ISSN:2249-5789

Review of E-mail System, Security Protocols and Email Forensics

Gurpal Singh Chhabra
Asst. Professor, Thapar University, Punjab (India)
gurpal.singh@thapar.edu

Dilpreet Singh Bajwa
Asst. Professor, CGC, Punjab (India)
dilpreetbajwa10@gmail.com

Abstract: E-mail is the most common mode of communication today. E-mail not only used for sending messages/text only but also to send audio, video and other files as attachment. It is main resource for business communication. As it is most popular and common mode of communication on internet, it also attracts criminals or persons having mischievous intent. Cyber criminals misuse it for sending spam, threats, Phishing-emails, propagating malicious software like virus and worms, distributing illegitimate material like child pornography, hoaxes and also used for other criminal activities. So it is necessary to secure our e-mail system and also to identify criminal, collect evidence against them and punish them under court of law. This paper review working and architecture of current email system and the security protocols followed generally to secure our email communications and the limitations they contained, further email forensics which is a process to analyze e-mail contents, header information, transit path for email, sender or receiver information and other details to collect evidence against culprit or to make our system more secure is discussed. It also discusses common email forensic investigation technique and tools used in email forensic process.

Keywords: E-mail, E-mail Architecture, E-Mail Security Protocol, E-mail Forensics, E-mail Investigation, E-mail Forensics Tools.

Introduction: E-mail is popular internet application. Millions of user daily uses e-mail for personal, business or for official purpose. As e-mail system influence our life so much today in a positive way, on the other way cyber criminals are also using e-mail as tool to fulfill their malafide intentions. So it is better to understand our email system and how it works and how it could be utilized to its maximum potential. It is also good to know what security protocols are commonly used to secure our communications while using email and up to what extent they are safe. Further what type of threats are facing from the cyber criminals while using e-mail as our mode of communication and how to protect ourselves or what e-mail forensic investigation tools and techniques are used by cyber forensic personnel to put these cyber criminals behind bars and prove their crime in court of law. This paper is an attempt to answer all these questions. This paper is broadly divided in to three parts, First part discussed current E-mail System Architecture, and how E-mail System works. It also discussed role of protocols like SMTP, POP3, IMAP and HTTP, in e-mail communication. In Second part the threats are discussed, facing during e-mail communication like eavesdropping, identity theft, Message modification, phishing attacks, E-mail Spamming, Repudiation, E-mail Spoofing, E-mail bombing etc.; Security protocols are also discussed which are currently in use like SSL/TLS, PGP and S/MIME, their limitations and role. In Third part of paper E-mail Forensics, E-mail Forensic Investigation techniques are discussed using by forensic personal like Header Analysis, Server Investigation etc. popular tools used by forensic personnel like Aid4Mail, Paraben EMX etc. discussed which help in these investigation and the features provide by them to collect evidences against the culprit.

Part I: Email Architecture and Working [1]

E-mail is the main mode of communication and very popular internet application in today's world. Its architecture consists of various components. The working and architecture of Email- Service is discussed. So that more easily define E-mail Forensics tools and techniques.

In the beginning, the messages sent through email are short and generally consisted of text only. Today Email is much more complex and can send messages which include text, audio and video also.

First, have a look at the general architecture of the email system in today's Scenario, after that the paper discusses the protocols which are used to implement the components of email system.

In the first case, email sender (Dilpreet) and receiver (Bajwa) are directly connected to the same/shared system. On this system there is one mailbox corresponding to each user, only the legitimate owner of the mailbox can access the mailbox after verifying his credentials.

User Agent (UA) program runs by Sender (Dilpreet) to compose the message and also store it in mailbox of receiver (Bajwa). The message contains the e-mail address of recipient and sender. Receiver (Bajwa) can

access his mailbox and read the contents of message using a User Agent (UA). Figure 1.1 above shows this case.

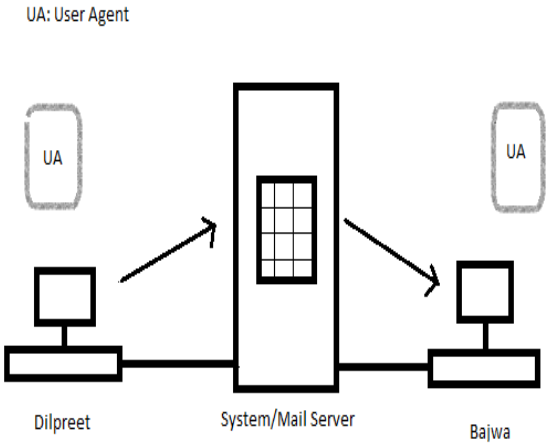


Figure: 1.1

In the second case, receiver and sender are working on two different systems. The message is transferred through internet as depicted in Figure 1.2 below:

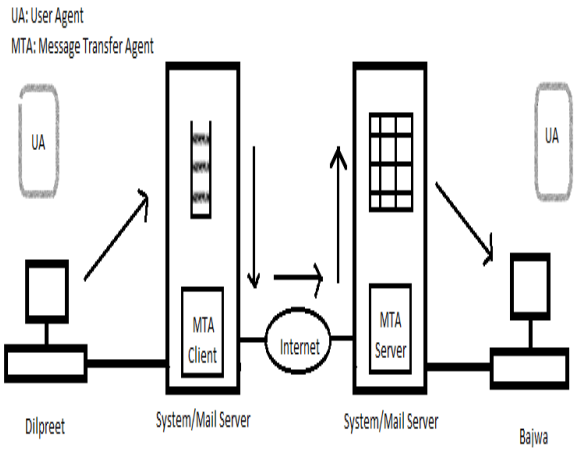


Figure: 1.2

Here as shown in above figure:1.2 we required one User Agent (UA) and one Message Transfer Agent (MTA) for both sender and receiver separately. Sender (Dilpreet) needs a UA program to prepare and send his message/mail to mail server available at his site. Similarly receiver (Bajwa) also needs a UA to access the messages/mails exist in the mailbox of the mail server available at his site. In addition to user agents, one MTA client required at sender's mail server and one MTA server required at receiver's end. The MTA server runs all time as it does not know when client will want to make a connection. On the other hand MTA client does not required to run all time because it is alerted by the system whenever any message is required to send.

In the third case, receiver (Bajwa) as mentioned in second case is connected directly to his mail server. Whereas sender (Dilpreet) is not connected directly to mail server. Dilpreet is either connected to the mail server/system via WAN or he is connected through a LAN. Figure 1.3 shown below clarifies the situation.

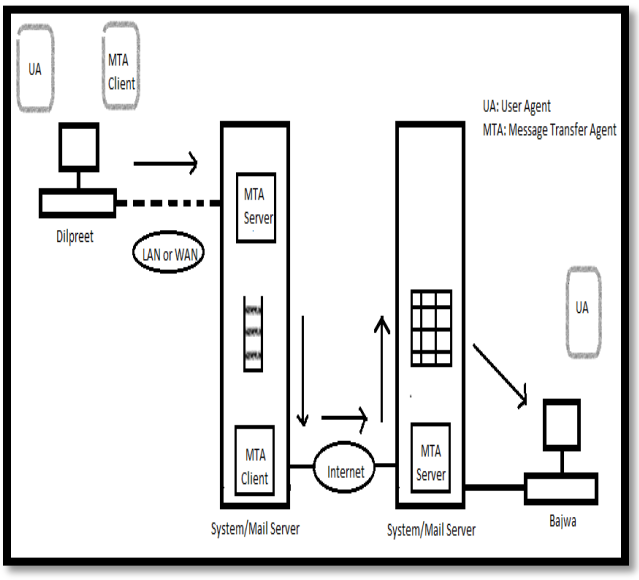


Figure: 1.3

In this scenario sender (Dilpreet) still require a UA to prepare his message. He then sends the message to mail server through LAN or WAN and this can be accomplished by using a pair of Message Transfer Agents (MTAs: MTA client/server). This all work like this, when sender (Dilpreet) wants to send some message, he calls the UA which further calls the MTA client. This MTA client makes a connection with MTA server which is running continuously on mail server at sender's site. The mail server system at sender's site queues all received message from Dilpreet and use its MTA client for transferring messages/mail to mail server system at receiver's (Bajwa) site. Mail server at Bajwa's site can receives the messages and stores it in Bajwa's mailbox. Bajwa uses his user agent to access and read messages from his mailbox.

In the fourth case which is most commonly used today. The receiver (Bajwa) is also not connected to his mail server directly but through some LAN or WAN. When message is reached at Bajwa's mail server, Bajwa wants to access it but as now Bajwa is not directly connected to the mail server at his site. So he needs one more pair of client/server agents called as Message Access Agents

(MAAs) for accessing messages from the mail server. Figure 1.4 given below clarifies the situation.

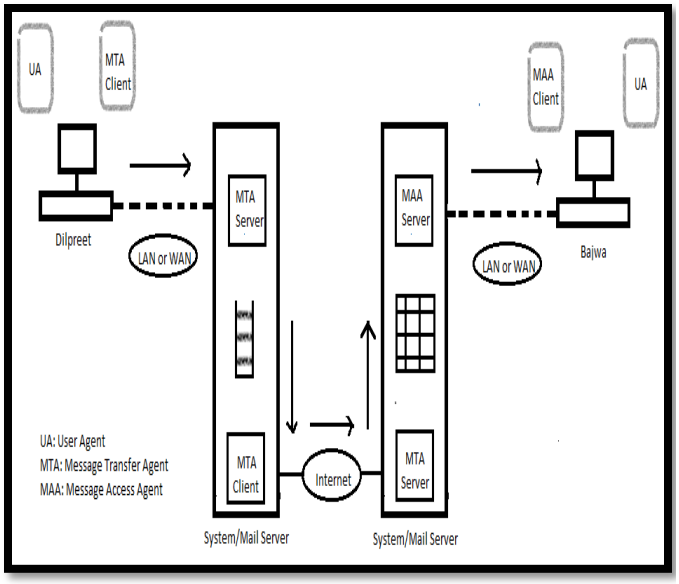


Figure: 1.4

Bajwa here needs another pair of MAA client/MAA server programs other than MTA Client/MTA Server to retrieve the messages because MTA Client/Server is called as a push program which pushes the message from client side to server side whereas Bajwa required a pull program i.e client require to pull messages from server side. Figure 1.5 given below shows difference between both pull and push programs.

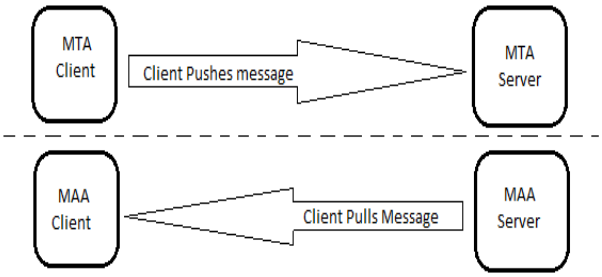


Figure: 1.5

Now take a look in more detail towards main components of E-mail System.

User Agent: It provides services like composing, reading, replying to messages, message forwarding and arranging/tackling mailboxes to make the process easy of sending and receiving messages for user. User agents are of two types: Command Driven and GUI Based. Generally we used GUI Based user agents because they are easy to use and it is easy to access services. Example of GUI based user agent is Netscape and Outlook.

User Agent works as follows:

Sending Mail: When user wants to send a mail, he creates mail through UA that is somewhat same concept as postal mail and has Envelope and Message. Generally envelope contains the address of sender and receiver and the message constitutes header and body. Here header defines sender, receiver, message subject and other related details like type of encoding used. The body includes the actual information sender wants to send and intended for recipient.

Addresses: To deliver a mail from sender to receiver, the mail handling system must use an addressing scheme to uniquely identify a user. In this addressing scheme, the address is divided in to two parts: a local part and a domain name. Both these parts are divided by @ sign for example: dilpreetbajwa10@gmail.com. Here dilpreetbajwa10 represents the local part and gmail.com specifies the domain name. The local part represents the user mailbox where emails received for user is stored and later on can retrieve by the user through message access agent. The Second part i.e the domain name represents the mail server or host. An organization generally has one or more mail servers for sending and receiving emails. The domain name uses for each mail server is either a logical name representing organization or comes from the DNS database.

Receiving Mail: At the receiving end, the UA is triggered by either the user or timer. The UA informs the user regarding availability of email. UA displays summary of messages in the mail box to the user. The user can select message he wants to read and the contents can display on screen.

MIME [1]: Electronic mail can use only NVT-7 bit ASCII format for sending messages. It is consider as a disadvantage because it cannot be used for those languages which don't support 7 bit ASCII format like German, Russian, Chinese etc. Secondly, it cannot be used for sending audio, video and binary files. To overcome this limitation Multipurpose Internet Mail Extension (MIME) is used, which is also a protocol and it converts non-ASCII data to ASCII data and vice versa. MIME uses five fields/headers containing information like MIME-Version, Content-Transfer Encoding, Content-Description, Content-Type, and Content-Id. These headers can combine with the original header section and helps in conversion.

<table>
<tr><th>Header</th><th colspan="3">Description</th></tr>
<tr><td>MIME-Version</td><td colspan="3">Display MIME Used version</td></tr>
<tr><td colspan="4"></td></tr>
<tr><td>Content-Type</td><td colspan="3">Shows type of data available in body of the message like Content-ID: < type/subtype: parameters>. Type/Subtype are defined below.</td></tr>
<tr><td></td><td>Type</td><td>Subtype</td><td>Description</td></tr>
<tr><td></td><td rowspan="2">Text</td><td>Plain</td><td>Unformatted</td></tr>
<tr><td></td><td>HTML</td><td>HTML format</td></tr>
<tr><td></td><td rowspan="4">Multipart</td><td>Mixed</td><td>Body contains ordered parts of different data types</td></tr>
<tr><td></td><td>Parallel</td><td>Same as above but not ordered</td></tr>
<tr><td></td><td>Digest</td><td>Similar to mix subtypes, but the default is message/RFC822</td></tr>
<tr><td></td><td>Alternative</td><td>Parts are different versions of the same message</td></tr>
<tr><td></td><td rowspan="3">Message</td><td>RFC822</td><td>Body is an encapsulated message</td></tr>
<tr><td></td><td>Partial</td><td>Body is a fragment of bigger message</td></tr>
<tr><td></td><td>External-Body</td><td>Body is a reference to another message</td></tr>
<tr><td></td><td rowspan="2">Image</td><td>JPEG</td><td>Image is in JPEG format</td></tr>
<tr><td></td><td>GIF</td><td>Image is in GIF Format</td></tr>
<tr><td></td><td>Video</td><td>MPEG</td><td>Video is in MPEG format</td></tr>
<tr><td></td><td>Audio</td><td>Basic</td><td>Single channel encoding of voice at 8kHz</td></tr>
<tr><td></td><td rowspan="2">Application</td><td>Post Stream</td><td>Adobe PostScript</td></tr>
<tr><td></td><td>Octet-Stream</td><td>Binary Data</td></tr>
<tr><td colspan="4"></td></tr>
<tr><td>Content-Transfer Encoding</td><td colspan="3">It defines the method used to encode the messages in to binary for transport. Like: Content-Transfer-Encoding: <type></td></tr>
<tr><td></td><td>Type</td><td colspan="2">Description</td></tr>
<tr><td></td><td>7 - bit</td><td colspan="2">NVT-ASCII Characters and short lines</td></tr>
<tr><td></td><td>8 - bit</td><td colspan="2">Non-ASCII Characters and short lines</td></tr>
<tr><td></td><td>Binary</td><td colspan="2">Non-ASCII Characters and unlimited short lines</td></tr>
<tr><td></td><td>Base – 64</td><td colspan="2">6 bit blocks of data encoded in to ASCII Characters</td></tr>
<tr><td></td><td>Quoted-Printable</td><td colspan="2">Non-ASCII characters encoded as an equals sign followed by their ASCII code.</td></tr>
<tr><td colspan="4"></td></tr>
<tr><td>Content-Id</td><td colspan="3">It uniquely identifies a message. Like: Content-Id: id=<content id></td></tr>
<tr><td colspan="4"></td></tr>
<tr><td>Content-Description</td><td colspan="3">It defines whether body is image, audio or video. Like: Content-Description: <description>.</td></tr>
</table>

Message Transfer Agent (MTA): SMTP: Actual transfer of mail is accomplished through MTAs. For sending an email, a system must have MTA client and for receiving email it must have MTA server. The protocol used to handle the working and proceedings of MTA client/server is Simple Mail Transfer Protocol (SMTP). SMTP is used at two stages, one between sender and his/her mail server and another between sender's and receiver's mail server. Between receiver's mail server and receiver either POP3 or IMAP4 protocol can be used because SMTP is a push protocol and at receiver's end, a pull protocol is required. SMTP simply uses command and responses for transferring of messages between MTA client and MTA server.

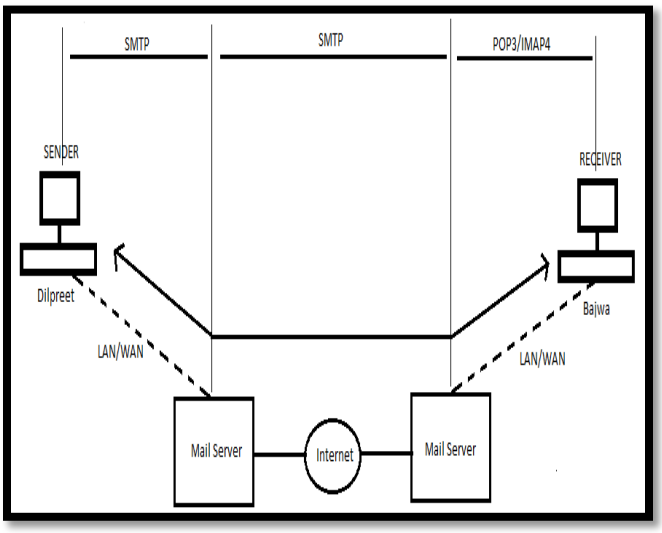


Figure: 1.6

MTA client send commands to server while MTA server sends responses to client. The message transferring procedure is divided in to three phases: connection establishment, message/mail transfer, connection termination.

Message Access Agent (MAA): POP3 and IMAP: As discussed earlier, the first stage (between sender and his mail server) and second stage (between mail server of sender and receiver) use SMTP. SMTP is not used in third stage (between receiver's mail server and receiver) because SMTP is a push protocol. Whereas the third stage required pull protocol i.e the client have to pull messages from server. So for this third stage Message Access Agents is used which implements generally one of two protocols POP3 or IMAP. Both are pull protocols.

POP3: POP3 stands for Post Office Protocol 3; it is quite simple with limited in functionality. POP3 client installation takes place at receiver's computer and POP3 server software installation must be done at mail server to which receiver is connected. When user wants to download mail from his mailbox residing on mail server then the MAA client establish a connection with MAA server of mail server on TCP port 110, verified its identity by passing username and password. After verification he can access his mailbox and retrieve his mails messages.

IMAP4: Internet Mail Access Protocol, Version 4 (IMAP4) is also a mail access protocol which is somewhat same as POP3 but with more features. POP3 has several limitations like it cannot allow user to arrange emails on server, user also does not categorize email in different folders on server, further it does not allow user to partially check email from server before downloading. IMAP4 overcomes all these limitations of POP3 and also add more functions like user can also create mailboxes hierarchy in a folder for email storage, a user can search mail by entering a particular string of interest before downloading from server.

In today's world email is so common that some websites like Yahoo etc. provide the service to anyone who accesses the site. The procedure is same but now the mail transfer from sender (Dilpreet) to his mail server is takes place by using HTTP protocol. The message transfer from sender's mail server to receiver's mail server is still accomplished by using SMTP. At last, the transfer from receiver's mail server to receiver's (Bajwa) browser is accomplished again through HTTP.

HTTP: Hyper Text Transfer Protocol (HTTP) is mainly used for accessing HTML documents on World Wide Web (WWW). It is somewhat mix of FTP and SMTP. It is like FTP because it is used also to transfer files and used services of TCP but it uses only one connection in contrast to FTP which uses two connections, one for control and one for data transmission. HTTP is similar to SMTP because the message format in both is almost same; further the format of headers is controlled by MIME like headers. Headers in HTTP are not intended for user as in SMTP, They are used by HTTP Server/Client for reading or interpretation. HTTP uses the TCP services at port 80. The request message from client to server contains the commands and requested contents from server to client are embedded in response message.

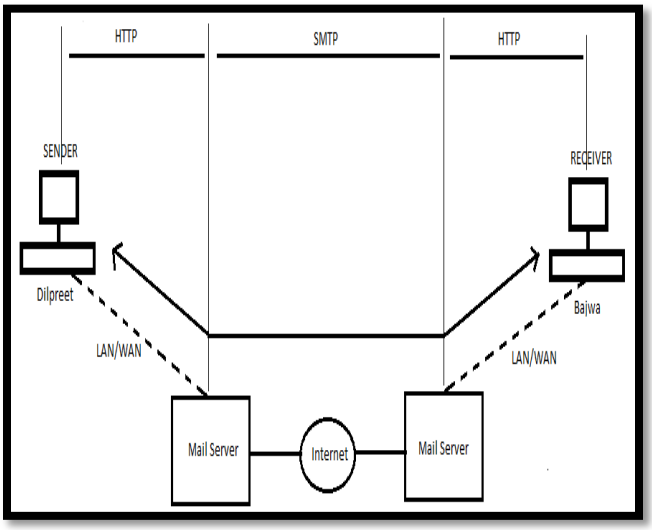


Figure: 1.7

Part II: Security Issues and vulnerability in Email System:

E-mail is one of the main modes of communication today but in the following section it can be seen how insecure it is. The importance of email is for corporate and private communication can be estimated by the summary presented by Radicati Group's report titled "E-Mail Market, 2012-2016" that the world wide each day total emails sent in 2012 was 144.8 billion, which is increased steadily with each passing year and in 2016 approximately 192.2 billion emails will sent each day. The report also states that corporate webmail clients grow from 629 million in 2012 to over one billion by the end of 2016 [23][24].

Threats in Email Communication:

Eavesdropping [25]: E-mail messages pass through networks which are part of big picture i.e Internet with a lot of people on it. So it is very easy for someone to track or capture your message and read it.

Identity Theft: Means someone pretend to be you on the network. It may be possible if not proper security protocols are followed that someone may steal or capture

your username/password and used to read your email messages. Further also send email messages from your account without your knowledge.

Message Modification: Anyone who captures your message can also alter your message contents if it is not encrypted. Further anyone having administrative rights on any of SMTP server your message visit can not only read your message but can also modifies it.

False Messages: Sender's name can easily be fabricated so it is very easy to send message that pretends to be send by someone else.

Unprotected Backups: Messages generally stored in plain Text on SMTP server and also backups can be created. Even if you delete the message they can be residing on the severs/backup-servers for years. So anyone who accesses these servers can also access or read your message.

Repudiation: As it is known that email messages can easily be forged so anyone sending you some message can later on deny regarding sending of message and it is very difficult to prove it. This has implications corresponding to emails use as contracts in business communications.

Email spoofing [18][19]: Sometime email that pretends to be received from an authentic source but in actual it is send from somewhere else.

Email Spamming [20]: Spam or junk mail refers to sending of email to no. of persons for any advertisement purpose or for some malicious intent. To send spam often lists are created by searching data from Internet, or by stealing mailing list from the internet.

Email bombing [25]: E-mail "bombing" is refers to sending identical mail repeatedly by abusers to a particular address/user.

Sending threats: Threatening mails are sending to users which disturb their state of mind or to provoke them to take some wrong step. Sometimes false statements are also forwarded to third parties or users to injure the reputation of some particular person. It is called as Defamation, a communication is not consider defamatory unless it is forwarded to someone other than the target.

Email frauds: Email Fraud is the intentional deception made for some personal or monetary gain.

Emails used as tools to spread malicious software: Emails are also used as tools to spread viruses, worms and other malicious software. They are attached to your emails as attachment, when you click on them they attack your computer or browser.

Phishing [21][22]: It is also most common attack through email. It is originally defined as an attack to steal your confidential information like passwords, ATM pin and other bank credentials. It works as some email coming to you that pretends to be from some trusted source you know like your bank. These emails entice you to click on some link present in email or to open some attachment or respond to some message and that click directed to you their site in actual but it appears like your trusted website of bank and ask to fill some confidential information like passwords which is actually stolen from you and use for any malicious intent later on.

Limitations exist in currently used protocols:

Any Network service like email system must provide following five services for security reasons:

Message Confidentiality: It promotes privacy that is the message transfer between sender and receiver is secure and no one can read or track the message while transferring.

Message Integrity: It says that the same message/data should arrive at receiver end as it can be send by sender. No alteration intentionally or accidentally takes place during transfer.

Message Authentication: It ensures that message can be received from the sender only or from the trusted source. In this receiver must be sure about the identity of sender.

Message Non-repudiation: It ensures that anytime sender should not be able to deny sending of message which originally sends by him/her.

Entity Authentication: It ensures identification of user; the user must be verified before accessing the resources and services. This is done by asking login-id and password.

SMTP: SMTP does not encrypt messages. So the communication between SMTP servers is in plain text so eavesdropping takes place. If you are login to SMTP server using your username and password that is also pass in plain text so again anyone stole your information during transfer. Messages sent through SMTP also contains information about sending computer and software used which when capture can be used for malicious intent. So SMTP lacks privacy concern. SMTP does not have any mechanism to authenticate source [3]. It also does not have functionality to check message integrity and so it is easy send phishing attacks. SMTP does not have any mechanism to control repudiation that would make sender to deny sending of emails. The

messages stored on SMTP servers as plain text and also their backups are also taken. Even if you delete the message they can be residing on the severs/backup-servers for years. So anyone who accesses these servers can also access or read your message easily.

POP and IMAP: POP and IMAP are pull protocols, Request is send to mail server to access the mailbox and for that login using username and password is required. These details are not encrypted before sending unless SSL is used. So our confidential information at stake.

To overcome all these limitations, the SMTP servers incorporate many security functions by using one or more add on email security protocols. These protocols provide several security features like symmetric or asymmetric encryption, digital signatures and IP address verification. These add-on protocols like PGP, S/MIME etc. provide security to a very good extent but still our systems are not fully secure and there may be chances of breaching the security. There are many add-on protocols but the following section discusses which are most commonly used. It discusses what benefits these add-on protocols provide and what their limitations are.

SSL/TLS [1] are almost same having slight differences; actually TSL is the latest version of SSL technology. They are meant to provide compression and security to the data generated by application layer. SSL/TLS are the protocols applied at transport layer for security. Generally SSL can work on data from any protocol of application layer, but most commonly the protocol used is HTTP. Data generated by application layer is digitally signed, compressed and also encrypted and after that data is transfer to transport layer reliable protocol like TCP. TLS/SSL plays a crucial role in the security and privacy needed for web commerce applications. For e.g. before your computer uploads your credit card information to some commerce website, your computer must make certain that the remote host is authentic. That's where a protocol like SSL/TLS comes in. This protocol is also widely used to protect email servers (running under SMTP, POP, and IMAP protocols).

PGP stands for Pretty Good Privacy and it provides security at application layer. It is designed to provide confidentiality and authentication to emails. The main features of PGP are that it is freely available and not controlled by any organization. It can run on different platforms. Sending of email is one-time activity so sender and receiver can send emails to each other independently and at different times. Session concept and handshaking concept is not there so PGP pass security parameters with message. Phases follow in PGP in this order Digital Signature, Compression, Encryption, Digital Enveloping and Base-64 encoding. PGP also has limitations: PGP mainly deals with private key, if it is lost all the data can lost, The main threat is imitation and tampering of public key.

S/MIME full form is Secure/Multipurpose internet mail extension. It also provides security at application layer level. It is most widely used security protocol. S/MIME is asymmetric cryptographic technique which is used to authenticate sender and provide strong signature semantics. [2][3] It also provides non-repudiation, message integrity and message security using encryption. S/MIME also has some limitations; a recipient can forward the email message with digital signature of sender without taking consent from sender, so it's a threat for privacy of sender. It is also unable to achieve non-repudiation through keys in situations like when keys are lost. Digital signatures work well for message integrity and sender's authentication but not full proof for spammers and phishers who can able to modify e-mail addresses to make believe recipients about source. S/MIME in addition to basic security services also provide different optional services which may differ according to each implementation of S/MIME, therefore it may not be consider interoperable and provide complete assurance to users. S/MIME and PGP does not generally sign the message headers which makes messages modifiable during transmission.

So from above discussion, it is clear that various threats and vulnerabilities are existing in E-mail System. Several security protocols are also there to protect user rights, their privacy and their information. Most are widely used and also are successful to a good extent but still no system is full proof. There are always chances exist to breach the security either due to user error, due to limitation in security protocol or due to lack of implementation of protocols or due to lack of compatibility of new protocols/software with old ones. Reason may be anyone but chances are always there for some breach, so there comes the Email Forensics to collect evidence against culprit in a manner so that it may be proved in court of law. E-mail Forensics more specifically is a branch of Network Forensics, which in turn part of broader scenario Digital Forensics. So our next section discuss about E-mil forensics, what approach it used, techniques and tools used to find the culprit or source of crime in case of e-mail system related crime or security breach.

Part III: E-mail Forensics and Investigation Techniques

Header Investigation [5][6][26]: It is the most common and popular technique to analyze the useful hidden personal information in E-mail. E-mail Header is important for investigation and collection of evidence.

Meta data present is present in the email Header as control information. They contain information about sender/receiver and the path followed by message to reach destination. So this is very crucial information from evidence point of view. Sometime this Meta data/control information is altered or spoofed. So in header analysis, authentication of the information present in the header is also checked. E-mail Header is like shown in figure 1.7.

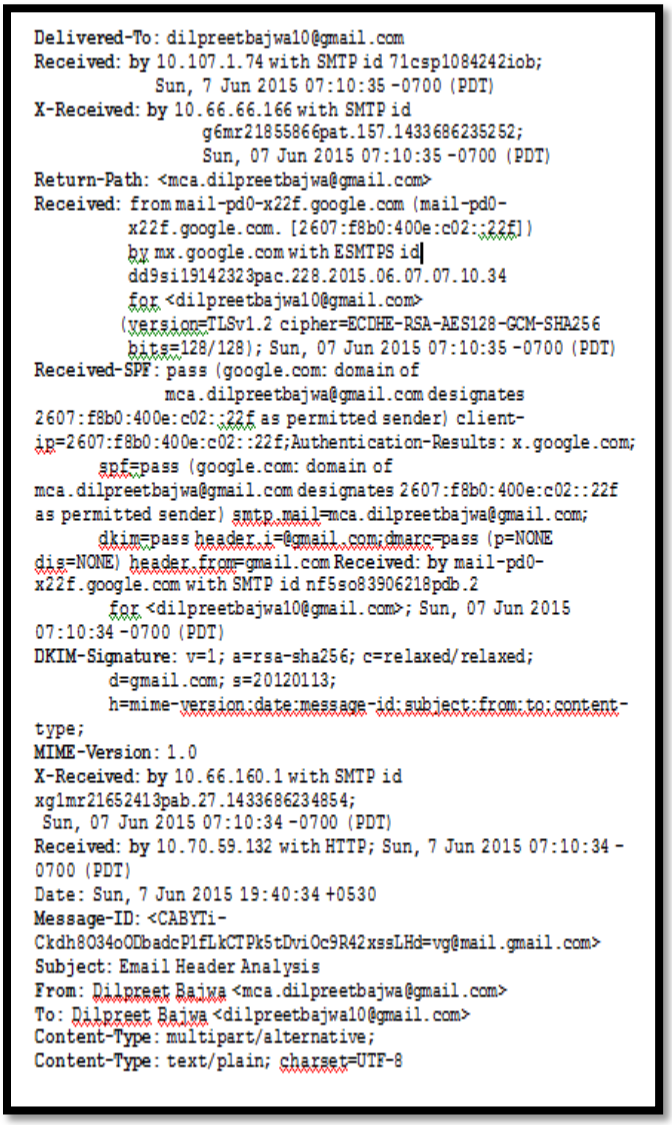
So there are many fields and provide lots of information regarding email, sender's IP address, return path, message-id, signatures field, server transit or path follows for transmission MIME and other security protocol information. The *Received* field in shows date and time when email is arrived at server. The *From* and *To* tells about sender and receiver. MIME version shows 1.0. So lots of useful information is there for analysis. Email Header analysis is used to collect crucial evidence to prove crime in court of law.

Some author proposed new techniques for header analysis as [5] proposed augmentations to the current trace header for implementation of digital forensic readiness. Inclusion of digital forensics readiness, also provide a level of integrity to the SMTP trace header which can be used for tracing spam's origin. Similarly digital forensic readiness is also included in envelope.

In order to determine the authenticity of email evidence, [6] says that methods to forge email are many, so it is mandatory to summarize the most common forms of forgery and analyze its implementation for identification.

Server Investigation: As it is mentioned earlier that server stores copy of our emails even if they are deleted from our mailbox. So investigation of mail servers on request can be done or through proper legal procedure can take back-up of emails and analyze the contents or other information related to email. Further logs are maintained by servers and they can be helpful for tracing the computer/server from where transaction takes place. When we are creating an email account than we are passing some information directly like name, phone number etc which are helpful in finding the person corresponding to some email address. If the information provided is fake then still server can store information like your location/area etc. that can be helpful in investigation.

Network and Network Device investigation: Some time it is not possible to take backup and log information from servers due to their non-availability, legal and other reasons in that case logs can be analyzed maintained by network devices like routers, switches and firewalls etc to find the source or authentic information. This type of analyses is complex.



```
Delivered-To: dilpreetbajwa10@gmail.com
Received: by 10.107.1.74 with SMTP id 71csp1084242iob;
          Sun, 7 Jun 2015 07:10:35 -0700 (PDT)
X-Received: by 10.66.66.166 with SMTP id
          g6mr21855866pat.157.1433686235252;
          Sun, 07 Jun 2015 07:10:35 -0700 (PDT)
Return-Path: <mca.dilpreetbajwa@gmail.com>
Received: from mail-pd0-x22f.google.com (mail-pd0-
        x22f.google.com. [2607:f8b0:400e:c02::22f])
        by mx.google.com with ESMTPS id
        dd9si19142323pac.228.2015.06.07.07.10.34
        for <dilpreetbajwa10@gmail.com>
       (version=TLSv1.2 cipher=ECDHE-RSA-AES128-GCM-SHA256
        bits=128/128); Sun, 07 Jun 2015 07:10:35 -0700 (PDT)
Received-SPF: pass (google.com: domain of
         mca.dilpreetbajwa@gmail.com designates
2607:f8b0:400e:c02::22f as permitted sender) client-
ip=2607:f8b0:400e:c02::22f;Authentication-Results: x.google.com;
      spf=pass (google.com: domain of
mca.dilpreetbajwa@gmail.com designates 2607:f8b0:400e:c02::22f
as permitted sender) smtp.mail=mca.dilpreetbajwa@gmail.com;
      dkim=pass header.i=@gmail.com;dmarc=pass (p=NONE
dis=NONE) header.from=gmail.com Received: by mail-pd0-
x22f.google.com with SMTP id nf5so83906218pdb.2
       for <dilpreetbajwa10@gmail.com>; Sun, 07 Jun 2015
07:10:34 -0700 (PDT)
DKIM-Signature: v=1; a=rsa-sha256; c=relaxed/relaxed;
       d=gmail.com; s=20120113;
       h=mime-version:date:message-id:subject:from:to:content-
type;
MIME-Version: 1.0
X-Received: by 10.66.160.1 with SMTP id
xg1mr21652413pab.27.1433686234854;
 Sun, 07 Jun 2015 07:10:34 -0700 (PDT)
Received: by 10.70.59.132 with HTTP; Sun, 7 Jun 2015 07:10:34 -
0700 (PDT)
Date: Sun, 7 Jun 2015 19:40:34 +0530
Message-ID: <CABYTi-
Ckdh8O34oODbadcP1fLkCTPk5tDviOc9R42xssLHd=vg@mail.gmail.com>
Subject: Email Header Analysis
From: Dilpreet Bajwa <mca.dilpreetbajwa@gmail.com>
To: Dilpreet Bajwa <dilpreetbajwa10@gmail.com>
Content-Type: multipart/alternative;
Content-Type: text/plain; charset=UTF-8
```

Figure: 1.7

Analysis of network traffic and packets for forensic purpose can also take place. [7] Proposed email forensics based network that analyzes different type of email data provided by network packets such as receiver, sender details and other information. In case of any doubt corresponding to emails coming from particular sender than forensic system can be implemented at place where it is must for illegal email packet to pass through it and all information which included specified sender can be collected.

[8] Study for organized crime investigation for email forensics. E-mail forensics is based on social network analysis and the main idea is to find relationship between communications corresponding to criminal organization.

Investigation of Software Embedded details: Software used to compose emails or to process emails at server side are also embedded some information like software version and software used or other details which are quite useful for forensic purpose. As forensic examiner can formulate theory or select tools acc. to particular software used and their version. Software used for composing email can include custom headers in the form of MIME, investigating these details may disclose some important information about sender's identity and his/her preferences [4]. Software information which is used to handle email can be collected from the received Header Field and information regarding software used at client side can be confirmed by using different set of headers like X-Mailer or equivalent [4].

Investigation and Discovery of Hidden Emails: A mail consider as hidden e-mail when it is an original email that has been quoted in at least one email in a folder, but cannot shows itself in the same folder; reason being it is deleted intentionally or unintentionally [9]. The reconstruction and searching for hidden emails is important for many purposes particularly for forensics.

Investigation of Anti Forensic Activity: Some time culprits also use anti forensic techniques which are used to counter cyber forensic investigation. Once a investigation technique is developed for cyber or email forensics, to counter it new defensive technique is developed. Hence in addition to forensics techniques, during investigation it is also necessary to take care of these anti-forensic activities [27].

Email Forensics Tools:

There are several E-mail forensics tools available for assisting in email forensics. Some work specifically for email forensics, some are part of broader general purpose framework of digital forensics; some are paid, free or open source.

Firstly E-mail Forensic Tools are discussed and their comparative study is also available by [10]. Later on other popular forensic tools are discussed. The order of tools is random and not prioritizes acc. to order.

MailXaminer[11]: MailXaminer is investigation tool for email forensics created by SysTools. The tool is used to examine email data files of MAPI / Non-MAPI desktop mail applications & Cloud mail services. It also analyzes emails for digital evidences, contacts, attachments, calendar entries, etc. which can be stored in the repository of different email services. It provides Advanced Keyword Search for better evidence collection even within the emails, Analyze & also recover deleted emails and also presented in form that can easily admissible in court, It also download mailboxes from Live Exchange server, Office365 and Google Apps admin without using IMAP, Provide Link analysis to establish link between different users, Provides filtering and analysis technique to easily distinguish the emails having objectionable or pornographic material/images, It can also export reports in HTML, PDF & CSV formats and also export carved evidences in to multiple formats, Creation, Case repository can also maintained and analyze them individually, also provide team collaboration feature so that multiple teams can work on same case at same time.

Aid4Mail [12]: Aid4Mail is an email forensics software solution. Aid4Mail supports approx 40 email formats and mail client programs, as well as remote accounts and webmail services through IMAP. It also processes easily local mail folders and files even after they get disconnected from their email client. Aid4Mail can also read the mbox files without prior conversion from Linux and Mac Systems. Aid4Mail easily tackle very heavy attachments and files, and huge mailbox files. It also extracts mails from corrupt mailboxes and also helps in fixing corrupt emails. It can also recover deleted emails from different type of mail boxes. It also exports reports and messages/data in several popular formats. It can convert messages to PDF or HTML. It also provides command line option.

Digital Forensic Framework (DFF) [13]: It is available free because it is a Open Source digital forensics software. It is built with the help of dedicated Application Programming Interface (API). It is used by many people including professional and non-experts. It facilitates easy, quick collection, preservation and analysis of digital evidences without compromise data and system. It provides access of remote and local devices, preservation of digital chain of custody, support several file formats like Raw, EWF, AFF 3 etc., also reconstruct virtual machine, works for both windows and linux operating system and file system, also recover deleted files, DFF can used for analysis of volatile memory, easy and efficient search criteria for regular and meta data. It can also export in to various file formats such as HTML, PST, PDF, MSG, EML and TIFF.

eMailTrackerPro [14]: EmailTrackerPro has ability to trace an email using the email header, an email headerworks as an footprint for finding from where actually email has travelled. Sometime criminals can also modified header, in that case emailTrackerpro recognize the inconsistencies and mark them as spam. It can also comes with a spam filter, which warns the use for each suspected email after scanning the email. Another important feature provided by eMailTrackerPro is Abuse Reporting. It automatically generates a email and report

abuse regarding spam emails and send to the email service provider.

Paraben EMX [15]: It is used to examine forensically various email formats which includes Windows mail, Outlook (PST and OST), Outlook Express, Thunderbird and more. It provides functionality to analyze attachments, message headers and bodies. It recovered emails even from deleted items and also export them to other formats. EMX supports reporting and exporting it in several formats. It is easy to use and analyze email thoroughly including header, provide comprehensive attachment analysis and sorting. It also supports all main types of email which can be stored on local computers for reporting, exporting/conversion and for analysis. Supports Microsoft Outlook Offline Storage (OST), America On-line (AOL), Microsoft Outlook (PST), The Bat! (version 3.x and higher), outlook Express, Thunderbird, Email file - RFC 833 Compliant(EML), Eudora, Windows mail databases, Plain Text mail, Maildir. Also Supports more than 750 MIME Types.

[10] also done comparative study on above email forensic tools which is based on nine different criteria and confirm that Aid4Mail can analyze emails stored at local computers as well as remote email servers. Aid4Mail has the highest capability to search and gather information. Among all the tools above Paraben EMX is the only one which also shows analysis details of attachments of emails other than header and body. The recovering capability of Paraben EMX and Add4Mail is consider somewhat better than other tools as they can also recover emails/files even from deleted folders. While considering email format support, paraben support more email formats and almost 750 MIME types. DFF support many output file formats. GUI support of MailXaminer is better. Almost all tools supports Windows platform and very few support Linux. DFF and Aid4Mail also supports extended device analysis such as pendrive.

Now let us have a look at other popularly used E-mail Forensic Tools:

MX Toolbox: This software contain many tools like MX lookup, Blacklist, Diagnostics, Domain Health, Header Analyzer, DNS lookup and many more, each have some specific purpose. The MX lookup is done directly against the domain's authoritative name server. You can use Diagnostics , which will connect to the mail server, verify reverse DNS records, perform a simple Open Relay check and calculate response time performance. It can also check each IP Address against 105 DNS based blacklists. The DKIM Record tool will test a domain name and selector for a valid published DKIM key record. Port Scan will tell you what ports are open and standard services are running on your server, we can type in an IP address or hostname. MailFlow is used to check your email system performance and even small delays that will be harmful for your business. E-mail deliverability tool requires you to send a test message to MxToolbox, it analyze the headers, the blacklist reputation of your outbound IP address, and your SPF records to generate a detailed deliverability report.

EmailTracer [16]: It is a product of Resource center for Cyber Forensics (RCCF) established by govt. of India. It tracks identity of email sender and can able to track any email you receive. It also analyze email header and also provides sender's IP Address and other details like his/her geographical area location. It also has facilities for doing NS LookUp, WhoIs Search and IPTraceBack. It can also be used for recover/access details from mailbox files with extensions .pst (Microsoft Outlook), .cnm(Pegasus), .dbx (Outlook Express), MailDir (Kmail), .mbox (Mozilla), .nsm (Netscape Messenger), .tbb (The Bat), .mbx (Eudora) and .imm(IncrediMail).

Conclusion: E-mail system is widely used and complex distributed internet application having several hardware or software components including services, protocols, server, and agents. Several threats are faced by user due to vulnerabilities present in the system. So there is need to make it more secure by overcoming the current security flaws. Further there is also need to adopt proactive forensics and to make our system to adopt forensic readiness.

References:

[1] Behrouz A Forouzan, Data Communications and Networking. New Delhi: Tata McGraw-Hill, 2011.

[2] N.Vijayalakshmi, E.Sivajothi, Dr.P.Vivekanandan, "Efficiency and Limitation of Secure Protocol in Email Services", International Journal of Engineering Sciences and Research Technology, pp-539-544, Nov-2012.

[3] M. Tariq Banday, "Effectiveness and Limitations of E-Mail Security Protocols", International Journal of Distributed and Parallel Systems, Vol-2, No.-3, May-2011.

[4] M. Tariq Banday, "Technique and Tools for Forensic Investigation of E-Mail", International Journal of Network Security and its Applications, Vol-3, No.-6, November-2011.

[5] F.R. Van Staden and H.S Venter,"Adding Digital Forensic Readiness to the Email Trace Header", Conference on Information Security of South Africa, IEEE, Aug-2010.

[6] Hong Guo, Bo Jin, Wei Qian, "Analysis of Email Header for Forensic purpose", International Conference on Communication Systems and Network Technologies", Computer Society, IEEE, 2013

[7] Wang WenQi, Liu WeiGuang, "The Research on Email Forensics Based Network", Ist International Conference on Information Science and Enginnering", Computer Society, IEEE, 2009.

[8] YanHua Liu, GuoLong Chen, Lili Xie, "An Email Forensics Analysis Method Based on Social Network Analysis", International Conference on Cloud Computing and Big Data", Computer Society, IEEE, 2014.

[9] Guiseppe Carenini, Raymond T. Ng, Xiaodong Zhou, " Scalable Discovery of Hidden Emails from Large Folders", KDD '05 Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining, pp- 544-549, ACM, 2005.

[10] Devendran, V.K., Shahriar, H. and Clincy, V., "A Comparative Study of Email Forensic Tools", *Journal of Information Security*, **6**, pp- 111-117, 2015.

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

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

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

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

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

[16] EmailTracer. http://www.cyberforensics.in/Products/emailtracer.aspx

[17] Ankur Dumka, Ravi Tomar, J.C.Patni, Abhineet Anand," Taxonomy of E-Mail Security Protocol". International Journal of Innovative Research in Computer and Communication Engineering, Vol-2, no.-4, April-2014.

[18] Kunal Pandove, Amandeep Jindal, Rajinder Kumar. "E-Mail Spoofing", International Journal of Computer Applications, Vol-5, No.-1, pp- 27-30, August 2010.

[19] P. Ramesh Babu, D. Lalitha Bhaskari, CH. Satyanarayana, "A Comprehensive Analysis of Spoofing", International Journal of Advanced Computer Science and Applications, Vol-1, No.-6, December 2010.

[20] Jitender Nath Srivastva, Maringati Hima Bindu, "E-Mail Spam Filtering using Adaptive Genetic Algorithm", I. J Intelligent System and Applications, pp-54-60, January 2014.

[21] Kim-Kwang Raymond Choo, "The Cyber Threat Landscape: Challenges and future Directions", Computer and Security, Science Direct, Elsevier, pp-719-731, 2011.

[22] Gori Mohamed .J, M. Mohammed Mohideen, Mrs.Shahira Banu. N, "E-Mail Phishing-An Open Threat to Everyone", International Journal of Scientific and Research Publications, Vol-4, No.-2, Feb-2014.

[23] Justin Paglierani, Mike Mabey, Gail-Joon Ahn, "Towards Comprehensive and Collaborative Forensics on Email Evidence", 9th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing, 2013.

[24] Radicati Group, Inc., "Email Market, 2012-2016," http://www.radicati.com/wp/wp-content/uploads/2012/04/Email-Statistics-Report-2012-2016-Executive-Summary.pdf, October 2012.

[25] Olalekan Adeyinka, "Internet Attack Methods and Internet Security Technology", Second Asia International Conference on Modeling & Simulation, May-2008.

[26] Satheesaan Pasupatheeswaran, "Email 'Message-IDs' helpful for forensic analysis", Proceedings of the 6th Australian Digital Forensics Conference, Edith Cowan University, Perth Western Australia, December 3rd 2008,

http://ro.ecu.edu.au/cgi/viewcontent.cgi?article=1048&context=adf.

[27] Anu Jain, Gurpal Singh Chhabra, "Anti Forensics Techniques: An analytical Review", Seventh International Conference on Contemporary Computing (IC3), IEEE, Aug – 2014.