



INTERNATIONAL
HELLENIC
UNIVERSITY

Computer and Network Forensics: investigating network traffic

Amarantidou Pinelopi

SID: 3305150002

SCHOOL OF SCIENCE & TECHNOLOGY

A thesis submitted for the degree of

Master of Science (MSc) in Communications and Cybersecurity

NOVEMBER 2017

THESSALONIKI – GREECE



INTERNATIONAL
HELLENIC
UNIVERSITY

Computer and Network Forensics: investigating network traffic

Amarantidou Pinelopi

SID: 3305150002

Supervisor:

Prof. Athanasios Papathanasiou

SCHOOL OF SCIENCE & TECHNOLOGY

A thesis submitted for the degree of

Master of Science (MSc) in Communications and Cybersecurity

NOVEMBER 2017

THESSALONIKI – GREECE

Abstract

The phenomenon of cybercrimes or crimes committed on the Internet has constituted the need of developing the science of *Digital Forensics*. Computers Forensics and Networks Forensics constitute two branches of Digital Forensics, which play an important role in case of ~~such a crime occurs~~, as important information can be extracted from them, used later as digital evidence.

Investigating network traffic, meaning investigating packets that travel across the Internet is a valuable aspect in case of analyzing applications, such as VoIP applications, as special information can be found. The scope of ~~that~~ dissertation is to ~~watch~~ what and if any content can be revealed while someone investigates network traffic, meaning ~~doing~~ packet sniffing, with a special monitoring and analyzing tool, Wireshark.

Amarantidou Pinelopi

18/12/2017

Contents

ABSTRACT	III
CONTENTS	V
1 INTRODUCTION.....	1
2 DIGITAL FORENSICS	3
2.1 INTRODUCTION OF DIGITAL FORENSICS SCIENCE	3
2.1.1 <i>Digital Forensics Definition</i>	3
2.1.2 <i>Applications and Uses of Digital Forensics</i>	4
2.1.3 <i>Digital Forensics Necessity and History</i>	5
2.2 DIGITAL EVIDENCE	6
2.2.1 <i>Digital Evidence Definition</i>	6
2.2.2 <i>Digital Evidence Sources</i>	8
2.2.3 <i>Data Types and Types of Digital Evidence</i>	8
2.2.4 <i>Nature and Forms of Digital Evidence</i>	10
2.2.5 <i>The Five Rules of Evidence</i>	11
2.3 BRANCHES OF DIGITAL FORENSICS	12
2.4 DIGITAL FORENSICS TOOLS.....	15
2.4.1 <i>Types of Digital Forensics Tools and Equipment</i>	16
2.4.2 <i>Tools Selection</i>	17
2.4.3 <i>Basic Tools</i>	18
2.5 DIGITAL FORENSICS PROCESS MODELS.....	19
2.5.1 <i>Introduction to Digital Forensics Process Models</i>	19
2.5.2 <i>Existing Models and their Categories</i>	21
2.5.3 <i>Presentation of the Four Major Digital Forensics Process Models</i>	
23	
2.6 CHALLENGES IN DIGITAL FORENSICS.....	36
2.6.1 <i>Technical Challenges in Digital Forensics</i>	36
2.6.2 <i>Legal Challenges in Digital Forensics</i>	39
2.6.3 <i>Resource Challenges in Digital Forensics</i>	40

3	COMPUTER FORENSICS.....	43
3.1	INTRODUCTION OF COMPUTER FORENSICS SCIENCE	43
3.1.1	<i>Computer Forensics Definition.....</i>	<i>43</i>
3.1.2	<i>Objectives of Computer Forensics.....</i>	<i>44</i>
3.1.3	<i>Primary Types of Computer Forensics Investigations.....</i>	<i>44</i>
3.1.4	<i>Applications and Uses of Computer Forensics.....</i>	<i>45</i>
3.2	THE GENERIC COMPUTER FORENSICS PROCESS AND INVESTIGATION MODEL 48	
3.2.1	<i>The Generic Computer Forensics Process.....</i>	<i>48</i>
3.2.2	<i>The Generic Computer Forensic Investigation Model.....</i>	<i>50</i>
3.3	MAIN INVESTIGATIVE PROCEDURES	52
3.4	KEY ELEMENTS, RULES AND PRINCIPLES OF COMPUTER FORENSICS.....	57
3.4.1	<i>Key Elements of Computer Forensics.....</i>	<i>58</i>
3.4.2	<i>Computer Forensics Rules.....</i>	<i>60</i>
3.4.3	<i>Principles of Computer-Based Electronic Evidence</i>	<i>62</i>
4	NETWORK FORENSICS.....	65
4.1	INTRODUCTION TO NETWORK FORENSICS SCIENCE	65
4.1.1	<i>Network Forensics Definition.....</i>	<i>68</i>
4.1.2	<i>Classification of Network Forensics Science.....</i>	<i>69</i>
4.1.3	<i>Challenges in Gathering Network Forensics Evidence.....</i>	<i>70</i>
4.1.4	<i>Recent Trends in Network Forensics.....</i>	<i>73</i>
4.2	THE GENERIC PROCESS MODEL FOR NETWORK FORENSICS	75
4.3	TECHNICAL FUNDAMENTALS ON NETWORKS	78
4.3.1	<i>The seven-layer Model (OSI).....</i>	<i>79</i>
4.3.2	<i>The Transport Control Protocol/ Internet Protocol (TCP/IP) Model</i>	<i>82</i>
4.4	LEARNING TO HANDLE THE EVIDENCE	87
4.4.1	<i>Identifying sources of network evidence.....</i>	<i>87</i>
4.5	DATA EVIDENCE ACQUISITION ON THE NETWORK.....	89
4.5.1	<i>Active- evidence acquisition.....</i>	<i>90</i>
4.5.2	<i>Passive-evidence acquisition.....</i>	<i>90</i>
4.6	PACKET SNIFFING	91
4.6.1	<i>Components of a packet sniffer</i>	<i>92</i>

4.6.2	<i>How packet analyzers work</i>	93
4.6.3	<i>Packet Sniffing Process</i>	93
4.6.4	<i>Network Sniffing and Packet Analyzing Tools</i>	96
5	INVESTIGATING NETWORK TRAFFIC	100
5.1	PROBLEM DEFINITION: INVESTIGATING NETWORK TRAFFIC IN VOIP APPLICATIONS	100
5.2	CONTRIBUTION.....	100
5.2.1	<i>Installing Wireshark on Windows 7</i>	101
5.2.2	<i>Wireshark user interface essentials</i>	102
5.2.3	<i>Establishing an Access Point</i>	104
5.2.4	<i>Capturing traffic from VoIP applications</i>	105
6	CONCLUSION	110
	BIBLIOGRAPHY	111

1 Introduction

As we all know, every technological improvement apart from its advantages in the specific field it is applied to, has also its drawbacks. The same exists in case of computers, which have totally changed and improved lives of people, but they “risk” a lot of their personal components at the same time. As a result, a computer shows a great possibility of “betraying” users, leaving trails about movements they (users) have done through cell phones, ATM transactions, web searches, e-mails, text messages, etc. Since computers became so “close” friends to people, ~~with the ability of almost everyone uses them,~~ a new type of crime, *computer crime* or *cybercrime* has emerged, which is defined as any criminal act involving computer as an instrument, target or by any other ~~mean~~ for criminal continuity.

In order ~~investigators~~ to deal with such crimes, they have developed the Digital Forensic science along with its branches, each one of which deals with a specific field. Two of these important branches are computer forensics, which deals with crimes that relate to computers in many ways (we are going to explain more at next chapters) and network forensics, which examines network traffic, usually ~~with~~ capturing packets and trying to analyze their contents for getting valuable information. Wireshark is one tool, which focuses on the capturing of packets and NetworkMiner is another one that focuses on ~~the~~ analyzing ~~of~~ them, both of which are going to be used in the scope of ~~that~~ dissertation.

So, let the investigation begin!

2 Digital Forensics

In this chapter we are diving into the science of digital forensics, capturing and presenting its core concepts, history, applications and necessity. Furthermore, process models and stages of digital forensics are displayed along with tools used and challenges that have to be faced.

2.1 Introduction of Digital Forensics Science

Digital forensics science, otherwise digital forensics, constitutes a branch of forensic science encompassing the uncovering, collection, examination and interpretation of electronic data evidence, found in computer systems, networks, wireless communications and storage devices or any other crime-related digital device, in a way that is admissible in a court of law.

Although the underlying logic behind digital forensics processes and procedures is based upon the traditional forensics mechanisms and methods, digital forensics establishes its own principles and standards. But what exactly is digital forensics and how is it defined?

2.1.1 Digital Forensics Definition

Since digital forensics is a relatively new, complicated, fast growing and evolving scientific field, many definitions derived in order to accurately interpret the term digital forensics. According to Digital Forensic Research Workshop (DFRWS), 2001, digital forensics can be defined as [1]:

“the use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation, and presentation of digital evidence derived from digital sources for the purpose of facilitating or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorized actions shown to be disruptive to planned operations”.

Another definition was given by Ken Zatyko (2007), who presented his documented opinion regarding the term of digital forensics in Forensic Magazine, stating that digital forensics science is [2]:

“the application of computer science and investigative procedures for a legal purpose involving the analysis of digital evidence (information of probative value that is stored or transmitted in binary form) after proper search authority, chain of custody, validation with mathematics (hash function), use of validated tools, repeatability, reporting, and possible expert presentation”.

Although originally used as a synonym for computer forensics, digital forensics was soon broadened such a way that it can encompass investigation of all devices capable of storing digital data, including future digital technologies.

2.1.2. Applications and Uses of Digital Forensics

Even though digital forensics investigations imply that the context is most often provided for use in a court of law, yet the truth is that digital forensics has a plethora of applications and can be used in other instances too. John Sammons (2014) [3] pointed out that digital forensics is applied mainly for criminal investigations, civil litigation, intelligence, and administrative matters. According to Watson D., Jones A. & Thornton F. (2013) [4], each sector is unique and comes with specific handling demands consistent to the scope of investigation.

Criminal Investigations

The most common use of digital forensics is in the context of criminal investigations as “electronic evidence can be found in almost any criminal investigation” [3]. Digital investigations aid mainly in the exposure of child pornography, identity theft, cyber bullying, cyber stalking, online predators, incidents of hacking, drug trafficking, fraud, homicide, sexual assault, robbery, and burglary crimes. Worth mentioning is the fact that all the above criminal actions are just indicative among the world of digital crime.

Apart from identifying direct evidence of a crime and its use before criminal courts, digital forensics can be held accountable for attributing evidence to specific suspects, confirming alibis and/or statements and determining criminal intent.

Civil Litigation

Also featured in civil cases, digital forensics constitutes a part of electronic discovery [3]. Electronic discovery, or most commonly called *e-discovery*, “refers to any process

in which electronic data is sought, located, secured, and searched with the intent of using it as evidence in a civil or criminal legal case. E-discovery can be carried out offline on a particular computer or it can be done in a network. Court-ordered or government sanctioned hacking for the purpose of obtaining critical evidence is also a type of e-discovery.” [5].

Intelligence

Digital age has also ~~intuites~~ terrorism and espionage providing advanced information technology for conducting numerous approaches for recruiting, communicating and planning disastrous attacks. These illegal actions are prevented and disclosed with the use of digital forensics.

Administrative Matters

Digital evidence can be also proven precious regarding incidents in the private sector, mostly when it comes to internal corporate investigations, intrusion investigations and reconstructions of computer security incidents. Last but not least, digital forensics can be applied for troubleshooting operational problems and recovering from accidental system damage and data loss.

2.1.3 Digital Forensics Necessity and History

It is widely known that digital age advanced and evolved the quality of human lives, however along with the advantages came many disadvantages too. Thus, digital crime emerged and is now well established. Consequently, the necessity for digital forensics derived, since digital devices very often leave traces and ‘betray’ their users.

As Watson et al. (2013) states [4], digital forensics generally encompasses the retrieval of digital evidence and data from any type of digital device that has storage memory. These data may have been accidentally or intentionally lost, concealed or erased either by a personal action or by a malware attack of any type.

Obviously, the power of digital forensics ~~relays~~ on the fact that someone can get access into the past. A great number of people maintains enormous quantities of information, either ~~because they want to~~ (in the form of log files and archives) or ~~without the intention of~~ (software that does not completely delete memory and files). As a consequence, ~~searchers~~ have the ability to recover old messages, chat logs, Google search terms and any other kind of data that has ~~previously~~ produced, days, months or even years before, using them to uncover the intentions of someone when a crime has ~~committed~~.

Unfortunately, law enforcement agencies were not prepared for the advent of digital crime and until 1980s the emerged computer related crimes were dealt with the existing legislation. The continuing increase of computer crimes over the next years led to the incorporation of computer offences in federal laws. In the following decade began the growth of the field, which started as computer forensics and later expanded so as to encompass all kinds of digital devices.

2.2 Digital Evidence

At this section, the meaning of digital or electronic evidence is presented along with the sources of evidence. Moreover, data types and types of digital evidence are explained followed by an overview regarding the nature and form of digital evidence.

2.2.1 Digital Evidence Definition

As ~~afore~~ mentioned before, digital forensics is dealing with a specific type of evidence which is referred as digital evidence or electronic evidence. In order a solid interpretation for the term digital evidence or electronic evidence to be adopted, it is essential for the reader to be familiar with the actual meaning of the words "*digital*", "*electronic*" and "*evidence*" in the context of the field we are studying.

By definition [6, 7], the word digital involves or relates to the use of computer technology and electronics, and deals with handling, generating, storing or processing information composed of data in the form of digital signals, expressed as series of ~~especially~~ binary digits of 0 and 1, which are typically represented by values of a physical quantity such as voltage or magnetic polarization. ~~The aforesaid data is processed and manipulated in the terms of two states: positive and non positive. Positive is declared by the number 1 whilst non positive by the number 0.~~ Thus, data transmitted or stored in digital electronic circuitry, is expressed as a string of 0's and 1's. Each of these state digits is referred to as a bit and a string of bits that a computer or/and a computer-based device can address individually as a group is referred as a byte. Examples of physical-world information that is converted to binary numeric form are digital images/photographs, digital audios and even digital broadcast (broadcasts using digital communications signals).

According to English Oxford Living Dictionaries site [6], the term electronic is commonly used to describe devices, circuits, or systems ~~developed through electronics~~, hav-

ing or operating with components such as microchips and transistors that control and direct electric currents.

Moreover, the same site provides the definition of evidence as "*the available body of facts or information indicating whether a belief or proposition is true or valid*". To explain further, evidence is anything (an indication or sign) which tends to prove or disprove something, providing ground for belief and/or proof. In the context of law enforcement, evidence is the information or data presented to a court or jury in order to establish and provide proof of the facts in the case of a legal investigation. Evidence is drawn from personal testimonies of witnesses and is obtained from records, documents, or material objects.

As far as the term digital evidence (or electronic evidence) is concerned, there are many definitions. To begin with, Casey Eoghan (2011) [8] mentioned that the Standard Working Group on Digital Evidence (SWGDE) defines digital evidence as "*any information of probative value that is either stored or transmitted in a digital form*", whilst the International Organization of Computer Evidence (IOCE) proposes the definition of "*information stored or transmitted in binary form that may be relied upon in court*". Moreover, as Casey states, the above definitions focus excessively on proof and neglect data that can be used for a further investigation. Thus, in his book provides two more definitions regarding digital evidence. A broader one which defines digital evidence as "*information and data of investigative value that is stored on or transmitted by a computer*", provided by the Association of Chief Police Officer and a more general one which was proposed by Brian Carrier (2006), defining digital evidence as "*digital data that supports or refutes a hypothesis about digital events or the state of digital data*".

According to Dr. Swarupa Dholam [9] the term "*digital*" is too broad, whilst the term "*binary*" is too restrictive and often inaccurate when used in the definitions of digital evidence, since it only defines one form of electronic data. Furthermore, Dr. Swarupa Dholam provides a definition regarding electronic evidence with the intention to cover *three* essential aspects: include all forms of evidence that can be created, manipulated or stored, include any form of device by which data can be stored or transmitted and include data that restricts relevant information to a particular investigation. Consequently, defines electronic evidence as "*data (comprising the output of analogue devices or data in digital format) that is manipulated, stored or communicated by any man-made device, computer or computer system or transmitted over a communication system, that*

has the potential to make the factual account of either party more probable or less probable than it would be without the evidence” [9].

2.2.2 Digital Evidence Sources

With the rapid evolution of technology the presence of digital devices is everywhere. Each type of digital device is capable of storing, manipulating and transmitting information and data. However, despite the fact that digital devices improved the quality of human’s life and help people communicate locally and globally with ease their major drawback is that they can easily be used in criminal ways. Any piece of digital technology that processes or stores digital data and is used in criminal action carries digital evidence which needs to be seized in an investigation. Thus, every digital device can become source of crucial and exclusive digital evidence.

The wide range of the unique sources of digital evidence requires the classification of different system types. As Casey (2011) presents [8], according to Henseler (2000), these types can be categorized into three groups: *open computer systems*, *communication systems* and *embedded computer systems*. Open computer systems are systems composed of hard drives, keyboards, and monitors such as Personal Computers (PC’s), laptops, desktops, servers or even game consoles. Communication systems include traditional telephone systems, wireless telecommunication systems, the Internet, and networks in general that can be a source of digital evidence. Embedded computer systems include mobile devices such as smartphones and wearables, smart cards and any other systems with embedded computers such as navigation systems and even home appliances that allow users to program them remotely via a wireless network or the Internet.

The above system types can be wealthy sources of digital evidence. As already mentioned, each system type contains numerous ~~of~~ digital devices every one of which requires a different evidence-gathering approach and ~~process~~ as well as different tools and methods for capturing that evidence.

2.2.3 Data Types and Types of Digital Evidence

Data Types

In any digital device, there are two major types of data, *volatile* data and *non-volatile* data. Volatile data is any data that is stored in memory or exists in transit and requires constant power in order to be retained, since when power goes out data is instantly lost. Some typical examples of volatile data are the random-access memory (RAM) storage

in a PC, ~~the computer's registry, computer history, deleted files~~, temporary files and ~~web browsing history~~. Nowadays, digital devices tend to have gigabytes of volatile storage and as a result data in the RAM is becoming more and more important. Because of the fact that volatile data can be key evidence and is ~~dependent~~ on the power of the device, thus becoming sensitive, the performance of a memory dump in any digital forensics investigation is necessary.

Non-volatile data is any data stored on a hard drive or another medium and is preserved when the power of the device goes out, meaning that data remains intact. Examples of devices that store non-volatile data are ~~except hard drives~~, disk drives and removable storage devices such as USB drives or flash drives, memory cards, optical discs, and ROMs. Non-volatile data can be conserved for a considerable long amount of time. The main difference between volatile and non-volatile data is that the latter in contradiction to the first, persists even without power ~~existence in~~ a device.

Types of Digital Evidence

Considering the plethora of the digital evidence sources in conjunction to the different data types it is safe to say that digital evidence can be of sundry types. Electronic mail messages, videos, audios, images/photographs, internet history and protocol information such as IP addresses, system and program files, temporary files and cache files, data extracted from GPS devices and smartphones are just some examples. The list can go on, however, some important and worth mentioning digital evidence, that is found merely at every device which is part of a criminal investigation, includes *metadata*, *slack space*, *swap files* and *unallocated space*.

Metadata is data that describes other data, meaning data embedded in the file itself, which contains information about that file. Basic information about data is summarized in metadata such as file name, size, location, file properties etc. and can be found in documents, spreadsheets, images, videos and even web pages. Metadata provides the ability to filter through it, providing the necessary information for locating and manipulating particular instances of data easier. There are two ways of metadata creation; either manually or by automated information processing. While manual creation allows users to input any information they feel it's necessary for a particular file, being in that way more accurate, an automated information processing creates the basic needed metadata including information about the file's size and extension along with information about when and who created the file [10].

Slack space, otherwise file slack space, as implied by its name, is the leftover storage space in a computer that is not occupied by an active file and yet is not available for use by the operating system. Every file in a computer fills a minimum amount of space which has been allocated by the operating system. Slack space results when file systems create a cluster (Windows) or block (Linux) but do not necessarily use the entire fixed length space that was allocated. Clusters are form of digital evidence because of collection of garbage and dangling references, constituting an important aspect of computer forensics [11].

The swap file (or swap space) is a hidden system file that is used for virtual memory when there is not enough physical memory for programs to be run. Space on the hard drive is temporarily swapped with the RAM as programs are running. This swap file contains portions of all documents and any other material a user produces while using the computer [12].

Unallocated space is the space flagged as no longer needed and be available for reuse when users delete or remove files. The original files remain on the system until they are overwritten, more specific until other files are stored in the same place. The remaining files are in unallocated disk space, where clusters/blocks are not assigned and can contain data of complete and/or partial files that may remain untouched for long periods of time [13].

2.2.4 Nature and Forms of Digital Evidence

Digital evidence can be derived from numerous sources and be of different types as mentioned in the previous section. However, it is important for readers to be familiar with both the form and nature of digital evidence. When it comes to the form of digital evidence it is essential for investigators to know that digital evidence can be present /active (documents, spreadsheets, images, email, etc.), archived (including backups), deleted (in slack and unallocated space), temporary (cache, print records, Internet usage records, etc.), encrypted or otherwise hidden and compressed or corrupted.

Digital evidence may be latent (hidden), like fingerprints or DNA evidence, or fragile, since it can be easily altered, damaged or destroyed with little effort or even not handled properly. Moreover, digital evidence can be time sensitive, since even the normal operation of the computer can destroy computer evidence that might be lurking in unallocated

space, file slack, or in the Windows swap file. Finally, in the context of law enforcement digital evidence crosses jurisdictional borders quickly and easily [14].

2.2.5 The Five Rules of Evidence

The findings of any digital forensics investigation must be based on proven techniques and methodologies, which are applied by the investigators and examiners. Similar to the traditional forensics science, extracted digital evidence must ~~be conformed~~ to the *five* rules of evidence in order to be fully utilized. These rules govern the way with which digital evidence is handled and must be followed by any forensics investigator from the moment a criminal action is identified until its settlement in courts. According to Braid (2001) [15], digital evidence must be admissible, authentic, complete, reliable and believable, and all these are explained in the following text:

Admissible

Admissible stands for the utilization of evidence and is the most significant rule. Evidence must be universally accepted in order to be used in courts or elsewhere. If the evidence is not admissible, then it is considered that it is not present at all and all processes, techniques and methodologies being used to collect and analyze that evidence had been conducted in vain.

Authentic

Evidence has to be also authentic, meaning that it is connected to the case in a sound and solid way. Digital forensics investigators and experts in their effort to provide proof, regarding an incident, must be capable of explaining that the evidence, they had extracted before, is related to a particular incident in a positive and pertinent manner.

Complete

The evidence collected must cover every aspect of the case under investigation. Therefore, evidence must be completed and all collected data and information has to be evaluated, examined and analyzed. Evidence should provide all the necessary information that is useful in proving attacker's actions, as well as the order in which these actions were performed. Indicating and demonstrating the attacker's involvement is not always enough, since in many cases evidence must be able to also prove the innocence of a potential suspect. Hence, it is also essential for investigators to collect exculpatory evidence that aids in eliminating alternative suspects, which is also a significant part of proving a case.

Reliable

Methods, techniques and procedures utilized for the recovery, collection and analysis of the evidence must not raise questions regarding the credibility, integrity and authenticity of the evidence.

Believable

The extracted evidence has to be presented in a court of law or other type of legal or administrative proceeding most of the times. Therefore, evidence must be believable by the individuals it is addressed to, meaning that it has to be clear, and easily understandable. Hence, evidence must be presented by experts who are able to provide all the appropriate details, comprehending the knowledge level of the jury.

Of course, ~~except~~ the aforementioned rules that have to be followed, it is of equal importance to maintain the ‘*chain of custody*’ or ‘*continuity of evidence*’. It is imperative that all evidence can be traced from the crime scene to the courtroom, and everywhere in between. For both physical and digital evidence, investigators must be able to prove that a specific piece of evidence was at a specific place, at a specific time and in a specific condition. All that vital information must be documented in detail in order to preserve the chain of custody. Failure in handling and ~~managing~~ properly evidence breaks the chain of custody and results in compromising, fatally in many cases, the fame of the investigator.

2.3 Branches of Digital Forensics

As mentioned before, digital forensics scientific field encompasses all kinds of digital devices that can be used either as a tool in enabling the crime or as a target of the crime. The devices have volatile memory, non-volatile memory or both, and the methodologies and processes for retrieving data and digital evidence are chosen based on the type of memory. According to Kumari and Mohapatra (2016) [16] digital forensics is divided into five branches depending on the type of devices, media or artifacts. The names of the different branches speak to the different areas which they focus on. The sub-branches of digital forensics are: Computer Forensics, Network Forensics, Mobile Device Forensics, Memory Forensics and Email Forensics.

A concise overview of the aforementioned branches is given in the section ~~below~~ followed by a presentation of another significant branch; database forensics. An in depth analysis regarding Computer Forensics and Network Forensics is given in the following chapters.

Computer Forensics

Computer forensics is the branch of digital forensics that focuses on applying computer science and technology for extracting evidence found in computers and digital storage media. The objective of computer forensics is not restricted in explaining the current state of a digital artifact but actually expands in order to "*examine digital media in a forensically sound manner with the aim of identifying, preserving, recovering, analyzing and presenting facts and opinions about the digital information*" [17]. This field comprises a wide variety of digital devices, from computer systems through to embedded systems (digital devices with rudimentary computing power and onboard memory) and static memory (such as USB pen drives). Computer forensics deals with a plethora of information, data and digital evidence with an extensive spectrum; from logs (such as internet history) to the actual files on the drive.

Network Forensics

Network forensics is the sub-branch of digital forensics that deals with the monitoring and analysis of computer network traffic, both local and WAN/internet, aiding in the recovery of information, legal evidence or intrusion detection [1]. Interception of traffic is usually achieved at packet level, and data evidence can be collected per network stack layer. The gathered data and information is either stored for later analysis or filtered in real-time. In contrast to other branches of digital forensics, which engage to stored or static data, and considering that network traffic is transmitted and then lost, network forensics usually deals with volatile and dynamic data that is rarely logged, thus leading to proactive investigations [18].

Mobile Device Forensics

As implied by the name, mobile device forensics is the sub-branch that relates to the recovery of digital evidence and data from a mobile device [19]. Although when it comes to mobile devices people usually think about mobile phones, this discipline covers a wide range of digital devices that have internal memory and/or communication ability including smartphones, tablets, GPS units, PDA devices and wearables. Kumari et al. (2016) points out [16] that the essential difference between computer forensics and mobile device forensics is that the latter deals with devices that have an inbuilt communication system and their own unique storage mechanisms.

Mobile devices can store a variety of data and information and mobile device forensics is usually applied to law enforcement investigations, military intelligence, corporate in-

vestigations, private investigations, criminal and civil defense, and electronic discovery. Despite the plethora of the above uses, mobile device investigations more often focus on simple data such as providing location information (either from inbuilt GPS/location tracking or via cell site logs, which track the devices within their range call), personal data and communications (SMS/Email) rather than in-depth recovery of deleted data [18].

Scott Polus (2016) [20] in his article “Mobile Device Forensics” mentions that the identification and collection of data from mobile devices are both problematic and perplexing not only because of the fragmentation of the operation systems, different versions within each operating system, and different variations between carriers but also due to the fact that mobile devices do not just operate as stand-alone data sources since they can constantly be synchronized with other devices and applications.

Memory Forensics

Memory forensics (otherwise memory analysis) is the forensic analysis of volatile data in a computer's memory dump [21]. Computers' memory (RAM) is explicitly reshaped from every function that is performed, either by an operating system or by an application. These modifications are preserved for a long time after each action occurred providing remarkable clarity into the runtime state of the system, such as current running processes, open network connections, and recently executed commands [22]. Thus, in a memory dump (a snapshot capture of computer memory data from a specific instant) valuable data is contained, such as encompassing disk encryption keys, memory-resident injected code fragments, off-the-record chat messages, unencrypted e-mail messages, and non-cacheable Internet history records. Consequently, the memory (RAM) is analyzed for forensic information in order to recover the aforementioned artifacts. The fact that this analysis does not depend on the system under investigation comes with the great advantage of the reduced chance that malware or rootkits interfere with the results. Memory forensics is mainly applied in investigating and identifying incidents of crash, security compromise and even advanced computer attacks or malicious behaviors which are so critical, avoiding leaving detectable tracks and data on the computer's hard drive.

Email Forensics

Email forensics is a relatively new sub-branch of digital forensics that relates to the extraction and analysis of information, data and digital evidence from emails. An email is

composed of many parts; header and its fields, email body, attachments and other properties each one of them provides unique information and data. Email evidence exists either in the email itself or is left behind as the email travels from sender to recipient (~~is~~ contained in the various logs and/or ~~is~~ maintained by system admins).

Email investigations primarily aim to the recovery of the emails' content, the identification of the email timestamp transmission and the tracing of both recipients and senders. Email forensics is applied mainly in criminal acts and serious improper actions, such as threats and frauds (phishing and spamming), including cases where emails are exploited with the intention to deceive [23].

Another Significant Branch - Database Forensics

As mentioned before, digital forensics is a wide and evolving discipline. So, taking into consideration that technology is continuously evolving, other branches ~~are~~ emerged and developed through the time. One of the most significant of them is the branch of database forensics which is worthy of mentioning.

According to Neeraj and Beniwal (2016) database forensics “*is another branch of digital forensics relating to the forensics study of database and its related metadata*” [24]. Database investigations require the examination of log files, and in-RAM data in order to build a timeline or extract consistent and pertinent information, starting from the investigation of the metadata in order to gain some clarity, regarding the entire database schema. Similarly to computer forensics, this branch follows the normal forensic process, applying examination methods to database contents and metadata, often requiring live analysis techniques as cached information might exist in a server’s RAM.

Database forensics investigations usually relate to the inspection and testing for validity of a database user's actions, focusing on the timestamps that had been applied at the time of alteration or updates in a relational table. The examination of such sensitive and critical data and information is the importance and essence of database forensics.

2.4 Digital Forensics Tools

At this section the various types of digital forensics tools are presented along with an overview regarding tool selection. Additionally, some widely known forensics tools are mentioned based on the branch of digital forensics they can be used.

2.4.1 Types of Digital Forensics Tools and Equipment

When conducting digital forensics investigations it is crucial to acquire, preserve, examine, analyze, and interpret data stored in the relevant digital device. To achieve this goal it is essential for the investigators to use the necessary, appropriate and specialized digital forensic tools. As mentioned by J. Wiles (2007) [25], “*the tools make the digital investigator*”, contributing to his/her work much more efficiently. There is a wide variety of digital forensics tools available, either focusing on specific purposes or serving a much broader functionality. They can be commercial tools that must be purchased or they can be open source items that are freely available. According to Nelson, Phillips and Steuart (2015) [26] there are two distinctive categories of digital forensics tools, *hardware tools* and *software tools*.

Hardware forensics tools are mainly designed and built explicitly for digital forensics and can be ranged from simple components that serve a unique and single-purpose to computer systems and servers. These tools encompass cloning devices, cell phone acquisition devices and kits write blockers, portable storage devices, adapters, cables and much more [3].

Building the right workstation depends on the type of the investigation, the available budget and the agency it is intended to. The more diverse the investigation environment is, the more options are needed and luckily there are many hardware vendors that provide a plethora of workstations that can be adapted in order the needs of a particular investigation to be met. The three major categories of the different workstations are the *stationary workstation*, which is basically a tower with several bays and many peripheral devices, *the portable workstation*, which is mainly a laptop computer with almost as many bays and peripherals as a stationary workstation and *the lightweight workstation*, which usually consists of a laptop computer built into a carrying case with a small selection of peripheral options [26].

Software forensics tools can be either open source or commercially produced tools. They exist in abundance and can be both general tools, serving a wide range of functions or can be more specialized, performing limited tasks. The latter type of these tools focuses mainly on a specific type of digital evidence, such as e-mail or Internet use. Software forensics tools are grouped into *command-line applications* and *GUI* (Graphical User Interface) applications.

Command-line forensics tools require few system resources and are designed to be run in minimal configurations. Such applications fit on bootable media (USB drives, CDs and DVDs) and can save both time and effort. Furthermore, most of these tools provide the ability to produce the needed text reports for the investigation at stake, where these reports can be easily stored on USB drives and other portable media. Command-line forensics tools are developed based on the operational system they are intended to be applied, thus there are different tools for Windows, UNIX/Linux and Mackintosh respectively.

On the other hand, GUI forensics tools simplify digital forensics investigations as they do not require the understanding of the operational systems or the knowledge of the commands that have to be used in command-line tools. Beginning examiners are most likely to use GUI forensics tools since limited training is required and the tools can be used easier. Such tools can perform multiple tasks as they are often developed as suite of tools. Of course, the use of command-line tools is still necessary because GUI tools may not work or may be unsuitable for specific situations or they can even miss critical evidence [26].

Except software and hardware forensics tools other equipment for performing digital forensics investigations are also needed. Preloaded crime scene kits with the appropriate and standard supplies required the collection of digital evidence from other equipment, such as pens, digital cameras, forensically clean storage media, evidence bags, evidence tape, report forms, permanent markers, gloves etc. [3].

2.4.2 Tools Selection

Wiles (2007) also states [25] that the selection of the proper hardware and software forensics tool for performing a digital forensics task and accomplish the extraction, preservation, collection, analysis and interpretation of the data on digital devices is critically important, since the findings are mainly presented in administrative, civil, or criminal proceedings and their introduction and applicability are governed by the (aforementioned) rules of evidence.

A digital forensics tool must be reliable, validated and above all must be the appropriate for the specific investigation under consideration. Locating as many tools as possible and using a number of tools for an investigation, is considered to be a good practice, taking into consideration that every tool comes with both advantages and disadvantages.

Keeping in mind that no single tool does everything or does everything accurately makes it essential for investigators to be familiar with as many tools as possible. Using multiple tools and obtaining the same results from two or more different sources is a great way to validate the findings, increasing the reliability of the evidence.

Since new forensics tools are emerging and the already existed ones are steadily being developed, updated, patched, revised and even discontinued makes it hard enough to select the right ones for an investigation. The criteria for the selection of the appropriate forensic tools include factors such as financial constraints, functionality, capabilities, support and of course finding the balance among disadvantages and advantages. Among the aforementioned criteria, it is of essence to acknowledge the fact that each specific branch of digital forensics is engaged with the examination of the particular data evidence.

2.4.3 Basic Tools

As stated before, the types of the digital forensic tools are the software and hardware used for gathering data from any digital device that is believed to be involved in a criminal action. Each branch of digital forensics deals with a specific type of digital devices and data types. Forensic tools are divided into various categories based on their specialization [16]:

Computer Forensic Tools, some of the basic and commonly used of which are Helix, Winhex, FTK Imager, CAT DETECT (Computer Activity Timeline Detection), Computer Forensics Timeline Visualization Tool, Encase, FTK (Forensic Tool Kit), Zeitline, and CFT (Cyber Forensic Timelab), Registry Recon, SANS Investigative Tool kit.

Memory Forensic Tools which are used to acquire and analyze a computers volatile memory such as CMAT (Compile Memory Analysis Tool) and Memoryze.

Mobile Device Tools that tend to have hardware and software components. Encase, PDA Seizure, Pilot-Link, OXYGEN Forensic KIT, Cellebrite Mobile Forensics device and MicroSystemation XRY.

Network Forensic Tools, which are designed to capture and analyze network packets either from LAN or Internet. Some basic and well known network forensic tools are Wireshark, NetworkMiner, Tcpdump, although there is a more detailed description of them in the next chapters.

Database Forensic Tools, that are related to the investigations applied on database and metadata such as HashKeeper and Arbutus.

Email Forensic Tools which are mainly used to extract information and data from multiple emails, instant messages and social networking, such as Digsby.

2.5 Digital Forensics Process Models

Here are presented the digital forensics process and its models definitions as well as a general description of the characteristics of the various models. Additionally, a chronological list is provided regarding the existing developed models followed by their clarification. In the last section is given a detailed description of the *four* critical and basic models.

2.5.1 Introduction to Digital Forensics Process Models

It is widely acknowledged that today's digital world is becoming an important, if not the most important, part of any criminal investigation. In order to reconstruct vital digital evidence from diverse digital sources, after a digital crime has been committed, the technique of digital forensics process has been developed. Technical skills and digital forensics tools are not usually enough when conducting a digital forensics investigation in order to investigate thoroughly and accurately a digital crime. Hence, the demand of a well-defined process model that goes beyond technical needs and provides the appropriate guidance for digital forensics investigators has derived.

Ashcroft (2001) [27] recognizes the digital forensics process as a scientific and forensics process used in digital forensics investigations. Casey (2004) [8] defines it as a number of consecutive steps from the original incident alert through the reporting of findings. According to Xiaoyu, Nhien-An and Scanlon (2017) [28], a process model in digital forensics is defined as “*the methodology used to conduct an investigation; a framework with a number of phases to guide an investigation*”. At that moment, a plethora of different models and techniques exists, ~~that~~ includes different phases or steps for investigation purposes, aiming in the implementation of the digital forensics process. This relies on both the wide range of different cases (such as cyber-attacks conducted by IT specialists, civil cases in a corporation, or criminal cases) and the fact that many of the process models were designed for ~~focusing on~~ a specific environment, such as law enforcement, meaning that they could not be applied with ease in other en-

vironments such as incident response. As a result there is not a unified and definite workflow in digital forensics investigations [28, 29].

Digital forensics process models focus on methodologies that provide a concrete sequence of phases and actions necessary in an investigation process. These methodologies aim at an accurate, robust and efficient digital forensics investigation. As G. Shrivastava, K. Sharma and A. Dwivedi (2012) [30] mention, digital forensic models have been developed in such a way that “*step wise or ordered inspection procedure of digital evidence can be made through it*”. These models provide a solid and meticulous principle of investigation process with the ultimate goal the outcome of reliable and admissible in courts digital evidence. Furthermore, digital forensics models constitute a step by step guide for digital evidence examiners and investigators, since they include detailed and concise information regarding particular aspects and phases that have to be taken under consideration during an investigation.

Each one of the models comes along with its advantages and disadvantages. Additionally, there are important similarities and differences among the various digital forensics process models. Frameworks with many phases and sub-phases might be more useful, despite the high existence of limitations regarding their usage scenario, whilst simplified frameworks with less phases and steps may lack the necessary guidance for an investigation process. Practically, every framework aims at the improvement and the enhancement of the standard methodology for every individual use case. On the whole, all process models agree on the importance of some phases and follow a wider similar approach. Moreover, most of the proposed frameworks accept some common starting points, giving this way an abstract frame that forensic researchers and practitioners apply on and use it to develop new research horizons with the scope of filling in continually evolving requirements [28, 30, 31].

Considering the rapid evolution of technology and digital forensics techniques, the optimal digital forensics process model has to be both general, meaning that it can be applied in almost every case, and capable of adopting and adapting new techniques in the investigation process [28]. However, an ideal, formal and globally accepted digital forensics process model that can be applied in every investigation has not been developed yet, despite the numerous attempts that already have been made [29].

2.5.2 Existing Models and their Categories

Since digital forensics is a constantly evolving field, numerous digital forensic process models have been developed. However, the number of proposed models and frameworks keeps on increasing and many refinements have been applied to the existing ones. The following list includes the developed models since 2000 which are presented in chronological order.

- A. *The Forensic Process Model (National Institute of Justice, 2001)*
- B. *The DFRWS Investigative Model (Digital Forensic Research Workshop, 2001)*
- C. *The Abstract Digital Forensic Model (M. Reith, C. Carr & G. Gunsch, 2002)*
- D. *The Integrated Digital Investigative Process (Carrier & Spafford, 2003)*
- E. *End-to-End Digital Investigation (EEDI) Process (Stephenson, 2003)*
- F. *Cyber Tools On-line Search for Evidence (CTOSE) (Hannan, Frings, Broucek, & Turner, 2003)*
- G. *An Extended Model of Cybercrime Investigations (Ciardhuain, 2004)*
- H. *The Enhanced Digital Investigation Process Model (Baryamureeba & Tushabe, 2004)*
- I. *The Digital Crime Scene Analysis Model (Rogers, 2004)*
- J. *A Hierarchical, Objectives-Based Framework for the Digital Investigations Process (Beebe & Clark, 2004)*
- K. *Case-Relevance Information Investigation (Ruibin, Yun and Gaertner, 2005)*
- L. *The Unified Modeling Language (UML) (Bogen and Dampier, 2005)*
- M. *Framework for a Digital Investigation (M. Kohn, J. Eloff, & Olivier, 2006)*
- N. *The Four Step Forensic Process (Kent, Chevalier, Grance and Dang, 2006)*
- O. *FORZA - Digital forensics investigation framework (Jeong, 2006)*
- P. *Process Flows for Cyber Forensics Training and Operations (Venter, 2006)*
- Q. *The Common Process Model for Incident Response and Computer Forensics (Freiling & Schwittay, (2007)*
- R. *Modeling Computer Forensic Process from a Workflow Perspective (Wang & Yu, 2007)*
- S. *The Two-Dimensional Evidence Reliability Amplification Process Model (Khatir, Hejazi and Sneider, 2008)*
- T. *Mapping Process of Digital Forensic Investigation Framework (Selamat, Yusof and Sahib, 2008)*
- U. *The Systematic Digital Forensic Investigation Model (SRDFIM) (Agarwal, M. Gupta, S. Gupta and S.C. Gupta, 2011)*

V. *The Advanced Data Acquisition Model (ADAM): A process model for digital forensic practice (Adams, 2012)*

According to Adams (2012) [29] the identified models are classified into three categories based upon the approaches they follow: ‘*ad hoc*’ models, ‘*process flow*’ and ‘*scientific*’ approaches.

Ad hoc digital forensics process models

In this category are included the models that do not conform to a standardized and recognized methodological approach. Each model is uniquely developed and presented, however there are models which were inspired and based on previously models. The following models belong into this category:

- *The Abstract Digital Forensic Model (M. Reith, C. Carr & G. Gunsch, 2002)*
- *The Enhanced Digital Investigation Process Model (Baryamureeba & Tushabe, 2004)*
- *The Digital Crime Scene Analysis Model (Rogers, 2004)*
- *A Hierarchical, Objectives-Based Framework for the Digital Investigations Process (Beebe & Clark, 2004)*
- *Framework for a Digital Investigation (M. Kohn, J. Eloff, & Olivier, 2006)*
- *The Four Step Forensic Process (Kent, Chevalier, Grance and Dang, 2006)*
- *The Common Process Model for Incident Response and Computer Forensics (Freiling & Schwittay, (2007)*
- *The Two-Dimensional Evidence Reliability Amplification Process Model (Khatir, Hejazi and Sneiders, 2008)*
- *Mapping Process of Digital Forensic Investigation Framework (Selamat, Yusof and Sahib, 2008)*

From approaches that include standardized and fundamental phases to approaches that include multiple phases and sub-phases, this category seems to encompass models that present a broad spectrum of approaches. Furthermore, whilst some authors focus on particular environments such as law enforcement or incident report others seek for the development of more generic approaches that can be applied into as many cases as possible.

Process flow approaches for digital forensics models

In contrast to ad hoc digital forensics process models, which are mainly concerned with the level detail of the investigative process, process flow approaches for digital forensics models focus on the workflow of the investigations. These approaches apprehend

and capture aspects of digital forensics processes that ad hoc models were unable to include. Hence, it is crucial for the investigators to embrace them although their practical use is finite. The models of this category are:

- *An Extended Model of Cybercrime Investigations (Ciardhuain, 2004)*
- *FORZA - Digital forensics investigation framework (Jeong, 2006)*
- *Process Flows for Cyber Forensics Training and Operations (Venter, 2006)*

Mapping the activities of the digital forensics practitioners to the information they generate and store for the purpose of the investigation, these three process flow models yield another angle of view in the digital forensics process than the one provided from ad hoc models.

Scientific approaches for digital forensics models

When developing the previously mentioned models, authors developed their terminology and definitions in their attempt to describe these models. However, even though it was quite handy, the major drawback was that models were not properly defined in order to be readily part of a scientific discipline, since they were not established under formal specifications. This category includes models that were developed focusing more on the modeling approach rather than the models themselves. Models that adopted the required formal method for their definition and description are:

- *End-to-End Digital Investigation (EEDI) Process (Stephenson, 2003)*
- *Cyber Tools On-line Search for Evidence (CTOSE) (Hannan, Frings, Broucek, & Turner, 2003)*
- *The Unified Modeling Language (UML) (Bogen and Dampier, 2005)*
- *Modeling Computer Forensic Process from a Workflow Perspective (Wang & Yu, 2007)*

2.5.3 Presentation of the Four Major Digital Forensics Process Models

In this section is presented a detailed description of the first four major models that were developed and constituted the core and the fundamental basis upon other models that were later built and developed.

A. The Forensic Process Model, 2001

The Forensic Process Model was proposed by the U.S National Institute of Justice (NIJ) in 2001. This model serves as a guide for law enforcement and other responders who are responsible for protecting an electronic crime scene. Additionally, it provides a

proper forensic procedure for the recognition, collection and preservation of digital evidence in order to overcome the challenges for its admissibility in court. The Forensic Process Model consists of four main and successive phases or steps: collection, examination, analysis and reporting of digital evidence. These phases are common for all process models which agree on the importance of them and are depicted in the following figure.

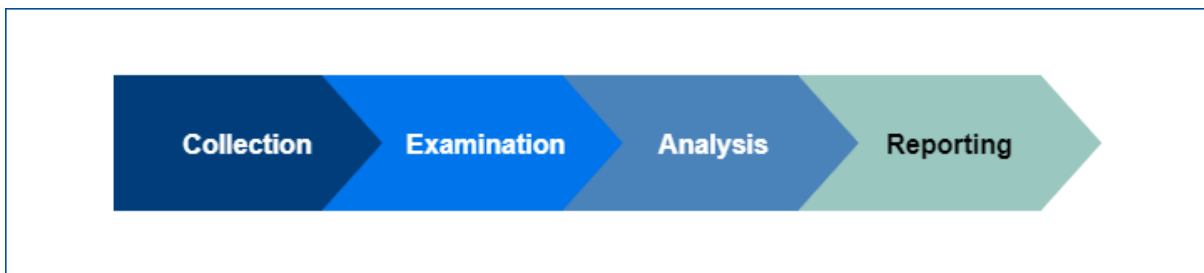


Figure 1: The Forensic Process Model (2001)

Collection: As implied by its name, the collection phase involves the search for, recognition of, collection of, and documentation of digital evidence, while following procedures that ensure and preserve the integrity of the data. At this step, data is identified, labeled, recorded and recovered from all possible sources. Data collection engages the extraction of both real-time and stored information, thus data should be collected in a timely manner ~~way~~ that no dynamic data will be lost, such as a list of current network connections or data collected in mobile phones and other devices.

Examination: The examination phase occurs directly after the collection phase. At this phase the collected data is forensically processed using a combination of automated and manual methods. Hidden data and obscured information of specific interest is identified, revealed and assessed, while its integrity is preserved. This phase aids in the visibility of digital evidence and its detailed documentation which includes the explanation of evidence's origin and significance. Additionally, the documentation allows all parties ~~discover~~ and comprehend the content and state of the evidence. When all data and information becomes visible, the process of data reduction begins which separates the relevant from the irrelevant data. This is a crucial process taking into consideration the vast amount of information and data that can be stored on the various storage media. The examination phase is considerably vital, since it bridges the gap between the collection of the data and its analysis which allows the use of the evidence in courts.

Analysis: In short, the analysis phase is the accurate inspection of the examination phase. The results of the examination are analyzed using well documented and legally justifiable methods and techniques, with the scope of the retrieval of essential information that answers the questions which constituted the kick off for the performance of the collection and examination processes. The aim of this process is the thorough studying of the examination results in order the importance and probative value of the evidence of the particular investigation case can be proved. The difference between examination and analysis phase is that the latter focuses on the product of the examination, highlighting its significance and probative value to the case whilst the first is a more technical review.

Reporting: In order the investigation to be completed, the reporting phase is necessary. It is basically a written report, which outlines the examination process, the results of the analysis phase and the relevant information acquired from the whole investigation. In brief, reporting is the conclusion of all the previously performed phases. The reported content includes the description of applied actions and an explanation regarding which and how tools and procedures were selected. Additionally, it provides more information determining what other actions need to be performed such as forensics examination of additional data sources, securing identified vulnerabilities and improving existing security controls. Furthermore, it includes recommendations regarding improvements of policies, procedures, tools, and any other related aspect of the forensic process. Since digital forensics technicians and examiners often have to be testified about the examination conduction and the validity of the procedure, preserving the reported notes is of a great essence [27, 28, 32, 33].

B. The DFRWS Investigative Model, 2001

In 2001, the first Digital Forensic Research Workshop (DFRWS) was held in Utica of New York. The attended audience consisted of civilian, military and law enforcement professionals, covering both academic and practical aspect of digital forensics science. The primary aim of the conference was the creation of a forum for the concerned community in order to share knowledge and forensics techniques on digital forensics science. During that workshop was agreed that digital forensics is a fundamental process with identified phases/steps, and proposed a general digital forensics investigative process, the DFRWS investigative model. This model indicates an investigative framework composed of six main consecutive phases: identification, preservation, collection, ex-

amination, analysis, and presentation. The processing of each phase requires various specialized methods and techniques to be used.

Identification phase is the first step of this model, where event or crime detection, resolving signature, profile detection, anomalous detection, system monitoring, audit analysis, etc. is conducted.

Preservation phase constitutes the second step of the DFRWS investigative model. At this point, the appropriate case management is set, imaging technologies are used, time synchronization is performed and all the required measurements are taken in order an accurate and acceptable chain of custody to be ensured.

The immediate following step is *Collection* phase, in which the collection of relevant data is performed based on approved methods, software and hardware. At this step assorted techniques for data recovery, sampling and reduction are encompassed along with techniques for lossless compression.

The two serial following phases, *Examination* phase and *Analysis* phase are considered to be critical and significant for the investigation process. Both phases are dealing with evidence traceability. During the examination phase discovery and extraction of hidden and encrypted data is performed, ~~as well as pattern matching is guaranteed.~~ This phase implicates the involvement of many validation and filtering techniques. On the other hand, in the analysis phase tasks such as data mining and timeline are involved.

The last step of the DFRWS investigative model is the *Presentation* phase. This phase includes documentation, expert testimony, clarification, mission impact statement, recommended countermeasures and statistical interpretation. Often, a seventh phase is added in the investigation process, ~~named decision phase~~, which is considered to be a pseudo phase and follows the presentation phase.

The figure below illustrates the DFRWS investigative model. In the first row the phases are depicted whilst the items in each column are the methods or techniques used in each phase.

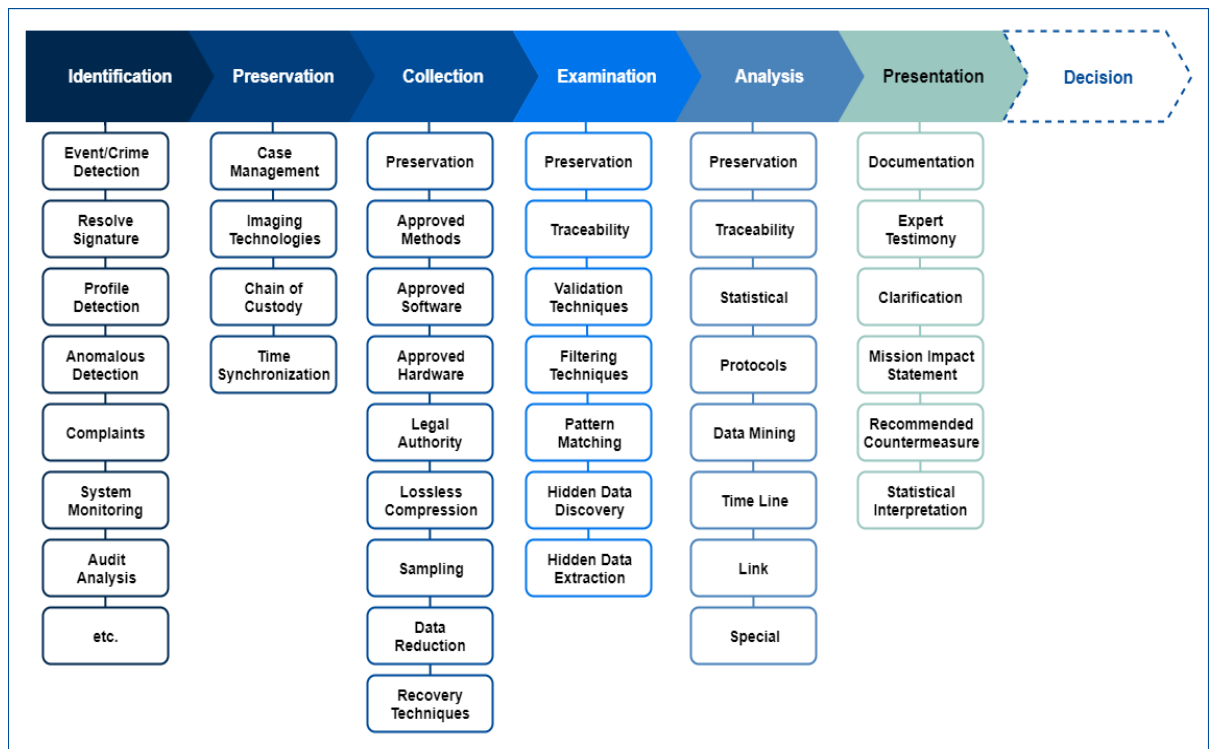


Figure 2: The DFRWS Investigative Model (2001).

The DFRWS investigative model pursues a linear process; however the incorporation of the feedback is of essence in order its efficiency to be advanced. Bearing in mind that real-time analysis aids in more effective detection, it is considered to be a highly important research objective. When conducting the analysis phase the construction of a digital forensics knowledge repository is needed, thus collaborative technologies are extremely helpful when it comes to digital forensics investigations.

The establishment of this model came with many advantages. First of all, the DFRWS model institutes the preservation phase as a guarded principle across all forensic phases. The advantage derives from the fact that the preservation phase is critical, since this phase ensures that the collected data is not contaminated and hence is reliable, authentic and admissible to courts. Another important advantage of this particular model is the capability of covering phases that previous developed models did not cover, such as the presentation phase. However, the main advantage of the DFRWS investigation model is that its development emerged from a universally accepted organization, which was lead mainly by academics, in preference to law enforcement professionals. Hence, this model was defined based on the scientific aspect of digital forensics and its challenges. Being comprehensive, consistent and standardized this model constituted the fundamental

basis for future works and enhancements regarding the digital forensics process models [32, 34, 35].

C. The Abstract Digital Forensic Model (ADFM), 2002

As seen previously, the DFRWS investigative model was built with the intention to be a generic “technology-independent” model. Based on the DFRWS investigative model M. Reith, C. Carr and G. Gunsch (2002) [36] developed and presented the Abstract Digital Forensic Model. This model is basically an enhancement of the DFRWS model which served as an inspiration. Using previous forensic protocols and frameworks Reith et al. (2002) defined [36] common abstract phases for this digital forensics model, imitating ideas and methods from traditional forensics approaches and evidence collection strategies which were already in enforce at that time. The developed model is composed of nine phases regarding digital evidence: identification, preparation, approach strategy, preservation, collection, examination, analysis, presentation and returning evidence, which are shown in the following figure:

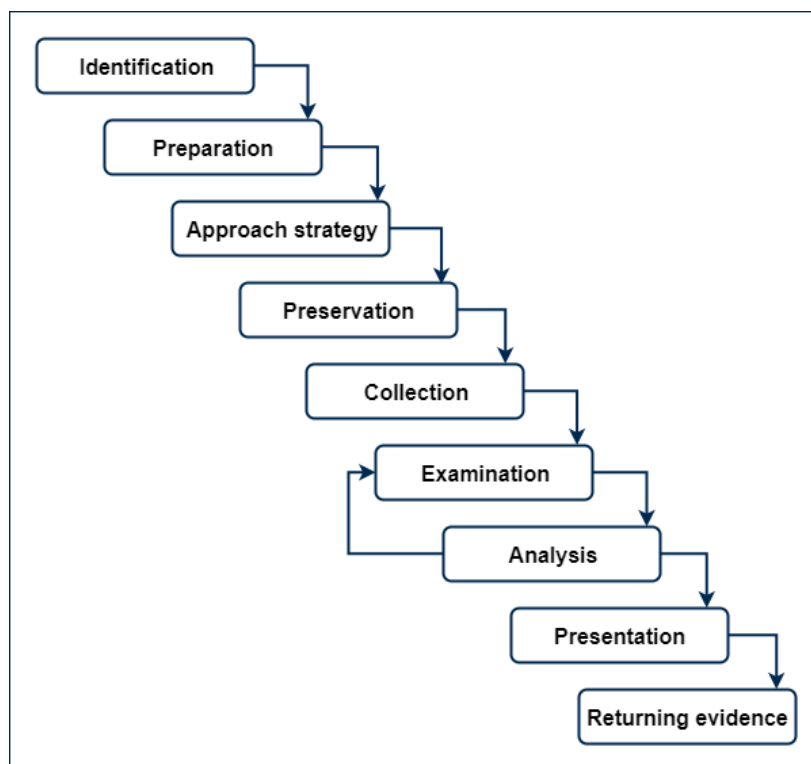


Figure 3: The ADFM Investigative Model (2002).

Identification: This phase ensures that the incident type is identified and determined properly. This phase is critical since all the upcoming phases depend on it, though it is explicit in the field of forensics.

Preparation: This phase constitutes the first introduced phase and encompasses the preparation of tools and techniques that will be used and required search warrants. Additionally, monitoring authorization and management support are groomed.

Approach strategy: During the approach strategy phase the different approaches and procedures that need to be followed are formulated. This constitutes the second introduced phase and aims in maximizing the collection of untainted evidence while minimizing the impact on the victim at the same time.

Preservation: Preservation phase involves the isolation, security and preservation of all acquired data while focusing on the conservation of the actual state of physical and digital evidence.

Collection: Under the collection phase the physical scene is recorded and all extracted and recovered digital evidence is duplicated using standardized and accepted procedures.

Examination: A thorough and meticulous systematic search of data and information that is connected to the particular suspected crime is conducted in this phase. The performed search focuses on identifying and locating potential evidence.

Analysis: The analysis phase deals with the product of the examination phase, meaning the results that occurred from the examination process, and aims in the determination of the significance and probative value of the evidence. Additionally, this phase involves reconstruction of data fragments and drawing of conclusions based on the evidence found.

Presentation: When it comes to the presentation phase, it deals with the development of a summary regarding the whole investigation process and the explanation of conclusions from all the previously performed phases.

Returning Evidence: Closing the investigation process all physical and digital evidence and property is returned to the proper owner. This phase constitutes the third introduced phase.

In contrast to the DFRWS investigative model, the Abstract Digital Forensic Model specifies a detailed description and representation of the complete process that has to be undertaken by digital forensics practitioners and examiners. Additionally, between the identification and preservation phases ~~are placed two more phases~~ the preparation phase

and the approach strategy phase. Furthermore, the pseudo phase of the DFRWS model is completely replaced with the returning evidence phase.

The most important value added by the Abstract Digital Forensic Model is that it consists of comprehensive pre and post investigation procedures. This model tried to solve the problems of the previous one and ~~partially~~ it succeeded. However, under a more close observation it is obvious that the second phase (Approach strategy) is to an extent a duplication of the first phase (preparation) and there is not a ~~distinguish~~ difference between them. Furthermore, it is practically more suitable if the preparation phase comes before the identification phase since digital forensics practitioners must be ready before any incident.

D. The Integrated Digital Investigative Process Model (IDIP), 2003

In 2003 Carrier and Spafford [37] developed and proposed the Integrated Digital Investigative Process (IDIP) model. Based on ~~the previous~~ ~~done~~ work their goal was to “integrate” all available models and investigative procedures. Their effort focused on mapping the digital investigation process to the physical investigation process. The IDIP model is quite extensive and is organized into five groups consisting of 17 phases which are illustrated in the following figure.

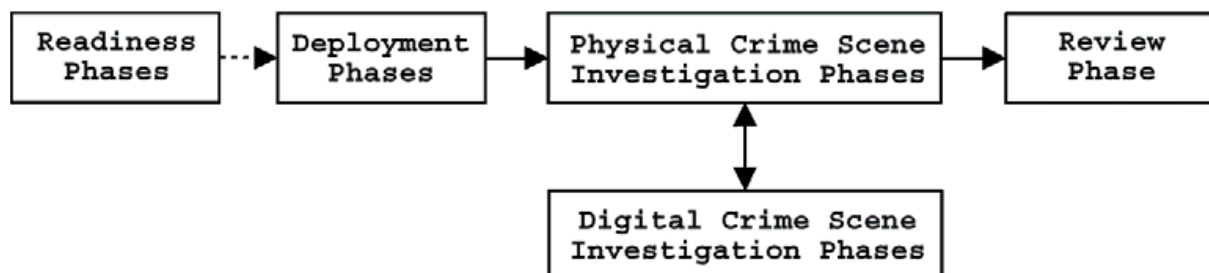


Figure 4: The five groups of phases in the IDIP model, (Carrier and Spafford, 2003).

Readiness Phases: The model begins with the readiness phases, which ensure that both operations and infrastructure are properly geared up in order ~~the~~ the entire investigation to be supported. Readiness phases include operations readiness phase and infrastructure readiness phase.

Operations readiness phase includes capacity training for every person who is involved in the investigation, such as first responders in the crime scene, investigators, lab analysts, etc. Additionally, in this phase ~~is~~ assured that all provided equipment for the personnel is functioning properly, is well maintained, up to date and ~~ready~~ ready when the incident data is delivered.

Infrastructure readiness phase ensures the existence of the required data in order a full investigation to take place. This phase is addressed exclusively to those who are responsible for the maintenance of the environment that could be the target of a crime. Physical examples of this phase include the deployment of sufficient infrastructure, like video cameras and card readers in order to record who was in the area at the time of the crime, whilst digital examples of this phase include sending server logs to a secured log host, or maintaining a change management database.

Deployment Phases: These phases provide the mechanisms for incident detection and confirmation, including the detection and notification phase and the confirmation and authorization phase.

The Detection and Notification Phase is, in short, the phase when an occurred incident is detected and appropriate people is notified. This phase is the impetus for the investigation process.

On the other hand, *the Confirmation and Authorization Phase* has different approaches depending on the situation at stake. The main goal of this phase is the confirmation of an incident and the obtainment of legal approval and authorization in order to proceed with the investigation. Whilst a search warrant is necessary in the context of law enforcement, when it comes to corporate incidents the verification of the incident from a response team seems to be followed by the approval from the appropriate supervisors who take under consideration the privacy policies that are in place.

Physical Crime Investigation Phases: Carrier and Spafford provided the clarification of the Physical Crime Scene as follows: “*Physical Crime Scene: The physical environment where physical evidence of a crime or incident exists. The environment where the first criminal act occurred is the primary physical crime scene and subsequent scenes are secondary physical crime scenes.*” [37].

The physical crime investigation phases, regarding digital investigations, aim in the collection and analysis of physical evidence in order to identify the responsible persons involved in an incident and reconstruct the actions that had been occurred during the incident. The encompassed phases are depicted in the figure bellow:

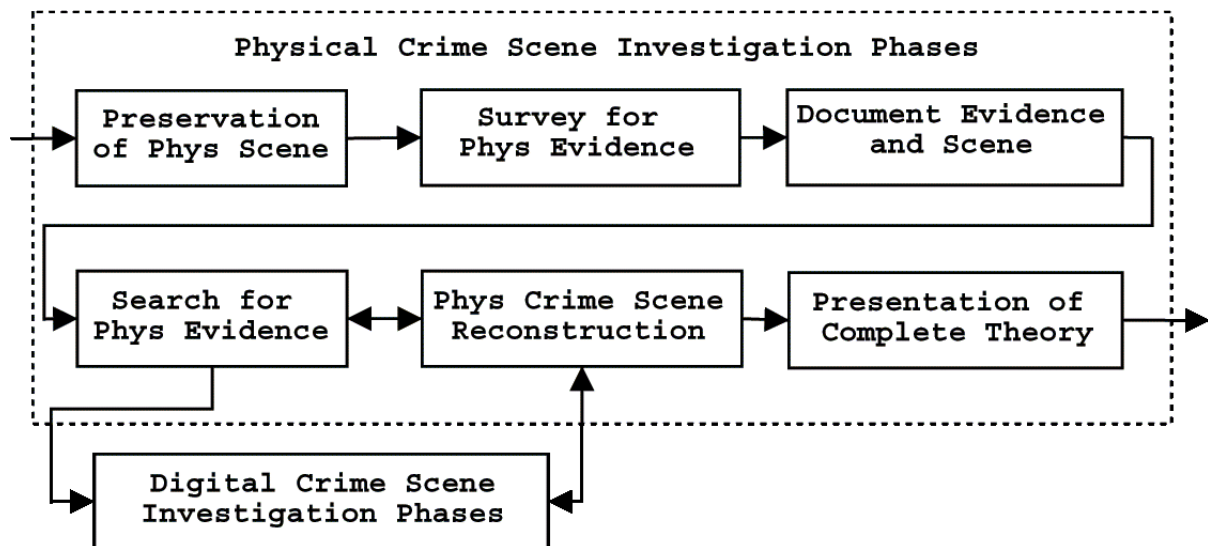


Figure 5: The six phases in the physical crime scene investigation and the interaction with the digital crime scene investigation phases, (Carrier and Spafford, 2003).

The Preservation Phase as implied by its name deals with the preservation of the physical crime scene so that evidence is later identified and collected by trained personnel and does not depend on the type of the crime. This phase includes actions such as witnesses' identification and suspect detention.

During *the Survey Phase of the physical crime scene* investigators walk through the physical crime scene and identify pieces of physical evidence. Fragile pieces of evidence are collected and documented immediately, so they cannot be damaged. Examples of physical evidence include computers, mobile and other removable media with computers considered as fragile evidence since evidence can be deleted with the use of remote systems.

The Documentation Phase of the physical crime scene aims in capturing as much information as possible from the crime scene in order to record and preserve its layout and all significant details. This phase encompasses processes such as taking photographs, videos and sketches of the crime scene and the physical evidence. It is also essential to document and take pictures of the state of the computers and their connections as well as information such as number and size of the amount of memory in case of hard drives. As Carrier and Spafford mention [37] it is highly important to notice that this phase does not include the generation of the final incident report.

The Search and Collection Phase of the physical crime scene represents a thorough search and collection of additional missing pieces of physical evidence. The Search and Collection Phase of the physical crime scene is basically the starting point of the digital

crime scene investigation. It is of essence to take under consideration that different evidence types require different procedures for their collection. The collected evidence is sent to labs for both processing and analysis and the results are available for use in the Reconstruction phase which follows.

The Reconstruction Phase of the physical crime scene deals with the organization of the results derived from the analysis and the development of a theory for the incident. Even though reconstruction is quite similar to the analysis phase, it mainly uses the results of the latter in order to put together the pieces, the link evidence and persons of interest to the incident.

The Presentation Phase of the physical crime scene is, in brief, the presentation of evidence (both digital and physical) and the developed theory to court or corporate management.

Digital Crime Scene Investigation Phases: The definition of the digital crime scene as proposed of Carrier and Spafford is: “*Digital Crime Scene: The virtual environment created by software and hardware where digital evidence of a crime or incident exists. The environment where the first criminal act occurred is the primary digital crime scene and subsequent scenes are called secondary digital crime scenes.*”[37].

The digital crime scene investigation phases begin the moment that digital devices are collected as physical evidence from the physical crime scene or when analysis for evidence involves the recorded network traffic. These phases use computer based approaches to search for evidence and each digital device is acknowledged as a unique and separate crime scene. The diversity of the digital devices allows the analysis to take place at different locations and its results are sent to the Physical Crime Scene Reconstruction Phase where the connection between the devices is identified. The phases included in this group are shown in the following figure:

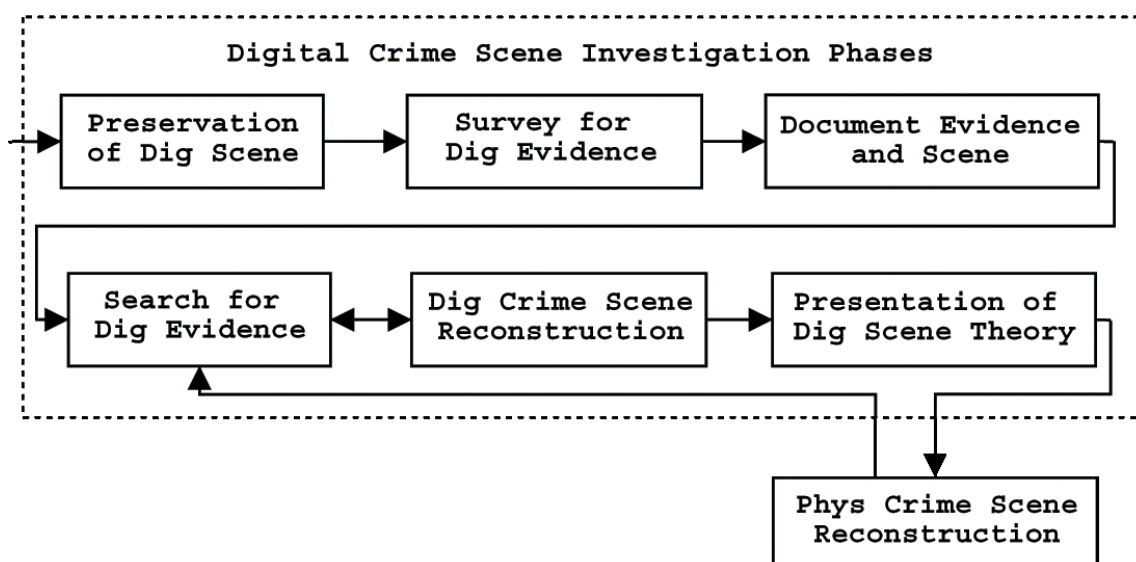


Figure 6: The six phases in the digital crime scene investigation. The results are fed back to the physical crime scene investigation, (Carrier and Spafford, 2003).

The Preservation Phase of the digital crime scene, similar to the physical crime scene, involves the preservation of the digital crime scene so that evidence is later collected by trained personnel. This phase deals mainly with the preservation and security of all types of digital evidence. In contrast to other models, in this one the preservation phase is about preserving the entire digital environment and not just the actual digital evidence. ~~More particular~~ the digital evidence has not been yet identified in this phase. A common process of this phase is the replication of the digital environment by making a complete forensics image backup of the system under investigation. Thus, the digital environment and evidence are preserved and critical systems can rebuilt while copies are used for analysis in the labs at the same time. However, there are cases where the original devices are required as physical evidence for examination and analysis.

The Survey Phase of the digital crime scene involves the use of the replicated images of the digital crime scene. This phase specifically occurs when investigators transfer relevant data to a controlled location. Even though it can occur in a live system as well, it is ~~preferred to be done~~ in the controlled environment of a lab. Sometimes it even occurs in the actual crime scene in order to identify if a system has to be fully analyzed in the lab. This phase aims in discovering the actual digital evidence that relates to a specific type of crime and can reveal the skill level of the perpetrator and the required analysis techniques for the investigation.

The Documentation Phase of the digital crime scene deals with the proper documentation of the digital evidence when it is found. All evidence found during the previous

phase is properly documented. At this phase the documentations is about each piece of evidence individually and does not constitute the final report which is generated in the Presentation Phase. In practice this phase is not explicit since in digital investigations evidence is documented as it is found.

In *the Search and Collection Phase of the digital crime scene*, an in-depth analysis of the digital evidence is performed. This phase uses the acquired results from the previous phase for further analysis. Some of the processes that take place in this phase are keyword search, extraction of unallocated space for processing, and analysis of low-level timeline of the activity and much more. The ~~time amount~~ that is spent in this phase is significantly higher ~~from~~ the other phases and there are numerous ~~of~~ different techniques that are used here. The search phase in conjunction to the survey phase is considered to be for the IDIP model whatever the examination phase is for the previously developed model.

The Reconstruction Phase of the digital crime scene is about connecting the dots in the digital riddle, developing the investigative hypothesis. In this phase methods and techniques are applied to identify every action of the crime, explain its existence, and test, approve or reject developed theories. Reconstruction Phase is similar to the analysis phase of other models.

The Presentation Phase of the digital crime scene is the final presentation of the retrieved evidence to the physical investigative team. This phase focuses on the presentation of the digital evidence that was found in a particular digital crime scene.

The final phase of the IDIP model is **the Review phase**. It involves the thorough review of the whole investigation process in order to identify aspects that their improvement is required to make the model more efficient. The review phase examines how accurate and effective ~~were~~ both physical and digital investigations, how well they worked together and identifies if the existence of physical and digital evidence was ~~enough~~ for solving the case. If everything worked out as expected there are no results for this phase. If did not, then the results themselves propose new required procedures and training.

The IDIP model simplifies the forensics process by grouping the phases into an abstract and manageable manner. Additionally, it highlights reconstruction and emphasizes ~~in~~ the review of the whole process, while putting the preparation phase before detection of

the incident. Furthermore, it differentiates between the digital and physical crime scenes.

On the other hand, although this model is generally a good reflection of the forensics process, it depicts the deployment phase which consists of confirmation of the incident as being independent of the digital and physical investigations. However, in practice the confirmation of a digital crime seems impossible ~~unless and until~~ some preliminary physical and digital investigation is carried out. Furthermore, it does not provide a sufficient specificity since it is not capable of drawing a clear distinction between investigations at the ~~victims~~ and ~~suspects~~ crime scene. Moreover, it illustrates the forensic process as linear and last but not least it contains two reconstructions that may sometimes contradict.

2.6 Challenges in Digital Forensics

Due to the significant evolution of the technology over the past years the use of more and more electronic/digital devices in criminal actions, cyber-attacks, frauds and malicious activities has emerged. As a result, gathering digital evidence and enabling investigations through technology has become an imperative need nowadays. Hence, in the undergoing digital age, forensics is constantly changing, thus digital forensics constitutes a crucial branch of forensics [16].

When it comes to digital forensics investigations and the acquisition of technology-oriented evidence, numerous ~~of~~ digital forensics tools, both hardware and software, ~~are~~ surfaced ~~and developed day by day~~ along with the improvement of methods and procedures. However, there are important challenges that have been ~~derived~~ and still remain in digital forensics in contrast to traditional forensics science. According to Fahdi, Clarke & Furnell (2013) [38], the challenges of digital forensics can be classified into three major categories: technical challenges, legal challenges and resource challenges.

2.6.1 Technical Challenges in Digital Forensics

There is a wide variety of technical challenges that obstruct the extraction of digital evidence in an investigation, such as the differing media formats and the live acquisition and analysis of data. According to Kumari et al. (2016) [16], a digital forensics process compatible with all digital devices seems to be impossible as each device is unique and can store various and different types of media and data formats. Forensics investigation

on a live system comes with momentous risks since the operating system can either hide or alter essential evidence. However, the real nightmare is the *anti-forensics*, which constitutes a critical technical challenge for digital forensics that has to be faced.

Anti-Forensics

Even though technology's primary intention is the invention of innovative things in favor of mankind's benefit, it also assists cyber criminals to accomplish their personal goals. Technology is being misused creating counter effects in digital forensics. Programmers design and develop anti-forensics tools in their effort to hide themselves and neutralize the effect of modern and advanced forensics tools. Anti-forensics programs can fool computers by changing the information in files' headers, divide files up into small sections and hide each section at the end of other files, change metadata attached to files, erase data if an unauthorized user tries to access the system, insert executable files into other kinds of files (such programs are called packers) and much more.

In contrast to many other sources of physical evidence, digital evidence can be easily modified, removed, deleted or even hidden often without leaving traces. Taking advantage of the fragile and sensitive nature of digital evidence, many forensics tools are designed to hinder information and data retrieval during an investigation or either to delay the digital evidence generation process [16].

The classification of anti-forensics techniques as presented by Rekhis & Boudriga (2010) [39], is encryption, steganography, covert channel, data hiding in storage space, residual data wiping, ~~tail~~ obfuscation, attacking the tools and attacking the investigators.

Encryption

According to TechTerms (2014), encryption is the conversion of electronic data into an unrecognizable or "encrypted" form called ciphertext [40]. Encryption prevents unauthorized parties from reading data as the latter becomes hard to be understood. Encryption is primarily used to protect the confidentiality of digital data stored on computer systems and storage devices or data transmitted via the Internet or other wireless networks. ~~Except~~ confidentiality, modern encryption algorithms provide security assurance of IT systems regarding authentication, integrity and non-repudiation. ~~Even though it can be considered as one of the most valuable anti forensics techniques, its efficiency can be avoided if the system under investigation is left on.~~

Steganography

Steganography is the technique of hiding any information inside a file carrier without modifying its outward appearance and the extraction of it at its destination. Steganography can be used along with cryptography, taking the latter a step further in data protection. Attackers use steganography to hide their hidden data (payloads) inside the compromised system. In modern digital steganography, data is inserted using special algorithms into redundant data that is part of a particular file format such as a JPEG image after it has been encrypted by standard means [5].

Covert Channel

As stated by Rekhis et al. (2010) [39], a covert channel in communication protocols is a type of computer attack that allows the communication of information by transferring objects through existing information channels or networks using the structure of the existing medium to convey the data in small parts. This makes conveyance through a covert channel virtually undetectable by administrators or users and allows attackers to hide data over the network and possibly bypass intrusion detection techniques. Covert channels have been mainly used to steal data from highly secure systems by maintaining a hidden connection between the attacker and the compromised system.

Data Hiding in Storage Space

Data hiding in storage space is the process of hiding data in storage areas making it invisible to the usual system commands and programs by using techniques such as rootkit. Rootkits are capable of hiding processes, files, system drivers, network ports, and even system services and what makes them unique is that they cannot be easily identified nor removed. Furthermore, rootkits delay significantly the data acquisition process making the investigation more complex.

Residual Data Wiping

When it comes to residual data wiping, it refers to the deletion of hidden processes that run on a computer such as temporary files and history of commands. Such processes often run on the background, without the knowledge of the attackers. To make a system work as if it has not been used for a malicious purpose requires high skills and intelligence of the perpetrator.

Tail Obfuscation—Attacking the Investigators

According to Rekhis & Boudriga, (2010) [39], the most common technique is the obfuscation of the source of the attack. In order to mislead investigators and make them miss data and clues of critical forensics value, attackers use false information such as false email headers and modified file extensions. Methods of trail obfuscation, which

used to thwart digital investigation process, are spoofing, log clearance, zombie accounts, Trojans and misinformation.

Attacking the Tools

As mentioned before, digital forensics tools must be reliable and validated. Attackers often target on forensics tools, blind spots or vulnerabilities in order to discredit the tools and as a result the reliability of digital evidence. Hence, digital evidence becomes contaminated, can be called into question, becoming worthless in a court of law and thus is rejected.

2.6.2 Legal Challenges in Digital Forensics

From the legal point of view, the challenges that have to be faced are mainly regarding jurisdictional issues, and privacy issues. However, there are more legal challenges in digital forensics such as the lack of standardized international legislation. According to Gal Shpantzer and Ted Ipsen (2002) [41], legal challenges in digital forensics encompass proof of scientific rigor, esoteric nature of digital evidence, volatility of evidence, lack of qualified personnel, jurisdiction and related issues, and finally admissibility of tools and techniques.

Jurisdictional Issues

The major query when it comes to jurisdictional issues is which law enforcement agency is in charge of the investigation of a cyber-attack, either it comes from across state lines or from abroad. As a result, there are cases, where “wars” are given between agencies, leading to inadequate interagency, as well as inadequate intra-agency cooperation. Situations such as competition for personnel, facilities, budgets and prestige are just examples that worsen relationships between agencies [41].

Privacy Issues

Privacy is a key aspect of legal challenges in digital forensics. As Sundar Narayanan (2015) mentions in his article in Forensics Magazine: "*evolving privacy and data protection regulations across geographies and maturing regulatory definitions/enforcements on such aspects may add to the complexity of gathering forensic evidence.*" [42].

Accessing data from a suspect's device can be considered to be a violation in certain countries as it can contain private and sensitive information. Moreover, investigators often come across and identify digital evidence related to a crime by accident. However, it can become a real challenge if this evidence is not allowed to be used in courts against

the attacker due to privacy issues. In ~~that~~ way investigators are constrained ~~to a point~~ and the whole investigation process can be affected.

2.6.3 Resource Challenges in Digital Forensics

Device diversity, volume of data, time taken to acquire and analyze forensics media and unspecialized tools are just some examples of the various resource challenges in digital forensics.

With the rapid evolution of technology, more and more types of digital devices capable of carrying digital evidence are emerging. PCs, laptops, PDAs, cell phones, smartphones, GPS devices, SD cards, USB sticks, game consoles and wearables are just some examples. Each of the aforementioned device type is capable of storing and processing an enormous variety of data. Depending on the scenario, a vast volume of data can be involved in a criminal case. ~~The way~~ the traditional forensics investigator must be in a position to examine and analyze every evidence of a criminal action, independently of its source, the same way a digital investigator must know how to handle every data ~~appeared~~ on any device anywhere on the planet. In case of ~~the volume of data~~ is large enough, the investigator has to go through all the collected data in order to gather evidence which is a very demanding and time consuming process. Since time is of essence, taking into consideration that data is time sensitive by its nature, it can become a highly limiting factor and another major challenge in the field of digital forensics. In volatile memory forensics, data is ephemeral as it is stored in the volatile memory where all users' activities are overwritten. Thus, investigators can retrieve and analyze only the recent information and data stored on the volatile memory. Under those circumstances, the forensics value of the data for the investigation is significantly reduced. When it comes to data collection from the various sources, it is crucial for the latter not to be damaged as it becomes worthless for the investigation. Additionally, extracting, preserving, processing and analyzing data is highly important for an investigator not only to ensure that none of the data is modified or missed during the investigation process but also to ensure its reliability and that the data is well secured.

Similarly to device diversity, there is a plethora of digital forensics tools both hardware and software. The major challenge regarding forensics tools is that they have to be established with the appropriate testing and validation. It is crucial to develop fully tested and validated forensics tools which require testing and validation all over again after an

update. Digital evidence reliability relies mainly on the digital forensics tool reliability being used in the investigation. For example, online disks been used to store data cause problems in an investigation hence making the latter an arduous process. As Kumari et al. (2016) states "*an effectively working forensics tool is required to execute the investigation in corresponding cybercrime branch*" [16]. Finally, it is critical for the investigators to understand both the forensics branch to which an attack belongs to and the capability of all forensics tools, in order to be able to use the proper combination of tools for an efficient investigation.



3 Computer Forensics

In this chapter we are diving into the science of computer forensics, its objectives and primary types, as well as the key elements, rules and principles that have to be followed at that field. Furthermore, process models of computer forensics and its main investigative procedures are displayed.

3.1 Introduction of Computer Forensics Science

At the beginning of the digital era and digital crime investigations, computer forensics and digital forensics were two interrelated concepts. ~~Howbeit~~, over the years digital forensics expanded in order to include all types of digital devices. As a result, computer forensics constitutes nowadays a branch of digital forensic science. Similar to all forms of forensics science, the discipline of computer forensics incorporates elements of law to computer science and applies investigation and analysis techniques to find and determine legal digital evidence from computer systems, networks, wireless communications, and digital storage mediums.

3.1.1 Computer Forensics Definition

Over the years many definitions were developed in order to explain accurately the term computer forensics. According to Caloyannides (2001) [43] computer forensics is defined as “*the collection of techniques and tools used to find evidence in a computer*”. A more complete and formal definition of computer forensics is provided by the CyberSecurity Institute as: “*The preservation, identification, extraction, interpretation, and documentation of computer evidence, to include the rules of evidence, legal processes, integrity of evidence, factual reporting of the information found, and providing expert opinion in a court of law or other legal and/or administrative proceeding as to what was found.*” [44].

In general, computer forensics is the application of investigation and analysis techniques and methods for gathering and preserving digital evidence from a particular computing device, preparing it for legal proceedings [45].

As implied by the definition, computer forensics focuses on unique and specialized methods and techniques for the extraction of digital evidence from specific platforms, namely computer systems [36].

Computer forensics is customarily used when a case involves issues relating to reconstruction of computer usage, examination of residual data, and authentication of data by technical analysis or explanation of technical features of data and computer usage. Additionally, it requires specialized expertise that goes beyond normal data collection and preservation techniques available to end-users or system support personnel.

3.1.2 Objectives of Computer Forensics

As aforementioned, the objective of computer forensics is not restricted in explaining the current state of a digital artifact but is broaden in order to include the examination of computer based systems, extracting this way digital evidence in a forensically sound manner. The primary goal of computer forensics is the performance of a structured computer investigation of digital evidence while maintaining a documented chain of the evidence. An additional aim is the identification of the digital evidence in a short amount of time since by its nature digital evidence is fragile and time sensitive. Computer forensics also aids in uncovering what happened on a computing device and who was responsible for it so as to solve a crime and provide evidence to support a case. ~~To make the long story short,~~ the overall objective of computer forensics is the detection of any computer related incident and crime, estimating the potential impact of the malicious activity on the victim, assessing the intentions and identity of the intruder, arresting and prosecuting the perpetrator in a court of law.

3.1.3 Primary Types of Computer Forensics Investigations

Computer forensics investigations are divided into *two* primary types. The first type deals with investigations where computer based systems are used as means to commit a criminal action or are involved in any other type of misuse and they are examined and analyzed.

The second type involves the use of computer systems as the target of the performed crime such as identity theft. The application of computer forensics methods and techniques to identify and assess this type of criminal activities is referred as incident response. This type of investigation includes the seizure and examination of volatile data,

such as information stored in RAM memory, pertaining running processes and network connections.

Because of the existence of the two types of computer forensics investigations the definition and interpretation of the term computer forensics can be altered. This change in the definition relies on the fact that each case requires different approaches, methods and techniques for conducting the investigation process. However, regardless of the situation, the methodologies, techniques and procedures used to be carried on a computer investigation are very similar and must conform to the general principles. Taking a step further, methodologies, techniques and procedures have to be standardized, sound and proven in order to guarantee to the courts the reliability and admissibility of the digital evidence that has been found previously[44].

3.1.4 Applications and Uses of Computer Forensics

As already mentioned computer based systems are used either as the tool to commit a crime or as the target of the committed crime. Since the use of computer based systems is widespread, nowadays, cybercrime occurs with a relative ease when information technology is used to commit or conceal an offence. The variety of the different cybercrime types is vast and involves criminal, domestic, security, internal and marketing affairs. Cybercrimes of many sectors are encompassed in the following list:

- Breach of computer security and hacker system penetrations (both external and internal attacks)
- Release, distribution and execution of malicious viruses and worms
- Damage of company service networks
- Financial frauds and theft, such as sales fraud, investment fraud and electronic fund transfer fraud
- Theft of intellectual property and copyright violations
- Industrial espionage and unauthorized disclosure of corporate information and data
- Computer based criminal fraud and deception cases
- Forgeries
- Inappropriate email and internet abuse in the work place from employees
- Regulatory compliance

- Phishing, data theft and unauthorized access of personal information, files and emails
- Credit card cloning and theft of bank account numbers
- Identity threats and theft
- Cyber stalking, harassment and sexual assault
- Cyber terrorism
- Child Porn
- More general criminal cases such as burglary, obscenity, homicide, suicide and narcotics investigations

The list can go on, since there are actually countless cybercrimes nowadays. Because of the exponential increase of the number of the cybercrimes and litigations, the need for computer forensics has emerged and became apparent and vital. It is widely known that criminal activity has no border, as it can be performed from anywhere in the world. Computer based systems often constitute a crime scene, withholding incriminating information from hacking attacks, emails, internet history, documents or other files relevant to more serious crimes such as murder, kidnap, fraud and drug trafficking. A computer forensics investigation can reveal all the hidden information that is of interest. The revealed digital evidence by computer forensics can be used in many types of criminal and civil proceedings. According to Judd Robbins [46] the revealed evidence can be used by:

- *Criminal Prosecutors*, who rely on digital evidence obtained from computers in order to prosecute suspects. The evidence can be used in a plethora of criminal cases where computer based systems are engaged, such as homicides, financial fraud and theft, and child pornography.
- *Civil Litigators*, who make use of personal and business data discovered on a computer in fraud, divorce, harassment, discrimination cases and much more.
- *Insurance Companies*, who use evidence discovered on computer to mollify costs (fraud cases, workers' compensation cases, arson cases, etc).
- *Private Corporations*, who hire computer forensics specialists to obtain evidence from employees' computers which can be used in harassment, fraud, embezzlement cases, and theft or misappropriation of trade secrets, as well as to obtain other internal/confidential information.

- *Law Enforcement Officials*, who rely on computer forensics to backup pre-search warrants and post-seizure handling of computer equipment.
- *Individuals*, who may obtain the services of professional computer forensics specialists to support claims of harassment, abuse, or wrongful termination from employment.

According to Hailey (2002) computer forensics are ~~primary~~ applied in three areas [44], public sector, private sector, and consulting.

Public Sector

Public sector encompasses the use of computer forensics by law enforcement agencies in order to investigate and prosecute crimes. Computers can be used in traditional crimes (e.g. homicide, rape, financial fraud and theft) and in cybercrimes (e.g. cyber terrorism, hacking and theft of personal data). Computers can be the target of the occurred crime, the ~~mean~~ to execute a criminal action or even as incidental to a crime. Law enforcement agencies have often been at the forefront of developments in the field of computer forensics, since they have been among the earliest and heaviest users of computer forensics.

Private Sector

Computer forensics are also applied in the private sector to aid in the investigation of improper use of computing resources by employees, embezzlement, improper use of company assets, and theft of trade secrets among others. More recently, commercial organizations have used computer forensics for their benefit in a variety of cases, such as intellectual property theft, industrial espionage, inappropriate email and internet use in the work place, employment disputes and much more.

Consulting

Computer forensics is also applied by experts in order to consult individuals or law firms. Even though this type of investigation is often thought to come under the private sector, Hailey (2002) states [44] that this area of computer forensics requires its own category because of the uniqueness of the performed work for these investigations.

3.2 The Generic Computer Forensics Process and Investigation Model

~~Here are presented~~ the Generic Computer Forensics Process along with the Investigation model, both of which are used separately in case of a computer forensics investigation.

3.2.1 The Generic Computer Forensics Process

Breaking down the definition of computer forensics, several technical aspects of the actual science of computer forensics have ~~been~~ emerged. The more in-depth definition includes the preservation of media and data, identification of computer-related evidence, extraction of the data and interpretation of the results. According to Krishnun Sansurooah (2006) [47] there is a fundamental and standardized process for examiners to conduct computer forensics investigations, regardless of the nature and variety of data sources and the approaches required from each of them. The generic computer forensics investigation process includes five phases: identification phase, acquisition of evidence, authentication of evidence, analysis phase and presentation phase.

The *identification phase* deals with developing the right strategy to be proceeded for the investigation. During this phase it is essential to identify what information is needed, what are the possible data sources that contain computer related evidence such as hard drives, floppy disks, log files etc., and what premeasure and acquisition actions are required to gather the data in the right order. In short, the identification phase is about ~~in-~~~~telligent~~ gathering, with the aim of predicting the forthcoming challenges and difficulties that have to be faced during the next phases, providing the appropriate solutions.

The next phase is the *acquisition of evidence*, where the developed strategy during the identification phase is applied. The main objective of this phase is the acquisition of the data and information, relevant to investigation at stake, in a forensically sound manner. At this phase, a standard set of procedures is required in order ~~the~~ integrity of the evidence to be maintained, ensuring that it is not contaminated either intentionally or accidentally. Preserving the original media and data is of essence when performing a computer forensics investigation. The first step is to isolate the device in question and make a forensic digital copy of the original device's stored media. Once the copies of the original media are obtained, the original media and data are secured in order ~~the~~ their safety to

be guaranteed, maintaining their pristine condition. The whole investigation is conducted on the digital copied media instead of the original.

Investigators use a variety of techniques for the extraction and acquisition of digital evidence, which includes snapshots and live datasets. All snapshot data has to be seized or forensically imaged, while live data has to be acquired. All the extracted digital evidence must be obtained, following proper and validated forensics proceedings, in order the chain of custody to be maintained, ensuring the reliability and admissibility of the evidence in a court of law. Of equal importance is the use of proprietary, specialized and sophisticated technological tools, applications and software programs in order the copied media to be examined, searching for hidden folders and unallocated disk space for copies of deleted, encrypted, or damaged files. As Krishnun Sansurooah (2006) [47] clearly states during this phase the following factors play a vital role, thus they have to be taken under serious consideration:

- Environmental Assessment and Documentation
- Drive Assessment and Documentation
- Evidence and Anti-Tampering Tagging and Documentation
- Drive Removal and Imaging Documentation
- Hardware and Software Tools Documentation
- Procedural Documentation.

During the *authentication of evidence* phase the collected evidence must be well documented in order to ensure that it has not been tampered or altered. The acquired evidence must prove its authenticity, so the documentation of the complete process is required. Evidence should be handled properly so as to ensure that it was left behind by the perpetrator. At this phase, techniques are used to timestamp and demonstrate the existence of evidence at a specific moment. Additionally, cryptographic techniques that calculate a value that functions as a sort of electronic fingerprint for an individual file or even for an entire hard disk are used so as to prove that the acquisition of evidence is identical to the original source by comparing the hash values of both the image and the original source. This phase is significant for the whole investigation, since if evidence is not authenticated the whole investigation process is compromised. If the reliability and authenticity of evidence is questioned then evidence may be rejected for use in the courts.

The *analysis phase* deals with the proper interpretation of the acquired evidence. During this phase the acquired data is actually turned into reliable evidence. The analysis phase is one of the most important elements of the computer forensics investigation, since at this phase conclusions, regarding evidence that emerged, are presented. Placing evidence in a logical and useful format (e.g. how did it get there, what does it mean, where did it come from) is of essence, as misinterpretation of the obtained evidence may blow the whole case. During the analysis phase the significance of the evidence to the case under consideration is determined. Techniques used in this phase include timeframe, data hiding, application and file ownership and possession. In the analysis phase a review of the request for service, legal authority for the search of the digital evidence, investigative leads, and/or analytical leads are also needed.

The last phase is the *presentation phase* which involves the creation of the final report regarding the whole investigation process including the obtained digital evidence. Findings and results of each phase have to be completely documented in an accurate and comprehensive manner. In brief, this report is a detailed documentation of each action taken during every single phase, which forensics tools were used, notes of the physical and digital evidence that were found, a step by step explanation of the investigation process and interpretation of the results. The produced report should be written for the intended audience and is often used as a guide for third party examiners to not only reproduce and validate every piece of evidence but also to re-perform the whole investigation process.

Even though the process framework indicates five successively phases, there are cases that backwards steps are necessary, such as when during the acquisition phase the acquisition plan should be reconsidered in order to include more data sources or when during the analysis phase the identification of references to data sources, that have not been yet acquired, emerges.

3.2.2 The Generic Computer Forensic Investigation Model

Yunus, Roslan & Zainuddin (2011) [34] studied and investigated the already developed digital forensics process models and extracted the basic and common phases among all models. Based on the previously developed models Yunus et al. [34] proposed in 2011 the Generic Computer Forensic Investigation Model (GCFIM) which is illustrated in the following figure:

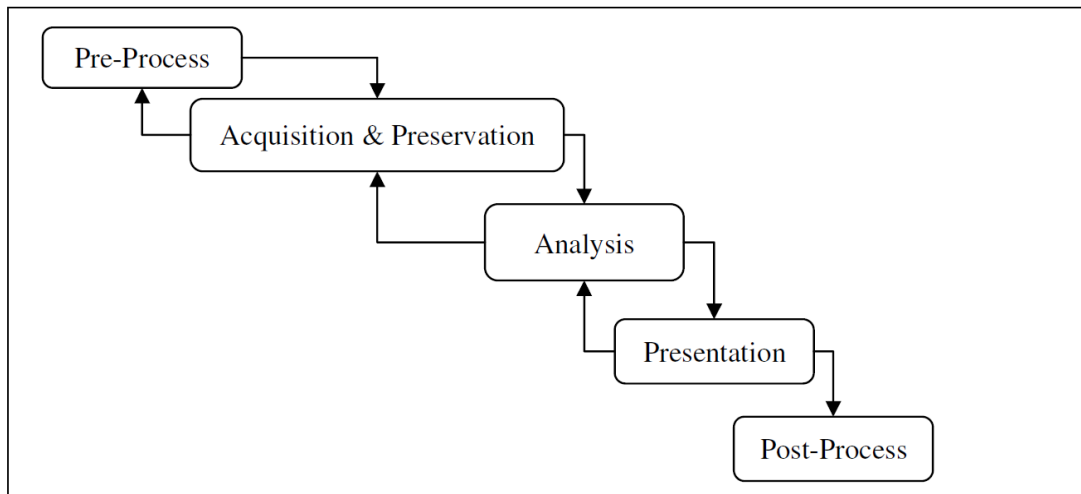


Figure 7: The Generic Computer Forensic Investigation Model (GCFIM)

As depicted in the figure above, the GCFIM is composed of five phases: pre-process, acquisition and preservation, analysis, presentation and post-process.

The *pre-process* phase includes all the required tasks that have to be ~~done~~ before the actual investigation process begins. Obtaining approval from the relevant authority, ~~de-~~velop the right strategy for the investigation, selecting the appropriate tools, making sure that they are geared up are some of the essential actions.

The ~~directly~~ next phase is the *acquisition and preservation* phase, where the official collection of data begins. During this phase, data is identified, acquired, collected, transported, stored and preserved. The relevant data and information that occur from this phase constitute the vital material for the next one phase.

The following phase is the *analysis phase*. It is the most extended phase and encompasses numerous sub-phases, techniques and methodologies for the right interpretation of the acquired data in order to set the timeline of the occurred crime, identify the source of crime and discover the identity and intentions of the perpetrator.

After the analysis phase occurs the *presentation phase*. At this phase, all findings derived from the analysis phase are ~~well~~ documented and presented to the authority. Evidence must be reliable, adequate and admissible in courts and furthermore presented in such a way that is well understood by the interested parties.

The last phase is the *post-process phase*, which sets the closure of the investigation process. Evidence, both physical and digital, is returned to the proper owner or stored in a safe place, if necessary. Additionally, this phase includes a review of the whole investi-

gation process in order to highlight vulnerabilities and difficulties of the investigation process and spot all the developments and improvements that need to be made.

The main advantage of this model is that it provides the ability to move backwards and is not restricted to moving only sequentially from one phase to another. Hence, correcting weaknesses and acquiring new information and data becomes a much easier task. After all, Yunus et al. (2011) [34] aimed in developing an efficient, robust and high level investigation model that can serve in any computer forensic investigation, aiding in any future development of new computer forensic investigation model.

3.3 Main Investigative Procedures

Gathering of digital evidence is crucial during computer forensics examination. There are multiple methods of discovering data on a computer system, such as recovering deleted, encrypted, or damaged file information, monitoring live activity and detecting violations of corporate policy. Collected information assists in arrests, prosecutions, termination of employment and preventions of future illegal activity.

According to Sansuroah (2006) [47] there is a fundamental and inherent methodology for computer forensics investigations, which can be outlined in the following *three* sentences:

1. Acquire the evidence without altering or damaging the source
2. Authenticate that ~~you recovered evidence in~~ the same as in the seized source
3. Analyze the data without altering it.

This methodology is known as *the three A's* and its enhancement and refinement are of essence, since technology evolves rapidly. Otherwise, the methodology will be rejected as outdated.

According to J. Mitchell [48] the main investigative procedures that are followed in every computer forensics investigation are the following:

Shut Down the Computer

This process includes unplugging the source power of the computer or terminating a computer network. The computer's power source is usually a direct source of power such as a wall outlet, power strip, or UPS (Uninterruptible Power Supply) if ~~someone is present~~. It is advisable to avoid shutting down a computer though the normal shutdown process, since depending on the operation system, it is likely cleanup procedures ~~to~~ exist

for deleting data working on the background or even password protected screen savers that can emerge at any time. When terminating a computer network, it is of essence to use the appropriate commands which are needed by the involved network. At any case, it would be ideal to record the shutdown procedure in real-time, using a digital camcorder. All the above factors contribute in complicating the process of shutting down the computer. By and large, the process of shutting down the computer system must be performed as fast as possible since time is of essence.

Document the Hardware Configuration of the System

As envisaged the computer system will be moved to a protected environment, where the processing and analysis of evidence will be carried out maintaining the proper chain of evidence. Before proceeding with computer disassembling, it is essential to document the computer's system hardware configuration, i.e. the computer system itself, its ~~surroundings~~ components and all the connections between them. Things to be documented include the computer's state (whether it was on or not, running programs on the system), manufacturer, vendor, model and serial number, as well as the surrounding environment which involves all attachments to the computer, such as external hard drives, speakers, cable modems, USB or network hubs, wireless network routers, and so on. Additionally, it is of equal importance to label each cable and wire in order to mark all connections, so as, at a latter point, to reestablish the system configuration to its original condition at the secured environment of the lab. Photographs and videos, of the computer's system hardware configuration, that can cover all aspects and angles, constitute good supplements to the handwritten notes.

Transport the Computer System to a Secure Location

It is fundamental to transport and store the seized computer hardware equipment to a controlled and secure location. When referring to a secure location, ~~it is meant~~ the lab environment where the evidence processing takes place. It is crucial to move the confiscated computer equipment to safe and well-equipped laboratories in order ~~the acquired evidence to be ensured that~~ has not been altered, damaged or even planted. Maintaining an appropriate chain of evidence ensures that any collected evidence is admissible. ~~Worth mentioning is the fact that~~ it is critical to treat the whole computer system, under investigation, as evidence.

Make Bit Stream Backups of Hard Disks and Other Media

It is ~~basic~~ to make bit stream images of all storage space, which include hard drives, disks, USBs and so on, before anything else. The processing of digital evidence is per-

formed on the copies of the bit stream backups as opposed to the original evidence. Hence, the latter is preserved and protected which is imperatively critical. Original evidence must be left untouched, since digital evidence is by its nature fragile and any alteration or destruction of data is usually irreversible. In any serious computer forensics investigation, the utilization of bit stream backups of all hard disk drives and other media is of vital importance.

Mathematically Authenticate Data on All Storage Devices

Investigators must be able to prove that they did not modify at any way the evidence since the time the computer system has come into their ownership. Computer forensics tools are utilized to verify that evidence has not been altered or modified. Such tools use a hash algorithm to generate a numeric expression and compare this to the same hash algorithm on the data that was backed up. This is used as proof that the file has not been changed. Present day software is capable of authenticating data using a 128-bit level of precision. Such a wide key provides a decent level of assurance that the data has not been in any way manipulated or adjusted. This process is highly required since it empowers investigators to refute allegations of original evidence modification by producing the necessary proof and verification.

Identify File, Program and Storage Anomalies

Information and data can be stored in binary format in encrypted, compressed and graphic files. Hence, text data stored in these file formats cannot be identified by a text search program. Such files require manual evaluation whilst when it comes to encrypted files a much more demanding and complicated process is involved. Additionally, the computer system and hard drives must be examined for potential hidden and/or formatted partitions which may include a vast amount of data and potential evidence. Furthermore, it is important to examine and evaluate all files that are contained on the Recycle Bin. Since all files were intentionally deleted, their selection for deletion emerges suspicions, regarding the relevance of the date to the case. All involved issues and relevant files that are found must be documented in detail.

Document the System Date and Time

It is vital to document the system's date and time in order to correlate events between two computers, or between the activities of a user and the time associated with particular files on the computer. Even though the accuracy of dates and time is essential, there are many obstacles in accomplishing it. Users can set intentionally the computer's date and time incorrectly or even utilize powerful programs and applications to change the

timestamp of files. Additionally, the system itself may timestamp incorrectly files or even slow one hour the system's clock due to daylight-saving time.

To overcome the aforementioned obstacles and inaccuracies the system's date and time must be captured and compared with a reliable time source (such as one synchronized with an atomic clock) once the computer system is seized. Date and time settings must be documented, and any discrepancies have to be noted.

Depending on the utilized computer forensics tool, it may be necessary to boot under controlled conditions the computer system under investigation. This is accomplished by using a USB or disc which contains a controlled version of an operating system to boot the computer. Since it can be alleged that any change in the computer's configuration can alter the content of the data, the need of bit stream image before even starting any other process is urgent and furthermore, the whole process must be recorded using digital camcorder.

Prepare a List of Key Search Words

Bearing in mind that hard drives have an increased storage capacity nowadays, reviewing and evaluating all files on a computer's hard drive by a computer forensics investigator is not possible by far. To assist the situation the involvement of automated forensics text search tools is required.

Firstly, investigators gather information, regarding the case under investigation, which is derived from allegations, computer users and any individual involved or is familiar of the case. After that, they generate a list of key words that pertain to the investigation. Finally, investigators search the contents of the hard drives for any incriminating evidence, utilizing the appropriate tools and the compiled list.

This process is very important since it can distinguish relevant from irrelevant data and even more helps in narrowing down some of the pertinent data. However, the list should be kept as short as possible and words that are common or constitute part of others must be avoided.

Examine the Windows Swap File

Systems have files that are used to cache information between memory and the hard drive. These files are known as swap files and may contain valuable data of evidence. Nowadays, with the use of sophisticated and automated tools, this process is performed with ease and takes only a few minutes. However, in the past, hex editors were utilized to perform this task and the process was humdrum, exhausting and time consuming. The

default settings of the swap file create it dynamically during computer's operation which means that when shutting down the computer the swap file is instantly erased. Recovering the erased swap file can be performed in the same way as any other erased file, however ~~it is not ensured~~ the recovery of the whole file.

Evaluate File Slack

File slack is the area of space in an occupied cluster that spans the point where file data ends to the end of the cluster. File slack occurs during the work session as files are closed and is beyond the reach or the view of the computer user. Usually, common users are unaware of this storage area, however, sophisticated users can utilize it to hide data. Furthermore, deleted data may remain in file slack. Hence, file slack can be the origin of vital evidence regarding the investigation. This source of data can provide relevant key words which should be added to the aforementioned search list. Taking under consideration the nature of the file slack, the use of specialized and automated computer forensics tools is required for its view and evaluation.

Evaluate Erased Files

Operating systems do not erase files completely. The allocated space just becomes available to be overwritten ~~with~~ new files. However, not all storage space associated with such files becomes unallocated and available. Hence, valuable data can be retrieved and restored from unallocated space often by using the operating system's undelete program. Similar to the Windows swap file and file slack, relevant key words can occur from the examination of the unallocated space and in this case the appropriate forensic tools are required.

Identify Email and Internet Storage Areas

Computers are widely used for accessing the Internet and communicating via emails. Thus, Internet folders, favorites, temporary Internet files, 'cookie' repository and email folders must be examined thoroughly since a vast amount of information and data of evidence can be revealed through this process.

Search All Areas for Key Words

The previously produced list is utilized with the appropriate text tool to search all relevant computer hard drives and other media. The output results of such tools should be reviewed and inspected, in order to identify pertinent information, data and evidence. All findings should be properly documented and examined so as to identify additional key words. When additional keys ~~are~~ emerged they must be added in the existing list and proceed with ~~the conduction of~~ a new search using the newly formed list.

Stenographic Awareness

Data and information can be also hidden in images using a process called steganography, which protects the data from being viewed or accessed with a specialized key. Since the only way to identify the use of steganography is the existence of such a stenographic application, investigators ought to be familiar with the names of such programs. However, this is not always enough since applications' names can be renamed to something innocuous with ease and furthermore, in some cases such programs can be identified only during their operation.

Document File Names, Dates and Times

Regarding evidence, information which includes file names, dates and time of creation, modification and deletion of files can be proved to be both pertinent and valuable. Hence, documenting everything, including allocated and 'erased files', with accuracy is essential.

Document the Findings

From the beginning notification of possible illegal activity to the very end of the investigation it is essential to document all findings and identified issues in every procedure. Additionally, a detailed documentation of the used computer forensics tools, including version numbers of the software used, serial numbers, manufacturers and so on is also important.

Retain Copies of Software Used

Retaining copies of the used software along with its output constitutes part of the documentation process. This process is essential to eliminate which version of the software was utilized during the investigation and is usually done on an external storage device such as an external hard disk. Before or even during trial it is often required to duplicate the forensics processing results. Taking under consideration that technology is rapidly evolved, computer forensics software tools are routinely upgraded. On the contrary, it can take years for a case to go to trial. Hence, if the original version of the used software is not retained, and the software is being upgraded the duplication of the results is an impossible mission.

3.4 Key Elements, Rules and Principles of Computer Forensics

Below, we are going to present the key elements, the computer forensics rules and the principles that have to be applied in order an investigation to be conducted efficiently.

3.4.1 Key Elements of Computer Forensics

In any kind of forensics investigation there are three key elements: the involved material, its relevance to the case under investigation and the validity of the conclusions drawn by the forensics examiners. Bearing in mind that computer forensics is similar to the traditional forensics science, these key elements are applied here too.

The Involved Material

As mentioned previously computers may be either the tool to commit a crime or the target of a criminal activity. In both cases, the involved material in a computer forensics investigation encompasses both physical and digital material. Physical material investigation is closer to the traditional forensics science where files, envelopes, boxes, weapons, etc. are examined in order to discover evidence. When it comes to computer forensics, physical material is referred to computer based systems such as laptops.

Since computer forensics focuses on computer based systems digital material encompasses information and data that can be extracted from computers and other device components that are part of the system, such as network nodes, printers, scanners etc. The amount of digital information that can be found in such systems is vast and includes various types of electronic data. The form of the data under investigation also varies and fluctuates from data that actually exist in hard copy such as e-mail text and headers, browser information and website log files, to deleted documents that have to be recovered and reconstructed, or even password protected or encrypted data.

Relevance

However, identifying the involved material is far from being enough for the investigation of a criminal case. Of equal and vital importance is ensuring the relevance of the material to the case at stake. The type of the case under investigation, the requesting agency and the nature of the request are the leading factors that define the pertinence of the material.

The requesting agency involves everyone who is in need of computer forensics and makes use of the recovered evidence, encompassing the victims, courts, law enforcement agencies, the government, private and insurance companies and even private individuals.

For ensuring the relevance of the involved material and the evidence that occurs after analysis, the first step that investigators have to take is making a thorough assessment of the situation. Identifying the nature, the environment and the requirements of the case

provides the context of the investigation and constitutes the stepping stone for its further proceeding. Having established specifics and details about the case, the purpose and the focus of the investigation process are defined. Investigators have to decide on what they have to work with, such as technical policies, permissions, and device logs and what is required to be monitored such as e-mails and chat rooms. Hence, they become aware of the types of evidence they are looking for (both physical and digital materials, and data that needs to be collected, examined and analyzed).

Validity

The relevance of the evidence is associated inseparably with their validity which depends mainly on the process of authentication of data. When dealing with evidence it is critical to extract data taking any possible precaution. Furthermore, it is of essence to define and maintain the chain of custody which includes packaging, storage, and transportation of data. Handling the evidence in a forensically sound manner is crucial in order its integrity to be maintained and guaranteed. Evidence is used to support or contradict a case, establish and reconstruct the facts in a criminal case. Hence, it is highly important to ensure the validity, reliability and admissibility of the evidence, since any kind of evidence contamination can result in the rejection of the whole case.

There are *four* primary questions that have to be answered in order the validity of the evidence to be certified:

- How was the evidence extracted and by whom?
- Who packaged the evidence?
- How and where was stored the evidence and by whom?
- Who transported the evidence?

Being able to provide all the answers to the above questions is a significant aspect of computer forensics, which involves *three* key factors: information about the case, the involved equipment and the emerged evidence.

Information about the case encompasses the case number (i.e. the assigned number to the case in order to be identified in a unique way), the nature of the case (basically a short description of the case, its specifications and requirements) and of course information about the involved investigator (including the investigator's name, expertise, connection to the company if any, etc.).

The involved equipment refers to the computer forensics tools used during the investigation process. The used tools can be both hardware and software and have to be validated. The required information includes the tool's serial number, manufacturer, vendor, model and even a small description of its operation.

Information regarding the derived evidence is often filled using the chain-of-evidence form and involves information about the person who recorded it, the location where the evidence was recorded and the exact date and time that the recording took place [49].

3.4.2 Computer Forensics Rules

Considering that the eventual outcome of the forensics process is liable to legal examination, it is fundamental that the rules, which were administered it, have been pursued. Although these standards are sufficiently general and adequately broad so as to be applied to any computer forensics investigation process, it is crucial to ~~obey to~~ them with fidelity in order the acceptability of the evidence to be secured and ensured n an official courtroom. Since the utilized strategy and methodology in association with the diverse techniques, procedures and frameworks are established by the individual forensics expert, the genuine ~~picked~~ process ought to be employed in a way that the pertinent rules are not compromised. The imperative rules that have to be applied in any kind of computer investigations are the following:

Minimal Handling of the Original

This rule can be marked as the most significant one in computer forensics science. When conducting a computer forensics investigation the utilized processes must be kept to the slightest possible level. Minimizing the probability of tampering with the original evidence is the primary aim in any computer forensics investigation. This is accomplished by replicating the original and authentic data and conducting the whole investigation process using the duplicate data, under the condition that this is possible. The duplicated data must be an explicit and identical ~~proliferation~~ of the initial data, and additionally, must be validated and approved, otherwise inquiries regarding the reliability and the integrity of the evidence can be raised.

Account for Any Change

As already mentioned, it is essential in any investigation to preserve the state of the original data and information, avoiding any alteration of them. However, under certain circumstances alterations to evidence, either on the original or the duplicated, may be unavoidable. For example, booting up or shutting down a computer system can result in

changes to the memory and temporary files. While the need to modify data happens occasionally, there are situations where the investigator is required to initiate changes in order to facilitate the forensics examination process. For instance, where access to data and information is limited by access control methods and techniques, the investigator might be compelled to change the data (either accessing a bit or even a whole string of binary data) in order to ease the access.

Where changes of data do occur, the nature, extent and reason for the change must be documented in an appropriate, accurate and detailed way. In such cases, investigators have the sole responsibility of correctly identifying and initiating the necessary changes. Investigators must be able to fully comprehend the nature of any change they have done and define perfectly its extension. Additionally, another obligation of the investigators is the correct and appropriate documentation of the occurred changes, which includes detailed information explaining the necessity of the change, its nature and extent, so as the integrity of the evidence to be ensured. The outcome of this process depends highly on the insight and expertise of the examiner, and constitutes a significant factor when the evidence is presenting in courts.

Even when the evidence is sound, its credibility along with the reliability of the applied process might be affected by the emerged questions regarding the examiner's skills and knowledge. Under sufficient doubt, the derived results of a computer forensics process can, in the most pessimistic scenario, be ruled as inadmissible and hence be rejected from the judicial procedures.

Comply with the Rules of Evidence

One of the cardinal rules in computer forensics is the imperative need to comply with the rules of evidence. Methodologies and techniques for handling and investigating evidence must be utilized and developed in compliance with the pertinent rules of evidence in order the admissibility of the collected data to be ensured and be used in a court of law. The same applies for the use and development of the computer forensics tools.

Another significant element, when conforming to the rules of evidence, is that the presentation of the collected evidence should be carried out carefully in order the notion of the evidence to be kept intact. Basically, this means that the introduced evidence and information are presented in a way that is as indicative to the original as possible.

Do Not Exceed Your Knowledge

Computer forensics experts and examiners should not proceed with an investigation if it requires knowledge and skills that exceed the investigator's level. It is vital for the investigators to be able to realize and recognize the boundaries and restraints of their insight and expertise.

When reaching their limits, investigators ought to seek alternatives in order to conduct properly a computer forensics investigation process. In such situations, it is wise to terminate any further examination and look for assistance from more specialized and experienced investigators. Another option, available to the investigators, is to direct the imperative research and obtain additional training in order to enhance their knowledge and develop more skills, so they can be able to proceed with the investigation on progress. This particular option is not always permitted, since an investigation shall conform to the applied time limitations. Proceeding with the investigation with the expectation that everything will go as expected is a significantly perilous choice and must be avoided.

The primary danger in proceeding with an investigation that is beyond the investigator's expertise relies on the fact there is an increased risk of harm and alterations that the inspector doesn't know about or does not comprehend and therefore may overlook. As a result, it is recommended not to continue with the examination, since it is imminent that the outcome of the case will be compromised. In any kind of computer forensics investigation it is vital for the inspector to be able to describe accurately the applied methodologies, techniques and processes. Inability to provide the necessary clarification of the utilized procedures may often raise questions, regarding the expertise and credibility of the investigator in any consequent legal proceeding. Summarizing, each computer forensics investigation must be conducted by the appropriate forensics expert who is properly qualified and has the required knowledge and skills [50].

3.4.3 Principles of Computer-Based Electronic Evidence

For evidence to be admissible it must be solid and not biased, implying that at all phases of a computer forensics investigation suitability ought to be at the forefront of the examiner's mind. A universally applied and respected set of rules which can direct the examiner in this area is the Association of Chief Police Officers Good Practice Guide for Digital Evidence (2011) [51], or ACPO Guide in short. In spite of the fact that the ACPO Guide is aimed at United Kingdom law enforcement, its primary principles are im-

perative to all computer forensics. The four main principles of computer-based electronic evidence as retrieved from this guide are as follows:

Principle 1: “No action taken by law enforcement agencies, persons employed within those agencies or their agents should change data which may subsequently be relied upon in court.”

Principle 2: “In circumstances where a person finds it necessary to access original data, that person must be competent to do so and be able to give evidence explaining the relevance and the implications of his/her actions.”

Principle 3: “An audit trail or other record of all processes applied to digital evidence should be created and preserved. An independent third party should be able to examine those processes and achieve the same result.”

Principle 4: “The person in charge of the investigation has overall responsibility for ensuring that the law and these principles are adhered to.”

Computer-based electronic evidence is liable to the same principles and laws that apply to narrative evidence, the tenet of which might be clarified along these lines: the onus is on the indictment to show to the court that the delivered evidence is no more and no less at this point than when it was first taken into the possession of law enforcement agencies. Operating systems, programs and applications often alter, add and delete contents of the electronically stored data. Users are not always aware of such changes, since the latter can happen automatically by the system itself. Keeping in mind the end goal is to conform to the standards of digital evidence, wherever practicable, an image ought to be made of the whole target device. In that way ~~is ensured~~ the preservation of the original data and third party re-examinations can be empowered achieving the same results. Examiners should be mindful, so as to guarantee that all evidence of probable value is identified and captured.

When data is stored at a remote location rather than locally, it is not always possible to obtain an image. Thus, it becomes imperative to directly access and recover the data in order to come in possession of the original. Bearing that in mind, the person who will initiate data access must be qualified to conduct the retrieval process of the data, and at a later point be able to present the evidence in an appropriate way in a court of law. An important factor that has to be taken under serious consideration is referred to legislation and jurisdiction matters that are applied to the case.

It must be noted that the applied principles do not imply nor exclude approaches and methodologies for conducting a computer forensics investigation efficiently. When deciding on the approach, methodologies and processes that have to be utilized in an investigation, an assessment of the case regarding its objectives and scope is necessary taking into consideration the available intelligence sources and investigative assets. This assessment is undertaken based on both technical and non-technical factors, the process must be transparent and all decisions must be reasonable and ~~rationale~~ recorded.

4 Network Forensics

In this chapter, we are going to ~~be inserted into~~ the details of the Network Forensics Science along with its difference from network security and computer forensics, its classification types, its generic process model, as well as challenges and trends that this field ~~is faced~~ nowadays. Furthermore, there is a paragraph that encompasses a basic background about the networks generally. Then, some rules are presented about how to handle the evidence gathered from the network forensics process, giving emphasis on the packet sniffing acquisition. Finally, we are going to present packet sniffing components and process, packet analyzers and how they work, and some ~~famous~~ network sniffing and packet analyzing tools, some of which we are going to use in the next chapter.

4.1 Introduction to Network Forensics Science

The need for Network Forensics is increasing, as more and more enterprises want to know who, what, when, why, where and how their services were being accessed and used [52]. Network forensics is not a way of preventing the occurrence of an attack, but it can reduce the impact by supplying analysis so that companies can respond to the ~~infection~~ faster. It decreases and simplifies the monitoring, reporting, analysis and remediation time required to defend against attacks. It aids prosecution through forensically complete evidence and supplies explanation about the root that causes the breach of security to provide quick, clever and successful response, preventing from harmful events and ongoing risks [35].

Network Forensics VS Network Security

Network forensics seems to have a lot of similarities with network security, however their contents are very different. Network forensics is a science that deals with ~~capture, record~~ and analysis of network traffic, ~~data of which~~ is captured using packet sniffers and alerts, while logs are gathered from network security tools that have been already installed. This data is examined to trace back the perpetrators and analyzed to estimate the attack characterization. Some deficiencies may appear in security products, which can be used to direct growth and development of these tools [35].

On the other hand, network security utilizes defensive mechanisms, like firewalls and intrusion detection systems (IDSs), the first of which is used for prevention whereas the second one for detection. Utilizing these mechanisms, someone can find out network vulnerabilities and block any malicious communications from outside. Firewalls manage traffic that enters and leaves a network based on source and destination addresses and port numbers. It filters malicious network traffic based on the firewall rules-signatures, which are nowadays difficult to be updated because more and more vulnerabilities always appear with their own signatures [35].

Intrusion detection systems (IDSs) are used for learning, detecting and reporting attacks, as they occur in real time [53]. IDSs contain two types: the first one, called *signature-based (misuse)* detection, uses specifically known patterns called signatures to detect malicious code, the main advantage of which is their ability to detect known attacks and the relatively low false alarm rate when rules are correctly defined, but it cannot detect new, unknown attacks. The second one, called *statistical-based (anomaly)* detection, does activity monitoring and detects abnormal behavior in the system. In contrast to misuse detection system, it can detect new unknown attacks, but has very high rate of false alarms, which leads to poor accuracy of such a system [35].

The network forensics process gathers all the evidence that is needed for incident response and investigation of the crime, whereas network security keeps system safe against attacks. Network security tools monitor the network in real time for any possible abnormal behavior without stopping, whereas network forensics plays its role only after a crime is noticed and its response is specific for each case, as crime scenario is different each time. There are cases that crimes do not breach network security policies but might be against the law and they can be only managed by network forensics [54].

Table 1: Comparison of network security and network forensics

Network security	Network forensics
System protection against attack	No system protection against attack
Usually in real time	Postmortem
Generalized – looking for any possible harmful behaviors	Case restricted – want to reconstruct the criminal scenario
Keep alert 24 h every day	After crime notification – notitia criminis
Continuous process	Time-bound process
Established field of computer science	Very immature and young science

Network forensics can be defined as the science which deals with ~~the~~ discovering and recovering valuable information in a networked environment about crime cases in a way that is legally accepted [55]. Such crimes can contain instances related to the security of a home-local network, corporate espionage, child pornography, traditional crime involving computer and network technology, employee monitoring or medical records, where privacy is considered as an asset of a great value [35].

Network Forensics VS Computer Forensics

Network forensics is considered as the next step of computer forensics, which was first presented by law enforcement and has a lot of directing concepts from the examined methodology of judicial system [56]. Computer forensics contains preservation, identification, extraction, documentation and interpretation of computer data, whereas network forensics developed as a response to the hacker community and contains capture, recording and analysis of network events so the examiners can find the source of attacks [35]. In contrast to computer forensics which deals with retrieving data from computer's disks, network forensics retrieves data on which network ports were utilized to access network [49].

In computer forensics, the examiner and the hacker are at different skill level, with the investigator ~~has the~~ advantage, whereas in network forensics they are at the same level, as the same tools are used by the hacker to drive the attack and by the examiner to re-search the attack. ~~However, network forensics examiner is more at a disadvantage position, since research is only one task he/she has to deal with, whereas the hacker has no time limit for enhancing his/her skills.~~ Some differences between computer forensics and network forensics are depicted below [35].

Table 2: Comparison of Computer forensics and Network forensics

Computer forensics	Network forensics
Introduced by law enforcement to handle computer data	Evolved as a response to the hacker community
The investigator and attacker are on two different levels	The investigator and the attacker are at the same skill level
The investigator and attacker use different tools, investigator has upper hand	The investigator and attacker use same tools and practices
Computer forensics contains preservation, identification, extraction, documentation, and interpretation of data	Network forensics involves the capture, record, and analysis of network events
It is about acquiring, providing chain of custody, authenticating, and interpretation	It is about investigation of packet filters, firewalls logs, and IDS logs

4.1.1 Network Forensics Definition

In general, network forensics is the branch of digital forensics that deals with the examination of events and operations related to digital networks gained through monitoring and capturing network traffic [57]. The goal of the related data is the collection of the necessary evidence in such a way that is legally acceptable, so the perpetrator can be prosecuted [58]. Network forensics tries to analyze inbound and outbound packets transmitted across the network connections, examining data logged through firewalls or IDS or routers and switches [35].

One definition of network forensics is “*the use of scientifically proven techniques to collect, fuse, identify, examine, correlate, analyze, and document digital evidence from multiple, actively processing and transmitting digital sources for the purpose of uncovering facts related to the planned intent, or measured success of unauthorized activities meant to disrupt, corrupt, and or compromise system components as well as providing information to assist in response to or recovery from these activities*” [1].

The definition of Markus Ranum about network forensics is “*the capture, recording, and analysis of network events in order to discover the source of security attacks or other problem incidents.*” [59].

Generally, network forensics contains the *CIA process*, which is described as:

- Capture (monitoring network traffic and capturing packets)
- Identify (identify if there is an anomaly in packets, based on certain criteria, so an attack is emerged)
- Analyze (in case of an attack, the nature of it should be determined through packets to understand what has happened) [57]

In contrast to other branches of digital forensics, network forensics examines data that are dynamic, unpredictable and volatile, in terms that traffic is transmitted and then lost, so most of the times network forensics is considered as a proactive energy [18]. Computer forensics deals with no dynamic data, ~~but data at rest~~. The planning process is to recognize the media that needs to be searched, produce and authenticate a forensic image, recognize the dissimilar artifacts to be examined, execute an in-depth analysis, ending with a report that gives attention to the findings. Most of the times, this report contains deleted, misnamed and hidden files and artifacts, registry entries, password-protected files, e-mail communications, carved data and much more. But, all these appear for the situation of the system at the collection and imaging time, which is the rea-

son we call it *postmortem* investigation, meaning that it does not contain live-memory forensics [57].

On the other hand, network forensics is dynamic, as we mentioned previously. Some agreements must be conducted to capture and store network traffic before the process begins. In order an analysis to be executed, a copy of the transmitted packets that come across the network is necessary, ~~that are logs~~. In particular, many network devices, appliances, operating systems in use and other devices on the network produce logs, each one of them directs the same incident in a different way. For instance, a login action will be called as a login by some operating systems in contrast to some other devices, which may call it as log on or user authentication incident. The message content and syntax of logs are vendor-specific and it can vary from appliance to appliance. Logs are time-sequenced and may be cryptic in nature [57].

In the case of computer forensics, logs still exist but they do not have this special meaning as they have in network forensics. So, all branches of digital forensics should work together and simultaneously, as most investigations usually contain all the branches of digital forensics in any case of a significant importance [57].

4.1.2 Classification of Network Forensics Science

The systems that examiners use to gather network data for forensics analysis are divided mainly in two basically types based on *packet capture* criterion [60]:

- *Catch-it-as-you-can* systems, which capture all packets that are transmitted across the network traffic, then analyze them, but this type requires large amounts of storage.
- *Stop-look-and-listen* systems, which analyze each packet in memory and only specific information is saved for future analysis, something that requires a faster processor to deal with inbound traffic.

However, there are some more system types of network forensics, based on other characteristics, such as [35]:

Purpose General Network Forensics (GNF) deals with increasing security. Data which is extracted from the network traffic is analyzed, aiming at discovering attack patterns. Strict Network Forensics (SNF) contains fixed legitimate demands, since the outcomes acquired will be utilized for evidence in order to prosecute the network crime [61].

Platform Network forensics can be a hardware appliance with pre-installed software or standalone software. In the first case, data is captured, analyzed, then the outcomes are displayed on a computer interface. In the second case, the software is installed itself on a host, examining packet captures, which are copied and stored in the host.

Time of Analysis Most of network forensics analysis appliances contain real-time network inspection, signature-based anomaly detection, data analysis and forensics research. However, there are some open-source software tools which are scheduled for postmortem research of captured packets. In particular, sniffer tools capture full data packets, store them in a host, then examined them later.

Data source There are flow-based-systems and packet-based-systems. For the first type, network equipment gathers statistical information based on some characteristics within the network traffic as it comes across the network, sending it to a flow collector which keeps it in reserve and analyzes this data. For the second type, full packet capture is executed at different points in the network, which are gathered and kept for deep observation.

4.1.3 Challenges in Gathering Network Forensics Evidence

Investigators encounter many challenges in network-based events in many areas, such as acquiring, storage, privacy, seizure, and admissibility. In particular [62]:

- *Acquisition* Locating specific evidence in network constitutes a difficult task. Sometimes, investigators know where exactly the essential object is located, which they want to use it for proven reasons, but they cannot gain access into it for other reasons.
- *Storage* Most network devices have limited storage capacity, due to the fact that they do not hold secondary or persistent storage. Obviously, this is something that brings a lot of difficulties in the investigators' work.
- *Privacy* Investigators must follow the legislation of every area they work. Sometimes, there may be legal issues that prevent them from acquiring the evidence they need, due to the 'personal privacy', even in the case of perpetrator's view.
- *Seizure* Obviously, grabbing and managing a hard drive is a much easier task than grabbing and managing a network device, where in many cases a total network part can be brought down for unlimited period of time.

- *Admissibility* In contrast to file system-based evidence, which is now accepted in both criminal and civil processes, network forensics constitutes a recent way of dealing for digital investigations, where there are often processes that are not legally agreed for admission or even they do not exist on dissimilar categories of network based digital forensics.

An important challenge in the total procedure of network forensics in order to be successful is the essential configuration and maintenance of network needs, as well as the appropriate updates that must be done by network manager. Another important thing is the essential infrastructure of network, meaning that all the necessary network forensics tools (hardware and software) can be supported. The main challenges that investigators should have on their mind about the network infrastructure are [62]:

- *Data sources* There are a lot of data sources, when speaking about network, from raw data to logs from network devices. It is not feasible to collect all data from all sources, especially in case of huge networks, even though this sounds the best solution. That is the reason why the choice of proper source of data is a crucial task.
- *Data granularity* This is another challenge in correlation with the choice of data sources, as investigators should decide how much details they want to collect, apart from the data source. In the case of a small network, the collection of entire packets (including packet headers, IP addresses, port numbers, rules of conduct, TTL, etc.) is much easier than in the case of a larger network where this is not practical at all.
- *Data integrity* This is the most crucial part of network forensics investigation method, as any change of collected data from the network drives to a potential dismissal of the whole forensics process or even to a not accepted as true evidence by the court case. Consequently, investigators need to guarantee about the integrity of data during and after the collection and analysis.
- *Data as legal evidence* Some data is collected as evidence for the court of law, which means that this data must be collected properly so it can be legally accepted through rigorous legitimate procedures.
- *Privacy issues* During the collection of data, investigators get access into some personal information, like emails, personal files, or more descriptive infor-

mation, such as picture data, GPS location, etc. This leads to the appearance of privacy issues ~~so for the victim side as for the enemy side.~~

- *Data analysis* It is very hard to discover, analyze and present all data that have been collected ~~before~~ due to the great amount of data at the collection phase and the difficulties that a network environment appears. That's why investigators use a variety of tools to make this task more convenient and more precise.

Network forensics investigators are looking for important clues inside collected data, so some other challenges that appear in almost every network forensics analysis are [62]:

- *Time* Sometimes investigators must submit the evidence data in a specific period of time, meaning that they must follow a deadline, or other times they just have to note the date information about the instances that have been analyzed. In both cases, time constitutes an aspect of a great value in every investigation.
- *Performance* People who work on network forensics should have the appropriate knowledge and skills in order for their results to be as fast and accurate as they can be. It is very important for an investigator to react fast on a security breakdown or policy violation when an alert is observed. However, nothing can work in that process ~~optimum~~ when both hardware and software have not ~~been~~ performed at top level, ~~even though human's knowledge and ability are well performed.~~
- *Complexity* As larger networks cause larger problems, this leads to more difficult complexity problems, too. This is due to the fact that larger networks use more network devices, more end users, more IDS and firewalls, servers in different geographical regions (which makes things much more difficult) and generally larger infrastructure, where the investigators can be placed depending to the size of data they want to examine.
- *Collection* As the size of collected data may be enormous, it is very important for an investigator to separate only the valuable [^] that is needed for the analysis.
- *Law* Many legal subjects arise through the performance of network forensics, maybe due to the fact that there is not an internal decision-making organization connected with the law for each government.
- *Hiding a Breach* There are a lot of organizations, which prefer to hide a network-failure incident, calling for help only when they have no other choice. But,

in the middle of this period of time a skillful attacker can cover his/her tracks, which is obviously something that needs to be avoided.

- *Network systems* Since network forensics systems focus on ~~audit trails~~, the price of these systems constitutes another challenge, too. These systems carry a lot of information, which means that large amounts of disk space are required with powerful CPU useful valuable supplies in order to analyze all the logs they have collected previously.

4.1.4 Recent Trends in Network Forensics

From its beginning, network forensics focuses on wired environments, especially dealing with the version 4 of the Internet Protocol (IPV4) and some other related protocols at the network layer of the TCP/IP protocol suite, which is described ~~more detailed~~ in next paragraph. However, there are some recent trends on this field. In particular:

Steganography Attackers usually utilize some kind of “light” forms of cryptography to make harder the identification of attack patterns, which otherwise it would be easier to be identified by any IDS [63].

Honeypot Forensics The use of honeypots is to be compromised, so important information can be gained about the attackers’ techniques and tools he/she has used earlier and after the intrusion on the honeypot. In that way, new kinds of rootkits, Trojans and potential zero-day exploits can be found, as well as an optimum comprehension of attackers’ field of interest [64, 65].

IP Version 6 Forensics IPv6 Internet supplies malicious users some advantages, in ~~terms~~ that instances are still poorly logged and monitored, and even a kind of anonymous connectivity is provided by some free tunnel brokers [66]. The passage from IPv4 to IPv6 needs time to be performed and for some period of time these two protocols will work collaboratively, which means that new security vulnerabilities and exploits will be emerged, which will require the appropriate forensics analysis [67].

Botnet Forensics Botnet is an army of infected computers that a botherder controls and gives instructions, sending spam e-mails or ~~disable~~ websites through a plenty of fake requests to other machines-victims. It is almost impossible to find the identity of spammers given just the electronic trail [68, 69].

Wireless Network Forensics Categorization of user activities is considered a demanding task, along with the network monitoring and content inspection, which are continually

growing, as wireless technology has been inserted more and more in our daily lives [70]. Obviously, ~~the level that people stand right now for handling efficiently wireless devices is not at the desired point, due to the shortage of necessary tools and procedures for forensic computing investigations, that's why there are a lot of misuse cases that leads to escape detection~~ [71]. Researches over attacks on wireless VoIP or VoIPoW, which is considered the most famous system for mobile communication nowadays, are still at a beginner level [72]. ~~In the field of Mobile Ad Hoc Networks (MANETs) there are some challenges, too, where the level of reliability is the feature with which the quantity of the evidence packets is checked~~ [73, 74].

Application Layer Forensics There is an observation that attacks have passed from the network and transport layer to the application layer of the TCP/IP protocol suite. Some types of attacks on Web security may be cross site scripting (XSS), SQL injection, buffer overflows, etc. The payload of the packets that are transmitted to and from the Web service can provide us with valid information which can be used as evidence [75]. Another trend that ~~comes up these period of time~~ is *Domain name service forensics* [76].

SCADA Network Forensics The recent Supervisory Control and Data Acquisition (SCADA) systems use TCP/IP to transfer sensor data and control signals, usually utilized for commercial activities and automation. Security subjects are brought up when ~~the use of TCP/IP~~ as a carrying protocol, as well as when IT is combined with SCADA networks, where a havoc can be caused in case of a successful attack on an IT network and its gateway devices [77, 78].

Grid Forensics Grid computing includes a total of all distributed resources, which demands a high level security to protect all that data and the appropriate security methods that can be applied to prevent the appearance of perpetrators. Since grid computing is a growing technology there is not ~~such~~ experience on this field [79].

Forensic Data Representation The extension of all these trends estimates a forthcoming crisis in digital forensics, requiring the design of new abstractions for data representation forensics processing, aiming at a more effective investigation [80].

Cloud Forensics Security policies have ~~been~~ emerged by cloud computing relating to remote access, use of data over a browser, privacy and audit mechanisms, reporting systems and management systems that recommend the way that data can be safe on a distributed computer system ~~which can be located anywhere~~. The complexity between the way that cloud provider and cloud consumer can take their own advantage of this ser-

vice constitutes a challenge for hackers and cyber-criminals and this ~~arises~~ the need for new investigative viewpoints [35].

Intelligent Network Forensics An intelligent network forensics system is more prepared to defend its data in a ~~next~~ potential attack through designing intrusion scenarios, making more difficult for the attackers to know about intrusion signatures, evidences, impacts and objectives. This constitutes problem-solving knowledge extremely important, as it gives instructions about how the system can utilize domain knowledge to examine malicious activities [35].

4.2 The Generic Process Model for Network Forensics

As it happens with the case of computer forensics tasks, the process of recovering and analyzing the evidence, which is collected from the network resources, is very important as it has great value in the court for legal reasons. Consequently, there are some steps (the generic process model for network forensics) that forensics investigators must follow in order to make this process more profitable. It is based on the already existing digital forensics models and ~~was~~ summarized ~~after a research was done, conducting at the following steps~~: preparation, detection, incident response, collection, preservation, examination, analysis, investigation and presentation. Below, a diagram is given which depicts the generic process model for network forensics, as well as a description for each one of these steps [81]:

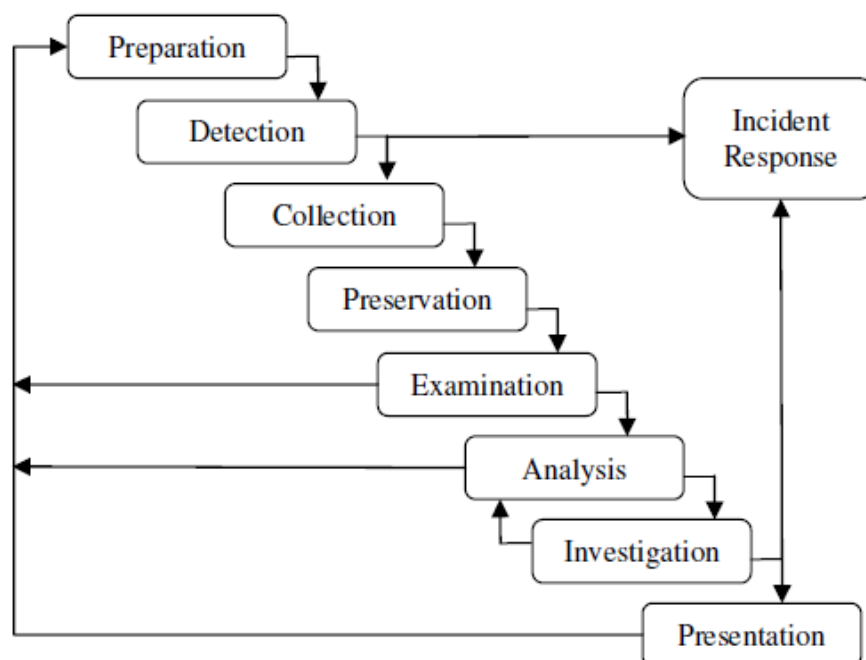


Figure 8: The Generic process Model for Network forensics, International Journal of Computer Science & information technology (IJCSIT), Vol 3, No 3, June 2011

Preparation: All the required security tools, such as IDS, firewalls, packet analyzers, traffic flow measure measurements, are placed at many different points on the network at this stage. Another task, which is contained in this stage, is the acquisition of all the necessary legitimate documents in order privacy to be applied and not to be violated. Also, people working in this stage must be trained appropriately, so the required knowledge can be gained and conducted in better results. Overall, a well-planned preparation stage reduces the total cost of the investigation.

Detection: This is where an alert is generated when a threat is detected. The anomalies and unlawful events are examined based on different factors, so as with a fast validation any suspected attack can be confirmed. After that, a significant decision is taken, whether to go on the investigation process, generating an alert or ignore the event. Also, this stage branches in two directions, incident response or collection.

Incident Response: The collected information is the one that determines the response to each crime, which is used for validation and assessment of the incidents. Also, the response is dependent to the type of the attack identified. At the same time, a decision is taken whether to go on with the investigation or collect more information. At this stage of response to an incident, it is very important all the data being collected not to be tampered or obstructed in order to perform network forensics analysis. Also, an organization policy is kept in place while responding to attack.

Collection: The traffic data, which is gathered at this stage, is collected by the sensors, which must be secure, fault tolerant, not accessed by everyone and able to stay away from compromising. All the hardware and software tools must be reliable, which are used by a well-defined procedure in order to gather maximum evidence causing minimum impact to the victim. This stage is very important as the traffic data changes rapidly and there is no choice to generate the same data which passed away when you lose it for the first time. The network must be monitored in order to recognize attacks that may appear in the future. Also, the integrity of data logged and network events recorded must be ensured. The amount of data logged is huge, so an enormous memory space is needed and the system must be versatile in nature, as well as able to manage various log data formats in the appropriate way. Tools which are used in this stage are TCPdump, Wireshark, Snort, etc., some of which are described at next paragraphs.

Preservation: The collected traces and logs are stored on a backup device with ~~the scope of reading only~~. This original data must be in a safe place and untouched, ~~as well as the hash of all traces in order~~ the integrity of data to be ensured and the chain of custody maintained. The analysis is done in the duplicate copy, ~~as when the process is repeated on original data the legitimate requirements will be easier.~~

Examination: ~~This is the stage that examines the previous one.~~ Traces are integrated and fused as a large data set on which the analysis is performed. Sometimes an appropriation is needed, when there are issues, such as unnecessary information or overlapping time zones or when alerts from different sources may be contradictory. The process of examination should be done in such a way that crucial information from important sources is not lost. The gathered data is classified and clustered into groups, so that the volume of data to be stored may be decreased to manageable chunks, as it is easier to analyze big groups of organized data. Searching of the gathered evidence is done methodically to export specific indicators of the crime. Minimum attack attributes selected must be so credible that the least information recorded holds the highest probable evidence that proves something. Feedback is given to improve the security tools.

Analysis: Data that has collected before is now analyzed, with the utilization of many methods, such as data mining (ANN, fuzzy and genetic algorithm GA) and statistical analysis to search if an invasion is matched to an attack pattern. In many cases, the crucial parameters are related to network connection establishment, DNS queries, packet fragmentation, protocol and operating system fingerprinting. The attack patterns are put together and rebuilt to understand the scope of the attacker, as well as his/her methodology of doing this. Feedback is also given to improve the security tools.

Investigation: The goal of this stage is to identify the attacker by determining the path from the victim/system network through any intermediate systems and communications pathways until reaching the attacker's source. Information that has collected at the previous analysis stage is used collaboratively with packet statistics in order to be utilized for attribution of the attack, which is a really difficult task, as attribution ~~is built~~ the identity of the attacker. The investigation stage presents data for incident response and prosecution of the attacker, which varies from case to case, depending on the type of the attack.

Presentation: This is the final stage of the generic process model for network forensics, where the outcomes are presented in readable and understandable format for legal per-

sonnel, along with a description of all the methods used to come finally at the termination. Everything that has been done must be satisfied with legal requirements and security policy, and systematic documentation is presented to authorities. Sometimes visualization of the outcomes is used, so the incident can be better understood and assimilated. A thorough review of the incident is conducted and countermeasures are suggested on how to prevent from similar attacks in the future. The whole instance is documented to have an impact on future investigations and to give feedback for the growth and the upgrade of the security products. Statistical data is also presented along with the outcomes and the network forensics process is ending here, since the information presented conducts to the prosecution of the attacker.

4.3 Technical fundamentals on Networks

Below, some basic fundamentals about the network will be presented. Initially, a *network* is a group of computers/devices that are interconnected, either wired or wireless. Each device on the network has a ~~one and only one~~ network address, which can be temporary or permanent. The problem is that these addresses cannot easily be remembered by people, as they are numeric quantities, but computers can work with them without any difficulty. These addresses are known as *IP addresses*. In order these IP addresses to be easily remembered by people, they are stored as *Domain Name Server* (DNS) servers, where they are stored as textual addresses. So, DNS servers do that work, translating numeric addresses into textual and vice versa [57].

As we mentioned, IP addresses determine a particular host machine working on a network, but a numeric *port number* is the one that shows the activity that is running on a host machine. Some of the most common port numbers, along with their applications in which they are applied to, are 20 and 21 for FTP, 23 for Telnet, 25 for SMTP (mail), 80 for HTTP, 110 for POP3 (mail), 443 for HTTPS, etc. [57]

Devices can communicate with each other, since they are interconnected and this is done by exchanging data. The messages are broken into packets and transmitted over the network, the so called *packet switching* procedure. Each of these packets has a specific maximum size and is divided into the header area and the data area. Everything that has to do with the details of the sending of the packet must be included in the header area, such as the sender and receiver addresses, as well as the required information in order the packets will be sent in the correct order [57].

So, computers can communicate with each other, when they are connected on the network, guided by some rules, which are known as *protocols*. Protocols determine valuable information, such as: addressing of messages, routing of messages, error detection, error recovery, packet sequence, flow controls [57].

At the beginning of computing, computers' role was to work as stand-alone machines, while all other work that required cross-computing was done manually. ~~What was actually happened was that files were moved on disks from computer to computer.~~ Obviously, there was a need for this work of cross-computing to be done by computers, where more than one computer could talk to others and vice versa [49].

~~As a result, a new movement was born, the open system movement, which was designed for computers' hardware and software manufacturers to execute that specific scope.~~ In order this ~~can~~ be fulfilled, standardization of equipment and software were required so that computers can communicate with each other. That's why the International Organization for Standardization (ISO) developed the *Open Systems Interconnection* (OSI) model, also known as *seven-layer model*. The OSI is an open layered architecture model, which operates as the network communication protocol standard, but it is not the most famous one, as there is the *Transport Control Protocol/Internet Protocol* (TCP/IP) model which is more widely used [49]. Below, both of these models are described in more details.

4.3.1 The seven-layer Model (OSI)

The evolution of the OSI model was based on the safe standard that a communication task over a network can be divided into seven layers, where in each one of them different tasks are performed. Different layers mean different services. The establishment is that each layer can communicate only with its own adjacent layers, which means that the protocols in each layer are based on the protocols of the preceding layers [49].

Tasks and information move down, beginning from the top layers to the bottom layers, where they are sent out over the network from the source system to the destination system, and then the tasks there work reversely, as they begin from the bottom of the layers until they reach the top. Each layer is created to receive work from the layer above it and pass it to the layer below it and vice versa. The interfaces between the layers are standardized for reasons of convenience of the communication. But, each layer stays

independent, can be created independently and its functionality should not influence the functionalities of others layers above and below it [49].

The OSI model is depicted in the table below, along with some brief description of each one of these layers underneath [57].

Table 3: The OSI protocol layers and corresponding services

Layer number	Protocol
7	Application
6	Presentation
5	Session
4	Transport
3	Network
2	Data link
1	Physical

- Layer 1: The *physical layer*. It is the physical infrastructure over which the data travels, through cables, hubs, etc. It is responsible for transmission and reception of raw data and unstructured bits and bytes in a physical medium. This layer is never concerned with protocols or other such higher-layer items.
- Layer 2: The *data link layer*. At this layer data encapsulation is performed in the form of packets, as well as the interception of them at the physical layer. The data link layer provides node-to node data (frames) transfer, initiating a logical link between two nodes on a network and terminating it. It detects and potentially corrects errors that may occur in the physical layer (error-free transfer) and it also establishes the protocol for flow control between the two connected devices.
- Layer 3: The *network layer*. This layer is responsible for the transmission of the packets, which contain variable length of data sequences (called datagrams), from a source to its destination in different networks. In case of the message is too large to be transferred from one node to another on the data link layer between those nodes, the network can perform message delivery by splitting the message into many fragments at one node, sending them separately, reassembling then after at the other node. There is not always reliable guarantee that the messages have been sent at this layer, sometimes a protocol can provide reliable messages but this is not necessary. At this layer is decided the route, mapping of the logical and physical addresses, and data traffic control.

- Layer 4: The *transport layer*. This layer is responsible for the delivery of the packets from a source to the destination, ensuring that the messages (segments) are delivered in a sequence without duplication or loss and is error-free. There are protocols that are state and connection oriented, which means that the transport layer can keep track of the segments and those that failed can be retransmitted again. This layer also includes the acknowledgment of the successful data transmission and in case of no errors occurred, it sends the next data. It produces packets out of the message received from the application layer. The process of dividing a long message into a smaller one is called *packetizing*.
- Layer 5: The *session layer*. This layer controls the network access, in terms of controlling the connections between computers. It sets up sessions among processes running on different nodes via different ports. It establishes, manages and terminates these sessions between the local and the remote application.
- Layer 6: The *presentation layer*. This layer sets up context between application-layer entities, where this layer can utilize dissimilar syntax and semantics in case of the presentation service provides a mapping between them. When the mapping exists, presentation service data units are encapsulated into session protocol data units and passed down to the protocol stack. The aim of that layer is to format the transmitted data into the form that the application layer accepts.
- Layer 7: The *application layer*. This is the closest to the user layer, in terms of both the OSI application layer and the user interact directly with the software application in order to communicate with each other.

In peer-to-peer communication, the two components, which communicate, can begin and get tasks and data, where the data goes from the top in the application layer of the protocol stack on each computer. Then, as we have mentioned before, data and tasks move down from the top layer until they reach the bottom layer, where they are forwarded over the network from the source system to the destination system. There, at the destination system, things work reversely, as task and data go up through the layers until they get the top. The rule is that each layer takes work from the layer above it and pass it to the layer below it. As these data and tasks travel between layers, each layer appends or takes away its own header to the data unit. On the other hand, at the destination, each header, which was added, is removed one-by-one until the receiving application gets the data that had to from the beginning. Each layer header includes information

about that layer's peer on the remote system, which gives some kind of directions about how to route the packet through the network or what should be done to the packets at the destination side [49].

The OSI model was designed to provide a standard for all proprietary models, including as many models as can be, but it never actually succeeded its goal, as it never succeeded to replace them. The reason for that is its complexity through this "all-in-one" concept it provides. Another reason is that it arrived late in the market, thing that made difficult its much expected interoperability across networks [49].

4.3.2 The Transport Control Protocol/ Internet Protocol (TCP/IP) Model

In contrast to the OSI model, TCP/IP model is less complex and it came first in the market, so it is more popular. ~~It has two to three less layers than the OSI model, which has seven, so it does not actual suit with the OSI model.~~ Initially, it was deployed for the US Department of Defense Advanced Research Projects Agency (DARPA), but it became so popular over the years that it is considered the 'de facto' standard for every Internet at this moment. As a result, it provides a simpler, effective, open communication infrastructure in academic and collaborative environment [82]. Basically, the attackers use the vulnerabilities ~~appeared~~ in the implementation of the TCP/IP protocol stack in order to exploit them, causing an attack [35]. Its name arose from the use of two basic protocols, the Transmission Control Protocol (TCP) and the Internet Protocol (IP). Below, there is a table, which depicts the layers and some protocols that are used in each of them. Then, there is a description of some of these protocols in these layers [49].

Table 4: TCP/IP protocol layers

Layer	Delivery unit	Protocols
Application	Message	Handles all higher-level protocols including File Transfer Protocol (FTP), Name Server Protocol (NSP), Simple Mail Transfer Protocol (SMTP), Simple Network Management Protocol (SNMP), HTTP, Remote file access (telnet), Remote file server (NFS), Name Resolution (DNS), HTTP, TFTP, SNMP, DHCP, DNS, BOOTP
		Combines application, session, and presentation layers of the OSI model
		Handles all high-level protocols
Transport	Segment	Handles transport protocols including Transmission Control Protocol (TCP), User Datagram Protocol (UDP)
Network	Datagram	Contains the following protocols: Internet Protocol (IP), Internet Control Message Protocol (ICMP), Internet Group Management Protocol (IGMP)
		Supports transmitting source packets from any network on the internetwork and makes sure they arrive at the destination independent of the path and networks they took to reach there
		Best path determination and packet switching occur at this layer
Data link	Frame	Contains protocols that require IP packet to cross a physical link from one device to another directly connected device
		It included the following networks
		WAN – wide area network LAN – local area network
Physical	Bit stream	All network card drivers

Application Layer

On the top of the TCP/IP model is the application layer, which is similar to the corresponding application layer on the OSI model, but it combines the functions of the OSI application, presentation and session layers. Its role is to guide how the host programs interface with transport layers services along with their related application protocols. ~~As it is shown~~ some application protocols are [57]:

- *FTP*: this is used for file transfer.
- *SMTP*: this is used for the transfer of electronic mail.
- *DNS*: it is for the network support.
- *SNMP*: it is used for the remote host management.
- *HTTP (Hypertext Transfer Protocol)*: it is an application-level protocol for distributed, collaborative, hypermedia information systems [83]. It is a generic, stateless protocol which can be used for many tasks beyond its use for hypertext, such as name servers and distributed object management systems, through extension of its request methods, error codes, and headers [35]. Users basically use

it for accessing data on the World Wide Web (WWW), which are transferred through HTTP messages between clients and servers. These messages are read and explained by both of them, the HTTP server and HTTP client (browser), the format of which is almost the same [35]. Particularly, a request message contains a request line, a header and a body, while the response message includes a status line.

The data structure of the application layer includes bit streams, as shown in the picture below [49]:

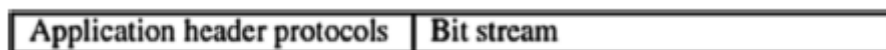


Figure 9: Application layer data frame

Transport layer

This layer controls the communication session between the host computers, as it does in the OSI model. It actually transports the application messages, which carry application layer protocols inside their headers, between the client and the server, using two standard protocols [49]:

- *TCP*: it supplies a connection-oriented service between the source and the destination [84]. Also, it guarantees that the application layer packets have been arrived in the destination in order, based on two mechanisms [49]: congestion control mechanism, which throttles the transmission rate of the source element in case of traffic congestion in the network, and the flow control mechanism, which attempts to suit the speeds of the sender and receiver in order to synchronize the flow rate and limit the packet drop rate. Other mechanisms that TCP uses are sequence numbers, acknowledgments, which are used in case of large data streams, where TCP breaks up this data stream into distinct data packets, each of which has its own sequence number stored in the header. On the destination system, these dissimilar packets are reassembled using the above mechanisms [57]. Also, there are the timers and 3-way handshakes mechanisms, which is a process that is required to be done before the sending of data begins, as the TCP needs a connection between the communicating parts to be established [57]. TCP also determines the port numbers of the source and the destination, which are elements of the header of the transport layer packets, along with the information from the mechanisms we mentioned above [57].

- *UDP*: in contrast with the TCP protocol, UDP is a connection-less protocol, with no guarantee of the delivery of the application layer packets. It is responsible for just the transmission of the data from one node to the other without acknowledgments and confirmations [57]. But, its main advantage is that because of the above characteristics is much more effective and faster in case of sending real-time data, such as streaming video and music, games, etc. [49].

Below, the data structure of TCP and UDP are shown [49]:

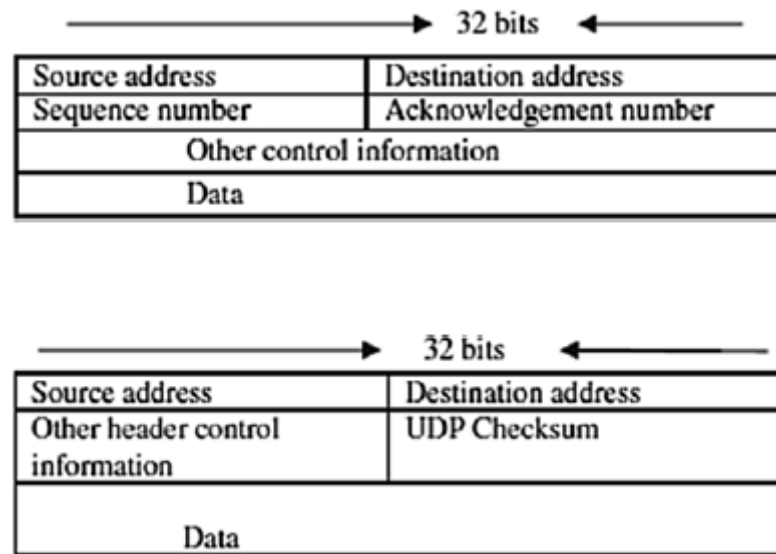


Figure 10: A TCP data structure above and a UDP data structure below

Network Layer

This layer is responsible for moving packets, which are called *datagrams* here, between routers along the path from a source host to the destination host [49]. Same basic protocols that it uses are:

- *ICMP*: it makes it easier to send one-way informational message to a host [85]. It is transferred in the payload of the IP packet (which we are going to mention below) and has many data structures of its own. ICMP is used by a router or a destination host to make the source host informed about errors in datagram processing, allowing routers to send error or control messages to other routers or hosts. Also, it offers communication between the two communicating machines at the network layer. Because the operations of ICMP protocol are mainly two: reporting non-transient error conditions and investigating the probing of network with request and reply messages, the messages of ICMP are divided into two types: ICMP error messages and ICMP query messages, where each one of these

messages is allocated with a number, the message type, which indicates the above type of message. There is another number, which stands for the specified ICMP type [35].

- *IP*: it is the most globally used protocol at the network layer [49]. Its main functions are two [57]: one is separating the data stream into standard size packets at the source, then placing them together reversely in the right sequence at the destination. The other is directing or routing a packet, beginning from the source device IP address to the destination one, through a series of intermediary routers, where the next hop of the packet is identified through routing algorithms [57]. IP utilizes header information from the transport layer protocols, which contain datagram source and destination port numbers from IP addresses and other TCP header and IP information to pass datagrams from router to router through the network [49]. Nowadays, the most famous form of IP address is the IPv4, which consists of a 32-bit addressing scheme [49]. However, because of the rapid growth in the number of devices connected to the Internet, there was a fear of running out of these addresses. As a result, a new version was created, IPv6, which consists of a 128-bit addressing scheme and it provides for much longer addresses, which means more Internet users [57]. All the powers of IPv4 are also included in IPv6 and if a server can support IPv6 version, it can also support IPv4, too [57]. As the previous layers do, the network layer also transports the network layer protocols to the next one, which is the data link layer [49].

The IP datagram structure is shown below [49]:

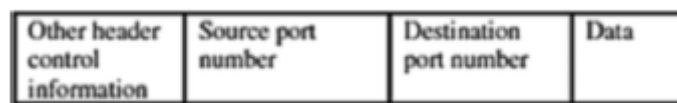


Figure 11: An IP datagram structure

Data Link Layer

As we mentioned in the OSI model, this layer moves packets between switches (routers), over connecting links, providing reliable delivery of network layer packets over them [49]. It is at the lowest level of communication, containing protocols such as the *network interface card* (NIC), *operating system* (OS), Ethernet, asynchronous transfer mode (ATM) and others like frame relay [49]. The *frame*, which is the data link layer protocol unit, can be moved over links from source to destination through dissimilar

link layer protocols at dissimilar links along the way [49]. A *local area network* (LAN) is a group of host devices in an adjacent area, which permits high data transfer rates among the devices that have the same IP address, while *wide area network* (WAN) is a computer network that extends over a larger geographical distance. Each node in the LAN has ~~a one and~~ only one MAC (*media access control*) address, a 48-bit serial number, defined to each NIC (*network interface card*) supplying a physical address to the host device. A NIC is a computer hardware, which is used in order to provide an interface between the host device and the computer network. ~~A MAC address has mainly two types of forms: the static address, which is permanent and it can be changed only if the NIC changes, the dynamic address, which is gained when the computer is on and connected to the internet [86].~~ Due to the fact that there are two addresses, the network-layer addresses (IP addresses) and the link-layer addresses (MAC addresses) there is an obvious need for them to be translated, and this is the task of ARP (*Address Resolution Protocol*).

Physical Layer

This layer actually moves data link datagrams bit by bit over the links and between the network elements. The protocols that are used at this layer depend on the features of the link medium and the signals above it [49].

4.4 Learning to handle the Evidence

After investigators identify the sources of the evidence, then they have to know how to manage that evidence. Since, the investigation process must follow rules, agreed with the law, investigators must ensure that all the processes followed by them do not expose the evidential value of the collected information [57].

4.4.1 Identifying sources of network evidence

In order an investigation to succeed its goal, ~~it is very important~~ the task of collection, preservation and analysis of the evidence that was captured before. So, investigators should have on their mind the ways with which they can identify the sources of network evidence, which can be divided into two categories, evidence ~~obtaining from~~ *within the network* and evidence from *outside the network*.

In the first case, evidence gathered from within the network, information can come from [57]:

- *Evidence from network and device logs*: log files are the recorders of all the activities and results performed by a device or by outside agents on a device. As a result all the incoming and outgoing events are logged on a system, that's why logs' role is so crucial at the investigation process. Some devices that generate logs are: firewalls, IDS, antivirus servers, etc.
- *Network traffic*: data is split up and transmitted across the network in the form of packets, requiring to be captured and be analyzed.
- *Memory of the individual computers under investigation*: volatile memory has great value for the purposes of investigation, as many malware often reside only in the memory of a computer. Data is required to be grasped from the suspect system's memory, when memory is involved at the investigation.
- *Evidence residing on the hard drives of individual computers under investigation*: important data that is used as evidence resides on the hard drives of compromised computers (such as traces of internet activity, web mail communications, attempts to cover tracks and obfuscate evidence, and others) all found at the investigation process, making the evidence unclear.

In the second case, evidence gathered from outside the network can be:

- *Internet service provider (ISP) logs*: these logs contain information about the access to different Internet resources which are provided by the ISP, such as information about the log on, log off, usernames, resources accessed, online content, online activity, IP addresses, date and time of usage, etc.
- *Domain name controller logs*: these logs can contain same information, like date and time, IP addresses, queried domain names, protocol used, etc. The obstacle here is that the duration of that data is very short, as the volume of data in logs is ~~at great height~~.
- *Internet archives (Wayback Machine)*: they are online resources that store websites and pages in an archive for a particular period of time, which give a direction to investigators about the state of an internet server offering websites before an attack is occurred. The URL of that machine is <http://archive.org/web/>.
- *Domain hosting provider logs*: servers that host a domain belong here. This is where all unauthorized efforts to log in to the domain host belong, so for exam-

ple if someone tries to hack a system, his/her activity will be found with this machine.

- *Evidence on mobile devices*: evidence of mobile phones or tablets interaction is produced on these devices.

Obviously, investigators should be concerned about the cases, where asking information from these sources is not an easy task, as privacy laws protect these sources and that it may be required to ask for permission from the law enforcement officers before requesting for them.

4.5 Data Evidence Acquisition on the Network

Network forensics is responsible for the analysis of the trace and log data of network intrusions captured by the network security products that already exist in the system in order to find potential misbehavior clues or an attack. It cannot stop the network crimes, but it can collect evidence, which can later be used for the investigation of the crime in order the attackers to be found. So, the challenge of the network forensics system is to recognize valuable network incidents and collect the least set required to use them as evidence [87, 88]. There are many security and forensic tools, which collect data about dissimilar features, as well as protocol characteristics, which log it in dissimilar ways. The information gained from the collected data can be fused into a file for later use and maybe for evidence reasons [35].

Network forensics investigators usually divide the evidence acquisition into two categories, the *passive* and the *active* one. Passive evidence acquisition includes the collection of forensic evidence from networks through traffic acquisition. On the other hand, active evidence acquisition is when someone collects evidence through interacting with stations on the network, which may be done by logging onto network machines via the console or through a network interface, or sometimes by scanning the network ports to define the situation at that moment [89].

Here we focus at the passive evidence acquisition, so a brief description of active evidence acquisition is followed, and after that we are going to expand into the passive form of data collection.

4.5.1 Active- evidence acquisition

In general, evidence exists almost in every place on the network. Sometimes, we have to decide which evidence we want to collect from network devices, containing firewalls, web proxies, logging servers and more in order to capture network traffic. The problem is that in some cases these devices can cause serious harm to business functions, if someone tries to abstract them from the production environment. Another scenario is when the evidence stored is volatile, so it must be gathered while the system is still running. That is the reason that investigators apply active evidence acquisition: interacting with network devices which are live and on the network. However, investigators must be fully concerned of the plenty of ways that active evidence acquisition process can modify the devices, as well as the environment under investigation, having on their mind that their goal is to cause as little impact as can be [89].

4.5.2 Passive-evidence acquisition

Network security and monitoring tools are not expected to know how to manage forensic investigations, but they can capture the whole data packets and analyze them in detail. The ways that the network traffic, actually the packets, can be captured are two. The first one is by running a packet sniffer, like TCPDump or Wireshark, which captures packets in libpcap (.pcap) files or WinPcap. The second one is the collection of them from the routers or switches. Capturing the packets helps the investigators to find which exactly the attack is, who hides behind it, when actually it starts, in which point the attacker entered the network and in what way the network defense was breached [35].

Again, we ~~emphasize~~ at the traffic acquisition software, particularly at packet sniffing, at the next paragraph, ~~not at the physical interception of data, although we present a brief description of it.~~

Physical interception

Although it is impossible to gain network traffic causing zero impact on the environment, capturing or sniffing traffic can be executed with very little impact. There ~~is a~~ plenty of ways of transferring data over physical media, as well as many ways to intercept it. The easiest scenario is a station connected to another station over a physical conduit, like a copper cable, where voltage can easily be strengthened and redistributed in a one-to- many configurations. The reason that hubs and switches are designed is to expand the physical media, so the baseband can be shared in other stations. So, network

forensics investigators can acquire network traffic by passively intercepting it, when it is transferred across cables, through the air, or through network equipment, like hubs and switches [89].

4.6 Packet Sniffing

Nowadays, the way that systems are connected with each other makes the whole process of investigation harder, as it can take place in more than one computer, each one of which working for different purposes. Some of them control for software updates, others collect email, tweets, or RSS feeds, others establish the connections through which an authentication process to a domain or access network resources is taken place. So, the role of investigation is very important, as it provides investigators with valuable information [90].

In order an investigation process to be efficiently managed, the whole data packets are captured, so they can be analyzed in more details when the time comes. In particular, a network communication is established by a set of packets that are sent across the network. ~~A device can send and receive a plenty of packets per minute and computer networks task is to send these packets to their destination [90].~~

Packet analysis is the process of capturing packets in order to define how a computer or device communicated with other devices on the network. Packet analysis is known as *packet sniffing* or protocol analysis and a tool through which this analysis is performed is called *packet sniffer* or packet capture tool [90]. Such a tool captures raw data across the wire in order to find which parties are communicating on the network, what and how much data is transferred, what network devices are in use and other details which help in the investigation process, such as what the attack is, who is responsible for the attack, when it was launched, where the attacker inserted into the network and how the network defense was breached [35].

Packets can be captured in libpcap (.pcap) files or WinPcap by running a packet sniffer. Libpcap is a UNIX C library that offers an API (application programming interface) for capturing and filtering data link layer frames from arbitrary network interfaces. The goal of libpcap was to supply a layer of abstraction so that programmers could design portable packet capture and analysis tools [89]. WinPcap is the corresponding library for Windows systems, based on the libpcap, as many people work on Windows system.

Some popular packet sniffing and analysis tools are tcpdump, Wireshark, Network-Miner, NetWitness Investigator, Kismet, EtherApe, Cain and Abel, and many others, some of which are going to be presented below. These tools may work in correlation, in terms of packets that have been previously captured, as one tool can read them and another tool can analyze them. So, some tools capture packets and store them in a file, and other tools can read these pcap files, while filtering the traffic, based on specific protocol information. Most of these tools are free to download and some of them ~~can be worked in~~ both command line program format and GUI format. Among them, Wireshark is the most widely used, because its installation is easy, as well as its use. It is an open source tool that can be downloaded for free. What is important from the investigators perspective is that Wireshark can delve deep in the packets and capture information as well as that it supports many operating systems, many protocols and media types [90]. More information about Wireshark is given below, as well as in the following chapter.

~~4.6.1 Components of a packet sniffer~~

A packet sniffer contains four components: hardware, driver, buffer and packet analysis. Although most packet sniffers use common adapters, there are some of them which require multi adapters, wireless adapters, etc. The first thing that someone should do in order to install a sniffer is to check if the specific system includes the required adapter for this sniffer and then he/she can proceed into the drive program, as without it no installation can be done. When all these are fulfilled, a buffer is required, as it is the storage device for capturing data from the network [90].

Data can be stored in the buffer, using two ways. The first one is when data can be stored in the buffer until it reaches its limits, meaning when it runs out of space. The drawback in this method is that no new data can be stored in such a “bad” scenario. The second way is when new data replaces the old one, when the buffer overflows. Which method will be used depends on the forensic investigator, who is the one that chooses the buffer storage. Of course, buffer storage depends also on the EMS memory (Expanded memory specification) of each device. The higher the EMS memory is, the bigger the buffer storage is.

Finally, the packet analysis is the most crucial part of the sniffer, since it captures and analyzes the data from the packets [90].

4.6.2 How packet analyzers work

Packet analyzers intercept network traffic, which travels across the wired and wireless network interfaces that they have access to. The type of the captured information depends on the form of the network and the way that network switches along with other tools are put together. Particularly, in a switched wired network, the sniffer has the ability to capture data from only the port it is connected to, except for the case, where port mirroring is executed on the switch, whereas in the case of a wireless network, the sniffer can capture data from only one channel, unless there are many interfaces that let data to be captured from more than one channel [90].

The form in which raw data is captured is unreadable to humans, so a conversion must be taken in order investigators can use them with the most efficient way. After that process, the analysis of this data takes place and meaningful information can be extracted. There are usually three types of packet sniffing: ARP sniffing, where the information is transported to the ARP cache of the hosts, driving from the administrator, the IP sniffing, where the information about a specific IP address filter is captured, and the MAC sniffing, which is responding to the IP sniffing but for a MAC address [90].

4.6.3 Packet Sniffing Process

Most of the times, the network interfaces present in the segment have ~~an one and only one~~ one hardware address and they have the ability to watch the transmitted data over the physical medium. Because this address is unique, a transmitted packet across the network will only be accepted from the host device it is meant to, although there is the case where hardware addresses can be altered in software through virtualization mechanisms [90].

Generally, every IP network has a subnet mask, network address and broadcast address. Each IP address is divided into two parts, the network address and the host address, which is done through the help of the subnet mask. The host address is also divided into the subnet address and the host address. The subnet mask determines the IP address of the system by executing AND operation on netmask, by transforming the network bits to 1 and the host bits to 0. Each network has two booked host addresses, 0 for network address and 255 for host address. Generally, subnetting a network aims at the division of huge networks into smaller ones. Network address recognizes a node in a network. They are unique within the network and there can be more than one network address

within any network. A broadcast address is utilized to transfer messages and data packets to network systems. Network administrators confirm successful data transmission through them. Dynamic Host Configuration Protocol (DHCP) and Bootstrap Protocol (BOOTP) clients use them in order to locate and transmit respective server requests [90].

After the arrangement of the NIC, it will reply to the specific network addresses, that are located in the same network as determined by the subnet mask and network address. Generally, that is the way that packet sniffing works and it can be categorized into three stages. Firstly, it is the collection stage, where software collects all data that have been transferred across the specific network each time. After that, this data is converted into a readable form by people, and finally, the presentation of this data is taken place in order this data to be analyzed, using some analysis methods on it [90]. Below, we are going to present more details about these steps.

Collection

This is the initial step of packet sniffing process technique, the collection of raw data from the packets that are transmitted over the network. The sniffer changes the particular network interface to a promiscuous mode, in which data packets ~~from host on that system~~ can be captured. When this mode is disabled, capture is done to only the packets, which are addressed to a specific interface, whereas when it is enabled, capture is performed to all the packets received on a specific interface. The receiving packets by the NIC are stored in a buffer and then processed [90].

The location where the packet sniffer is placed, known as tapping the wire or getting on the wire, is also a very important task in order for investigators to work in the most efficient way. This task of finding the right physical place for the sniffer to be located is considered as difficult as the analysis of packets task is, as in case where this is placed in wrong location no packets can be fetched. As we mentioned before, the NIC should be in promiscuous mode for the process of capturing, but this is something that most operating systems do not give the permission to forensic investigators to execute, unless they possess specific privileges, so no packet sniffing can be performed on that network. That is why sniffing is much easier when hubs are installed in the network, as in that case, when traffic is sent over a hub, it crosses to every port that is connected to the hub. As a result, if someone connects the packet sniffer to an empty port of the hub, he/she can receive packets which are transmitted across the hub [90].

The most famous form of network is a switched network, which offers broadcast, multicast and unicast traffic, as well as full duplex communication, where the host system has the ability to send and receive packets at the same time. This makes things more complex as far as the setting up of the packet sniffing tool in such an environment, as well as the visibility window for the packet sniffer, which is far lower here, as the traffic that can only be captured is the one that is sent to the broadcast address and the host machine [90].

So, in a switched environment data can be captured in three ways: port mirroring, ARP Cache poisoning and hubbing out. In the first case, which is the simplest above all, what is required is the access to the command line interface of the switch by the forensic investigator or the network administrator, who will enter a command through which the switch can copy traffic from one port to another [90].

Hubbing out is another way of capturing data in a switched environment, where the target device and the analyzer are placed within a network by connecting them directly to the hub, so a hub and some cables are required for that connection from the investigator or administrator. The procedure is the following: initially, unplug the host from the network and after that plug the target and the analyzer to the hub. After that, connect the hub to a network switch, given that way the data to be captured to the hub and the analyzer at the same time [90].

As we have already mentioned, IP addresses and MAC addresses are used together in order data can be transferred and the ARP protocol is used for the translation between these two addresses. When a computer wants to send data to another device, an ARP request is sent to the switch, followed by an ARP broadcast packet sending to the systems that are connected to the computer. The device that possess the same IP address responds to the request and sends its MAC address, which is stored in the cache in order not to send a new request in case this data is used in the future. That is the last way for capturing network traffic, which is called ARP cache poisoning or ARP spoofing [90].

Conversion

The data that is captured in the collection process is at a form which is not readable by human, so it must be converted into a form that network administrators can understand, extracting important information. The job of the most packet sniffers stops here, leaving the rest to be executed by the forensic investigator or network administrator [90].

Analysis

This is the last step of packet sniffing method, where the converted data can be analyzed (as stated above) to help administrators or investigators to collect valuable information by comparing many packets to watch the moves on the network, emphasizing at the right packets which are surely captured with all the above methods. The network problems can be analyzed and required actions must be taken by the network administrators in order to stop further appearance of other network problems. Analysis of data has the goal of recognizing and investigating the digital content in order to maintain and regain the original data that is present, helping network administrators in many cases [90].

4.6.4 Network Sniffing and Packet Analyzing Tools

Below, some basic network sniffing and packet analyzing tools are presented, although the list is expanded with many more.

Tcpdump

It is a free software packet analyzer tool for capturing, filtering and analyzing network traffic through the command line [89]. It is used to view TCP/UDP connection settlement and expiration, as well as to permit the user to expose TCP/IP and other packets being transferred or received over a network to which the device is connected. It supports the most Unix-like operating systems, some of which are Linux, Solaris, BSD, macOS, HP-UX, Android and AIX, where tcpdump utilizes the libpcap library for capturing packets, whereas the Windows version is called WinDump, which uses WinPcap library, the Windows port of libpcap [89].

Most of the times tcpdump is used for two mainly reasons. First, it is utilized to facilitate on-the-fly analysis for troubleshooting network issues in a tactical way, which contains the procedures of capturing, filtering and analyzing being executed at the same time. But, this option is suitable for cases that only a quick glance at the data is sufficient [89]. The second use of tcpdump is capturing ~~interested~~ traffic passing on a target segment over a longer period of time, storing it for offline analysis with the possibility of correlation with other data. However, an important drawback of tcpdump is the size of that collected data, as its volume can be huge, depending on the throughput and usage of the network segment and the amount of each packet maintained [89].

Wireshark

Wireshark, originally known as Ethereal, is an open-source GUI-based packet and protocol analyzer tool [91], used for capturing, filtering and analyzing network traffic [89],

although there is a terminal-based version (non GUI), called Tshark. It is a cross-platform, which runs on Linux, Windows, mac OS, BSD, and other UNIX-like OS, using the libpcap (or WinPcap), as the Tcpdump does, for capturing packets from the network. Wireshark gives the user the opportunity to browse packet data from a live network, though it can also read a capture file that was previously saved [35].

Wireshark's vital advantage is that it supports many platforms, operating systems, media types and protocols, which are grown in time, as new updates are released [90]. The main reason why the number of protocols is increasing is the open source nature of the tool, since a developer has the chance to add his/her protocol into Wireshark, after the approval of this code from the Wireshark development team [90].

Another reason why Wireshark is the top packet sniffer tool is ~~how easy it can be used~~, as its graphical interface is very simple and available to the public [90]. The menus are understandable with an easy layout and raw data ~~is appeared~~ graphically, which makes it easier for a new unexperienced user to deal with this tool. Moreover, Wireshark's forensic investigator or network administrator community is considered as the best among other open source projects, as it offers a suitable program support with all the latest updates as well as FAQs, too [90]. We are going to get into details of that tool in the next chapter.

NetworkMiner

NetworkMiner is a Network Forensic Analysis Tool (NFAT) mainly for Windows, but also for Linux, Mac OS and free BSD. It has the ability to detect Operating Systems, host names, IPs, open ports, sessions, etc. without putting any traffic on the network [92]. It can take out files transferred over the network, but it has also the ability to parse pcap files for off-line analysis, regenerating transmitted files and certificates from pcap files. NetworkMiner is also easy to be installed and ~~be~~ used by its providing interface [90].

~~In contrast to Wireshark, which is considered as an active sniffing tool, NetworkMiner is a passive packet sniffing tool. The difference between these two is that in active sniffing, the sniffing tool sends the request over the network and uses the response to capture packets, whereas a passive sniffing tool just scans the traffic without getting noticed in the network, without sending a request for receiving a response [90]. Another difference between these two is that a passive sniffing tool gathers data about the hosts, known as host centric method, rather than gathering data about the traffic [90].~~

NetWitness Investigator

The NetWitness Investigator is a packet sniffing tool, which was originally used only with critical environments but after long the free version of the software was released, giving to public the opportunity to use it. The investigator captures every packet traveling through the network from both wired and wireless network interfaces, ~~emphasizing into~~ the data that is contained inside the packets [90]. It works together with most of the crucial packet capture systems but it also has some more interesting characteristics as it organizes the report in a way that users can quickly ~~reference~~. In particular, it analyzes the data in layers of networking, from users email addresses, files, full content searching, exporting the information collected in PCAP format, IPv6, which is the replacement of IPv4 and is very important for a tool to support every new requirement, and others [90].

One crucial characteristic of NetWitness Investigator tool is that it does not alert forensic investigator or network administrator for troubles in network based on familiar threats, but it captures packets in real time, examining the network for differences in behavior and reports the same to them that exact time. ~~A quite configuration support is needed in order someone installs that tool.~~ It mainly works in Windows OS for free, but its commercial version supports Linux, too, which has more advantages compared to the free one. Moreover, some of the characteristics only existed in enterprise version, such as support for Linux platform, remote network monitoring, informer, decoder and automated reporting engine [90].

Kismet

Kismet is a wireless network detector, packet sniffer and IDS, which works with any wireless adapter that supports raw monitoring mode [93]. It is also free software and supports ~~the most OS, like~~ Linux, BSD, MacOS, Windows. In contrast with other wireless network detectors Kismet passively gathers packets without interfering in the network traffic [35] that's why it is ~~the most common used one~~. Particularly, it can detect the existence of both wireless access points and wireless clients, making the connection between them. Some of its characteristics are that it contains main wireless IDS features, like detecting active wireless sniffing programs, as well as a number of wireless network attacks. It also has the capacity to log all sniffed packets, saving them in a pcap file format.

Ngrep

It is a free software network packet analyzer tool that can look for traffic based upon specific regular expressions or particular strings or patterns anywhere within the payload of the packet [35]. It has the ability to write out the packets that suit to a separate file. Ngrep supports many protocols, such as IPv4-IPv6, TCP, UDP, ICMP, etc., printing out synoptic details like IP addresses or port numbers for matching reasons.

EtherApe

It is a graphical packet sniffer/network traffic monitoring tool for Unix. It exposes network traffic operation graphically by featuring link layer, IP and TCP modes with a color coded protocols display. It has the ability to filter traffic to be shown and can read traffic from a file as well as live from the network. EtherApe is a free, open source software [86].

Snort

Snort is an intrusion detection system (IDS) used for IPbased networks [94]. It examines network traffic in order to detect worms, vulnerability exploits, port scan or any other suspect movements. Mainly, Snort works three stages: the first one is the sniffer mode, where network packets are read and exposed on a console, the second one is the logger mode, where those packets are logged and stored to the disk and the last one, the intrusion detection mode, where network traffic is analyzed based upon specific rule sets. These rules can be designed by users, too, checking many attributes of packets by that way to conclude if the traffic should be allowed or not [35].

5 Investigating Network Traffic

As we have explained in previous chapters the importance of analyzing network traffic, we can now get into practice. In this chapter, we are going to define initially the scope of that dissertation, which is the investigation of network traffic in VoIP applications (~~famous or not~~), sharing our results, giving a general conclusion. Wireshark was used as the main tool for that scope, so there is a brief description about how we have installed it, as well as its user interface essentials. Then, some examples on specific VoIP applications are presented along with their results, considerations, problems.

5.1 Problem definition: Investigating network traffic in VoIP applications

Voice over IP or VoIP is the transmission of voice and multimedia context (such as writing messages or sharing images, videos, links, etc.) over Internet Protocol (IP) networks. VoIP is enabled by a group of technologies and methodologies responsible for the delivery of voice communications over the internet, LAN or WAN, as well as other services, as we stated above. The main benefit of VoIP is that people can communicate cheaper with each other (sometimes even totally free) all over the world. They can send messages or perform video calls, with only the requirement of having a high speed internet connection. There are a lot of VoIP protocols, however the two most important have been the SIP and the H.323, although the use of the second one is limited due to its complexity in contrast to others ~~newly~~ protocols. SIP (or Session Initiation Protocol) is a connection management protocol developed and supported by the IETF, while H.323 is one of the first VoIP call signaling and control protocol developed by the telecommunications companies. Its use is limited to carrying existing long-haul network traffic, as new protocols have been arisen, less complex, such as MGCP and SIP [95].

In its contribution to offer lower cost services, VoIP has ~~appeared~~ some vulnerabilities, leading to exploitations, some of which have been remote eavesdropping, VoIP spamming, VoIP phishing, risks when attacks on the network occur, etc.

The target of ~~that~~ dissertation is capturing network packets while using these applications with Wireshark, and analyzing them to see what and if any content can be revealed.

5.2 Contribution

Before beginning the capturing and analyzing of packets, it is essential to present some details about Wireshark, as well as some actions that we have taken in order to set up our environment. After that, we will present the results from each one VoIP application we have tested.

5.2.1 Installing Wireshark on Windows 7

Wireshark, as stated previously, is the most famous packet sniffing and analyzing tool, that's why this is going to be used in order packets can be captured during the scope of that dissertation, while VoIP applications are used. Its installation is simple, since it is almost the same as any other software installed in Windows and the demanding system configuration is minimal [91]. Let us start our journey to network analysis using Wireshark.

First of all, we need to set up the Wireshark environment on our system, and particularly on a 64-bit Windows 7 operating system. The steps that we have followed for this scope are the following [96]:

- Step 1: Wireshark needs enough free *disk space*, 60 MB at least, which also uses for storing capture packets. It also requires free *memory space* of at least 256MB and a minimum of 400 MHz of processor speed. The system should also have WinPcap capture driver, which is the corresponding library for Windows as we've stated in previous chapters and a network interface card (NIC) that supports promiscuous mode. So, we used the version 4.1.3 of the WinPcap library.
- Step 2: The most convenient way of downloading Wireshark on Windows is through downloading a compressed package from the official website (<http://www.wireshark.org>), after checking the version of the owned OS.
- Step 3: Now that we are sure about the configuration of our system, we can proceed into the next step, which is the installation and it takes only few minutes to be completed. We got the *Version 2.4.2 (v2.4.2-0-gb6c63ae086)* of Wireshark. After that, all we had to do was to double click the executable file for the installer to open up, accept the terms and conditions and select the components that needed to be installed along the tool, such as the WinPcap option, installation of which started sometime after the Wireshark main installation.

5.2.2 Wireshark user interface essentials

Before we get into more details, let's see some essentials options about the Wireshark interface. Once the installation is completed, someone can open the tool and select interfaces from which he/she wants data to be captured, as it is shown in the figure below:



Figure 12: Wireshark's capture option

With that way someone can perform his/her data capture using Wireshark, where as we can see in the below figure, the main window is full of data that have been collected and presented to the forensic investigator or network administrator. So, it is necessary some descriptions to be given for that interface in order to get into more details for the analysis process. There are eight sections or elements of the default Wireshark user interface, as shown in the figure, which we are going to present in details below [97]:

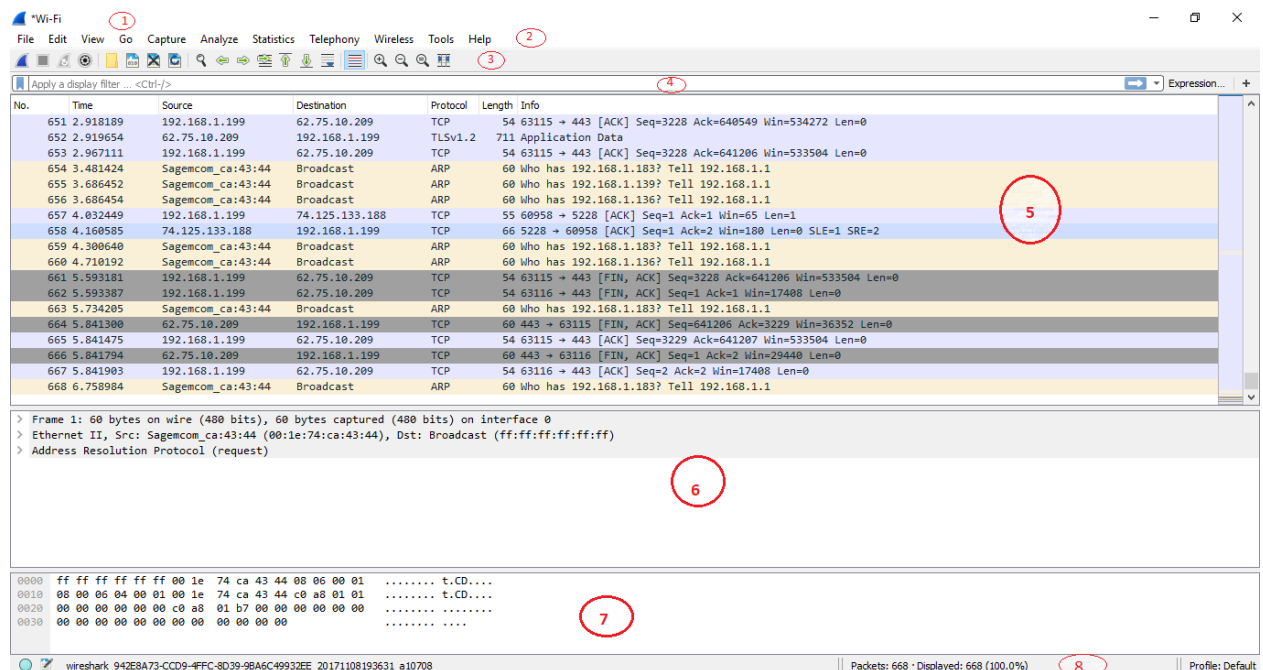


Figure 13: The eight sections or elements of Wireshark interface

Title (1): In this area the default title of Wireshark is depicted along with the current version of the tool and the interface from which the capturing has been made. In case an

already stored pcap file is opened with the tool, someone can see the name of that file in this area.

Menu (2): This is the fixed row of main features that Wireshark offers, all categorized under suitable titles.

Main toolbar (icons) (3): This area provides a quick access to the most frequently used functions of the tool. In case the capturing has not been started yet some icons may be grayed out, but after the capturing they are available for use.

Display filter toolbar (4): Filters help investigators to isolate the packets of their interest, getting the information that only matters for their analysis. There are two ways of filtering in Wireshark, *display filters* and *capture filters*. Display filters determine which frames are displayed and ~~be~~ shown in the packet list after they are captured, whereas capture filters act on the capture process, dropping only packets that correspond to the specific rules supplied, displayed them later in the packet list [98]. The syntax of the two types of filters is also different, as display filters use a logic syntax, ~~most famous programming languages will recognize~~, while capture filters use the so-called Berkeley Packet Filter (BPF) syntax, like some other tools use too, such as tcpdump, Tshark. In general, someone uses capture filters, when he/she needs limited amount of network data been processed, displayed and later saved, and display filters to drill down into only the packets he/she wants to analyze after the data has been processed [98]. The *Expression* option helps also in the creation of filter expressions easily, without the need of memorizing them.

Packet list pane (5): This list provides the investigator with a full list of all the captured packets in a sequential order along with valuable information for each one of them. Each row on that list corresponds to a single captured packet, while each column offers additional information about each one of that packet. Particularly, the columns correspond to the following information [96]:

- No.: this is the packet sequence number and is used for identifying the packets uniquely.
- Time: it provides the timestamp when a packet is captured by the tool and the time difference between the reception of the packets.
- Source: this is the IP address from which the packet is coming.
- Destination: this is the IP address where the packet is going to.

- Protocol: this is the protocol that each specific packet used.
- Length: this offers the size of the packet.
- Info: this is a brief summary of the context of each packet.

Packet details pane (6): This pane provides valuable information for the selected packet in the packet list pane, which is divided into sections for the various protocols contained in that packet.

Packet bytes pane (7): This pane shows the raw data in the actual form that was originally captured in hex bytes and ASCII form, which sometimes may be helpful.

Status bar (8): This section shows information, such as the count of packets, the file location where the captured pa

ckets are stored, current configuration profile. It also provides an expert info indicator along with options such as editing or adding capture comments.

5.2.3 Establishing an Access Point

First of all, we have set up an Access point, using the wireless interface which supports AP mode, connected on a laptop. The laptop had already an Ethernet interface connected to an ADSL router, bridging the Internet access to the wireless access point. Two Android mobile phones with version 4 and 5 respectively, were used in our set up, which were connected to the wireless access point.

In order to do this, we followed the next steps: clicked on the *Start* button and entered `cmd.exe` in the *Search Programs and Files area*, then right click on `cmd.exe` and selected *Run as administrator* from the menu. This opened a DOS prompt with administrator privileges, necessary to execute the CLI command. A single command is required to create the Windows 7 access point and here it is: `netsh wlan set hostednetwork mode=allow "ssid= Wireless Network Connection" "key=12345" keyUsage=persistent`, where the *ssid* parameter configures the ssid that will be broadcast by the Windows 7 operating system and the *key* parameter defines the WPA2 Personal key (password) that the clients need to enter in order to connect to the Wi-Fi network. Next step is to start the hosted wireless network. The command to start the hostednetwork is `netsh wlan start hostednetwork` and needs to be run as administrator, run in the same DOS prompt previously used. When executed, the above command created the required Microsoft Virtual WiFi Miniport adapter, setting up the hostednetwork. The

new Microsoft Virtual WiFi Miniport adapter is visited now in the *Network Connections* panel as shown below.

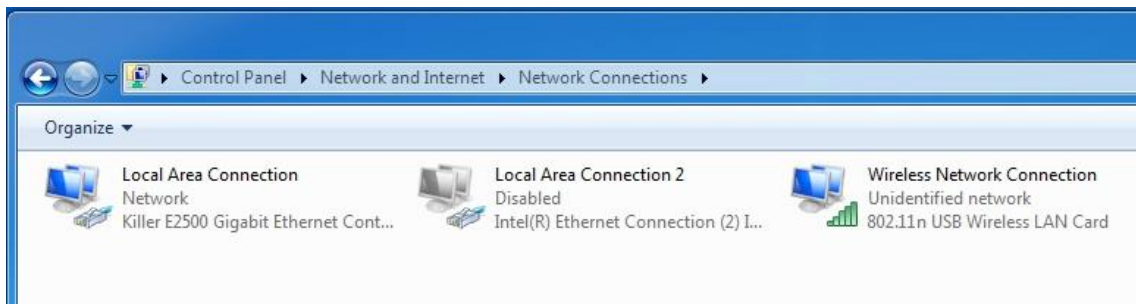


Figure 14: Network Connections

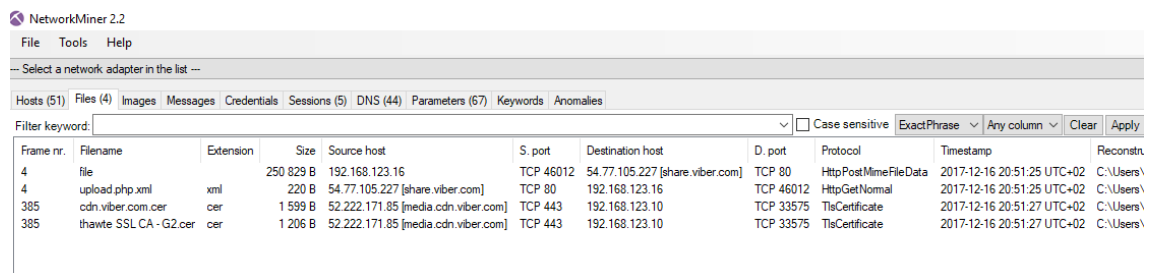
With our hosted network initiated, all that's required was to enable Internet Connection Sharing on Windows 7. This forced our newly created hosted network (access point) to provide Internet and DHCP services to our wireless clients. To enable Internet Connection Sharing, we went to the *Control Panel > Network and Internet > Network and Sharing* and selected *Change Adaptor Settings* from the left menu. Right-click on the computer's LAN network adaptor (usually Local Area Connection) and selected *properties*. Next, selected the *Sharing tab* and enabled the *Internet Connection Sharing option*. Under Home networking connection selected the newly created wireless network connection, in our example this was Wireless Network Connection, and untick *Allow other network users to control or disable the shared Internet connection setting*. At this point, our Windows 7 system has transformed into an access point and is ready to serve wireless clients. So, now we can proceed to capturing from VoIP applications.

5.2.4 Capturing traffic from VoIP applications

Below, there are some examples of VoIP applications from where we got traffic, capturing packets, trying to reveal their contents. Generally, almost all of the famous VoIP applications we have tried transmit their packets encrypted, as they want to provide privacy to their clients, so they are the only one who can see what they have sent or received. Even in the case of applications that did not use encryption for the transmission of their packets in the previous years, nowadays they have inserted encryption, which means that packets' content cannot be easily revealed, as cryptography is strong enough in order this privacy to be ensured. The results from the research we have done are shown below.

Viber

It is one of the most popular cross-platform instant messaging and VoIP application, offered as freeware for almost all platforms, Windows, MacOS, Linux, Android, iOS, given people the ability to chat, make calls, exchange images, videos, etc. Its desktop version uses TCP and UDP ports 5242, 4244, 5243, 9785 and the standard HTTP/HTTPS ports 80 and 443. In 2016, end-to-end encryption was added in the latest platforms, meaning that data of all types is encrypted from the point it is sent until it reaches the recipient and it cannot be picked by anyone else in the middle. Particularly, the platforms that support encryption are: Windows 10, iOS/Android and Desktop version 6.0 and newer, according to the official web page on Viber. So, in our environment, Windows 7, there are actually some results that can be revealed with an older version of Viber used, version 4.3.0.712. First, we have done the capturing with Wireshark, while one user has sent an image to the other. The content of this picture was not shown in the Wireshark, but when we inserted that pcap file in the NetworkMiner tool, there were some results actually. Particularly, we went to the files options, as it is shown below:



Frame nr.	Filename	Extension	Size	Source host	S. port	Destination host	D. port	Protocol	Timestamp	Reconstr
4	file		250 829 B	192.168.123.16	TCP 46012	54.77.105.227 [share.viber.com]	TCP 80	HttpPostMimeFileData	2017-12-16 20:51:25 UTC+02	C:\Users\
4	upload.php.xml	xml	220 B	54.77.105.227 [share.viber.com]	TCP 80	192.168.123.16	TCP 46012	HttpGetNormal	2017-12-16 20:51:25 UTC+02	C:\Users\
385	cdn.viber.com.cer	cer	1 599 B	52.222.171.85 [media.cdn.viber.com]	TCP 443	192.168.123.10	TCP 33575	TlsCertificate	2017-12-16 20:51:27 UTC+02	C:\Users\
385	thawte SSL CA - G2.cer	cer	1 206 B	52.222.171.85 [media.cdn.viber.com]	TCP 443	192.168.123.10	TCP 33575	TlsCertificate	2017-12-16 20:51:27 UTC+02	C:\Users\

Figure 15: NetworkMiner analysis of Viber

So, the interesting file is shown and if we right-click on it and select open folder, then another window opens that contains that file, as shown below:

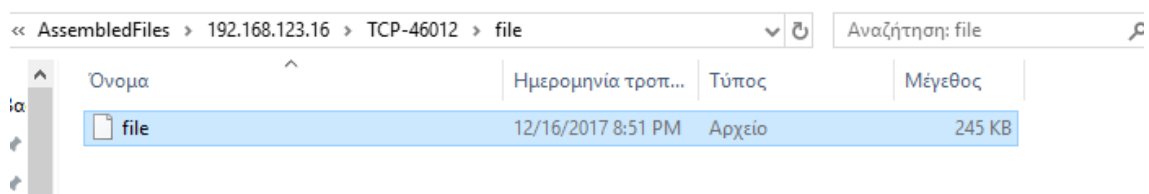


Figure 16: File location

Select it and right-click on it and choose Windows photo viewer and the image appears, as we can easily see:

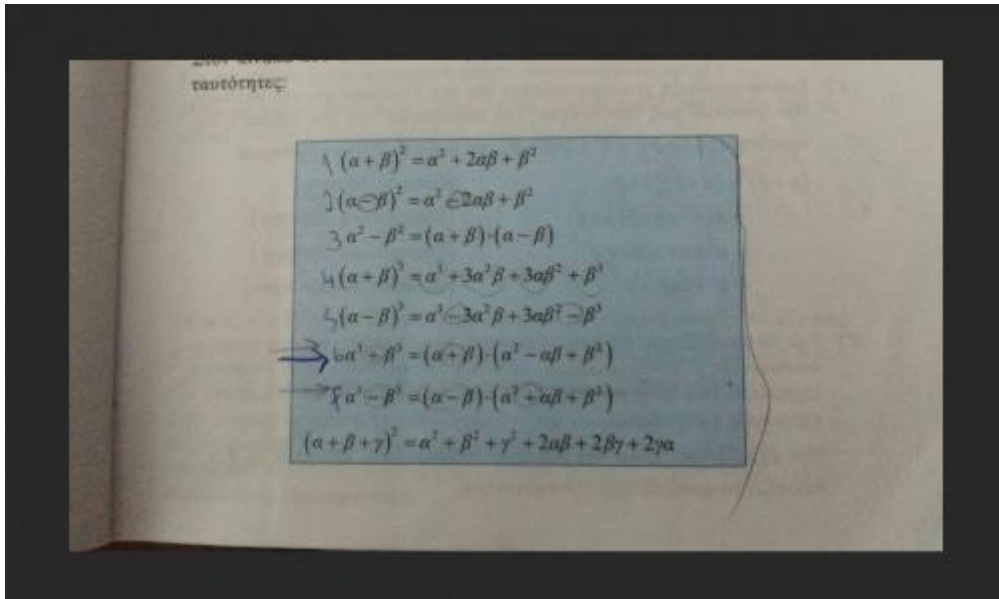


Figure 17: The sending image

Except for images, other unencrypted data that can be sent with that version of Viber are: doodles, videos, locations images.

So, Viber encrypts its messages for clients that use version 6 and so on, that's why it's so important for users to update this application, so they gain its advantages and progresses it has inserted in the app, feeling more secure while using it.

Skype

Skype is another VoIP application, offering video chat and voice calls between computers, mobile devices, tablets, etc. across the Internet, while offering instant messaging services, too. It uses a proprietary Internet telephony network, called Skype protocol.

According to Skype's own website, "All Skype-to-Skype voice, video, file transfers and instant messages are encrypted", preventing a potential eavesdropping by malicious users. It uses AES (Advanced Encryption Standard) to encrypt conversations, while RSA for key certification.

So, when capturing Skype conversations with Wireshark, the only information that we could extract was the IP addresses of the two parties that communicated, as we can see below:

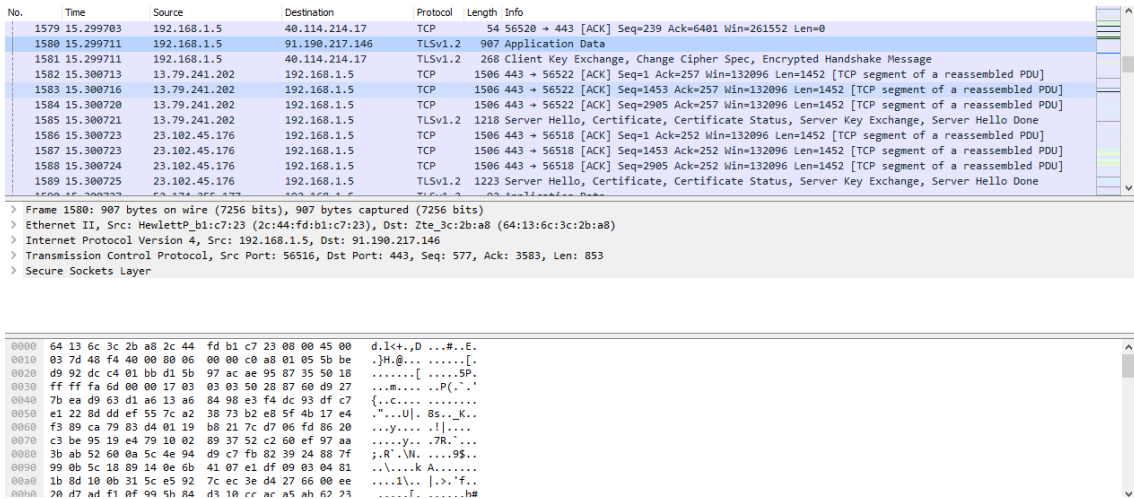


Figure 18: Wireshark on Skype

Facebook-Messenger

It is another instant messaging service, which also supports voice and video calling, one of the most popular globally used. It also supports end-to-end encryption, so again no information can be revealed with the Wireshark tool, as shown below:

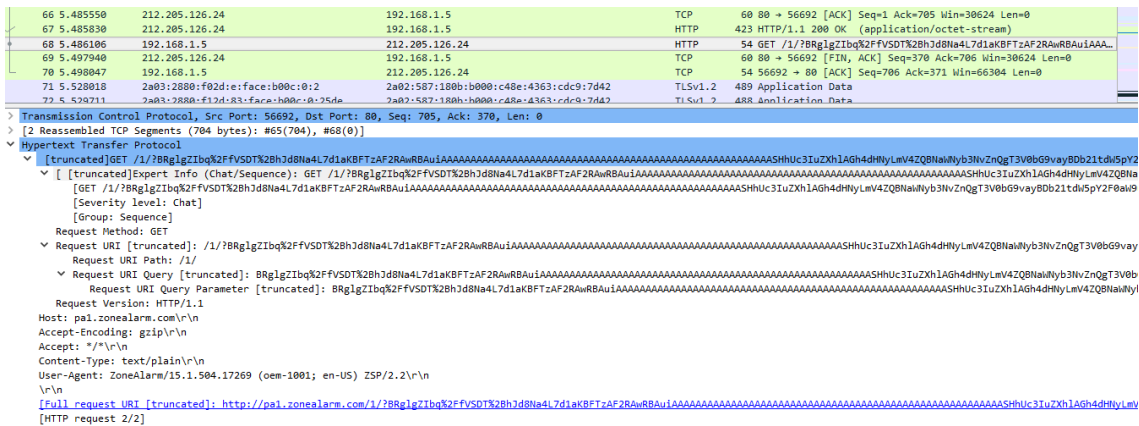


Figure 19: Wireshark on messenger

Teamtalk

TeamTalk is a conferencing system which consists of a server and a client application and people use it in order to communicate on the Internet using VoIP and video streaming. Most of the times users only need to install the client application, unless they want to run their own TeamTalk server. The TeamTalk client application supports various platforms, such as Mac OS X, Windows, Linux, CentOS, Debian, Raspbian, iOS and Android. TeamTalk is not so secure at that moment, meaning that the messages that two users have exchanged can be revealed through Wireshark. A pcap file extracted from one TeamTalk conversation through this tool is shown below:

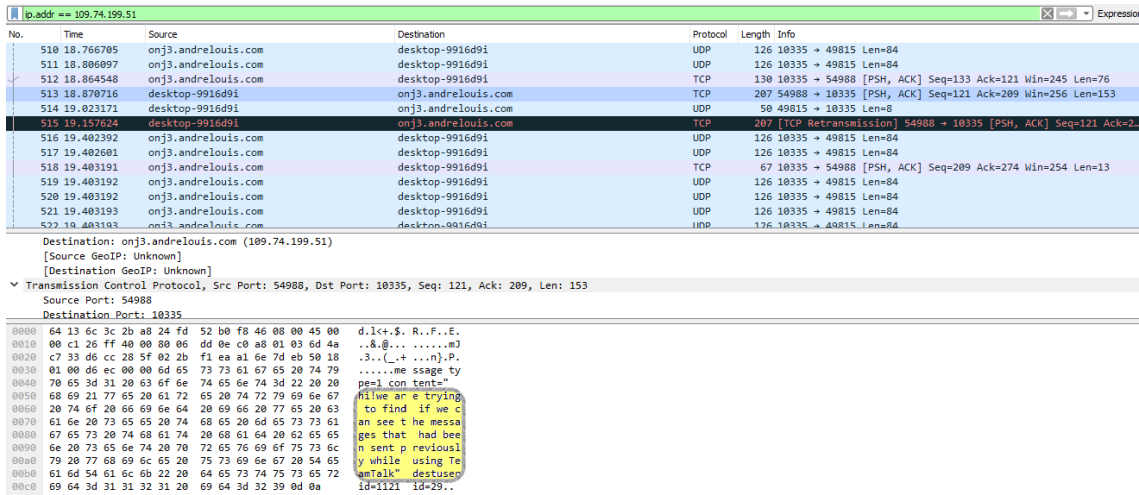


Figure 20: Capturing of TeamTalk with Wireshark and results

As we can see, a message revealed in the packet bytes pane, no 513, as well as another one at the packet no 815, which is depicted below.

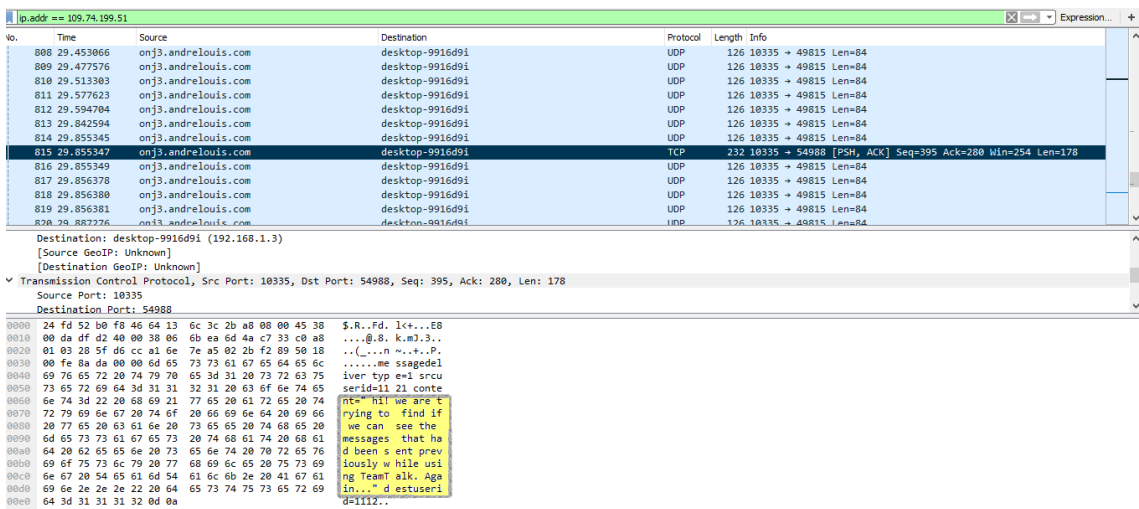


Figure 21: Capturing of TeamTalk with Wireshark and results (2)

6 Conclusion

After that dissertation has finished, we can say that cryptography is a very important aspect applied to almost all VoIP applications nowadays, making the whole concept of network forensics investigations more difficult. Manufacturers have inserted encryption to almost all VoIP applications in order their clients to feel more comfortable to exchange data over the Internet and share their personal information. With that way, the privacy of clients is ensured, strengthening the perception of more and more users to use these applications and general the Internet.

On the other hand, that movement make things more complicated in case of a cyber-crime occurs, as even the network investigators themselves face difficulties when they try to decrypt the content of the interested packets. That's why they need to use more specialized tools and other techniques to enhance their scope. As we can see the content of the packets cannot decrypt with the common used tools, such as Wireshark or NetworkMiner, nowadays.

In conclude, our initial trial was to reveal any content that is found from these VoIP applications, but most of them send/receive their packets encrypted. We found a still "open" unsecured application TeamTalk to view to our readers the results as they are without encryption. Also, we did an experiment in an older version of Viber and we observed that the results were different from the latest up to now version, highlighting how important is these apps to be updated.

Clients should always have on their mind that they need to use the latest version of all these applications in order to be as safer as it can be, as the most recently versions provide more facilities and more protected features. In case someone uses an older version, he/she might jeopardize his/her privacy, as we saw above in our experimental with the older version of Viber.

Bibliography

- [1] [Gary, P. \(2001, November\). *A road map for digital forensic research*. In *Digital Forensics Research Workshop*.](#)
- [2] [Zantko, K. \(2007\). *Commentary: Defining digital forensics*. *Forensic Magazine*, 20.](#)
- [3] Sammons, J. (2014). *The basics of digital forensics: the primer for getting started in digital forensics*. Elsevier.
- [4] Watson, D. L., & Jones, A. (2013). *Digital forensics processing and procedures: Meeting the requirements of ISO 17020, ISO 17025, ISO 27001 and best practice requirements*. Newnes.
- [5] TechTarget (2007), DEFINITION electronic discovery (e-discovery or ediscovery), retrieved from: <http://searchfinancialsecurity.techtarget.com/definition/electronic-discovery>
- [6] English Oxford Living Dictionaries, retrieved from: <https://en.oxforddictionaries.com/definition/digital>
- [7] TechTarget (2005), DEFINITION digital, Retrieved from: <http://whatis.techtarget.com/definition/digital>
- [8] Casey, E. (2011). *Digital evidence and computer crime: Forensic science, computers, and the internet*. Academic press.
- [9] Dr. Swarupa Dholam (2015), “Electronic evidence and its challenges”, retrieved from: <http://mja.gov.in/Site/Upload/GR/Article%20on%20Electronic%20evidence.pdf>
- [10] TechTarget (2014), DEFINITION metadata, retrieved from: <http://whatis.techtarget.com/definition/metadata>
- [11] TechTarget (2016), DEFINITION slack space (file slack space), retrieved from: <http://whatis.techtarget.com/definition/slack-space-file-slack-space>
- [12] TechTarget, (2005), DEFINITION swap file (swap space or pagefile), retrieved from: <http://searchwindowsserver.techtarget.com/definition/swap-file-swap-space-or-pagefile>
- [13] David Dunning, Techwalla, “What Is Unallocated Space?”, retrieved from: <https://www.techwalla.com/articles/what-is-unallocated-space>
- [14] National Forensic Science Technology Center (NFSTC), (2009), “A Simplified Guide to Digital Evidence”, retrieved from: <http://www.forensicsciencesimplified.org/digital/>
- [15] [Braid, M. \(2001\). *Collecting electronic evidence after a system compromise*. Australian Computer Emergency Response Team.](#)

- [16] Kumari, N., & Mohapatra, A. K. (2016, March). *An insight into digital forensics branches and tools*. In Computational Techniques in Information and Communication Technologies (ICCTICT), 2016 International Conference on (pp. 243-250). IEEE.
- [17] Noblett, M. G., Pollitt, M. M., & Presley, L. A. (2000). *Recovering and examining computer forensic evidence*. Forensic Science Communications, 2(4), 1-13.
- [18] Casey, E., Blitz, A., & Steuart, C. (2014). *Digital evidence and computer crime*.
- [19] Mahalik, H., Tamma, R., & Bommisetty, S. (2016). *Practical Mobile Forensics*. Packt Publishing Ltd.
- [20] Polus Scott, (2016), "Mobile Device Forensics", LTT (Law Technology Today), Retrieved from: <http://www.lawtechtoday.org/2016/12/mobile-device-forensics/>
- [21] Lord Nate, (2017), "What is memory forensics? A definition of memory forensics", Retrieved from: <https://digitalguardian.com/blog/what-are-memory-forensics-definition-memory-forensics>
- [22] Ligh, M. H., Case, A., Levy, J., & Walters, A. (2014). *The art of memory forensics: detecting malware and threats in windows, linux, and Mac memory*. John Wiley & Sons.
- [23] Alexa Jackson. (2016), "Techniques and Tools for Forensic Investigation of Email", Thechlila, Tech tips and tricks, Retrieved from: <https://www.techlila.com/techniques-and-tools-for-forensic-investigation-of-email/>
- [24] Neeraj & Beniwal .(2016). "Digital Forensics: Analyze and Monitor Network Traffic Using Sniffer (Application Software)", International Journal for Research in Applied Science & Engineering Technology (IJRASET)
- [25] Jack Wiles. (2007). "Guide to E-Discovery and Digital Forensics" ISBN 13: 978-1-59749-223-2
- [26] Nelson, B., Phillips, A., & Steuart, C. (2015). *Guide to computer forensics and investigations*. Cengage Learning.
- [27] J. Ashcroft. (2001). "Electronic Crime Scene Investigation, A Guide for First Responders". NIJ (National Institute of Justice) Guide, Office of Justice Programs, U.S. Department of Justice
- [28] Du, X., Le-Khac, N. A., & Scanlon, M. (2017). *Evaluation of Digital Forensic Process Models with Respect to Digital Forensics as a Service*. arXiv preprint arXiv:1708.01730.
- [29] Adams R. (2012). "The Advanced Data Acquisition Model (ADAM): A Process Model for Digital Forensic Practice". Thesis presented by Richard Brian Adams for the degree of Doctor of Philosophy of Murdoch University
- [30] Shrivastava, G., Sharma, K., & Dwivedi, A. (2012). *Forensic computing models: Technical overview*. CCSEA, SEA, CLOUD, DKMP, CS & IT, 5, 207-216.
- [31] Jafari, F., & Satti, R. S. (2015). *Comparative Analysis of Digital Forensic Models*. Journal of Advances in Computer Networks, 3(1).
- [32] Ademu, I. O., Imafidon, C. O., & Preston, D. S. (2011). *A new approach of digital forensic model for digital forensic investigation*. Int. J. Adv. Comput. Sci. Appl, 2(12), 175-178.
- [33] Sammons, J. (Ed.). (2016). *Digital Forensics: Threatscape and Best Practices*. Syngress.

- [34] Yusoff, Y., Ismail, R., & Hassan, Z. (2011). *Common phases of computer forensics investigation models*. Int. J. Comput. Sci. Inform. Technol., 3(3).
- [35] Joshi, R. C., & Pilli, E. S. (2016). *Fundamentals of Network Forensics: A Research Perspective*. Springer.
- [36] Reith, M., Carr, C., & Gunsch, G. (2002). *An examination of digital forensic models*. International Journal of Digital Evidence, 1(3), 1-12.
- [37] Carrier, B., & Spafford, E. H. (2003). *Getting physical with the digital investigation process*. International Journal of digital evidence, 2(2), 1-20.
- [38] [Al Fahdi, M., Clarke, N. L., & Furnell, S. M. \(2013, August\). *Challenges to digital forensics: A survey of researchers & practitioners attitudes and opinions*. In Information Security for South Africa, 2013 \(pp. 1-8\). IEEE.](#)
- [39] [Rekhis, S., & Boudriga, N. \(2010, May\). *Formal digital investigation of anti-forensic attacks*. In Systematic Approaches to Digital Forensic Engineering \(SADFE\), 2010 Fifth IEEE International Workshop on \(pp. 33-44\). IEEE.](#)
- [40] TechTerms. (2014). Encryption. Retrieved from: <http://www.techterms.com/definition/encryption>
- [41] [Shpantzer, G., & Ipsen, T. \(2002\). *Law enforcement challenges in digital forensics*. In Colloquium on Information Systems Security Education.](#)
- [42] Sundar Narayanan. (2015). "Emerging Challenges in Digital Forensics". Retrieved from: <https://www.forensicmag.com/article/2015/12/emerging-challenges-digital-forensics>
- [43] Caloyannides, M. A. (2001). *Computer forensics and privacy*. Artech House Publishers.
- [44] S. Hailey. (2002). "What is computer forensics?". Cyber Security Institute. Retrieved from: <http://www.cybersecurityinstitute.biz/forensics.htm>
- [45] ComputerForensicsTraining101.com. (2011). "Computer Forensics Definition". Computer Forensics Training101, The complete guide to learning computer forensics. Retrieved from: <http://www.computerforensicstraining101.com/what-it-is.html>
- [46] [Robbins, J. \(2008\). *An explanation of computer forensics*. National Forensics Center, 774, 10-143.](#)
- [47] Sansurooah, K. (2006, April). *Taxonomy of computer forensics methodologies and procedures for digital evidence seizure*. In Australian Digital Forensics Conference (p. 32).
- [48] Mitchell, J. (2010). *Computer Forensics: Finding & Preserving the Hidden Evidence*. Retrieved on April, 22.
- [49] Kizza, J. M. (2015). *Guide to computer network security*. Springer.
- [50] McKemmish, R. (1999). *What is forensic computing?*. Canberra: Australian Institute of Criminology.
- [51] ACPO. (2011). "Good Practice Guide for Digital Evidence", ACPO (Association of Chief Police Officers), Version 5, October 2011. Retrieved from: <https://www.acro.police.uk/home.aspx>
- [52] Infosecurity (2010) Network forensics helps bolsters confidence in cloud computing security. Available: <http://www.infosecurity-us.com/view/13252/network-forensics-helps-bolstersconfidence-in-cloud-computing-security/>. [31 Mar 2016]
- [53] Axelsson, S. (2000). *Intrusion detection systems: A survey and taxonomy (Vol. 99)*. Technical report.

- [54] Broucek, V., & Turner, P. (2001). *Developing a conceptual approach for an emerging academic discipline*. In Proceedings of the 5th Australian Security Research Symposium (pp. 55-68).
- [55] Guan Y (2009) Network forensics. In: John RV (ed) Computer and information security handbook. Morgan Kaufmann, Boston, pp 339–347
- [56] Berghel, H. (2003). The discipline of Internet forensics. *Communications of the ACM*, 46(8), 15-20.
- [57] Datt, S. (2016). *Learning Network Forensics*. Packet Publishing Ltd.
- [58] Yasinsac, A., & Manzano, Y. (2001, June). Policies to enhance computer and network forensics. In *Proceedings of the 2001 IEEE workshop on information assurance and security* (pp. 289-295).
- [59] Ranum, M. J. (1999). Intrusion detection and network forensics. In *USENIX Security*.
- [60] Garfinkel, S. (2002). Network forensics: Tapping the internet. *O'Reilly Network*. Retrieved on January, 25, 2014.
- [61] Ren W, Jin H (2005) *Modeling the network forensics behaviors*. In: Workshop of the 1st international conference on Security and Privacy for Emerging Areas in Communication Networks (SecureComm' 05), Athens, Greece, pp 1–8
- [62] M.Elavarasi, Assistant Professor,Sri Ram College Of Arts and Science, Trivellore-602001, (2016) *Network Forensics And Its Investigation Methodology*, International Journal of Emerging Trends in Science and Technology, ISSN 2348-9480
- [63] Forte D (2002) The future of computer and network forensics. *Network Security* 2002(10):13–15
- [64] Raynal F, Berthier Y, Biondi P, Kaminsky D (2004) *Honeypot forensics, Part I: analyzing the network*. IEEE Secur Priv 2(4):72–78
- [65] Raynal F, Berthier Y, Biondi P, Kaminsky D (2004) *Honeypot forensics, Part II: analyzing the compromised host*. IEEE Secur Priv 2(5):77–80
- [66] Nikkel BJ (2007) *An introduction to investigating IPv6 networks*. Digit Investig 4(2):59–67
- [67] Govil J, Govil J, Kaur N, Kaur H (2008) *An examination of IPv4 and IPv6 networks: constraints and various transition mechanisms*, In: IEEE Southeastcon 08, Huntsville, Alabama, USA, pp 178–185
- [68] Vural I, Venter H (2010) *Mobile botnet detection using network forensics*. In: Berre A, Gómez-Pérez A, Tutschku K, Fensel D (eds) Future internet – FIS 2010, vol 6369. Springer, Berlin/Heidelberg, pp 57–67
- [69] Vural I, Venter HS (2010) *Using network forensics and artificial intelligence techniques to detect bot-nets on an organizational network*. In: Seventh international conference on Information Technology: New Generations (ITNG' 10), Las Vegas, Nevada, USA, pp 725–731
- [70] Qureshi, A. (2009). 802.11 Network forensic analysis. *White Paper*, SANS Institute InfoSec Reading Room, 3-48.
- [71] Turnbull B, Slay J (2008) *Wi-Fi network signals as a source of digital evidence: wireless network forensics*. In: Third international conference on Availability, Reliability and Security (ARES 08), Barcelona, Spain, pp 1355–1360
- [72] Pelaez JC, Fernandez EB (2006) *Wireless VoIP network forensics*. In: Fourth LACCEI international Latin American and Caribbean conference for Engineering and Technology (LACCET' 06), Mayagüez, Puerto Rico, pp 1–12
- [73] Otaka A, Takagi T, Takahashi O (2008) *Network forensics on mobile Ad-Hoc networks*. In: Lovrek I, Howlett R, Jain L (eds) Knowledge-based intelli-

- gent information and engineering systems, vol 5179. Springer, Berlin/Heidelberg, pp 175–182
- [74] Yinghua G, Simon M (2010) *Network forensics in MANET: traffic analysis of source spoofed DoS attacks*. In: 4th international conference on Network and System Security (NSS' 10), Melbourne, Australia, pp 128–135
- [75] Guo R, Cao T, Luo X (2010) *Application layer information forensics based on packet analysis*. In: International conference of Information Science and Management Engineering (ISME' 10), Xian, China, pp 206–209
- [76] Nikkel BJ (2004) *Domain name forensics: a systematic approach to investigating an internet presence*. Digit Investig 1(4):247–255
- [77] Kilpatrick T, Gonzalez J, Chandia R, Papa M, Sheno S (2008) *Forensic analysis of SCADA systems and networks*. Int J Secur Netw 3(2):95–102
- [78] Kilpatrick T, Gonzalez J, Chandia R, Papa M, Sheno S (2006) *An architecture for SCADA network forensics*. In: Olivier M, Sheno S (eds) Advances in digital forensics II, vol 222. Springer, Boston, pp 273–285
- [79] Naqvi S, Massonet P, Arenas A (2006) *Scope of forensics in grid computing – vision and perspectives*. In: Min G, Di Martino B, Yang L, Guo M, Runger G (eds) ISPA' 06 workshop on frontiers of high performance computing and networking, vol 4331. Springer, Berlin/Heidelberg, pp 964–970
- [80] Garfinkel SL (2010) *Digital forensics research: the next 10 years*. Digit Investig 7(Supplement 1):S64–S73
- [81] Pilli ES, Joshi RC, Niyogi R (2010) *Network forensic frameworks: survey and research challenges*. Digit Investig 7(1–2):14–27
- [82] Leiner, B., & Rekhter, Y. (1993). *The multiprotocol internet* (No. RFC 1560).
- [83] Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P., & Berners-Lee, T. (1999). *Hypertext transfer protocol--HTTP/1.1* (No. RFC 2616).
- [84] Postel, J. (1981). Transmission control protocol specification. RFC 793.
- [85] Postel, J. (1981). RFC 792: Internet control message protocol. *InterNet Network Working Group*.
- [86] EC-Council. (2010). *Computer Forensics: Investigating Network Intrusions and Cyber Crime*. Nelson Education.
- [87] Mukkamala S, Sung AH (2003) *Identifying significant features for network forensic analysis using artificial intelligent techniques*. Int J Digit Evid 1(4):1–17
- [88] Shanmugasundaram K, Memon N (2006) *Network monitoring for security and forensics*. In: Bagchi A, Atluri V (eds) Information systems security, vol 4332. Springer, Berlin/Heidelberg, pp 56–70
- [89] Davidoff, S., & Ham, J. (2012). *Network forensics: tracking hackers through cyberspace* (Vol. 2014). Upper Saddle River: Prentice hall.
- [90] Joanna Kretowicz, “Learn packet analysis with Wireshark and PCAP analysis tools”, eForensics magazine, Issue 10/2014 November ISSN 2300-6986
- [91] Combs G (2007) Wireshark. [Online]. Available: <http://www.wireshark.org/>, 14 Dec 2014
- [92] NetworkMiner (2008) Network Forensic Analysis Tool (NFAT). [Online]. Available: <http://www.netresec.com/?page=NetworkMiner>, 23 Apr 2015
- [93] Kershaw M (2004) Kismet readme. [Online]. Available: <http://www.kismetwireless.net/>, 08 Nov 2013
- [94] Roesch M (1999) Snort: lightweight intrusion detection for networks. [Online]. Available: <https://www.snort.org/>, 09 July 2013

- [95] Psaroudakis, I., Katos, V., Saragiotis, P., & Mitrou, L. (2014). A method for forensic artefact collection, analysis and incident response in environments running session initiation protocol and session description protocol. *International Journal of Electronic Security and Digital Forensics*, 6(4), 241-267.
- [96] Singh, A. (2013). *Instant Wireshark Starter*. Packt Publishing Ltd.
- [97] Baxter, J. H. (2014). *Wireshark Essentials*. Packt Publishing Ltd.
- [98] Bullock, J. (2017). *Wireshark® for Security Professionals: Using Wireshark and the Metasploit® Framework*.