	**Digital Forensic Framework**	February 28, 2013	It provides us the details like the message, date, time and other details of the mail	It can just find certain strings which the user search for
8.	**eMailTrackerPro**	N.A	It provides the IP address that sends the message along with geographical location of the former to determine the threat level or validity of an email	It would be unable to find a spammer which is not blacklisted into its database

IV. THE CURRENT PROBLEM

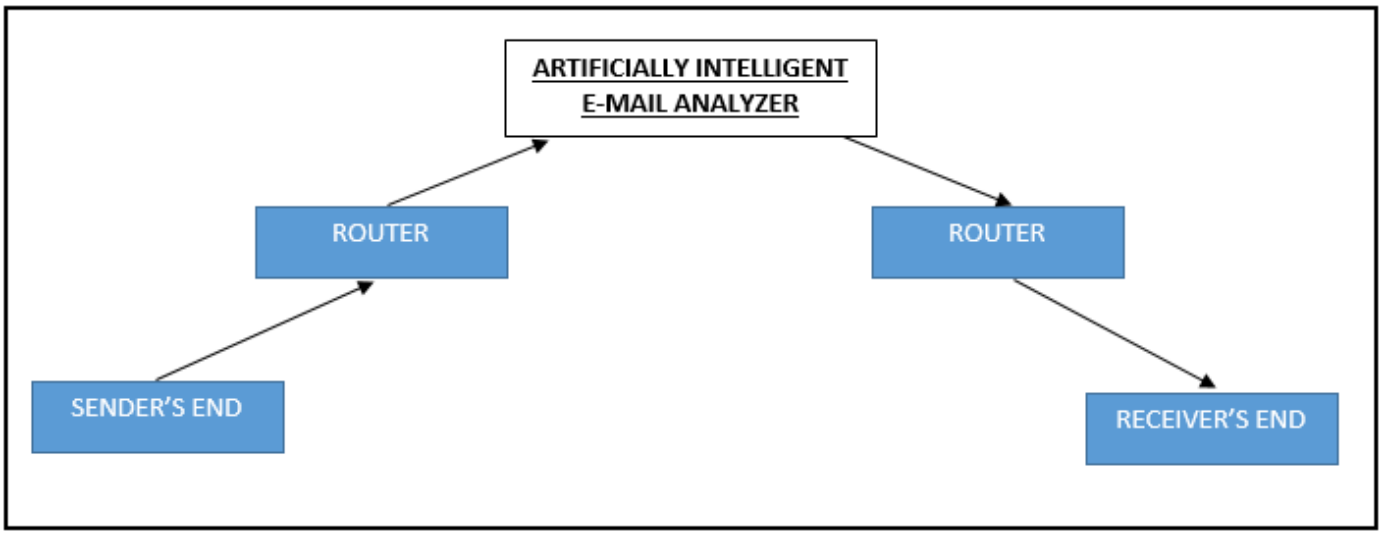
Although we have all the above listed software but still we cannot feel safe as none of the tool available to is real-time. So it means that these tools can be used post crime for the investigation purpose. There is no software so formed that can work real-time and check all the mails going through the server and can push the mail to the network forensic expert if it find any suspicious activity in the same. So, a software is needed which can serve all the purposes that a good network analyzer should possess. It must indicate all suspicious activities, must trace the location, IP address, ISP of the sender.

V. PROPOSED SOLUTION

Here is an idea of what can be a solution to all these problems. The idea of a miscellaneous software which would be artificially intelligent, will analyze all the mails going through the internet and will gather the notice of the network forensic analyzer according to the threat level of the mail so identified.

VI. PROPOSED ARCHITECTURE OF THE SOFTWARE

Here, we have the proposed architecture of the automatic email analyzer as follows.



Here between the sender end and receiver end, we have an artificially intelligent email analyzer between the routers of the network connections.

This email analyzer would be powered by artificial intelligent which would analyze all the emails going through the network connection.

The emails can be analyzed in many ways but there can be two most feasible ways to solve this problem which are as follows:

- The first way is that it would check all the strings of the mails using a string filter embedded at the central node of the network. As soon as some similar message from same sender to same receiver is encountered, it could be marked as a suspicious activity and can be pushed to the forensic analyst after particular number of attempts of that type of similar messages.
- Another way of analysis is that it can check for the header (subject) of the mail and the content of the mail. If no relation is established between the subject and the content, it can mark it as a suspicious mail and again after many similar mails, it can push it to the forensic analyzer for further deep analysis.

VII. CONCLUSION

Finally we can conclude that due to lack of technology needed to form this software program to monitor all the emails, it is currently just an idea as what can be alternate solution to the problem of security in the email forensics that can be implemented in near future with the advent of technology.

VIII. REFERENCES

[1] https://en.wikipedia.org/wiki/Forensic_Toolkit
[2] https://en.wikipedia.org/wiki/EnCase
[3] https://en.wikipedia.org/wiki/Sawmill_(software)
[4] https://www.technibble.com/repair-tool-of-the-week-dbxtract-45/
[5] https://en.wikipedia.org/wiki/MailXaminer
[6] http://www.aid4mail.com/
[7] https://en.wikipedia.org/wiki/Digital_Forensics_Framework
[8] http://www.emailtrackerpro.com/
[9] Research paper by Vamshee Krishna Devendran on "A Comparative Study of Email Forensics"
[10] A research paper by Sobiya R. Khan on "Email Data Analysis for Application to Cyber Forensics investigation using data mining"
[11] https://fenix.tecnico.ulisboa.pt/downloadFile/1970943312267438/csf-13.pdf
[12] https://www.google.co.in/search?hl=en&site=imghp&tbm=isch&source=hp&biw=1366&bih=638&q=ftk&oq=ftk&gs_l=img.3..0l10.8317.8912.0.9114.3.3.0.0.0.0.193.193.0j1.1.0....0...1ac.1.64.img..2.1.191.zoCdFiLcSGw#
[13] https://www.google.co.in/imgres?imgurl=http%3A%2F%2Fwww.ndm.net%2Fediscovery%2Fimages%2Fstories%2Fimages%2F01encase_forensic.jpg&imgrefurl=https%3A%2F%2Fwww.ndm.net%2Fediscovery%2FGuidance-Software%2Fencase-forensic&docid=kYRnFJvEJsRKkM&tbnid=riOYXYC1xXhysM%3A&vet=10ahUKEwiq6sX9yaDTAhXFq48KHUr3Cl0QMwg-KAwwDA..i&w=1280&h=705&hl=en&bih=638&biw=1366&q=encase&ved=0ahUKEwiq6sX9yaDTAhXFq48KHUr3Cl0QMwg-KAwwDA&iact=mrc&uact=8
[14] https://www.google.co.in/imgres?imgurl=https%3A%2F%2Fwww.sawmill.co.uk%2Fassets%2Fimg%2Fslider%2Fvital_signs.png&imgrefurl=https%3A%2F%2Fwww.sawmill.co.uk%2F&docid=m_E3Up42jd6-SM&tbnid=At4UGWKy3faogM%3A&vet=10ahUKEwi5pO2hyqDTAhVDwI8KHfMDDZwQMwgkKAQwBA..i&w=2680&h=890&hl=en&bih=638&biw=1366&q=samwill%20group%20wise%20log%20analyzer&ved=0ahUKEwi5pO2hyqDTAhVDwI8KHfMDDZwQMwgkKAQwBA&iact=mrc&uact=8
[15] https://www.google.co.in/imgres?imgurl=http%3A%2F%2Fwww.extractdbx.com%2Fimg%2Fextract-dbx.gif&imgrefurl=http%3A%2F%2Fwww.extractdbx.com%2F&docid=-huUGqEVQVCPRM&tbnid=pfTglw1NFyReCM%3A&vet=10ahUKEwj--IK1yqDTAhUJQo8KHX-9AJAQMwggKAAwAA..i&w=380&h=335&hl=en&bih=638&biw=1366&q=extract%20dbx&ved=0ahUKEwj--IK1yqDTAhUJQo8KHX-9AJAQMwggKAAwAA&iact=mrc&uact=8
[16] https://www.google.co.in/url?sa=i&rct=j&q=&esrc=s&source=images&cd=&cad=rja&uact=8&ved=0ahUKEwjmlOTvy6DTAhVEPY8KHX6sDH0QjRwIBw&url=http%3A%2F%2Fwww.dataforensics.org%2Ftop-cyber-forensic-investigation-analysis-tools%2F&psig=AFQjCNEQohguTPRjCnTqtQnBpkSBOIG66w&ust=1492144044160089
[17] https://www.google.co.in/imgres?imgurl=http%3A%2F%2Fhtml.scirp.org%2Ffile%2F5-7800270x7.png&imgrefurl=http%3A%2F%2Ffile.scirp.org%2Fhtml%2F5-7800270_55520.htm&docid=DLt1JPDODVSFBM&tbnid=-qskvAonGtzP7M%3A&vet=10ahUKEwiAiYniyqDTAhVJK48K

HTYaCw4QMwggKAAwAA..i&w=902&h=563&hl=en&bih=638&biw=1366&q=add4mail&ved=0ahUKEwiAiYniyqDTAhVJK48KHTYaCw4QMwggKAAwAA&iact=mrc&uact=8

[18] https://www.google.co.in/imgres?imgurl=https%3A%2F%2Fi.ytimg.com%2Fvi%2F02uZv72KS88%2Fmaxresdefault.jpg&imgrefurl=https%3A%2F%2Fwww.youtube.com%2Fwatch%3Fv%3D02uZv72KS88&docid=0zFGT01EJ4n-sM&tbnid=hlwPjZ5vYZh3JM%3A&vet=10ahUKEwiRhqaEy6DTAhXLr48KHcBtBQgQMwghKAEwAQ..i&w=1280&h=720&hl=en&bih=638&biw=1366&q=digital%20forensics%20framework&ved=0ahUKEwiRhqaEy6DTAhXLr48KHcBtBQgQMwghKAEwAQ&iact=mrc&uact=8

[19] https://www.google.co.in/imgres?imgurl=http%3A%2F%2Fwww.niharsworld.com%2Fwp-content%2Fuploads%2F2009%2F12%2FEmail-Tracing-using-eMailTrackerPro-v8.gif&imgrefurl=http%3A%2F%2Fwww.niharsworld.com%2F2009%2F12%2F10%2Fdownload-full-version-emailtrackerpro-v8-free-valid-registration-code%2F&docid=IOua-yFLSyHReM&tbnid=RBxCb-weMb7a0M%3A&vet=10ahUKEwih_ayhy6DTAhXJKo8KHZhRDjIQMwghKAEwAQ..i&w=496&h=351&hl=en&bih=638&biw=1366&q=email%20tracker%20pro&ved=0ahUKEwih_ayhy6DTAhXJKo8KHZhRDjIQMwghKAEwAQ&iact=mrc&uact=8