Computer Network Security

Lesson 10

Forensics

Lecturer: Dr. Msagha J Mbogholi, PhD

Flashback from Lesson 9

- SSL provides 2 key security functions: confidentiality and authenticity; it provides a 'tunnel' between two communicating devices on the Internet. TLS is and advanced version of SSL.
- Secure shell (SSH) protocol was designed to replace remote login protocols like telnet and rlogin which were deemed to be insecure.
- Wireless networks are protected using WPA or WPA2.
- In order to protect email there are four main standards that are used: Secure MIME (S/MIME), Pretty Good Privacy (PGP), Privacy Enhanced Mail (PEM) and Domain Keys Identified Mail (DKIM)
- Internet Protocol Security (IPSec) is more a framework for secure communication than a specific protocol or standard

Content

- Introduction to Digital Forensics
- E-mail and Web Forensics
- Data Forensics
- Mobile Forensics
- Network Forensics

Part 1

Introduction to Digital Forensics

Introduction

- Have you watched any of the episodes of the series CSI? The original series had such a global fanatic following that the producers had to split into different states…CSI: New York, CSI: Miami, CSI: Cyber.
- The thing that made the series so interesting was the forensics part…it was just awesome; the collecting, analyzing and concluding (inferencing) of evidence and data from crime scenes would get any inquisitive mind hooked. In real life though can you imagine how much time it would take to collect and analyze all that stuff? Uh huh.
- This gives us an idea of what forensics is all about.
- Of course you can already see that it is a wide area of study, since I haven't even mentioned computers yet….which is what this course is all about right?

Definitions

- The Google dictionary defines forensics as "scientific tests or techniques used in connection with the detection of crime."
- The Miriam Webster dictionary defines it as "relating to or dealing with the application of scientific knowledge to legal problems".
- The American Academy of Forensic sciences recognizes 11 forensic disciplines: anthropology, criminalists, digital and multimedia sciences, engineering and applied sciences, general, jurisprudence, odontology, pathology/biology, psychiatry and behavioral science, questioned documents and toxicology.
- Our interest in this course falls under digital forensic science. The term computer forensic science is sometimes used interchangeably with digital forensics; in a strict sense they are two different terms.
- While computer forensics is limited to the investigation of computers, digital forensics will include computers, media, networks, mobile devices, and so on. Now this is what we are interested in!

Definitions

- In this lesson we are interested in understanding email and web forensics, data forensics, document forensics, mobile forensics and network forensics.
- The field of digital forensics is very wide and it is not possible to go into extreme details in one lesson…it needs a full course on its own!
- However, it is expected that the learner will be able to have a fair grasp of the issues involved in digital forensics. Further reading is encouraged.
- Let us conclude this part by defining digital forensics.
- NIST defines digital forensics as " the application of computer science and investigative procedures involving the examination of digital evidence - following proper search authority, chain of custody, validation with mathematics, use of validated tools, repeatability, reporting, and possibly expert testimony." (Source: https://csrc.nist.gov/glossary/term/digital_forensics)

Part 2

E-mail and Web Forensics

Internet Architecture

- From networking knowledge we are aware that a packet moves from its source (sender) to its destination via several routers over the Internet.
- We are also aware that the internet is built up on a client-server architecture such that most direct exchanges are between a client and a server.
- This is the case with email as well. When a client wishes to send an email to another client the email will be sent to the email server and then retrieved by the recipient via the same email server.
- Each message from a client machine is broken down into packets and then sent along its way across the Internet. Each packet will contain a source address (where it came from) and a destination address (where it's going) among other pieces of information including its payload (the data itself).
- Bearing this in mind how does Email architecture work on the Internet?

E-mail Architecture

- All emails are sent to or received from a certain domain on the internet, such as Gmail, ymail, yahoo, and company specific domains (for example I am in the pu.ac.ke domain). An email from a particular domain will have two parts: the_username @ the_domain name. An example would be jmsagha@gmail.com, or Mbogholi.m@pu.ac.ke for the Gmail and pu domains respectively.
- Email has literally become such a part of our lives that we rarely give thought to the process behind sending and receiving an email.
- Fig 1 captures the key terminology and the process behind sending and receiving an email over the Internet.

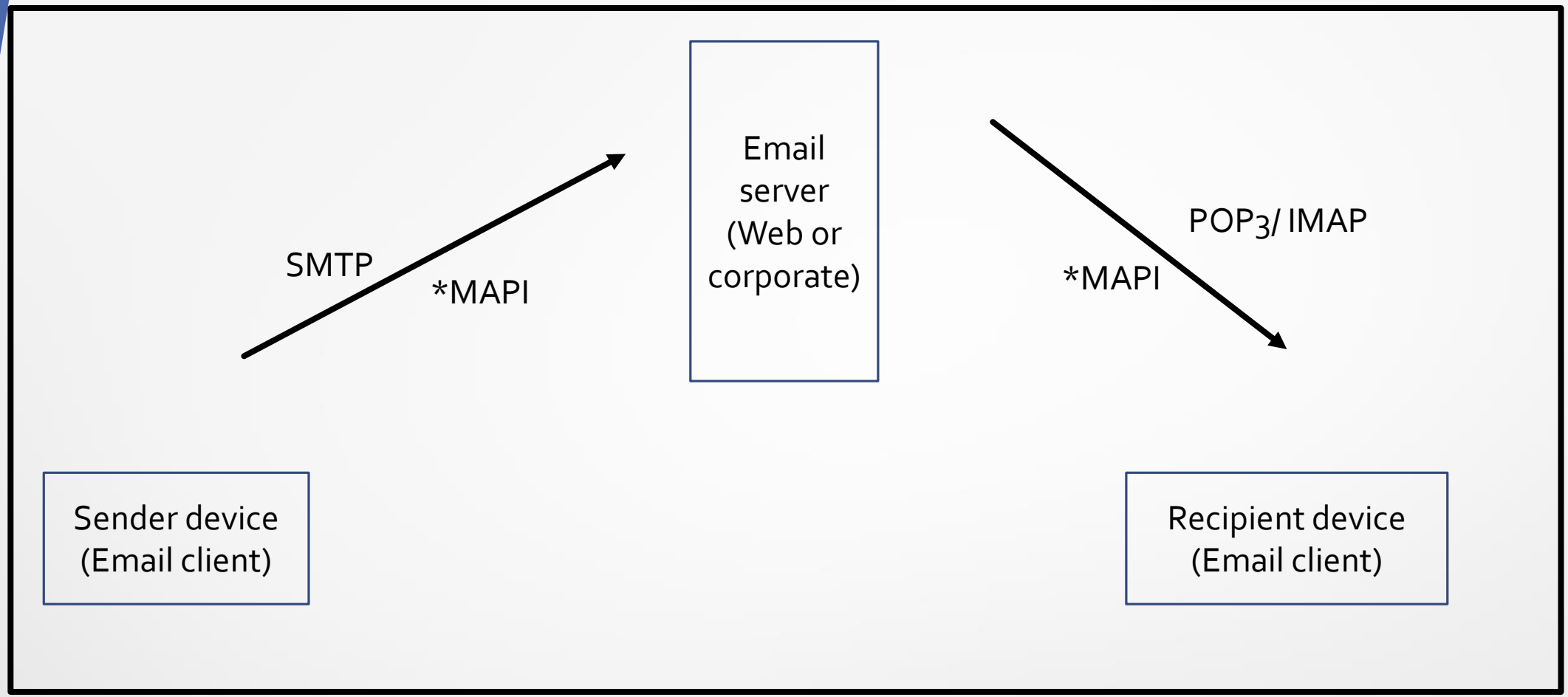

Fig. 1 Key e-mail terminologies

Email Architecture (cont'd)

- Fig 1 shows the path an email takes from a sender to the recipient. What is of most importance is the path and the protocols involved.
- The sender types out the email using a client software such as outlook or webmail. The mail is sent out using the simple mail transport protocol (SMTP) or message application programming interface (MAPI) (in the case of Microsoft outlook).
- The mail is received at the email server which maybe a web server or a dedicated corporate server.
- The mail is retrieved by the recipient via either post office protocol 3 (POP3), internet message access protocol (IMAP) or MAPI (in the case of Microsoft outlook)
- These protocols all work differently in terms of how emails are stored and retrieved.

Email Architecture (cont'd)

- With SMTP the message is sent to the server. If the server needs to forward the message to another server then it will use SMTP to forward the email; the term used is that SMTP "pushes" the email from email client to the recipient's email server. Without SMTP you can't send email.
- The recipient's email server will receive the email; this is where differences come in the retrieval process. With IMAP the recipient sees the message headers (think of how your Gmail or yahoo email interface looks like) and then chooses which messages to download. POP, however, allows the client to "pull" all the messages to their email client; thereafter it deletes the emails from the server (however, during setup one can choose to leave a copy of the email on the server). MAPI is a Microsoft proprietary protocol; it works with the settings that are provided with Outlook and by extension the Exchange server.
- Where these emails are saved is crucial to us in forensics.

Email and Forensics

- When performing an investigation an email can prove to be a very useful source of information. Let us examine how this is so.
- By examining an email we can tell where the email came from, when it was sent, and to who it was delivered. An email message will contain a header and a body; the former has the source and destination of the email, making it easy to trace the origin of the email. The latter will contain the actual information that was being sent by the sender.
- By expanding an email a lot more information can be obtained. To give you an idea of how much information a raw email contains please perform the following steps (assuming you have a Gmail account, which most people do ☺):

Email and Forensics

- Open a random email in your web browser. Next locate the 3 dots at the extreme right corner of the email and click on them. Choose the option "show original". This will open the original message in a new tab (assuming you're using google chrome, but it should be the same with any browser). You will see the header showing the message ID, where it came from, who created it, who it was sent to, the subject, security information such as SPF (sender policy framework – used as an authentication framework) and DKIM (from lesson 9 remember?) and so on. Additionally you will also see the path that the message took to reach you (if you are able to follow all the jargon!)
- Further more information can be obtained by seeing who the email has been copied to, whether it has attachments (and what is contained in them), and also whether it was forwarded from somewhere else.
- Additionally and more importantly an email trail can be established.
- All this information is crucial to a forensics expert; however, the message can still be tampered with if for example, the user changed their clock or if it has been forwarded from another source.

Client – based Email

- As described earlier emails may be sent from a client based email such as Microsoft Outlook, or from a web based email such as Gmail (the emails are accessed directly via a web browser).
- As a forensics expert your interest is in establishing the path the email took to get to this client. Normally using a client based email is a bit of a less challenging process as most of the time the emails have been downloaded to it. From here other details from the message header can help to establish that the message did indeed come from a particular server (through the message ID).
- To emphasize this the server also needs to be examined; this information will be found in the logs. The logs will show the details of the email through the message ID. In so doing the path the email took from the source, the network(s) that were used, the server(s) it went through, and the delivery details can all be obtained.
- With this information you can now download and/or print the email (and attachments) together with all the evidence you have gathered.
- As an investigator this should give you cause to rub your hands in glee!

Web-based Email

- The biggest challenge with webmail is the fact that emails remain on the server (unless a user specifically asks to download) until deleted. This means that there is no copy on the local machine. Consequently it may mean finding a way to access the server. The lawful way of doing this is to obtain a court order compelling the provider to give this information.
- However, all is not lost. An examination of the local cache might give some desired results (normally in temp folder). One can also check unallocated space where the files were before being deleted. How? There are forensic tools that can help you achieve this.
- Another way is to search for files with .html extension to see if the email might be found locally. Using search tags can help to narrow down the scope.
- One can also search within the browser by examining the browser local cache and also the history and downloads parts.
- Two good forensic software tools that can assist with investigations are EnCase and FTK.

Part 3

Data Forensics

Introduction

- Data forensics is the branch that deals with finding data for the purposes of investigation.
- A lot of times the data that is needed for investigation can be found by simply searching for it from Windows explorer as long as you know the file name or have a tag(s) to work with. But is it really the case?
- More often than not the suspect will have hidden the files in the computer or even deleted them; does this mean that they can't be found? No!
- What is required is a proper understanding of the operating system in question in order to understand where and how files can be hidden, and a good forensic software tool to perform the actual extraction.
- Luckily there are just 3 major operating systems used by the majority of computer users globally and chances are more than 90% that you will be working with one of these.
- The operating systems being referred to are Microsoft Windows, Mac OS, and Linux (or a variant of it). Our discussion in this part does not cover mobile operating systems (these are covered in a different part of this lesson)

Hard Disk

- Most information is stored in some form of media. The media in use is mostly magnetic storage and hard disks use this technology.
- Hard disks come in different formats (internal and external) and size (from GBs to TBs); nonetheless the design and structure is all the same.
- A sector is the smallest unit of storage; several sectors make up a track; several tracks are found on a platter; several platters stacked one on top of the other make up a cylinder. The sectors and tracks are read using read/write heads (mostly with one pair per platter, either for reading the upper and lower surfaces.
- Data is written to the disk in the forms of 0s and 1s, based on the polarized magnetic material found on the disk.
- It is important to understand these terminologies in the terminologies of finding data on the disk.

File Deletion

- When a file is deleted regardless of the operating system it is not deleted from the hard disk.
- What is usually deleted is reference to it in the file directory of that operating system.
- Since it is no longer referenced in the directory the operating system can overwrite it or just leave that portion of space as unallocated space.
- Different file systems deal with these deleted files in different manners but they can still be found (unless an expert can physically delete them from the hard disk).

Windows

- Most users globally work with this operating system. For some it is because of ease of use, for others it is a default (for example at the workplace) and yet for others it is because of compatibility with most applications (most applications are designed for Windows first then other operating systems).
- Our interest is in the working of windows file system. Windows has evolved through 2 file systems namely FAT (the earlier versions) and NTFS (current versions). The file system recognizes the hard disk by organizing it into clusters (a group of sectors) and partitions (logical subdivisions). The MBR (master boot record) is the first partition in the hard disk and contains information (code) that tells BIOS where to find the operating system in order to load it to RAM (random access memory).
- When looking for deleted files is it first important to identify whether the OS (operating system) version running is based on FAT or NTFS as both handle them differently.

Windows (cont'd)

- When a file is deleted in the FAT system it is allocated a special hexadecimal character (E5) in the FAT table and this tells the OS that this cluster is available for use. A forensic tool will therefore look for this character in order to find where deleted files are located. Since the space is declared as unallocated all the deleted files will be found here.
- NTFS introduces the Master File Table (MFT) which works like the FAT table (well, almost) and stores more metadata than FAT. Deleted files are moved to the recycle bin but metadata about each of these files is stored in info2 record file; thus accessing this file will give access to deleted files. When the recycle bin is emptied NTFS handles the deleted files like FAT does (declares the clusters available and they become part of unallocated space).

Apple

- Uses the Hierarchical File System (HFS).
- This file system uses volumes to identify files and their locations.
- Normally information from this file system can be extracted using forensic tools mentioned earlier such as EnCase.
- Nonetheless Mac OS X uses the Apple File System (APFS) which features better security and improved file system features.

Unix / Linux

- Linux is a 'child' of Unix and therefore has inherited Unix's file system structure.
- Simply put everything is considered a file in Unix and by extension Linux. This file system also makes Linux installations less prone to virus attacks; since the virus only affects the user account where they are installed and not the super user account.
- Linux stores information about bad sectors in the Bad Block Inode (BBI); however, this file can only be accessed by the super user account.
- An expert may try and list a good sector as a bad one in the BBI in an effort to hide data, and therefore this file needs to be thoroughly analyzed when performing forensics on a Linux installation.

Where is the Data Hidden?

- The first step towards finding data is to know what you are looking for ☺
- Recall that when a file is deleted the OS informs the file system that the area is now available for use, i.e. it is now unallocated. However, the actual act of deleting the file from the hard disk is yet to happen. If the hard disk is sufficiently big that file will remain on it for a while to come! Thus by analyzing the unallocated space several deleted files can be found and retrieved using computer forensic software.
- Computer forensic software can be used to retrieve files that have been deleted as well as those that are in caches.
- If the data being looked for is to be found in RAM then it presents challenges, especially for Unix based installations (including Linux of course) since they can get rid of the entire contents of RAM. Nonetheless a program like WinHex or one with similar functionality can help.
- These same software can be used to analyze information in the registry, as well as extract and rebuild it.

Part 4

Mobile Forensics

Introduction

- There are an estimated 7.1 billion individual mobile phone users in the world as of 2020 representing close to 90% of the world population.(https://www.bankmycell.com/blog/how-many-phones-are-in-the-world, accessed 9th November 2021).
- This includes users with smart phones as well as feature phones (those without internet, we call them 'kabambe' in my country ☺)
- The study of mobile forensics purposes to capture forensics for both types of phones and clearly the challenge is evident.

Introduction (cont'd)

- This is because of the different features and operating systems that these phones come loaded in them.
- Nonetheless android is undoubtedly the most used operating system globally boasting an average share of 70% of the market.(https://gs.statcounter.com/os-market-share/mobile/worldwide, accessed 9th November 2021)
- There are many crimes that can be analyzed and captured via the use of mobile devices and the evidence is found right there in the phone.
- Mobile forensics aims to capture this information for analysis and evidence of crimes.

Communication Means

- Mobile devices communicate using different communication means. Commonly the following are used:
- Bluetooth
- Wireless (802.11x)
- Popular file sharing applications (Nearby Share, SHAREit, Xender, and so on)

Evidence Locations

- There are several sources of information on the phone that can be used to find evidence:
- Subscriber Identification Module (SIM) – will verify the identity of the owner and tie them to a particular service provider
- Call logs – these can be obtained from the phone and the service provider. They show who called the phone, the location of the phone at any given time, any missed called calls and also calls made.
- Phone book – a list of all saved contacts together with their data such as email and addresses
- Texts (sms) – these will provide evidence of details of communications between parties (even deleted texts can be retrieved just like in the hard disk case)
- Calendar – will give clues to appointments and recorded events which can be used as evidence.

Evidence Locations (cont'd)

- Photos – by examining the gallery photos can be found which can assist in investigations.
- Audio and videos – these can also be found in the gallery
- Applications – more information can be found by analyzing the different applications on the phone, for example word processing software.
- SD card – for phones that have and SD card slot an analysis of the contents of the card might also produce the evidence being sought.

Extracting the Data

- Once you have the phone in your possession the next step is to determine its details namely, make, model, IMEI number (the number that uniquely identifies the phone on the network), OS, RAM and ROM.
- The make will also assist in determining the kind of software to use in order to extract the desired information.
- Also note that the phone maybe password protected and therefore it may be necessary to have unlocking software at hand (be careful that the software does not wipe out (flash) the contents of the phone!)

Mobile Forensic Tools

- Note that the tool you use should not tamper with any of the existing information; further it should not add anything on top of the existing data.
- Another key point is that it should be able to prove the above, that is, that the data has not been tampered with (integrity is intact).
- Some of the mobile forensic tools available in the market today include:
- Open source android forensics
    - FTK
    - Andriller
    - Encase
    - Cellebrite
    - Elcomsoft iOS Forensic Toolkit

Part 5

Network Forensics

Introduction

- For a proper understanding of network forensics it is imperative to have some basic knowledge of terminologies used with computer networks.
- Thankfully we have had mentioned and had a look at these in this course.
- So it is safe to say that you understand how a router and a switch work, right?
- How about the OSI model? Yes? Good.
- In a network environment agents are used to monitor the goings-on and also to report anything wrong happening in the network. They send these reports to a designated server.

Data Sources

- The sources of data on the network include:
- Agents – collect data for forensic analysis (events and so on) and send them to the forensic server (the one dedicated to collecting the info)
- Hosts – individual devices on the network (these can provide information as data devices like in data forensics)
- Routers – they contain a ton of useful information such as the logs showing the packet information, ARP information, and so on.
- Information can also be collected from the firewall, switches, Intrusion Detection System (IDS – more of this in lesson 11), Intrusion Prevention System (IPS – more of this in lesson 12) and network printers (and copiers).

Analysis

- Most good network forensic tools will reconstruct the event (s) that occurred over the network for you (hooray).
- They will even timestamp the event(s) so that a sequence can be easily demonstrated.
- Tools that can be used include:
    - Network Test Access Port (TAP)
    - Port mirroring – moving data from a port (s) under investigation and mirroring it to a forensic port.
    - Wireshark
    - NetDetector
    - Xplico
    - NetScarab

Summary

- Digital forensics includes the forensics of computers, media, networks, mobile devices, and so on.
- Email and web forensics is about using email details (obvious and hidden) to find information that can be used for investigative purposes.
- Data forensics purposes to find hidden data in media using forensic tools; this data can then be used as evidence in court.
- Mobile forensics is the use of mobile information from both user or/and provider in order to extract information that can be used in an investigation.
- Network forensics uses hosts and agents on the network to get information on the goings-on in the network. Network forensic tools can then string together events to give them meaning.
- A good software forensic tool should preserve the integrity of the data it is investigating.

References

- EC - Council (2010). *Computer forensics* (Vol. 4). Course Technology Cengage Learning.
- Volonino, L., & Anzaldua, R. (2008). *Computer Forensics for dummies*. Wiley.