


Book 8 - Emerging Trends and Special Topics in Digital Forensics

1.	SOCIAL NETWORKS FORENSICS.....	2
2.	BIG DATA FORENSICS	24
3.	INTRODUCTION TO CLOUD COMPUTING	61
4.	THE CHALLENGES OF CLOUD COMPUTING IN DIGITAL FORENSICS	79
5.	BITCOIN AND CRYPTOCURRENCIES	94
6.	DIGITAL FORENSICS REVERSE ENGINEERING FUNDAMENTALS.....	116
7.	STEGANOGRAPHY	145

1. Social Networks Forensics

Scope															
Number	2														
Title	Social Networks Forensics														
Introduction	The growth of online social network applications has resulted on connecting huge number of users worldwide. In addition, the massive use of such applications and media sites has resulted in many types of online criminal activities. Therefore, social networks generated data need to be analyzed for forensic purposes when criminal activities take place. This chapter reviews social network forensics, social network data acquisition and retrieval, social networks tools, social networks security analysis and associated challenges.														
Outcomes	<p>At the end of this unit you should be able to:</p> <ul style="list-style-type: none"> • Explain social network forensics and investigations on digital media; • Explain social network forensic procedure; • Conduct acceptable digital investigations using potential social network forensic tools and based on the investigative process such as identification, preservation, examination, analysis and reporting; • Demonstrate clear understanding of Social Networks Security, analysis and relevant challenges. 														
Topics	<ul style="list-style-type: none"> - Introduction to Social Networks Forensics - Social Networks Forensic Procedure - Social Networks Forensics Tools - Social Networks Evidence Acquisition and Retrieval - Social Networks Security and Analysis - Challenges of Social Networks Forensics 														
Study Guide	<p>Instructions on how to study this unit:</p> <ul style="list-style-type: none"> • Required study time: <p>You should plan to spend approximately 25 hours studying this unit. You may find it convenient to break up your study as follows:</p> <table border="1"> <thead> <tr> <th>Activity</th><th>Time</th></tr> </thead> <tbody> <tr> <td>Preparation and Content Review</td><td>2 hours</td></tr> <tr> <td>Set Textbook Content</td><td>1 hours</td></tr> <tr> <td>Software/Hardware Review</td><td>10 hours</td></tr> <tr> <td>Thinking (Review questions, MCQs):</td><td>2 hours</td></tr> <tr> <td>Tutorial and Related Course Work</td><td>10 hours</td></tr> <tr> <td>Total</td><td>25 hours</td></tr> </tbody> </table> <ul style="list-style-type: none"> • Required hardware/software: <ul style="list-style-type: none"> ✓ Digital Forensics Lab. ✓ Windows OS 7 or higher version ✓ Current Web Browser. • Required external resources including links and books: <ul style="list-style-type: none"> ✓ E- Library. 	Activity	Time	Preparation and Content Review	2 hours	Set Textbook Content	1 hours	Software/Hardware Review	10 hours	Thinking (Review questions, MCQs):	2 hours	Tutorial and Related Course Work	10 hours	Total	25 hours
Activity	Time														
Preparation and Content Review	2 hours														
Set Textbook Content	1 hours														
Software/Hardware Review	10 hours														
Thinking (Review questions, MCQs):	2 hours														
Tutorial and Related Course Work	10 hours														
Total	25 hours														

Content Template	
Section Number	2.1
Section Title	Introduction to Social Networks Forensics
Introduction	<p>The large use of social networks has resulted in many benefits and user satisfactions. However, this huge success was also associated with many kinds of network crimes. Therefore, different forms of digital investigations needed. As such, social network forensics deals with collecting, investigating online social media contents and sources, continually analyzing, and preserving it as an evidence.</p> <p>In this type of investigation, the examiner will be dealing with huge amount of live data environment. In addition to the collection and classification of data from live environment, the most important part is on how to find useful and relevant data without violating the law.</p> <p>Generally, there are two types of data collection techniques. The simple one is to collect data manually such that by visiting the social network website and taking live contents and screenshots. However, this type of data collection is time consuming and may not give complete result. On the other hand, digital examiner may use aided tools or commercial solutions for the task. However, good knowledge on the capability of the evidence extraction tool is also a question. This chapter presents social network forensics, procedure, tools and demonstration of live evidences.</p>
Content	<p>Social network applications afford plenty of user-friendly services and facilities to the end-user. Most of these services tend to be social; however, commercial services and exchanges afforded by most of the social networking applications. Therefore, there has been a rapid growth in the use and development of social networking application at individual and organizational aspects.</p> <p>The wide-ranging usage of social network applications and the advancements in cloud computing, IoTs, has derived us to a sharp evolution of social networking, and one of the main concerns of such evolution is security.</p> <p>To provide security, network administrators must be aware of the local network activities and must be able to provide continuous and useful reports. The following demonstrates some common social networks environment as in Figure 1 below.</p> 

	<p style="text-align: center;">Figure 1. Common Social Network Environment</p> <p>Social network forensic is a form of digital investigation which deals with the collection, analysis, interpretation and demonstration of digital evidences from social events, social web sites and social applications that can be presented in a court of law. Social networks forensics exists as an art of collection, interpretation, preserving, and demonstration of social networks digital data as evidence using potential network forensic tools.</p> <p>Investigating social network applications content is not simple task and time consuming for many reasons such as:</p> <ul style="list-style-type: none"> - Collection and preservation of the evidence need deployment of potential techniques and tools. As such, identifying technology and other needs are potential for the effectiveness of the acquired evidence. - The dynamic nature of the environment, which means the data is changing constantly - The lack of control - Recovery of information challenge, for example the data and information are retrieved from live environment - Legal constraints of retrieving evidence from social networks <p>The following list some of the popular social networking websites and associated features:</p> <ol style="list-style-type: none"> 1- Facebook A social networking site that helps to connect and make friends, upload photos, post videos, get news, tag friends, and view friend's status. 2- Twitter A social networking site that helps to keep people informed on what you are up to do next. Updating personal status with limited characters and following others. 3- LinkedIn LinkedIn is a business-oriented social networking site. It is usually used to build professional network, and update career profile. 4- MySpace A social networking site, which helps to create user profile page to meet new friends. User can post videos, movies, news, and blog. 5- Instagram Instagram is a photo and video-sharing social networking application. The app supported by the email and used for different purposes such as business and marketing needs, advertisement and social networking. <p>Misuse of social networks applications may occur in many different forms such as misuse organizational and individuals' data violate privacy, defamations, and inappropriate dissemination of contents. In some cases, social networks investigators may face limited access to the data sources due to the visibility of the server hard derive or leveraging from the service operator directly. As such, access to the evidence may be restricted because of Term of Use and of the impossibility, excluding in case of exceptional gravity (e.g. International Terrorism) to get any kind of support directly from the social network provider. Furthermore, with the vast amount of social networks data, forensic investigator may on the other hand use tools to visualize collected social networks data, and derive visualization graphs to answer questions of interest. Although it is hard to read the basic social visualization graphs, it may just give clue to the investigator on particular sources of data. In addition, it is</p>
--	--

	<p>necessary to examine the contributing devices used to upload material on social applications.</p> <p>As there are many entities involved, social networks forensic examiner must plan and follow the formal procedures in retrieving and analyzing the evidence. The following section describes social networks forensic procedures.</p>
--	--

Content Template	
Section Number	2.2
Section Title	Social Networks Forensic Procedure
Introduction	<p>This section introduces the implementation of social network forensic procedures. Although there is no common standard for such implementation, digital forensic investigation team may start collecting volatile data, capture system image, network history and logs, video, time, hashes, screenshots, consider talking to witnesses, preserve a chain of custody and be familiar with big data analysis.</p>
Content	<p>There exists a number of different process models and frameworks for social network forensics. Overall, it seems there are no standard procedures in digital forensics investigation for social forensics. However, social networks forensic procedure for digital evidence involves a number of informal steps. These include the data acquisition and retrieval, examination, and presentation of data. As such, the U.S. Department of Justice published an investigative process model in the electronic crime scene investigation.</p> <p>Accordingly, the process consists of four major phases: collection, examination, analysis and reporting. Generally, social network investigation procedure includes the following steps:</p> <ul style="list-style-type: none"> - Timeline construction - Formal and informal interviews - Documentation - Collection of suspect records which may include, personal computer records, mobile phone records, and social media site records - Evidence interpretation and demonstration. <p>In Selamat, Yusof, & Sahib study, the research presented the existing investigation frameworks and merged the same activities and processes that provide the same result into following five common phases as per the following:</p> <ol style="list-style-type: none"> 1- Preparation for capturing social network evidence: this phase includes planning, authorization, warrant, notification, and confirmation. 2- Collection and Preservation of social network evidence: this phase produce details of the crime type, potential evidence sources, media, devices, and events. 3- Examination and analysis of social network evidence: Examination and analysis phase has the files, log files, events log, data, and information 4- Presenting the social network evidence: Presentation and reporting phase has a list of evidences and relevant reports. 5- The Disseminating phase: Provide evidence explanation, new policies and investigation procedures, evidence disposed, and closing of investigation.

Content Template					
Section Number	2.3				
Section Title	Social Networks Forensics Tools				
Introduction	<p>The use of technology and tools is very important in this sort of investigations. Using forensic tools provide digital examiners an effective, and solid investigation results. Although there are some basic forensic tools exists since decades, digital investigation without proper and up to date tools is no longer satisfactory. This is due to the advances in the computing hardware and software technologies such as the hard drive capacity and processing speed are always on action.</p> <p>This chapter introduces an example of digital forensic tools that used to support evidence extraction in social networking environment.</p>				
Content	<p>Social network investigation is unpredictable and treated as discontinuous process. This fact is due to the contrast number of the computing devices potentially used to update material on social networks applications. Therefore, social network forensic examiner needs to take into account range of devices for examining an incident and deriving relevant and accurate evidence list.</p> <p>On the other hand, the purpose on which the examination carried on may require particular formal procedure to follow. As such, the investigation procedure would include devices such as personal computers, laptops, mobile telephones, tablets or even server computers. The investigation may initially focus on the information, which derived from the cache and Internet history files and cookies.</p> <p>Although most of the web browsers save enough information about the browsing activities, they provide challenge to the forensic examiner when it comes to choosing the most relevant system and network based artifact and artifact possible location. Therefore, the forensic examiner must be able to determine the primary acquired data locations such as the browsers cache and use well-known and reliable forensic tools.</p> <p>The following table present an example of digital forensic tools that can be used and support evidence extraction in social networking environment. Some of the given tools below not specifically designated for social network forensic as such, however, they can be used in the overall network based forensic and help in supporting the relevance and the nature of the gathered social network forensic evidence and as per the description of each tool below.</p> <p>Table 1. Social and Network Forensic Tools</p> <table> <tr> <th>Forensic Evidence Extraction Tools Name</th><th>Description</th><th>Platform</th></tr> </table>		Forensic Evidence Extraction Tools Name	Description	Platform
Forensic Evidence Extraction Tools Name	Description	Platform			

	CacheBack	CacheBack is an offline forensic investigation tool which can support the investigation of social networking events such that lets the examiner to rebuild cached web pages and analyze Internet histories all in one very intuitive and productive interface.	Windows	
	Internet Evidence Finder	Internet Evidence Finder is a digital forensics software solution, which can support the investigation of social networking events, used by thousands of forensics professionals around the world to find, analyze and present digital evidence found on computers, smartphones and tablets.	Windows	
	EnCase	EnCase Forensic can support the investigation of social networking events and enables forensic examiners to quickly search, identify, and prioritize potential evidence, in computers and mobile devices.	Windows	
	Forensic Tool Kit	FTK is a well-known forensic tool, which used by government agencies and law enforcement around the world.	Windows	
	SafeBack	Digital media collection and create bit by bit backup	Windows (DOS)	
	GraphViz	Graphviz is open source graph visualization software.	Linux, macOS, Windows	
	Helix3 Pro	Provides incident response and computer forensic tools. Allows to make forensic images of all internal devices and physical memory	Linux	
	Gephi, Socilyzer, Netlytic, UCINET	Visualization and exploration software for various kinds of graphs and networks.	Linux, Windows, macOS	
	Paraben	Provides digital forensics solutions for portable devices such as mobile phone or PDAs. Also supports hard drive and network evidence acquisition.	Windows	

	Oxygen forensic suite	Mobile forensic software that goes beyond standard logical analysis of cell phones, smartphones and PDAs.	Windows	
	Cellebrite	Provides extraction and analysis of invaluable evidentiary data from mobile devices.	Windows	
	XRY	XRY is a complete digital forensics system for mobile devices that used on any Windows PC. XRY can recover data from thousands of different mobiles including deleted information.	Windows	

Content Template	
Section Number	2.4
Section Title	Social Networks Evidence Acquisition and Retrieval
Introduction	<p>Data gathering is first initial step in any sort of digital investigation. The investigation process may take several steps including data acquisition from online sites targeting potential evidence. The analysis includes mapping the gathered evidence based on certain relationship or connection. During the evaluation period, the investigator validates the input and make sure it just meets the requirement. In fact, such operation is time consuming and difficult due to the diversity of data and the relevant network connection. Therefore, digital network investigator must follow very formal and sophisticated method for social networks evidence acquisition and retrieval. This section illustrates social networks evidence acquisition and retrieval.</p>
Content	<p>The following illustrate common and best practices to retrieve reliable data and information from social networks sites.</p> <p>A. Data Acquisition and retrieval</p> <p>Prior analyzing the gathered data, the data should be normalized and cleaned. With the existence of huge data on social networks sites and the use of ad-hoc methods and tools and the lack of standards for data extraction, it become necessary to perform investigation, analysis and evaluation to extract reliable data and information. Traditionally, data based on collected artifacts from local cache or based on the server hard drive griped data. The following describes the artifacts that need to be collected from social network sites:</p> <ol style="list-style-type: none"> 1- User location 2- Web history 3- Web cache 4- Cookies and sessions 5- Images and video 6- Comments and reply 7- Wall post and status update 8- Messages and chat 9- Emails and Notifications <p>On the other hand, several reliable LAN data acquisition methods used such as passive sniffing using potential software tools. For example, Wireshark is a free and open-source packet analyzer, which can be used for network troubleshooting, analysis, protocol development, and more. The following list the well-known techniques that used for social networks data acquisition.</p> <ul style="list-style-type: none"> - Network traffic analysis - Ad-hoc applications - Cloud and data crawling - Application programming interface and php data retrieval. <p>B. Social Networks Data Pool</p> <p>The number of users on the social media is huge. In addition, the number of users is of increasing mode all time. This behavior can lead to variation in both features and architectures of any two or more relevant social networking sites. In addition, the more increase in the data pool the more increase in the relevancy of data. This type of data usually called big data. However, being dependent on this type of</p>

	<p>implementation, digital investigators must identify generic data sources to be of interest during social network forensic investigation.</p> <p>C. Interpretation</p> <p>The direct and safest way of data retrieval usually happen upon the consent of the social data operator. However, this is not always the case. Thus, during social network analysis, digital investigator may come out with facts and different perspectives and understanding of the investigation case. Therefore, the role of the forensic digital investigator is to interpret the acquired data from the social network events understood and used in the court of law.</p> <p>D. Inference</p> <p>After the analysis and interpretation of social media data, it is important that these data presented in an understandable format. Generally, inference refers to the presentation of any social media data in a form that understood by a common person.</p>
--	---

Content Template	
Section Number	2.5
Section Title	Social Networks Security and Analysis
Introduction	<p>Various social and personal information-sharing practices generally supported by different social networking sites. Overall, this kind of practices have led success in many ways. On the other hand, such practices controlled by many concerns and fear from security point of view. Furthermore, as most of the personal and social information kept in the databases of the social networking sites, security will remain as one of the most concerns for any social networking site. Therefore, an updated preventive and security measures will always needed.</p>
Content	<p>The following are the common different type of attacks, which will always need an up to date prevention measures on social network sites.</p> <ol style="list-style-type: none"> 1- Malware There are many different types of malware such as ransomware, which designed to control or damage a computer system. 2- Phishing Social networks are actually one of the media across which phishing campaigns are distributed. This type of attack is usually associated with email system such that a fake official email link to fake website where victims log in and giving up their own or the organization official information. 3- Man-in-the-middle attack In the Man-in-the-middle attack, hackers attach themselves between the computer system and the webserver. Hence, allow leaking social media personal and sensitive data and information. 4- Denial of service (DoS) and Distributed DoS This attack is generally associated with huge network of computer systems that overload servers with data and shutting it down. Therefore, it will disturb the service and may refute several social media sites from dynamic response. 5- Cross-site scripting This type of attack work after injecting malicious code into a social media website targeting visitor's browsers. 6- SQL injection SQL injection cause corruption of server data that on the other hand make a server reveal real data such that in using credit cards numbers and username. <p>In addition to the above mentioned, risks are also associated data and information sharing or social network sites. The highest risk is associated with the site, which are vulnerable to any of the major social network attack. The following illustrate the common threats to social networking services:</p> <ol style="list-style-type: none"> 1- Viruses Social networks are good channel to distribute malware. As such, Viruses considered as mixture of malicious software codes, which has the ability to self-replicate themselves into another computer programs targeting online users and unprotected systems. Generally, such code designed with a malicious intent to destroy the uniqueness of the operating system running programs and copy/delete or even encrypt data. 2- Software tools and application development

	<p>The availability of software tools and programs make it easy for attackers to use potential tools to target social networking sites. From this point of view, attackers can play different roles on behalf of the users, targeting different type of users on the lane.</p> <p>3- Social engineering attacks Social engineering is common term used for wide range of malicious activities using different type of psychological techniques such as pretexting and phishing. As such, the attacker generally prepares a ground for the user by the deceiving the victims intent to gain personal data or information.</p> <p>4- Identity theft Social networking sites usually considered as an open platform for this type of attacks. This type of attacks explains the use of the victim identify to gain some sort of benefit such as financial advantage. The common attack starts at the point where the attacker commences to collect some common and personal information about the victim and use it as potential user. This considered as one of the most complicated and challenging attacks specially during the detection phase. It is very hard to determine or know that personal information targeted or used to launch an attack.</p> <p>5- Third party applications Social networking site provide services to use variety of third party applications. Although these applications might not contain any sort of malicious codes, they generally used as an advantage by the attacker to access potential information from user's profiles.</p> <p>6- Public comments Public comments commonly used to enhance social media sites policy and provide better service. On the other hand, one has to be very careful as it may contain false contents used to destroy or negatively affect the structural performance of an organization.</p>
--	---

Content Template	
Section Number	2.6
Section Title	Challenges to Social Networks Forensics
Introduction	<p>The number of users of social networking sites are increasing very rapidly. This problem itself is a great challenge to store, maintain and secure data on these web based applications and services. During social network forensic process, huge data and information will be gathered and generated. In order to make use of these data, the data must be normalized and understood. This will help the digital investigator to predict the attacker behavior and even protect against any future criminal activity. However, this operation is not that easy due to many challenges. This section illustrate the most common challenges associated with social networks forensics.</p>
Content	<p>Generally, the challenges associated with social network analysis are categorized into two main categories as per the following:</p> <ul style="list-style-type: none"> A- Technical Challenges <p>The technical challenges are associated with data collection, data categorization and normalization, data storage, and the software tools used during the forensic operation.</p> B- Legal Challenges <p>The legal challenges are associated with issues related to presetting the digital evidence in the court of law. This include permissible data, permissible tools, time and date, and people associated with evidence gathering process.</p> <p>The following illustrate various common challenges to social networks forensic analysis form technical and legal perspectives.</p> <ul style="list-style-type: none"> 1- Social network evidence collection and detection <p>Collecting forensically sound evidence from social networks is critical. In the first glance, the collected evidence is always messy, has duplications, and has many connected events. In most of the cases the accumulated data representing logs of different occasions such as Firewall logs, data captured from network sites or sniffers and may even represent different geographical borders. As such, providing high standard data and relevant events is always a challenge in this case. Therefore, admitting mistakes is one of the common ethical practices for the majority of the digital forensic investigators. In addition, some challenges exist due to the access of the servers hosting social sites physically exist in different country, which means different forensic laws, thus, collecting such data or using it is difficult.</p> 2- Data fusion and Examination <p>Data fusion and examination is another critical part for social network investigator. This include the validation of data and the tools prior the investigation and presentation for evidence.</p> 3- Analysis and investigation <p>Prior the data analysis phase, the data must be cleaned and understood. This include the categorization of data and clustering different network events. Any sort of mistake or irrelevant connection may fabricate the results of the analysis and misunderstanding of the attack behavior. Therefore, the recommendation is to filter data with reference to time, date,</p>

	<p>location, and preference critical attribute prior any sort decision made. The most important part of this step is the ability to generate permissible evidence in the court of law.</p> <p>4- Response</p> <p>The modern requirement of social networking sites is the ability to update and take some sort of real-time action against any intruder in order to prevent future attacks and provide safe environments. As such, providing security, safe and responsive network environment remains a challenge.</p>
--	---

Activity	
Number	2.1
Title	Introduction to Social Networks Forensics
Type	Review questions
Aim	To explain social network forensics and investigations on digital media
Description	1- Discuss the role of social network forensics investigator. 2- Differentiate between three famous social networking sites and explain the associated features. 3- Discuss different types of data collection techniques used in social network forensics.
Timeline	Two Hours
Assessment	Classroom discussion

Activity	
Number	2.2, 2.4
Title	Social Networks Forensic Procedure, Evidence Acquisition and Retrieval
Type	Research and reflection questions
Aim	To explain social network forensic procedure
Description	<ol style="list-style-type: none"> 1- Explain briefly how to implement basic forensic procedures? 2- Discuss how to maintain the chain of custody in the case of social network forensics? 3- What do you understand by the order of volatility? Discuss how the order of volatility help the digital investigator in decision making? 4- How the forensic examiner should convince the court of law on the integrity of the acquired evidence? 5- Discuss the difference between live and post mortem social network evidence acquisitions in the case of social network forensics. Illustrate your answer with example and demonstrate when each is necessary and how it can be done?
Timeline	Four Hours
Assessment	Classroom and Lab discussion

Activity	
Number	2.3
Title	Tools
Type	Research questions
Aim	Conduct acceptable digital investigations using potential social network forensic tools based on well-known investigation process such as identification, preservation, examination, analysis and reporting;
Description	<ol style="list-style-type: none"> 1- Conduct a research to conclude to what extent social network security analysis, techniques and tools used to assess and identify potential groups on any social network site. Show practical implementation using at least three different social networking forensic tools to support your conclusions. 2- Select three different digital forensic tools under each of the following category. Compare these tools from the purpose, capability, platform, and license point of view. <ol style="list-style-type: none"> a- Computer forensic tools b- Mobile device forensic tools c- Social networking sites forensic tools
Timeline	Eight Hours
Assessment	Lab discussion

Activity	
Number	2.5, 2.6
Title	Social Networks Security and Analysis
Type	Research and reflection questions
Aim	Demonstrate clear understanding of Social Networks Security, analysis and relevant challenges.
Description	<ol style="list-style-type: none"> 1- What are the various types of web attacks associated with social networking sites? 2- How can you examine intrusion and security events on social networking sites? 3- What are the challenges involved in controlling online social network crimes? 4- There are many common threats to online social networking services. Research and discuss what security preventive measures taken to reduce social network security risks? Illustrate your answer with examples.
Timeline	Four Hours
Assessment	Classroom and Lab discussion

Think Template (MCQs)	
Number	2.1, 2.2, 2.3, 2.4, 2.5, 2.6
Title	Social Networks Forensics
Type	Choose correct answer
Question	<p>1. What term describe the chronological documentation in which records in details social network physical or electronic evidence.</p> <p>a) User based activities b) Chain of custody c) System based activities d) None of the above.</p> <p>2. What best describes social network site metadata?</p> <p>a) Dump data b) Social media data c) Operating system data d) Data about data.</p> <p>3. Social networks forensics procedure include:</p> <p>a) Evidence acquisition b) Evidence Interpretation c) Evidence Interference d) All of the above.</p> <p>4. EnCase forensic can support social networking investigation events such that enable examiners to quickly search, and prioritize evidence in:</p> <p>a) Computer devices b) Mobile devices c) Compute and mobile devices d) None of the above.</p> <p>5. The following tool can be used to detect and present live social network evidence.</p> <p>a) Intrusion Detection and Protection Systems b) Snort c) Wireshark d) All of the above.</p> <p>6. Phishing on any social networking sites such as Twitter resembles any other form of phishing.</p> <p>a) True b) False</p> <p>7. Challenges to social networks forensics include:</p> <p>a) Technical challenges such as data collection and analysis b) Legal challenges such as the state law c) Resource challenges such as the storage of data d) All of the above.</p> <p>8. Spamming on social networks is considered as:</p> <p>a) Phishing attack type b) Continuous annoying and harmful messages c) Unwanted messages sent over the Internet</p>

	<p>d) All of the above.</p> <p>9. A social engineering practice whereby attackers take advantage to access to confidential information is commonly referred to as:</p> <ul style="list-style-type: none"> a) Shoulder surfing b) Phishing c) Penetration testing d) All of the above.
Answers	<p>1. What term describe the chronological documentation in which records in details social network physical or electronic evidence.</p> <ul style="list-style-type: none"> a) User based activities b) chain of custody c) System based activities d) None of the above. <p>2. What best describes social network site metadata?</p> <ul style="list-style-type: none"> a) Dump data b) Social media data c) Operating system data d) Data about data. <p>3. Social networks forensics procedure include:</p> <ul style="list-style-type: none"> a) Evidence acquisition b) Evidence Interpretation c) Evidence Interference d) All of the above. <p>4. EnCase forensic can support social networking investigation events such that enable examiners to quickly search, and prioritize evidence in:</p> <ul style="list-style-type: none"> a) Computer devices b) Mobile devices c) Compute and mobile devices d) None of the above. <p>5. The following tool can be used to detect and present live social network evidence.</p> <ul style="list-style-type: none"> a) Intrusion Detection and Protection Systems b) Snort c) Wireshark d) All of the above. <p>6. Phishing on any social networking sites such as Twitter resembles any other form of phishing.</p> <ul style="list-style-type: none"> a) True b) False <p>7. Challenges to social networks forensics include:</p> <ul style="list-style-type: none"> a) Technical challenges such as data collection and analysis b) Legal challenges such as the state law c) Resource challenges such as the storage of data d) All of the above. <p>8. Spamming on social networks is considered as:</p> <ul style="list-style-type: none"> a) Phishing attack type

	<ul style="list-style-type: none"> b) Continuous annoying and harmful messages c) Unwanted messages sent over the Internet d) All of the above. <p>9. A social engineering practice whereby attackers take advantage to access to confidential information is commonly referred to as:</p> <ul style="list-style-type: none"> a) Eavesdropping b) Phishing c) Penetration testing d) All of the above.
--	---

Extra	
Number	2
Title	Social Networks Forensics
Topic	2.1, 2.2, 2.3, 2.4, 2.5, 2.6
Type	<ul style="list-style-type: none"> • Book/Chapter (ISBN) <ol style="list-style-type: none"> 1- Choo, K. R., & Dehghantanha, A. (2017). Contemporary digital forensic investigations of cloud and mobile applications. Amsterdam: Elsevier. 2- Computer Forensics: Investigating Network Intrusions and Cybercrime (CHFI). Cengage Learning; 2nd edition (May 6, 2016). ISBN-10: 9781305883505 ISBN-13: 978-1305883505 • Offline content (Full reference) <ol style="list-style-type: none"> 1- Karabiyik, Umit; Canbaz, Muhammed Abdullah; Aksoy, Ahmet; Tuna, Tayfun; Akbas, Esra; Gonen, Bilal; and Aygun, Ramazan S. (2016) "A Survey of Social Network Forensics," Journal of Digital Forensics, Security and Law: Vol. 11 : No. 4 , Article 2. 2- Karabiyik, Umit; Canbaz, Muhammed Abdullah; Aksoy, Ahmet; Tuna, Tayfun; Akbas, Esra; Gonen, Bilal; and Aygun, Ramazan S. (2016) "A Survey of Social Network Forensics," Journal of Digital Forensics, Security and Law: Vol. 11 : No. 4 , Article 2. 3- A. Abraham (ed.), Computational Social Networks: Security and Privacy, Springer-Verlag London 2012. 4- NIJ. (2008). Electronic crime scene investigation: a guide for first responders. Washington, DC: U.S. Department of Justice 5- Selamat, S. R., Yusof, R., & Sahib, S. (2008). Mapping Process of Digital Forensic Investigation Framework. IJCSNS International Journal of Computer Science and Network Security, 8(10), 163-169. • Online content (URL) <ul style="list-style-type: none"> - https://www.infosecinstitute.com/ - https://techtalk.gfi.com/ - https://www.cellebrite.com/en/home/ - https://digital-forensics.sans.org/ - https://www.sans.org/ - https://www.getcybersafe.gc.ca/

2. Big data Forensics

Scope Template	
Number	3
Title	Big data Forensics
Introduction	This chapter entitled "Big data Forensics", is an introduction chapter about big data in the field of digital forensics. Section one defines big data terms and general usage. The second section explains the big data architecture. The big data risks and threats have been discussed in section three. Section four, explains the relation between evidence data and the big data. In Section five the reader can find a review of the latest study and big data forensic techniques. In the last section, we tried to explain the digital forensic process when the investigation case is related to big data and its applications.
Outcomes	<ol style="list-style-type: none"> 1. Gain basic knowledge about big data definition and techniques in the field of digital forensics. 2. Use big data tools and techniques to analyze evidence data. 3. Define and understand the digital forensic steps when the investigation case involves big data.
Topics	<ol style="list-style-type: none"> 1. Introduction to Big data 2. Big data Architecture 3. Big data Risks and Threats 4. Digital Evidence Data as Big data 5. Big data Forensics Techniques 6. Big data Forensic Process
Study Guide	<p>Instructions on how to study this unit.</p> <ul style="list-style-type: none"> • Required study time: 12 Hours. • Unit comprehensive reading. • Refer to external resources for more details such as the references appeared in the text • You are required to have a PC or laptop to install Apache Hadoop to be able to try the examples and the activities.

Content Template	
Section Number	3.1
Section Title	Introduction to Bigdata
Introduction	This section is dedicated to give an overview of the big data. After reading this section, the reader can define the big data and describe its characteristics and differentiate between large-scale database systems and big data applications.
Content	<p>Too many applications we use every day that manage our business life such as banking systems, airlines, hotels reservation systems and more. In addition to these applications, social media applications and websites are used every second by millions of users at which they share their posts and photos. All of these applications generate terabytes of data stored in several forms and servers using variety of technologies and tools.</p> <p>The question now is what size the data should exceed to consider it big data? The answer is that big data does not refer to the size of data only, but it has other factors and characteristics to be considered big data such as the speed at which the data is generated, the number of variety of data sources and other factors as described and introduced by Doug Laney and other data researchers.</p> <p>Doug Laney introduced three characteristics of big data called them the "Three Vs" of big data. The introduced characteristics are volume, velocity and variety (Laney, 2001).</p> <p>The volume describes how much the data is large in size, but the size itself is not important as how can we handle the large volume of data. Nowadays, the storage media is too cheap comparing with the last decades, but the challenge is how to retrieve and extract meaningful information from the large size of data and find new ways and tools to do.</p> <p>Velocity, the second "V" of big data. Data is generated in very high speed. For example, Google processes about 40000 search queries every second on average as I find on www.internetlivestate.com on 30 November, 2018 at 21:55. Now, imagine how many emails are sent every second around the world in addition the number of posts, photos are uploaded on the social media or blogs applications and websites. Despite of the small size of each post, photo, email and search query, the high speed of its generation creates a real challenge for data scientists to find new methods and tools to collect, process and extract meaningful information from this huge number of small size data pieces.</p>

Content Template	
Section Number	3.2
Section Title	Introduction to Big data
Introduction	This section is dedicated to give an overview of the big data. After reading this section, the reader can define the big data and describe its characteristics and differentiate between large-scale database systems and big data applications.
Content	<p>The Third "V" of big data according to Doug Laney is variety which refers to the different forms of data. Data can be structured such as database records and comma separated files, and it can be unstructured or semi structured such as emails, social media posts, audio and video files. According to this factor, the challenge is to design new methods and tools to collect, manage and analyze these various forms of data in one solution.</p> <p>In addition to the "Three Vs" of big data discussed before, veracity was introduced to be one of the influencing factors on data. Veracity highlights the uncertainty problem of the data and the data quality. So, a new challenge is raised here which is find a way to determine the quality of the data and decide if the data can be trusted or not.</p> <p>In conclusion, big data is a term used to describe huge volume of structured and unstructured data of variety forms and comes from variety of sources and changes constantly.</p>

Content Template	
Section Number	3.3
Section Title	Big data Architecture
Introduction	In this section the reader can find the situation where using big data in better than using large-scale database systems. In addition, this section explains the big data architecture which is very important to be understood to be able to identify related data to the investigation case and where the investigators can find it in the big data applications.
Content	<p>There are many situations enforce us to use big data architecture instead of using file systems or relational databases. The following list gives some examples of the situations at which we need to use big data architecture.(Sremack, 2015):</p> <ul style="list-style-type: none"> • Having multiple large structured and unstructured data sources. • Analyzing massive unstructured dataset. • Transforming large size of unstructured data into structured format. • Extracting data from social networks or web blogs. • In is the case where the structure of the data to be stored or processed changes frequently. <p>In this section, we will discuss the main four logical layer of big data architecture.</p> <ul style="list-style-type: none"> • Source Layer: This layer is the entry point of any big data system, at this layer the data arrives from many sources such as social media, blogs, emails, data base management systems, and other data sources. • Data Massaging and Storage Layer: This layer receives the data from the source layer, do some reformatting on the data if necessary, like converting unstructured data to structured format. After that, the structured data will be stored in RDBMS, and the unstructured data can be stored in a specialized file system like HDFS "Hadoop Distributed File System" and NoSQL databases.

Content Template	
Section Number	3.4
Section Title	Big data Architecture
Introduction	In this section the reader can find the situation where using big data is better than using large-scale database systems. In addition, this section explains the big data architecture which is very important to be understood because to be able to identify what is the related data to the investigation case and where the investigators can find it in the big data applications.
Content	<ul style="list-style-type: none"> • Analysis Layer: This layer interacts with the data and data storage using several analysis tools including machine learning techniques and algorithms in order to extract valuable information from the collected data. • Consumption layer: This layer is responsible to display the results coming from the analysis layer in appropriate format and reports.

Content Template	
Section Number	3.5
Section Title	Big data Risks and Threats
Introduction	This section is dedicated to give an overview of the big data risks and threats specially when the data is being collected from the internet and open data sources.
Content	<p>As we introduced in this chapter, the main sources of data in the big data systems and applications are blogs, emails and social media. We all will know that these sources contain a lot of sensitive personal data, which raise a real concern about privacy and security issues.</p> <p>Because the big data may contain sensitive personal information, new challenges become real in the big data field. In the following list you can find some examples of the challenges you have to take them in the consideration when working in the big data field (Sremack, 2015; Sharma and Navdeti, 2014):</p> <ul style="list-style-type: none"> • Identifying the information that can be collected, find how much this information is sensitive and private, then decide with whom this information can be shared or hidden. • Protecting the datasets and control the datasets accessibility. • Having the appropriate tools, roles and a professional team to manage and control the privacy and security issues. <p>In addition to the privacy and security issues we have review in this section, we have to take concern about the other threats and security issues known in the computing systems such as threats and risks of file systems, relational database systems, NoSQL and networks vulnerability and others.</p>

Content Template	
Section Number	3.6
Section Title	Digital Evidence Data as Big data
Introduction	This section highlights the relation between the big data application and digital forensic process. In particular, this section talks about the data that has been collected during the investigation and its characteristics to decide whether it is a big data or not.
Content	<p>In section one of this chapter we have discussed the characteristics of data to be considered big data, and why do we need special tools and methods to manage and study big data. Volume, velocity and variety are the main characteristics we have discussed about big data. The evidence data can be found as a part of the data, so evidence data is growing in the same speed as the data is growing. Therefore, if the data is big data then the evidence data might be big data also. While big data analysis needs sophisticated tools and experience, the digital forensics on big dataevidence big datarequires such tools and experience. Therefore, the digital forensics investigators have to adapt the big data techniques and tools to be used in big data digital forensics, and they should have the experience to work on big data and data scientists.</p> <p>It is become clear now that there are common characteristics between big data and digital evidence data. Digital forensics using big dataevidence faces the same challenges as the big data analysis, and the digital forensics investigators need to follow the data scientists to deal with such that challenges. In the following we will describe the main challenges that the digital forensics face.</p> <ul style="list-style-type: none"> • The size of data evidence associated to a case is growing from time to time. For example, according to FBI Regional Computer Forensics Laboratories (RCFLs) reports, the RCFLs in 2008 processed 27% more data that it was processed in 2007. Another from RCFLs in 2010 shows that the average size of digital evidence of a case is about 0.4 terabytes. This challenge is related to the velocity and volume of big data. According to a recent survey in 2013 on Forensic Focus, half of the cases involve more than on Terabyte of data, with one in five over five Terabytes in size.

Content Template	
Section Number	3.7
Section Title	Digital Evidence Data as Big data
Introduction	This section highlights the relation between the big data application and digital forensic process. In particular, this section talks about the data that has been collected during the investigation and its characteristics to decide whether it is a big data or not.
Content	<ul style="list-style-type: none"> • Another challenge is related to the source of data evidence. data evidence may be collected from several devices, can have several formats, and it could be structured and unstructured. <p>The digital evidence may have the same characteristics of big data. In addition, the same challenges in big data systems also challenges for data evidence. in the case that the digital evidence is big, then we should collect, preserve, analyze and generate report about the evidence using the big data tools and methods, because using conventional tools and methods will not work well in the case where digital evidence is big.</p>

Content Template	
Section Number	3.8
Section Title	Big data Forensics Techniques
Introduction	<p>In previous section we have discussed the challenges of big data examination and analysis. This section reviews some techniques that can be used to overcome those challenges discussed previously.</p> <p>Big data forensics techniques that can be used in the digital investigation where the big data is part of the investigation case are discussed in this section.</p>
Content	<p>As we already know, the forensic process goes through four main steps which are identification, collection, analysis and presentation. These steps are common among all the digital forensics investigations including big dataevidence.</p> <p>(XU et al., 2013) proposed a system consists of rule engine and finite state automaton to maintain big data and find the problems and the reason of any breakdown. This proposed system can be used in the big data acquisition step to verify the correctness of the collected data.</p> <p>(Noel and Peterson, 2014) proposed a method that can be used to identify the relevant information to the investigation case from the acquisitioned big data. The proposed method by Noel and Peterson uses keywords search and matching to find the relevant information and documents after applying files and content summarization.</p> <p>Data analysis is very important step in digital forensic, and because we are dealing with big data, therefore special analysis tools are required to perform the analysis step. One of the techniques that can be used in this case is data clustering. Data clustering in used in this filed to find the useful information and documents for the investigation. For more detail about data clustering and find how it can be used in the big data forensic you can read the following research articles as examples:</p> <ul style="list-style-type: none"> • Document clustering for forensic analysis: an approach for improving computer inspection. (da Cruz Nassif and Hruschka, 2013) • Clustering digital forensic string search output. (Beebe and Liu, 2014)

Content Template	
Section Number	3.9
Section Title	Big data Forensics Techniques
Introduction	Big data forensics techniques that can be used in the digital investigation where the big data is part of the investigation case where discussed in this section.
Content	<p>In addition to the techniques that have been discussed, some useful techniques and methods were proposed to deal with big data forensic challenges in particular the size of the data being examined. Finding the uninteresting data and unrelated files to the investigation, help in reducing the size of the data to be examined then reduce the time and efforts needed in the investigation. In general, we can call the methods used to minimize the data set by eliminate the irrelevance information data reduction methods. You can find some examples of using data reduction methods and techniques in the following research articles as examples:</p> <ul style="list-style-type: none"> • Content triage with similarity digests: The M57 case study.(Roussev and Quates, 2012) • A new approach for creating forensic hashsets. (Ruback et al., 2012) • Identifying forensically uninteresting files using a large corpus. (Rowe, 2013) • Big forensic data reduction: digital forensic images and electronic evidence. (Quick and Choo, 2016)

Content Template	
Section Number	3.10
Section Title	Big data Forensics Techniques.
Introduction	Big data forensics techniques that can be used in the digital investigation where the big data is part of the investigation case where discussed in this section.
Content	<p>The data from multi heterogenous data sources is very important issue need to be take in the consideration while investigating big data. (Zhenyou et al., 2011) proposed a system that use Hibernate technology and query optimization to link different data sources together. The following list contains some examples of researche papers that have been done to solve the multi-heterogenous sources of big data:</p> <ul style="list-style-type: none"> • Analysis and design of heterogeneous bioinformatics database integration system based on middleware. (Liu et al., 2010) • A semantic big data platform for integrating heterogeneous wearable data in healthcare. (Mezghani et al., 2015) • Intrusion detection and big heterogeneous data. (Zuech et al., 2015) <p>In conclusion, in several cases digital forensic has to work with big data. Sometimes the investigators have to collect the evidence from big data, and sometimes the evidence itself is big data. In other words, we are working on big data, therefore the methods and techniques that are used to process and create any big data system will suitable to be used in big data digital forensics. In this section we tried to mention some techniques and researches of big data as part of digital investigation.</p>

Content Template	
Section Number	3.11
Section Title	Big data Forensic Process
Introduction	In this section, the reader can find the difference between the steps of the digital forensic on regular data the steps of the digital forensic on big data.
Content	<p>Identifying the big data evidence, collecting data, data analysis and the result representation are the main steps on any digital forensic process as we have discussed previously.</p> <p>(Sremack, 2015) discussed the forensic steps and explained how to perform them using Hadoop as an example. In the following you can find an overview of the steps as explained by (Sremack, 2015).</p> <p>Step1: Identifying Big data Evidence</p> <p>In this step the investigators study the system to find the data sources, the data and sources owner, the data source content to be able to determine what is the relevant data to the investigation case.</p> <p>Step2: Collecting Data</p> <p>The data should be collected from the Hadoop file system data and Hadoop application data as proposed by (Sremack, 2015).</p> <p>Hadoop distributed file system contains metadata, configuration and user files which contain a lot of information such as how the system was operated and how it was used.</p> <p>In addition to the data exists in Hadoop distributed file system, the investigator should also collect data from the Hadoop application data in the case when the data that is relevant to the evidence exists in the content.</p>

Content Template	
Section Number	3.12
Section Title	Big data Forensic Process.
Introduction	In this section, the reader can find the difference between the steps of the digital forensic on regular data the steps of the digital forensic on big data.
Content	<p>Step 3: Data Analysis</p> <p>In this step the investigators analyze the data that has been collected from the Hadoop file system and Hadoop application data.</p> <p>Analyzing the data that has been extracted from the Hadoop file system requires reconstructing the collected data into its original format, analyzing the logical files, analyzing the log and configuration files contents. in this part of the step the investigators can find how the clusters was configured and what events occurred.</p> <p>In this step also, the investigators should prepare and setup the analysis environment, which is almost a large-scale database system. In such environments the investigators can apply different digital forensic techniques and use different digital forensic tools depending on the nature of the investigation case.</p> <p>After preparing the analysis environment, the data that has been extracted from the Hadoop applications should be transformed into a form that can be used in the prepared analysis environment.</p> <p>Step 4: Results Presentation</p> <p>Depending on the nature of the investigation case and to whom the results will be presented, the investigators should prepare the findings reports in a clear and accurate format. The presentation step was discussed in details in the introductory chapters which discussed the definition and the digital forensic concepts.</p> <p>We can use ELK Stack to collect sata, analyse data and result presentation. ELK Stack is a collection of four open source products Beats, Elasticsearch, Logstash and Kibana.</p> <p>Beats can be used to collect data. Logstash is responsible of data aggregation and processing. Elastichsearch provides the indexing and storage service for the system. The responsibility of data analysis and data presentation are for Kibana.</p>

Activity Template	
Number	3.1
Title	What is the relation between big data and digital forensic?
Type	The student required to do a research about big data and its relation to the digital forensics.
Aim	After completing this activity, the student will be able to define the big data, and he will be able to explain what is the relation between big data and digital forensics.
Description	In this activity the student is required to write a report that explain how and why big data influence digital forensic.
Timeline	<ul style="list-style-type: none"> • Find and collect the required references and resources such as books and journal. • Find and summarize the related information required to accomplish the report. • Report writing and editing using the academic writing criterions. • Required time 3 hours.
Assessment	This activity will be assessed based on: <ul style="list-style-type: none"> • The completeness. • The correctness. • The overall quality. • The followed process.

Activity Template	
Number	3.2
Title	Discuss three potential threats that may influence big data.
Type	The student required to do a research about risks and threats that may influence big data applications.
Aim	After completing this activity, the student will be able to evaluate some threats in the big data field.
Description	In this activity the student is required to write a report that explain some threats of big data applications.
Timeline	<ul style="list-style-type: none"> • Find and collect the required references and resources such as books and journal. • Find and summarize the related information required to accomplish the report. • Report writing and editing using the academic writing criterions. • Required time 2 hours.
Assessment	<p>This activity will be assessed based on:</p> <ul style="list-style-type: none"> • The completeness. • The correctness. • The overall quality. • The followed process.

Activity Template	
Number	3.3
Title	Hadoop architecture
Type	The student required to do a research about Hadoop architecture.
Aim	After completing the activity, the student will understand the architecture of big data application.
Description	In this activity the student is required to explain the big data architecture of a Hadoop application.
Timeline	<ul style="list-style-type: none"> • Find and collect the required references and resources such as books and journal. • Find and summarize the related information required to accomplish the report. • Report writing and editing using the academic writing criterions. • Required time 3 hours.
Assessment	This activity will be assessed based on: <ul style="list-style-type: none"> • The completeness. • The correctness. • The overall quality. • The followed process.

Activity Template	
Number	3.4
Title	Big data investigation study case
Type	The student is required to find a real case of digital investigation where a big data is part of the investigation case.
Aim	After completing the activity, the student will understand the big data forensic steps.
Description	In this activity the student is required to explain each step of big data forensic steps as was applied in the selected case.
Timeline	<ul style="list-style-type: none"> • Find a real big data forensic case. • Find and summarize the related information. • Explain each step that the investigators did during the investigation • Report writing and editing using the academic writing criterions.
Assessment	<p>This activity will be assessed based on:</p> <ul style="list-style-type: none"> • The completeness. • The correctness. • The overall quality. • The followed process.

Think Template (MCQs)	
Number	3.1
Title	Introduction big data
Type	Choose correct answer
Question	In the comparison between RDBMS and big data application like Hadoop we can say that Hadoop _____ .
Answers	A. Works better on unstructured and semi-structured data. B. Is suitable to read and write many times. C. Does ACID transactions. D. Has higher data integrity.

Think Template (MCQs)	
Number	3.2
Title	Big data Architecture
Type	True / False
Question	Data messaging and storage layer is responsible of converting all the coming data into structured format.
Answers	A. False. B. True.

Think Template (MCQs)	
Number	3.3
Title	Big data Forensic Process
Type	True / False
Question	The first step of big data forensic process is responsible to extract evidence data from the data sources.
Answers	A. False. B. True.

Think Template (MCQs)	
Number	3.4
Title	Big data Forensic Process
Type	Choose correct answer
Question	In which step the data reduction technique should be used.
Answers	<ul style="list-style-type: none"> A. Step 1: Identifying big data evidence. B. Step 2: Collecting data. C. Step 3: Data analysis. D. Step 4: Result presentation.

Extra Template	
Number	3.1
Title	3D data management: Controlling data volume, velocity and variety
Topic	Introduction to Bigdata
Type	Journal Article: Laney, D., 2001. 3D data management: Controlling data volume, velocity and variety. META Group Res. Note 6, 1.

Extra Template	
Number	3.2
Title	Big Data Forensics–Learning Hadoop Investigations
Topic	Bigdata Architecture
Type	Book: Sremack, J., 2015. Big Data Forensics–Learning Hadoop Investigations. Packt Publishing Ltd. / ISBN: 9781785281211

Extra Template	
Number	3.4
Title	Securing big data hadoop: a review of security issues, threats and solution
Topic	Bigdata Risks and Threats
Type	Journal Article: Sharma, P.P., Navdeti, C.P., 2014. Securing big data hadoop: a review of security issues, threats and solution. Int J Comput Sci Inf Technol 5, 2126_2131.

Extra Template	
Number	3.5
Title	A big data acquisition engine based on rule engine
Topic	Bigdata Forensics Techniques
Type	Journal Article: XU, X., YANG, Z., XIU, J., Chen, L.I.U., 2013. A big data acquisition engine based on rule engine. J. China Univ. Posts Telecommun. 20, 45–49.

Extra Template	
Number	3.6
Title	Applicability of Latent Dirichlet Allocation to multi-disk search
Topic	Bigdata Forensics Techniques
Type	Journal Article: Mezghani, E., Exposito, E., Drira, K., Da Silveira, M., Pruski, C., 2015. A semantic big data platform for integrating heterogeneous wearable data in healthcare. J. Med. Syst. 39, 185.

Extra Template	
Number	3.7
Title	Document clustering for forensic analysis: an approach for improving computer inspection
Topic	Bigdata Forensics Techniques
Type	Journal Article: da Cruz Nassif, L.F., Hruschka, E.R., 2013. Document clustering for forensic analysis: an approach for improving computer inspection. IEEE Trans. Inf. Forensics Secur. 8, 46–54.

Extra Template	
Number	3.8
Title	Clustering digital forensic string search output
Topic	Bigdata Forensics Techniques
Type	Journal Article: Beebe, N.L., Liu, L., 2014. Clustering digital forensic string search output. Digit. Investig. 11, 314–322.

Extra Template	
Number	3.9
Title	Content triage with similarity digests: The M57 case study
Topic	Bigdata Forensics Techniques
Type	Journal Article: Roussev, V., Quates, C., 2012. Content triage with similarity digests: The M57 case study. Digit. Investig. 9, S60–S68.

Extra Template	
Number	3.10
Title	A new approach for creating forensic hashsets
Topic	Bigdata Forensics Techniques
Type	Conference Paper: Ruback, M., Hoelz, B., Ralha, C., 2012. A new approach for creating forensic hashsets, in: IFIP International Conference on Digital Forensics. Springer, pp. 83–97.

Extra Template	
Number	3.11
Title	Identifying forensically uninteresting files using a large corpus
Topic	Bigdata Forensics Techniques
Type	Conference Paper: Rowe, N.C., 2013. Identifying forensically uninteresting files using a large corpus, in: International Conference on Digital Forensics and Cyber Crime. Springer, pp. 86–101.

Extra Template	
Number	3.12
Title	Big forensic data reduction: digital forensic images and electronic evidence
Topic	Bigdata Forensics Techniques
Type	Journal Article: Quick, D., Choo, K.-K.R., 2016. Big forensic data reduction: digital forensic images and electronic evidence. Clust. Comput. 19, 723–740.

Extra Template	
Number	3.13
Title	Research on the integration and query optimization for the distributed heterogeneous database
Topic	Bigdata Forensics Techniques
Type	Conference Paper: Zhenyou, Z., Zhang, J., Shu, L., Zhi, C., 2011. Research on the integration and query optimization for the distributed heterogeneous database, in: Computer Science and Network Technology (ICCSNT), 2011 International Conference On. IEEE, pp. 1533–1536.

Extra Template	
Number	3.14
Title	Analysis and design of heterogeneous bioinformatics database integration system based on middleware
Topic	Bigdata Forensics Techniques
Type	Conference Paper: Liu, Y., Liu, X., Yang, L., 2010. Analysis and design of heterogeneous bioinformatics database integration system based on middleware, in: Information Management and Engineering (ICIME), 2010 The 2nd IEEE International Conference On. IEEE, pp. 272–275.

Extra Template	
Number	3.15
Title	A semantic big data platform for integrating heterogeneous wearable data in healthcare
Topic	Bigdata Forensics Techniques
Type	Journal Article: Mezghani, E., Exposito, E., Drira, K., Da Silveira, M., Pruski, C., 2015. A semantic big data platform for integrating heterogeneous wearable data in healthcare. J. Med. Syst. 39, 185.

Extra Template	
Number	3.16
Title	Intrusion detection and big heterogeneous data: a survey
Topic	Bigdata Forensics Techniques
Type	Journal Article: Zuech, R., Khoshgoftaar, T.M., Wald, R., 2015. Intrusion detection and big heterogeneous data: a survey. J. Big Data 2, 3.

Extra Template	
Number	3.17
Title	What Is Big Data and What Does It Have to Do with IT Audit?
Topic	Digital Evidence Data as Big data
Type	Journal Article: "What Is Big Data and What Does It Have to Do with IT Audit?", ISACA Journal, 2013, p.23-25

3. Introduction to Cloud Computing

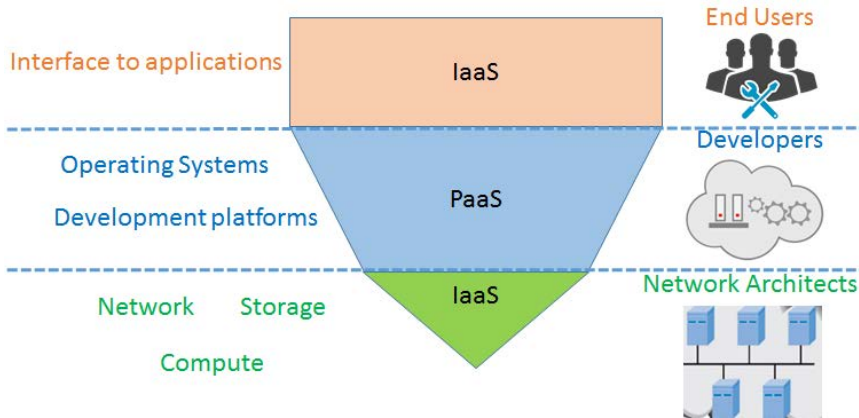
Scope Template	
Number	4
Title	Introduction to Cloud Computing
Introduction	This chapter introduces cloud computing as it is very popular and widely used nowadays. Recently, the number of successful cyber-attacks and fraudulent activities in clouds have dramatically increased. Additionally, there has been a serious need to not only protecting clouds from attacksbut also investigating the attacks to know who is doing them and why so that future attacks can be predicted and prevented. Therefore, much research has focused on cloud forensic to mitigate cloud digital crime which caused the appearance of cloud computing forensic.
Outcomes	After studying this chapter, student will have the ability to: <ul style="list-style-type: none"> - Define cloud computing and know its benefits - Understand how cloud computing works based on its architecture - Determine the deployment model of a cloud system - understand virtualization for cloud computing - Explain the security and protection methods of a cloud system
Topics	1- Overview 2- Cloud computing service models 3- Cloud computing deployment models 4- Cloud data centers 5- Cloud virtualization 6- Cloud security
Study Guide	

Example of study guide

Task	Time
Preparation (Introduction and On-line Planning):	1hr
Disk-based Content:	2.5hrs
Set textbook Content:	1
Thinking (On-line discussions, Review questions)	1hr
Tutorial Work:	2.5hrs
*Related Course Work:	1hrs
Total	09 hours

Content Template	
Section Number	4.1
Section Title	Introduction to cloud computing
Introduction	This section discusses cloud computing and its benefits. This section defines cloud computing and highlights why forensic computing is necessary for cloud.
Content	<p>4.1.1 Overview</p> <p>Cloud computing is very popular and widely used nowadays. Recently, the number of successful cyber-attacks and fraudulent activities in clouds have dramatically increased. Additionally, there has been serious need to not only protecting clouds from attacks, but also investigating the attacks to know who is doing them and why so that future attacks can be predicted and prevented. Therefore, much research has focused on cloud forensic to mitigate cloud digital crime, which caused the appearance of cloud computing forensic.</p> <p>Cloud computing forensic is defined as "the application of scientific principles, technological practices and derived and proven methods to reconstruct past cloud computing events. This is done through identification, collection, preservation, examination, and interpretation and reporting of digital evidence." Also, it was defined as "the application of digital forensic science in cloud environments as a subset of network forensics." The second definition borrows the term digital forensic which is defined as "an applied science to identify an incident, collection, examination, and analysis of evidence data."</p> <p>To know how to apply digital forensic on cloud computing, it is necessary to understand what cloud computing is and how it is working. This chapter focuses on explaining cloud computing, its services, its deployment models, and its security issues.</p> <p>4.1.2 Definition of Cloud Computing</p> <p>The term cloud computing has no standard definition but it simply means the delivery or access of computer services over the internet. One standard definition is "Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction." By using cloud computing, users can access or deploy cloud applications from anywhere in the world. In a company's point of view, cloud means the delivery of services because companies offer the service to users and these companies are called service providers. Here, users can access cloud services and can pay for these services based on usage in "pay as you go" billing model. Computer services include servers, storage, databases, networking, software, analytics and more.</p> <p>Cloud computing is an example of a distributed computing system. All computations in cloud applications are distributed to servers in data centers called virtual machines. These virtual machines support elastic and on demand utilization of hardware and software resources. By this concept, the desktop computing is moved to service-oriented platforms that contain clusters of servers and huge databases. Also, cloud computing leverages the dynamic resources to deliver a large number of services in a high-throughput computing paradigm (HTC). Cloud computing borrows the multitasking approach to achieve high throughput by serving multiple users and application at the same time.</p> <p>4.1.3 Benefits of cloud computing</p> <p>Since the evolution of the internet in 1990, the world has witnessed several computer technologies that affect the overall business cycle. For example,</p>

	<p>the way people use computer to do business at the time of the uni-processor computer is totally different compared to the use at the time of parallel computing, grid computing, distributed computing, pervasive computing or cloud computing. We can say cloud computing has the following general benefits:</p> <ol style="list-style-type: none"> 1- Cost: Cloud computing eliminates the money spent on buying hardware and software for establishing data centers. Data centers are used by companies to store or process a huge amount of data and they usually cost much money. To run onsite data center, a company needs to have racks of servers, round-the-clock electricity for power and cooling, and IT experts for managing the infrastructure. 2- Speed: users of cloud computing can provision a vast amount of computing resources in minutes. In fact, most cloud computing services are provided as self-service and on demand which enables users to manage their business with few clicks. Speed is guaranteed as all hardware and software resources in cloud computing data centers are regularly upgraded to the latest versions. 3- Global scale: cloud computing services are provided on demand from the right geographic location with the right amount of computation resources e.g. storage, time, power and bandwidth. Therefore, one can say that cloud computing services has the ability to scale elastically. 4- Productivity: with cloud computing, users or IT specialists can spend more time doing necessary tasks and therefore they can achieve more important business goals. Before cloud computing, such time was spent on managing onsite servers including hardware and software setup and maintenance. 5- Reliability: in cloud computing data can be mirrored at multiple redundant sites on the cloud provider's network. This makes data backup, disaster recovery and business continuity easier and less expensive.
--	---

Content Template	
Section Number	4.2
Section Title	Cloud computing service models
Introduction	This section introduces the cloud computing service models that are: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS).
Content	<p>Cloud computing contains three service models: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). Figure 4.1 below shows the three service models and their components. These three models allow users to access services over the internet. To make use of the services of infrastructure, platform and software, users need to subscribe to the service provider and pay based on usage. Usually the cost of service depends on the availability, data protection, security and performance.</p>  <p>Figure 4.1: the three layers of cloud computing and their components</p> <p>4.2.1 Infrastructure as a Service (IaaS) By subscribing to this model, users can access the virtualized IT resources such as storage, networking, and computing. Thus, the IaaS model includes the storage as a service in which users can manage their storage resources, compute instances as a service in which users can manage computation resources such as the operating system and the deployed applications, and network as a service in which users can select the network components.</p> <p>4.2.2 Platform as a Service (PaaS) In this model, the cloud provides programmers and developers with the hardware and software required to develop, deploy and execute applications. Therefore, the platform should have the required resources including operating systems programming libraries. Examples of companies that provide programming platforms are Google App Engine, Amazon Elastic MapReduce, Microsoft Azure, and Salesforce. Some programming languages that are supported by the cloud are java, python, R, C++ Ruby, Perl, and .net. In this model, developers from different parts of the world can cooperate to develop one application. Developers can target different applications such as Web applications, Business applications such as CRM, e-commerce, and data processing. Cloud platforms should satisfy some requirements to enable efficient applications development. They should enable the user to share the development capability with other users. Also, the platform should enable load balancing among different user developing different applications. Further, the platform should support distributed scheduling mechanism for tracking event at particular times.</p>

	<p>4.2.3 Software as a Service (SaaS)</p> <p>This allows users to connect to and use cloud-based applications over the Internet using a browser. SaaS offers the resources required to utilize the services offered by PaaS and IaaS. Examples of SaaS are email, calendaring, and office tools. All of the underlying infrastructure, middleware, application software, and application data are located in the service provider's data center. The service provider manages the hardware and software, and with the appropriate service agreement, will ensure the availability and the security of the app and your data as well.</p> <p>The advantages of using SaaS can be summarized as gaining access to sophisticated applications without the need to purchase, install, update, or maintain any hardware, middleware, or software. Also, saving money because the SaaS service automatically scales up and down according to the level of usage. Further, mobilizing workforce easily as SaaS makes it easy to "mobilize" workforce and users can access SaaS applications and data from any Internet-connected computer or mobile device. Furthermore, accessing applications data from anywhere because data are stored in the cloud.</p>
--	--

Content Template	
Section Number	4.3
Section Title	Cloud computing Deployment models
Introduction	The benefit of the aforementioned service models depends on the way how the cloud is deployed and who can access it. Cloud computing can be public clouds, private clouds or hybrid clouds.
Content	<p>4.3.1 Public Clouds In general, the term public cloud is used interchangeably with cloud computing. Public cloud refers to a computing deployment model used for the provisioning of storage services and computational services to the general public over the internet. In the public cloud any user who pays for the service can access it. Examples of public cloud are storage services from Amazon, worksheets from Google, and Azure programming environment from Microsoft. The aim of the public cloud is to free up the customers from the responsibilities of taking care of infrastructure, software and other computer architectural issues. Although public clouds offer application flexibility among different users, security is an issue that users worry about when accessing or using public clouds.</p> <p>4.3.2 Private Clouds The term private cloud refers to the delivery of cloud services in a way similar to the public cloud but only to users from a single organization. The services are provided within the organization domain regulated by the organization standards and controlled by the owner. For example, a university can provide its employees with cloud services behind a firewall. In this way, the organization can take advantages of virtualization and other cloud benefits without dealing with security issues arising from the public cloud. Private clouds maintain the security and privacy within the organization and achieve customization of services with high efficiency. However, private clouds increase the cost of cloud deployment and reduce the scalability.</p> <p>4.3.3 Hybrid Clouds A hybrid cloud refers to an integrated cloud service utilizing both private and public clouds to carry out different functions within the same organization. With hybrid clouds, an organization can maximize the efficiencies by employing public cloud services for all non-sensitive operations, only relying on a private cloud where if it is required. For example, a university can build a private cloud for the research centers particularly to process research operations. At the same time, the university can buy other services like email from public clouds.</p>

Content Template	
Section Number	4.4
Section Title	Cloud data centers
Introduction	A data center is a major part of cloud systems. There are several types of data centers and each type has several requirements.
Content	<p>4.4.1 data center types</p> <p>The first type of data centers is the warehouse-scale data center which can be large with 500,000 to 1 million servers per one center or small with 1000 servers per one data center. A server may contain one disks or more with 1 TB disk drive and 8 GB DRAM. The network among the servers should support fault tolerance, low latency, high bandwidth, and low cost. These requirements are designed to support distributed operations and huge communications with the servers. The high number of servers produce huge heat which needs to be cooled down by cooling systems.</p> <p>The second type of data centers is the modular data center in shipping containers (trucks). The need for cheap electricity and cheap cooling systems forces companies to put servers in shipping trucks. These trucks can move to places where electricity and cooling systems are cheap. Each container can contain 30 PB storage and 46,080 computing cores. The container has an interconnection network system to connect servers inside the container, and inter-module network that connects containers with each other. However, some problems may arise when using these containers such as data integrity, security, and server monitoring.</p> <p>4.4.2 requirements of efficient data center</p> <p>The basic requirements for efficient data centers are as follows:</p> <ul style="list-style-type: none"> • Data centers should be scalable to allow for the growth in storage, computations, Input/output (I/O), power, and cooling system. • Data centers should be reliable as the virtual machines have to be integrated with fault tolerance to enable recovery from failure or disaster • Data centers should have a low cost for users and providers • Data centers should be secure and maintain user privacy • Data centers should have high availability to allow users to obtain the service at any time

Content Template	
Section Number	4.5
Section Title	Cloud virtualization
Introduction	<p>Virtualization means that cloud resources are virtualized in a way that users or developers do not need to know or care about the physical resources used in the infrastructure. In cloud computing systems, a special kind of software (virtual machine) which simulates the execution of hardware is used to virtualize hardware. This software can also be used to run operating systems and as a platform for developing cloud applications.</p>
Content	<p>4.5.1 virtual machines</p> <p>Previously, computers used to run Operating System and application on top of the Operating System, but now, with the help of virtualization software like Hypervisor, multiple Virtual VMs on a single computer can be created. The Operating Systems can be installed on VMs and run all of them at the same time as in figure 4.2.</p> <div data-bbox="472 808 1340 1196" data-label="Diagram"> </div> <p>Figure 4.2: the difference between old computer system and cloud computing in terms of virtualization.</p> <p>Virtual machines (VMs) are the units of virtualization as they are the containers of cloud services. Every virtual machine has virtual devices that provide the same functionality as physical hardware and have additional benefits in terms of portability, manageability, and security. When running any cloud service, the provisioning tool will initially find the corresponding physical machines and then run the VMs on them. VMs are usually used to host third-party programs, and they provide flexible runtime since developers do not worry about the system environment. VMs also support flexibility for users. The resources are shared by many users and at the same time each user privileges are maximized and separated. In other words, a user can control his own data without the interference of others.</p> <p>4.5.2 Virtualization of IaaS, PaaS, and SaaS</p> <p>Virtualization can be applied to cloud computing on different levels. In the IaaS level, many small physical servers are combined into one large physical server, so that the processor can be used more effectively. The operating system that is running on a physical server gets converted into a well-defined OS that runs on the virtual machine. The network can be also virtualized by the management and monitoring of a computer network as a single managerial entity from a single software-based administrator's console. It is intended to allow network optimization of data transfer rates, scalability, reliability, flexibility, and security. It also automates many network administrative tasks. Network virtualization is specifically useful for</p>

	<p>networks that experience a huge, rapid, and unpredictable traffic increase. The storage is also virtualized as multiple network storage resources are present as a single storage device for easier and more efficient management of these resources.</p> <p>In the PaaS level, the operating system utilizes the hypervisor technique to connect to the memory pool and makes that pooled memory available to applications. Meanwhile, applications running on connected computers that are directly connect to the memory pool through an API or the file system.</p> <p>In the SaaS level, as users can access cloud remotely and are able to work from any location. Virtualization provides a lot of flexibility for users to work from home or on the go. It also protects confidential data from being lost or stolen by keeping it safe on central servers. Also, users can manipulate data and know how it is formatted or where it is physically located. Virtualization decreases data errors and workload.</p>
--	---

Content Template	
Section Number	4.6
Section Title	Cloud security
Introduction	Cloud computing poses new security threats that are more difficult to deal with than the traditional computing environment. Therefore, new security models have to be used to ensure protection, overcome security challenges and increase users trust.
Content	<p>Cloud computing requires security protection at all levels. Examples of cloud components that need security protection are:</p> <ul style="list-style-type: none"> • The protection of servers from attacks by worms, viruses, and malwares • The protection of VMs from software-based attacks and the denial of service attack (DoS) • The protection of data from theft, corruption, and natural disasters <p>Some of the new security techniques that are applied to cloud computing nowadays are software APIs to protect data integrity. These APIs are used for authenticating users and protecting data from alteration, deletion or copyright violation. Another technique is the data coloring in which each data object is labeled by a unique color, and the corresponding user ID is also colored. A matching process is implemented to identify which user belongs to which data. Coloring and color matching consumes less computational resources than the traditional encryption-decryption techniques.</p>

Activity Template	
Number	4.1
Title	Design cloud computing system for car rental companies
Type	<ul style="list-style-type: none"> • Reflection
Aim	<ul style="list-style-type: none"> - Use the service models of cloud computing for designing a cloud system
Description	<p>Imagine a cloud system for a car rental company,</p> <ul style="list-style-type: none"> - What will be the components for IaaS, PaaS and SaaS? - Draw a brief design of that cloud system - write a report about the cloud system design for car rental companies
Timeline	One week
Assessment	Report based evaluation

Activity Template	
Number	4.2
Title	Write a report about the cyber-attacks to cloud systems.
Type	<ul style="list-style-type: none"> • Research • .
Aim	<ul style="list-style-type: none"> - To know the types of cloud computing security threats, attacks and solutions
Description	<p>Students need to write a report about the attacks to cloud systems</p> <ul style="list-style-type: none"> - What are the most serious attacks? - What are the methods used to protect the cloud from these attacks? - What are the attacks that have no solution until this moment?
Timeline	One week
Assessment	Report based evaluation

Think Template (MCQs)	
Number	4.1
Title	Choose correct answer.
Type	Multiple Choice Question
Question	Cloud computing borrows the _____ approach to achieve high throughput a- Infrastructure b- Service c- Multitasking d- networking
Answers	c- Multitasking

Think Template (MCQs)	
Number	4.2
Title	Choose correct answer.
Type	Multiple Choice Question
Question	<p>_____ makes data backup, disaster recovery and business continuity easier and less expensive.</p> <ul style="list-style-type: none"> a- Productivity b- Reliability c- Availability d- Vulnerability
Answers	b- Reliability

Think Template (MCQs)	
Number	4.3
Title	Choose correct answer
Type	Multiple Choice Question
Question	Cloud computing can not be attacked because it employs high protection methods. a- True b- False
Answers	b- False

Think Template (MCQs)	
Number	4.4
Title	Choose correct answer
Type	Multiple Choice Question
Question	<p>Data center should have high _____ to allow users to obtain the service at any time</p> <ul style="list-style-type: none"> a- Storage capacity b- processor speed c- reliability d- availability
Answers	d-availability

Extra Template	
Number	4.1
Title	Distributed and Cloud Computing: From Parallel Processing to the Internet of Things
Topic	Topic 1, Topic 2, Topic 3, Topic 4, Topic 6
Type	Book Hwang, K., Dongarra, J., & Fox, G. C. (2011). Distributed and cloud computing: from parallel processing to the internet of things (First). Morgan Kaufmann.

Extra Template	
Number	4.2
Title	Cloud Computing Bible
Topic	Topic 5
Type	Book (Li, 2016) Li, X. (2016). <i>Advanced Design and Implementation of Virtual Machines</i> . CRC Press.

4. The Challenges of Cloud Computing in Digital Forensics

Scope Template	
Number	5
Title	The Challenges of Cloud Computing in Digital Forensics
Introduction	This chapter illustrates the challenges of cloud computing in digital forensics
Outcomes	At the end of this chapter, students will be able to: <ol style="list-style-type: none">1. Detail the cloud computing from several aspects such as the cloud computing infrastructure, its characteristics, its services models, and its deployment modality.2. List the challenges that will be added to digital forensics investigators because of using the cloud environment.3. Illustrate the challenges of applying digital forensics in cloud environment.
Topics	5.1 cloud computing 5.2 digital and cloud forensics 5.3 cloud forensics challenges
Study Guide	Instructions on how to study this unit. <ul style="list-style-type: none">• Required study time:<ul style="list-style-type: none">○ Preparation: 1 hour○ Disk-based Content: 2 hours• Practical activities: 2 hours

Content Template	
Section Number	5.1
Section Title	Cloud computing overview
Introduction	An introductory paragraph on cloud computing will be covered in this section.
Content	<p>Cloud computing is defined as an information technology (IT) principle that empowers ubiquitous and pervasive approach over the Internet to services that are considered as higher-level services and be rapidly provisioned. Cloud computing also allows access over the Internet to shared and configurable system resources. In order to achieve economies of scale and coherence, sharing resources is a crucial in cloud computing. Another definition of cloud computing is using numerous services over the cloud (i.e. over the Internet), such as platforms servers, software platforms, and storage services. All cloud services are provided by cloud Service Providers (CSPs) upon signing agreements with consumers. For supplying the various cloud services, a CSP preserves and maintains the infrastructure needed for computing (the infrastructure are computer systems with high availability in data centers or clusters), it runs the cloud software and transfers the cloud services to the consumers via the Internet.</p> <p>As shown in the previous chapter, virtualization techniques are essential in cloud computing to supply and provide platform support, software, and equipment as remote services. Generally, broad network access, on-demand self-service, measured service, rapid elasticity, and resource pooling are the five major characteristics that distinguish the cloud model. CSPs provide their "services" according to three basic models: Platform as a Service (PaaS), Software as a Service (SaaS) and Infrastructure as a Service (IaaS).</p>

Content Template	
Section Number	5.2
Section Title	Digital and Cloud Forensics
Introduction	An introductory paragraph on digital and cloud forensics will be covered in this section.
Content	<p>According to McKemmish [1], the definition of Digital Forensic is the “process of identifying, preserving, analyzing and presenting digital evidence in a manner that is legally acceptable”. US-CERT [2] provides a comprehensive definition: “The discipline that combines elements of law and computer science to collect and analyze data from computer systems, networks, wireless communications, and storage devices in a way that is admissible as evidence in a court of law”.</p> <p>Generally, digital forensics is considered as the domain that gathers, maintains and examines data professionally in a way considered as evidence in court. The main goal of a forensic investigation is to recognize, preserve and maintain the evidence, derive the information, document every process, and examine the extracted information in order to discover answers with respect to the 5Ws (Why, When, Where, What, and Who) [3]. The forensic investigation process starts upon occurring an incident such as:</p> <ul style="list-style-type: none"> • Cases of Criminal Damage (CD) such as threats to damage property or destruction of another’s belongings and [4]. • Cases of Industrial Espionage (IE) include inventions, patents, and trade secret theft, which is considered as a highly profitable crime. • Cases of Financial Investigations (FI): These cases usually are related to economics such as credit card fraud, insurance fraud, and money laundering. • Cases of Corporate Policy Violation (CPV) include misconduct, email abuse, and employment termination investigations. • Cases of Child Abuse (CA): These cases are criminal offences such as possession of indecent child media content and child grooming. • “Defense-in-depth” is considered as a method or an approach to network security. The capability of conducting forensic investigations would improve the general integrity and survivability of the infrastructure related to business [2]. <p>We can conclude from the last example that digital forensics can be conducted not only in law enforcement agencies but also private organizations. Currently, a lot of private organizations are including forensic departments in their organization in order to enhance the overall infrastructure security. It is important to note that if digital forensics practiced incorrectly, the evidence may be damaged by the incorrect analysis; thus, the evidence will be improper in a court of law [2]. Thus, the correct methodologies and procedures in digital forensics are essential. For example, e-mail forensics investigates e-mail related cases, network forensics handle investigations in the infrastructure related to</p>

networks and mobile forensics specializes in handheld devices.



Figure 1. digital forensics categories

Cloud forensics is considered as a multidiscipline since it integrates both digital forensics and cloud computing in one domain. In order to illustrate the science of cloud forensics, the intersection between both the science of cloud computing and the digital forensic analysis should be explained. The definition of Cloud forensics is the digital investigation and analysis in cloud environments with the consideration of forensic foundations, principles, methodologies, and procedures.

Digital forensics employs the concepts of computing and computer science to regain electronic evidence for presentation and displaying in a court of law. Practically, cloud forensics is classified as a part of digital forensics that focuses on forensic investigations of networks.

Since cloud computing relies on wide network access, cloud forensics adopts the basic steps of network forensics with methodologies and techniques modified to cloud computing environments. Regarding the National Institute of Standards and Technology (NIST), the science cloud computing forensic is defined as "the application of scientific principles, technological practices and derived and proven methods to reconstruct past cloud computing events through identification, collection, preservation, examination, interpretation and reporting of digital evidence."

The process of Data collection is defined as the process of recognizing, labeling, and gathering forensic data that includes data related to client-side (i.e. data exist on client premises) and data related to provider-side artifacts that reside in the infrastructure for the provider. Cloud environments investigators should

	<p>take into account several additional issues in comparisons with computer investigators (digital forensics). Based on the cloud computing services models, several tools and procedures are used to gather forensic data. The data integrity must be maintained and preserved in the collection process with certainly defined separations of tasks between both the client and provider. Laws or regulations should not be breached in the jurisdictions where data is gathered, collected, or compromise the confidentiality of other tenants that share the resources.</p> <p>Moreover, evidence identification in cloud environments is considered as a hard task and process since there are multiple services and deployments models. Additionally, in cloud environments, there is a limitation of seizing (physically) the computer device including the evidence. Next section details the challenges of the cloud forensics.</p>
--	--

Content Template	
Section Number	5.3
Section Title	Cloud Forensic Challenges
Introduction	This section presents the challenges that face the cloud forensics.
Content	<p>In this section, the challenges that face the cloud forensics which conducted from literature review will be illustrated.</p> <p>Accessing to evidence in log files</p> <p>The first priority for the investigators is identifying an incident which can be performed by having access to log files. In the cloud, accessing log files is very hard since data are stored in unknown locations because of the geographical allocation of the cloud hardware devices. According to the procedures conducting by digital forensics, there is a need to physically approach the hardware devices [5]. In cloud forensics, there is no way to recognize the hardware device including sensitive data such as log files because these data are preserved and maintained in systems that are distributed and located in various areas. This difficulty and challenge employed to all the service models of cloud computing.</p> <p>Additionally, cloud service models affect the detection of log files. For example, there is no way for detecting the system log files and status in PaaS and SaaS model because the limitation in clients' access (i.e. access is provided via the API or the predesigned interface). However, in the IaaS cloud model, those files are partly applicable since IaaS service model employs the Virtual Machine (VM) technique which allows clients to behave the same as using an Actual Machine [6]. Generally, most of Cloud Services Providers (CSPs) limit the services of gathering logs and in some cases intentionally hide all details from customers.</p> <p>Volatile data</p> <p>In the IaaS cloud model, in a Virtual Machine instance, all data will be lost upon rebooting the VM or turning it off. Therefore, several important pieces of evidence could be lost such as processes registry entries, and temporary files on the Internet. Suppose that an attack on a VM was initiated. This attack is designed with no constant storage synchronization. Upon completing the attack, attackers could turn off the Virtual Machine instance arising in losing volatile data (i.e. assuming there is no additional countermeasures were sat up and installed) [7].</p> <p>Distribution - collaboration.</p> <p>As mentioned before, a cloud environment (in all three service models), computer systems are distributed. Thus, investigators challenge problems with</p>

	<p>laws and jurisdictions. Investigators have to wait for authorization in order to access the information that they need. This procedure is considered as costly and time consuming for investigators. Thus, global collaborations among law administrations and cloud service providers should be considered. Guidelines must be updated and re-written globally by all countries for the aforementioned reasons.</p> <p>Dependence on CSP - Trust</p> <p>CSPs aim at assisting both clients and investigators to acquire all information and evidence from their cloud infrastructures. Because of the fear that all information and evidence may be used against CSPs, they refuse to supply information residing in their premises. Generally, investigators depend on CSPs to recognize, maintain and collect all the evidence that can approach to the incident in all CSPs' models, particularly in PaaS and SaaS. Additionally, sometimes CSPs depend on other CSPs (third parties) to be capable to use their services. Thus, all parties that have an effect on the chain of custody should be covered by investigations. This challenge is considered as a major one in the identification phase, the preservation phase, and the collection phase.</p> <p>Identification of the Client side</p> <p>As known, evidence can be detected in both interfaces the providers' interface as well as the clients' interface. In SaaS and IaaS models, the only application that allows the client system to communicate with cloud services is the user agent. Therefore, in a comprehensive forensic investigation, the acquired evident from the browser environment should not be excluded [7].</p> <p>Privacy</p> <p>clients privacy is affected in IaaS service model because of the cloud virtualization. Thus, all regulations and standards must be held by investigators to collect the needed pieces of evidence without breaching the privacy of clients.</p> <p>Time synchronization</p> <p>In PaaS, SaaS, and IaaS models, time synchronization is vital to be able to conduct correct results since data are preserved in different geographical regions with multiple time zones. All the time stamps from the devices should be acquired by investigators in order to be able to initiate and establish a precise events timeline.</p> <p>Internal Staffing</p> <p>A team that consists of investigators that have technical experience, legal advisors and external experts who have in-depth knowledge and skills must be involved to perform a cloud-based forensics investigation.</p>
--	---

	<p>Chain of custody</p> <p>Preserving and maintaining the chain of custody is considered as a vital key in order to be able to present the evidence in a court of law. Thus, preserving the chain is recognized as a huge challenge for cloud service providers because of the multi-jurisdictional laws. Suppose that an investigation should be conducted. In this investigation, the cloud service provider has to provide the investigators with all the submitted data. The person who is responsible for gathering the data is not trained to maintain evidence in accordance with specific forensic techniques. In that case, there will be no preserving in the chain of custody. For each case in which will be stood in the court, investigators must ensure that the chain of custody contains data and information such as the name of the person who collected the evidence, information related to how and where the evidence was gathered, how the evidence was preserved, who can access the evidence and who already has accessed the evidence, etc.</p> <p>Multi-jurisdiction and Multi-tenancy.</p> <p>Because of the presence of multiple jurisdictions and multi-tenancy in cloud computing, significant challenges to forensic investigations are proposed. Each jurisdiction forces different requirements regarding data access and retrieval, evidence recovery without breaching tenant rights, evidence admissibility and chain of custody. The absence of a worldwide regulatory body or even a federation of national bodies significantly impacts the effectiveness of cloud forensic investigations.</p> <p>Lack of forensic tools</p> <p>Suitable forensic tools for analyzing data are needed in cloud environments. A lot of tools have been developed for investigations in regular digital forensics. Currently, these tools are also used in cloud investigations. Because the cloud infrastructures are distributed and with no physical access, these tools cannot fully cap the investigations in IaaS, PaaS and SaaS models. Thus, new tools should be developed to be applied in acquiring data in the preservation phase "collection stage". Additionally, new certified tools should be implemented to be used in data examination and analysis.</p> <p>Volume of data</p> <p>Because of the increasing amount of data maintained in CSPs' centers, finding useful evidence for the investigation is considered as a hard process.</p> <p>Encryption</p> <p>In PaaS, SaaS, and IaaS service models, cloud users encrypt their data before saving and storing to protect these data from criminal activities. To conduct an investigation, all encrypted data will be considered as un-useful if the encryption keys were not able to be acquired. If the owner of the data is the</p>
--	---

	<p>only one who can provide the key, or if the key is destroyed, the evidence can be compromised. Additionally, encryption methods are applied to store clients' data in the cloud via the CSPs.</p> <p>Reconstruction</p> <p>To perform an investigation in a regular way, the scene of the crime should be reconstructed. In cloud, data are distributed across different locations with time difference so that reconstructing the scene of the crime could be a hard work [6]. Furthermore, when forcing a virtual machine instance to turn off, all potential pieces of evidence and data will be lost and the reconstruction phase cannot be performed.</p> <p>Unification of log formats</p> <p>In order to analyze the acquired data from CSPs, investigators must deal with multiple log formats. Thus, this process is considered as a time-consuming process.</p> <p>Identity</p> <p>In regular digital forensics, a user is associated with the data that is stored in his/her personal computer device. This method is straight forward since the computer device is found in his/her house. The investigation of cloud forensics is complicated since data are maintained and preserved in different and multiple remote locations in a multi-tenant environment. Thus, determining the owner of the data among a large number of cloud users is considered as an intricate process.</p> <p>Another situation is upon a person involved in a criminal movement through his/her VM from a veiled IP address and then claims that his/her credentials have been compromised from another user.</p> <p>Complexity of testimony</p> <p>Investigators should be ready to handle a case where the jury consists of persons with only the basic knowledge in computer systems. Thus, investigators must be prepared to provide a complete explanation of cloud computing (how cloud computing operates), and cloud forensics (how the evidence collected maintained and reported during the investigation). This is considered as an essential issue towards the advances of the trial.</p> <p>Documentation</p> <p>Investigators must ensure that all principles and methods have to be followed by all people and parties involved in the investigation to maintain the chain of custody related to the gathered evidence. All stages should be documented for digital evidence.</p>
--	---

Activity Template	
Number	5.1
Title	Review a research paper entitled "Cloud storage forensics: ownCloud as a case study"
Type	<ul style="list-style-type: none"> • Review
Aim	This activity aims at providing students with an in-depth understanding of the artefacts required to undertake cloud storage forensics. This activity is related to outcome #3.
Description	The storage as a service (StaaS) cloud computing architecture is showing significant growth as users adopt the capability to store data in the cloud environment across a range of devices. Cloud (storage) forensics has recently emerged as a salient area of inquiry. Review the paper entitled "Cloud storage forensics: ownCloud as a case study" that aims at providing forensic researchers and practitioners with an in-depth understanding of the artefacts required to undertake cloud storage forensics. Report their experiments that focus upon client and server artefacts, which are categories of potential evidential data specified before the beginning of the experiments.
Timeline	2 hours
Assessment	Brief description of how the activity will be assessed.

Think Template (MCQs)	
Number	5.2
Title	digital and cloud forensics
Type	Fill in the blanks
Question	Evidence identification in cloud environments is a difficult process because of ----- and -----.
Answers	Evidence identification in cloud environments is a difficult process because of the different deployment and service models, and also the limitation of seizing (physically) the computer device containing the evidence.

Think Template (MCQs)	
Number	5.2
Title	Cloud Forensic Challenges
Type	Choose correct answer
Question	In ----- and ----- models, the only application that allows the client system to communicate with cloud services is the user agent (e.g. the web browser)
Answers	<ul style="list-style-type: none"> a. SaaS b. IaaS c. PaaS and SaaS d. SaaS and IaaS

Extra Template	
Number	5.1
Title	<p>References</p> <ol style="list-style-type: none"> 1. McKemmish, R. What Is Forensic Computing?; Australian Institute of Criminology: Canberra, Australia, 1999. 2. United States Computer Emergency Readiness Team (US-CERT), Computer Forensics. Available online: https://www.us-cert.gov/sites/default/files/publications/forensics.pdf (accessed on 14 May 2016). 3. Kruse, W.G., II; Heiser, J.G. Computer Forensics: Incident Response Essentials, 14th ed.; Pearson Education: Indianapolis, IN, USA, 2010. 4. UK Legislation, Criminal Damage Act 1971. Available online: http://www.legislation.gov.uk/ukpga/1971/48/contents (accessed on 8 May 2016).
Topic	5.2
Type	<p>Could include:</p> <ul style="list-style-type: none"> • Book/Chapter (ISBN) • Offline content (Full reference required) • Online content (URL)

Extra Template	
Number	5.2
Title	<p>References</p> <ol style="list-style-type: none"> 5. Grispos, G.; Storer, T.; Glisson, W.B. Calm before the storm: The Challenges of cloud computing in digital forensics. <i>Int. J. Digit. Crime Forensics</i> 2012, 4, 28–48. 6. ISO/IEC 27043:2015. Incident Investigation Principles and Processes; The International Organization for Standardization (ISO); The International Electrotechnical Commission (IEC) ISO/IEC: Geneva, Switzerland, 2015. 7. Catteddu, D. Cloud computing: Benefits, risks and recommendations for information security. In <i>Web Application Security</i>; Springer: Berlin/Heidelberg, Germany, 2010; p. 17.
Topic	5.3
Type	<p>Could include:</p> <ul style="list-style-type: none"> • Book/Chapter (ISBN) • Offline content (Full reference required) • Online content (URL)

5. Bitcoin and Cryptocurrencies

Scope Template	
Number	6
Title	Bitcoin and Cryptocurrencies
Introduction	The chapter will cover all aspects of crypto currencies. The technical aspects, its uses, how its generated, the future and its unregulated nature will all be covered in this course. This chapter focuses on the bitcoin as an example of a crypto currency. It demonstrates the bitcoin mining, and illustrates how to use and manage a bitcoin.
Outcomes	At the end of this chapter students will be able to: <ul style="list-style-type: none">1- Understand a bitcoin and its usage.2- Explain the bitcoin mining.3- Manage a virtual currency and transact via a wallet.4- Illustrate the blockchain technology
Topics	6.1. Introduction to cryptography and crypto-currencies 6.2. Blockchain 6.3. Bitcoins 6.4. Bitcoins Mining 6.4.1. Finding a valid block 6.4.2. Creating new bitcoins 6.5. Using bitcoins 6.6. Managing your bitcoin
Study Guide	Instructions on how to study this unit. <ul style="list-style-type: none">• Required study time:<ul style="list-style-type: none">○ Preparation: 1 hour○ Disk-based Content: 2 hours○ Practical activities: 2 hours

Content Template	
Section Number	5.1
Section Title	Introduction to cryptography and cryptocurrencies
Introduction	This section gives a brief introduction about cryptocurrencies
Content	<p>A cryptocurrency (or crypto currency) is a type of digital asset, digital currency, alternative currency or virtual currency. A crypto currency works as a medium of exchange that uses cryptography and decentralized controls without a central bank or single administrator. Because of cryptography, all crypto currency transactions are secure. Additionally, cryptography manages the creation of additional units and validates the assets transfer. Through a blockchain, which is a public transaction database functioning as a distributed ledger, the decentralized control of each cryptocurrency is maintained. More details about the blockchain will be given in section 2.</p> <p>According to Jan Lansky, a cryptocurrency can be formally defined as a system that meets six main conditions:</p> <ol style="list-style-type: none"> 1- The system does not require a central authority, but a distributed consensus on its state. 2- The system keeps an overview of cryptocurrency units and their ownership. 3- The system defines whether new cryptocurrency units can be created. If so, the system defines the circumstances of their origin and how to determine the ownership of these new units. 4- Ownership of cryptocurrency units can be proved exclusively cryptographically. 5- The system allows performing transactions in which the ownership of the cryptographic units is changed. A transaction statement can only be issued by an entity proving the current ownership of these units. 6- If two different instructions for changing the ownership of the same cryptographic units are simultaneously entered, the system performs at most one of them. <p>In 2009, the first cryptocurrency, namely Bitcoin, was launched by an individual or group known under the pseudonym Satoshi Nakamoto. Satoshi detailed how the cryptocurrency would work. According to him, bitcoins were to be mined by specific computer software and hardware (Section 4 details the bitcoin mining). After mining, the currency would be transferred directly from one user to the other without any intermediary. This movement of bitcoins would then be recorded in a blockchain. The success of Bitcoin has generated several competing cryptocurrencies, such as Altcoin, Ethereum, Litecoin, Namecoin and PPCoin.</p> <p>As mentioned before, cryptocurrencies rely on a peer-to-peer decentralized system to conduct transactions, such as transferring funds between two parties. Therefore, all transactions must be extremely secure via cryptography. Cryptography is a method of using advanced mathematical principles in storing and transmitting data in a particular form so that only those for whom it is</p>

intended for can read and process it. Two main cryptographic technologies are used to build Bitcoin: **public-key cryptography** and **hash functions**.

- 1- Public-key cryptography, or asymmetrical cryptography: this type of cryptography uses two keys; one is the public key, which may be disseminated widely, and the other is the private key. Public-key cryptography can be used for confidentiality, authentication, or both.

When a transaction is created, each coin is associated with the public key (ECDSA or Elliptic Curve Digital Signature Algorithm) of its current owner. Suppose Alice wants to send a bitcoin to Bob (see Figure 1). Alice creates a message (transaction), attaching Bob's public key to a number of bitcoins, and signs it with her private key. Upon broadcasting this transaction to the Bitcoin network, Alice's signature on the message confirms to everyone that the message is authentic. The complete history of transactions is kept by everyone (blockchain), so anyone can verify who is the current owner of a certain group of coins.

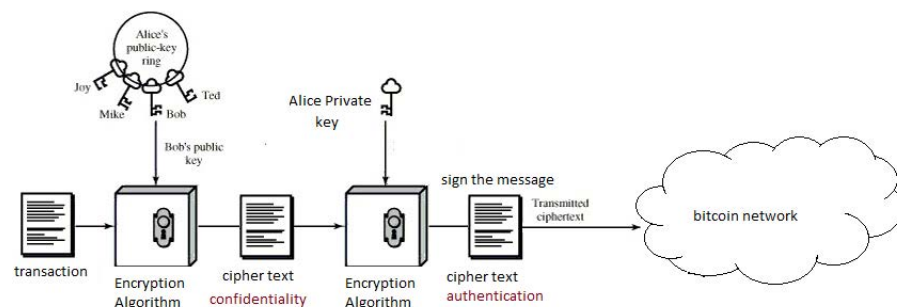


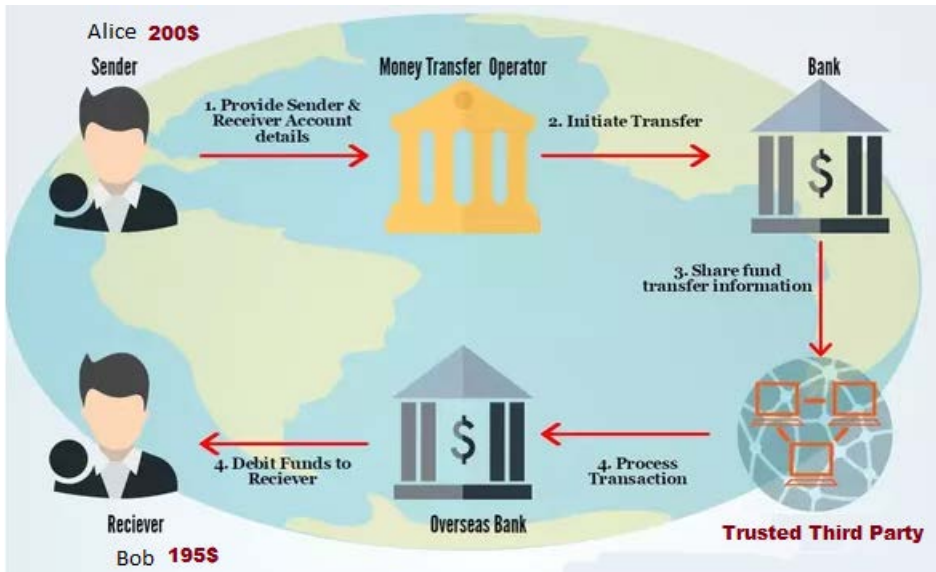
Figure 1. Bitcoin transmission

- 2- Hash Function: a hash function is another cryptography technology used to generate a bitcoin. A cryptographic hash function essentially takes input data, which can be of practically any size, and transforms it, in an effectively-impossible to reverse or to predict way, into a relatively compact string ($H(X)=h$, where X is a variable-length message and $H(X)$ is the fixed-length hash value). In the case of SHA-256, the output of the hash function is 32 bytes or 256 bits.

Actually, the purpose of a hash function is to produce a "fingerprint" of a file, message, or another block of data. However, to be useful for authentication, a hash function H must have the following properties:

1. H can be applied to a block of data of any size.
2. H produces a fixed-length output.
3. $H(x)$ is relatively easy to compute for any given x , making both hardware and software implementations practical.
4. A hash function should meet the one-way property. For any given value h , it is computationally infeasible to find x such that $H(x) = h$.
5. A hash function should meet the weak collision resistance. For any given block x , it is computationally infeasible to find $y \neq x$ such that $H(y) = H(x)$.
6. It is computationally infeasible to find any pair (x, y) such that $H(x) = H(y)$. This property is referred to as strong collision resistance.

	<p>Making the slightest change to the input data changes its hash unpredictably, so nobody can create a different block of data that gives exactly the same hash. Therefore, by giving a compact hash, you can confirm that it matches only a particular input, and in Bitcoin the input data being added to a blockchain is significantly larger than the SHA-256 hash. In this way, Bitcoin blocks don't have to contain serial numbers, as blocks can be identified by their hash, which serves the dual purpose of identification as well as integrity verification. An identification string that also provides its own integrity is called a "self-certifying identifier".</p>
--	--

Content Template	
Section Number	5.2
Section Title	Blockchain
Introduction	This section introduces the blockchain technology.
Content	<p>A blockchain is a technology that enables transferring digital currencies or assets from one individual to another. Now, I will explain the difference between traditional money transfer and digital money transfer in order to ease understanding of the blockchain.</p> <p>Assume that Alice resides in U.S.A and wants to transfer 200\$ to Bob, who resides in Japan (See Figure 2). Money transfer is accomplished via a trusted third party. First, Alice via her bank sends 200\$ to Bob. The bank transfers the transaction to the trusted third party in order to identify the receiver Bob (the verification will cost some fees). After verifying, the money will be transferred to Bob's account and Bob will receive 195\$. This transaction will be accomplished by 3 to 5 working days. The idea behind blockchain is to send money without a trusted third party, to transfer money faster (it will be transferred immediately by the blockchain) and to eliminate the third party fees.</p>  <pre> graph LR Alice[Alice 200\$ Sender] -- "1. Provide Sender & Receiver Account details" --> MTO[Money Transfer Operator] MTO -- "2. Initiate Transfer" --> Bank[Bank] Bank -- "3. Share fund transfer information" --> TTP[Trusted Third Party] TTP -- "4. Process Transaction" --> OB[Overseas Bank] OB -- "4. Debit Funds to Reciever" --> Bob[Reciever Bob 195\$] </pre> <p>Figure 2. Typical Money Transfer</p> <p>A blockchain is an open ledger (chain of transactions).</p> <p>Figure 3 details the concept of open ledger. Assume that we have a network of four users. User A has 10\$. Suppose that user A wants to transfer 5\$ to user B. A transaction will be added to the ledger and money will be transferred. If user B wants to send 3\$ to user C, another transaction, linked to the previous one, will be added to the ledger and money will be sent. When user C wants to send 1\$ to user D, a new transaction will be added to the ledger and linked to the previous two. The ledger is open that all users in the network can see where the money is and how much a user has in his/her wallet?, this allows to decide whether the transaction is valid or not. Suppose that user A wants to send 15\$ to user D. All users in the network can know</p>

that A has only 5\$ in his/her wallet and this transaction is not valid. Therefore, the transaction will not be added to the ledger.

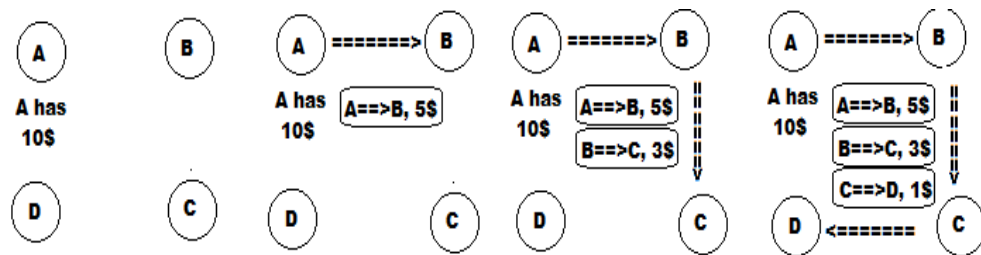


Figure 3. Open ledger concept

The blockchain is distributed: any user of the network can have a copy of the ledger (see figure 4).

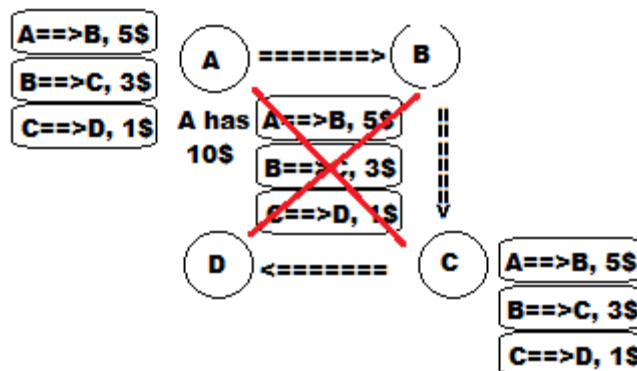


Figure 4. blockchain is distributed

The users that have a copy of the ledger are called "minors". For example, user A has a copy of the ledger, as does user D. Thus, there is no need to keep a centralized copy of the ledger. All copies should be synchronized. A user who has a copy of the ledger is the only user who can verify new transactions. Suppose now that user B wants to send 1\$ to user D. Only user A and user C have a copy of the ledger; thus, only the two of them can verify the new transaction. User A and user C will compete in verifying the transaction (a transaction verification is called "mining", see section 4). The winner will verify it and will add a new transaction to the ledger. He will then distribute it to the network with a key to allow other miners adding the transaction to their copy of the ledger (see Figure 5).

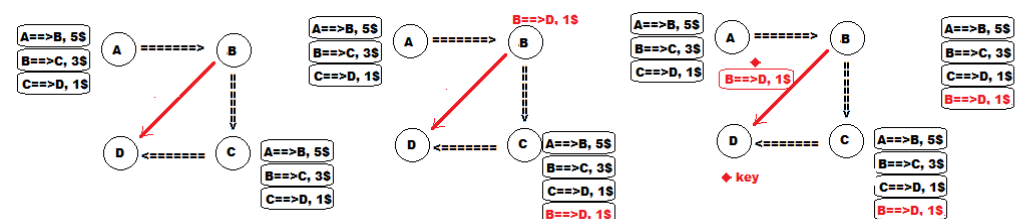


Figure 5. transaction verification

In summary, a blockchain is a sequence of records called "blocks". Each block contains a group of transactions that have been sent since the previous block. All computers in the Bitcoin network have a copy of the blockchain (see figure 6), which they keep updated by passing along new blocks to each other. Integrity is preserved by the blockchain since each block confirms the integrity of the previous one, all the way back to the genesis block which is the first block of the blockchain. Record insertion is costly because each block must meet certain requirements that make it difficult to generate a valid block. In this way, no party can overwrite previous records by just forking the chain.

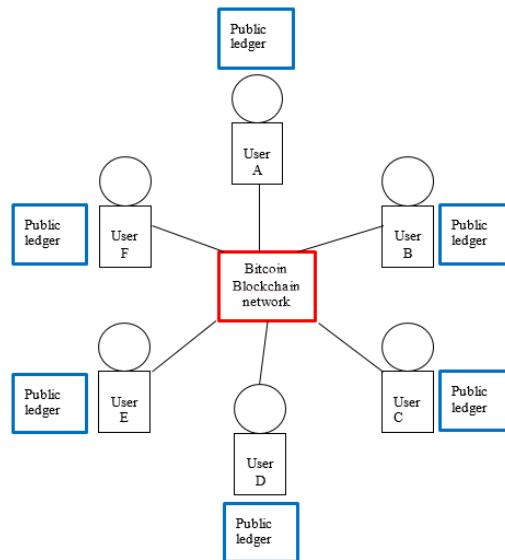


Figure 6. blockchain network

Content Template	
Section Number	5.3
Section Title	Bitcoin
Introduction	The section covers basic properties of Bitcoin and the mechanics behind it.
Content	<p>Bitcoin is the first worldwide decentralized digital currency or cryptocurrency, which was invented by Satoshi Nakamoto as a reward for mining and released as open-source software in 2009. Bitcoin is also the name of a peer to peer protocol that allows a network of computers to supervise all the rules of the bitcoins without a central bank or a single administrator. It allows users to directly perform transactions without an intermediary. Bitcoins are created through a process called "mining" (section 4)</p> <p>The Bitcoin network consists of: users with wallets containing keys (a wallet is a place for users to store the bitcoins they own; the term also refers to the bitcoin client which must be installed to become part of the Bitcoin network), transactions that are propagated across the network, and miners who produce (through competitive computation) the consensus blockchain, which is the authoritative ledger of all transactions.</p> <p>A blockchain, as described in section 2, is a database containing historical records of all the transactions that ever occurred in the network. Every full node in the network has a copy that he keeps up to date and verifies. Some nodes extend the blockchain, they are called "miners". The blockchain is defined as a decentralized or distributed digital ledger (database) of transactions that can be implemented to record both financial transactions and virtually everything of value. The blockchain is distributed and stores identical blocks (a Bitcoin block is a record of transactions that are grouped together) of information across the entire network. Thus, there is no control of a single entity and it cannot have any single failure point. The first block, block 0, is called "Genesis Block" (see figure 7).</p> <div data-bbox="496 1391 1340 1583" data-label="Diagram"> <pre> graph RL B0[Block 0 Genesis Block] <-- B1[Block 1] B1 <-- Dots[...] Dots <-- BN1[Block N-1] BN1 <-- BN[Block N] </pre> </div> <p>Figure 7. blockchain overview</p> <p>A Bitcoin client software simultaneously allows users to manage and spend bitcoins. It preserves the blockchain in order to record all transactions confirmed by the Bitcoin network, which consists of thousands of machines running the Bitcoin software. The Bitcoin network broadcasts transaction information and verifies these transactions to ensure that the same bitcoins cannot be spent twice.</p> <p>Bitcoins can be exchanged for other currencies, products and services. As of February 2015, over 100,000 merchants and vendors accepted bitcoin as payment. A research study conducted by the University of Cambridge</p>

	estimates that in 2017 there were 2.9 to 5.8 million unique users owning a cryptocurrency wallet, most of them using Bitcoin.
--	---

Content Template																															
Section Number	5.4																														
Section Title	Bitcoin Mining																														
Introduction	This section covers the mining process of the bitcoins																														
Content	<p>Bitcoin mining is a term that describes the process of solving mathematical problems in order to verify transactions. The reward for calculating the right answer is a fixed number of bitcoins that are released into the Bitcoin network. The bitcoin mining is defined in the Bitcoin protocol, implemented in the Bitcoin software, and it is considered as an essential function in managing the Bitcoin network, securing the Bitcoin system and enabling the lack of a central authority.</p> <p>Mining is the process of validating new transactions and recording them in the blockchain. Mining verifies transactions by evaluating them against the ones that happened previously and prevents "double-spending". This term refers to transactions with non-existing bitcoins or bitcoins that were already spent. They must send bitcoins to valid addresses (a Bitcoin address is like a physical address in the real world or an email address) adhere to every rule defined by the protocol, collect transaction fees, create the money supply, and protect the network by piling tons of processing power on top of past transactions. Miners additionally verify blocks generated by other miners to enable the entire network to continue to build on the blockchain.</p> <p>On average, new blocks from the latest transactions are generated every 10 minutes, and the number of bitcoins defined by the current block reward is produced. In other words, a block in a blockchain is mined every 10 minutes. When miners, based on a cryptographic hash algorithm (SHA256), solve difficult mathematical problems, the mining output is called "proof-of-work", which means the miner spends a lot of time and resources to solve the problem. The transaction is considered as "confirmed" when a block is solved and the bitcoin involved in the transaction is spent. That means a transaction needs about 10 minutes to be confirmed.</p> <p>In general, new bitcoins and transaction fees are two types of rewards that miners can gain. Every 210,000 blocks (approximately 4 years), the number of issued bitcoins is decreased (it is divided by 2). Nowadays, a newly created block issues 12.5 bitcoins. According to table 1, this number will keep decreasing until no more bitcoins will be issued. By 2141, when approximately 21 millions of bitcoins will have been issued, no more bitcoins will be created. "Satoshis" are the smallest unit of bitcoin. 50 bitcoins = 5000000000 satoshis</p> <p>Table 1. Number of bitcoins created with each mined block</p> <table><tr><th colspan="10">Amount of bitcoin created with each mined block (dates are indications)</th></tr><tr><th>Before November 2012</th><th>Nov 2012 - Jul 2016</th><th>Jul 2016 - 2020</th><th>2020 - 2024</th><th>2024 - 2028</th><th>2028 - 2032</th><th>2031 - 2036</th><th>2036 - 2040</th><th>2040 - 2044</th><th>2044 - 2048</th></tr><tr><td>50</td><td>25</td><td>12,5</td><td>6,25</td><td>3,125</td><td>1,5625</td><td>0,78125</td><td>0,390625</td><td>0,1953125</td><td>0,09765625</td></tr></table>	Amount of bitcoin created with each mined block (dates are indications)										Before November 2012	Nov 2012 - Jul 2016	Jul 2016 - 2020	2020 - 2024	2024 - 2028	2028 - 2032	2031 - 2036	2036 - 2040	2040 - 2044	2044 - 2048	50	25	12,5	6,25	3,125	1,5625	0,78125	0,390625	0,1953125	0,09765625
Amount of bitcoin created with each mined block (dates are indications)																															
Before November 2012	Nov 2012 - Jul 2016	Jul 2016 - 2020	2020 - 2024	2024 - 2028	2028 - 2032	2031 - 2036	2036 - 2040	2040 - 2044	2044 - 2048																						
50	25	12,5	6,25	3,125	1,5625	0,78125	0,390625	0,1953125	0,09765625																						

Content Template	
Section Number	5.4.1
Section Title	Finding a Valid Block
Introduction	This section illustrates the process of finding a valid block during the mining process
Content	<p>A miner creates a list of recent transactions to find a valid block. Additionally, the miner collects some information about the proposed block (previous block hash). Then, the information will be integrated with a nonce in order to create a block header and calculate the hash of the block. A check of this hash will be made to see if it is small enough to win the current complexity level. Actually, when the calculated hash does not win the problem, the nonce is changed and a new hash is calculated and tested.</p> <p>We can infer that there is no way to find a valid block. Thus, a brute force technique is used. "Brute force" means trying one nonce, then another if the first one does not win, and another one, repeating the process until the miner wins and a valid block is created. In a brute force technique, there is no way to predict that the next nonce will have a smaller hash than the last. Therefore, the only way to increase the chance of winning is to boost the speed of nonces testing. The more processing power you have at your disposal, the faster you can search and the more likely you will be able to find a valid block.</p> <p>A valid block will be then broadcasted to the Bitcoin network and will be verified by the other nodes in the network. The effort to find a winning number is adjusted every 2016 blocks, therefore blocks are generated every 10 minutes on average.</p> <p>To illustrate the Bitcoin mining process, we will follow the lifecycle of one block from its construction to its final validation through an example.</p> <p>Alice is a miner who was competing to validate the block 502446, but unfortunately someone else solved the math problem before her. Thus, once block 502446 was mined, Alice's computer (node) updated her local blockchain copy and started creating a new candidate block, block 502447. Alice's computer is now searching the proof-of-work for the previous block and at the same time it is listening for new transactions. The new transactions will be added to the node "memory pool" (or "transaction pool"), where they wait until they can be included in a new block once validated.</p> <p>When Alice's node is notified that block 502446 has a valid Proof-of-Work, a new candidate block will be constructed. All transactions added previously in the memory pool will be gathered, and all the transactions related to the previous block are removed, others will be used to construct a candidate block. The block is called "candidate" since it does not have valid proof of work yet.</p> <p>The first thing that Alice's node does is creating a coinbase transaction, which means the node gets rewarded for the block mining. The coinbase transaction says, "pay Alice's wallet address with xxx BTC to reward her of the valid proof of work". Additionally, Alice's node calculates the transaction fees in the block. Alice's reward = mining reward + transaction fees.</p> <p>Alice's node must create a block header for the block that is mining. Three different set of data construct the block header: the previous block hash, the Merkle tree root (Merkle trees are binary trees of hashes. In Bitcoin they use a double SHA-256, i.e. the SHA-256 hash of the SHA-256 hash of something), and the data for mining competition (the version number to track the software and/or protocol upgrades, the timestamp, i.e. the seconds from Unix epoch</p>

when the block was created, the target, that is the proof of work algorithm targeting the block, the nonce, i.e. the counter used for the proof of work algorithm).

Everything is now ready for Alice's computer to mine the block by finding a value for the nonce that will result in a hash lower than the target. Thus, the mining node will try billions or trillions of nonce values before it gets a valid hash. After approximately 10 minutes, Alice's node discovers a valid hash. Alice's node immediately transmits the block to all its peers that need to validate the new block before broadcasting it to its peers. The validation process means ensuring that:

- 1- the hash of the block header is less than the target
- 2- the block size is within acceptable limits
- 3- the timestamp is valid
- 4- a coinbase transaction is the first and the only transaction and it has a valid reward
- 5- all transactions within the block are valid

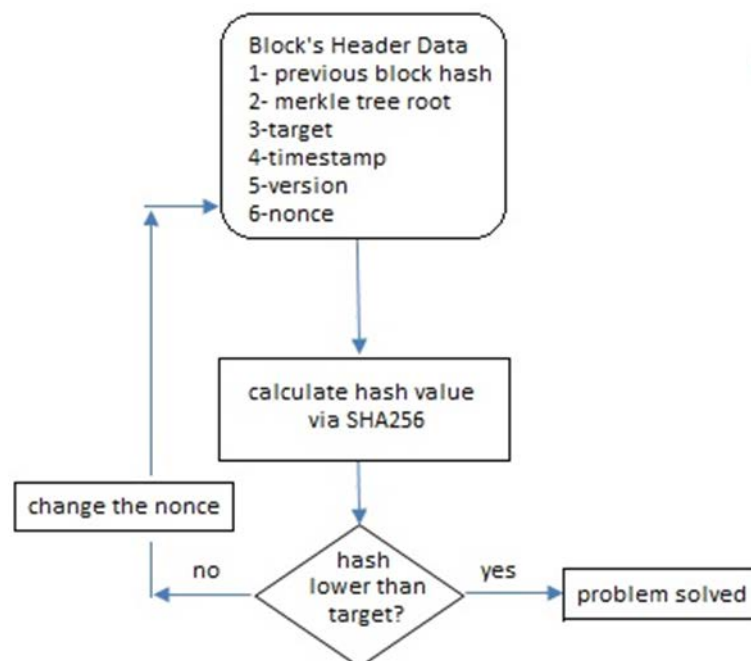
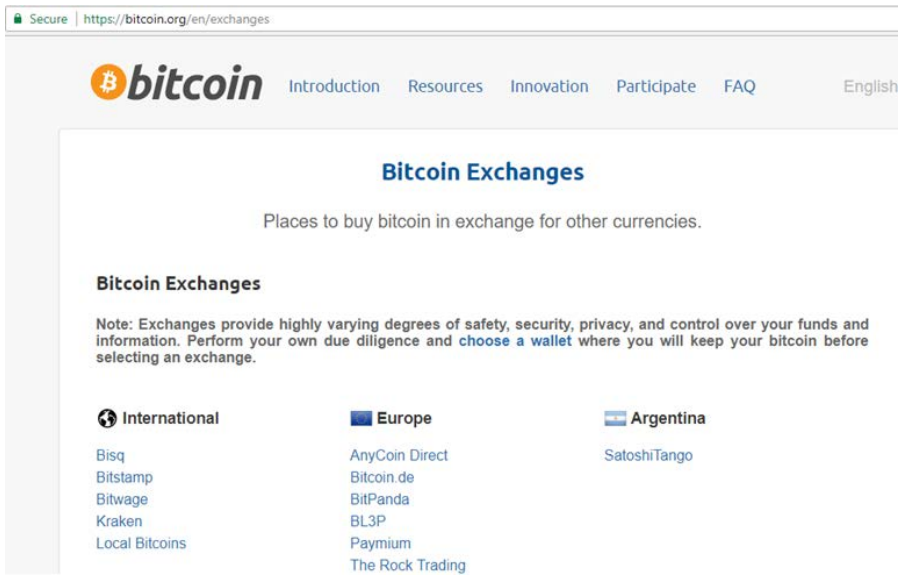


Figure 8. transaction validation

Every node in the Bitcoin network validates independently new blocks following the same previously mentioned rules (the key idea of decentralization). If a block is valid, then all the other miners will update their own copy of the blockchain with the new block 502447. Alice's block hash is now used by the other miners to mine block 502448.

Content Template	
Section Number	5.4.2
Section Title	Creating New Bitcoins
Introduction	This section shows the creation of new bitcoins
Content	<p>Once a miner finds a new valid block, the miner includes the new address to reward (rewards = new bitcoins + some transaction fees). This reward is the monetary incentive for users like you and me to run miners.</p> <p>As mentioned before, 50 bitcoins are awarded to a miner who finds each valid block. This will continue until block 210,000 is found. After four years, the block reward will be halved to 25 bitcoins. The reward will be halved again 210,000 blocks thereafter. This means the maximum number of bitcoins that can be created is 21 million (it is estimated to occur in near 2041).</p> <p>You may ask, "where do these bitcoins come from?"</p> <p>They are created by the Bitcoin network as part of the Bitcoin protocol. This is the same process that created any bitcoin you will ever own or use.</p>

Content Template	
Section Number	5.5
Section Title	Using Bitcoins
Introduction	This section illustrates how to use bitcoins
Content	<p>Bitcoins are virtual currencies (they are money and can be used as such) or units of currency used to store and transmit value among the participants of the Bitcoin network. Users can transfer bitcoins over the network to do everything that can be done with conventional currencies, including buy and sell goods, send money to people or organizations, or extend credit.</p> <p>Nowadays, a growing number of services and merchants accepts bitcoins all over the world. For those who only accept dollars, euros, yen and other national currencies, a user needs to exchange his/her bitcoins through an online Bitcoin exchange.</p>  <p>Figure 9. Bitcoin Exchanges</p> <p>In order to use online Bitcoin Exchanges to sell bitcoins, a user has to create an account, send them the bitcoins and place an order to sell. Once an order to buy these bitcoins appears, the user's bitcoins will be traded for his/her preferred currency. By a money transfer method that is compatible with Bitcoin, the user will get his/her money.</p> <p>Another important feature of Bitcoin is irreversibility, which means that once bitcoins are sent to an address, there is virtually no way to reverse the transaction.</p>

Content Template	
Section Number	5.6
Section Title	Managing your Bitcoins
Introduction	This section shows how to manage bitcoins
Content	<p>Managing Bitcoins is defined as the process of handling the bitcoins properly. At the first run, the Bitcoin client software creates a Bitcoin wallet. The wallet includes the user private keys (similar to the secret PIN, or signature on a check) that consist of random letters and numbers. Owning the private keys means owning the bitcoins stored in the wallet. For each private key, there is a corresponding public key (similar to the bank account) and a corresponding Bitcoin address. Note that usually the Bitcoin address is created from the Bitcoin public key. Assume that Alice would like to send some bitcoins to her friend Bob. First, Alice creates a message (transaction), attaching Bob's public key and address to the number of bitcoins, and signs it with her private key (via her Bitcoin client software). Upon broadcasting this transaction to the Bitcoin network, Alice's signature on the message confirms to everyone that the message is authentic. The complete history of transactions is kept by everyone (blockchain), so anyone can verify the current owner of the bitcoins.</p> <p>With Bitcoin "being your own bank", securing your bitcoins is a key issue. When handling significant quantities of bitcoins, one mistake can have considerable consequences. Several users lose bitcoins because of technical failures or thefts from online exchanges and wallets. Most of these failures can be prevented by following these basic rules:</p> <ul style="list-style-type: none"> • Enabling a "two-factor authentication" on the user account when using an online service such as a Bitcoin exchange. <p>A two-factor authentication (2FA) is a verification process consisting of two steps. It is an extra security layer that demands in addition to a user password and username some information that only that user knows or has immediately at hand, such as a physical token. Enabling 2FA makes it harder for intruders to access and steal user personal data or identity. Enabling 2FA protects the user account from theft even if someone (attacker) steals the user password.</p> <ul style="list-style-type: none"> • Encrypting the wallet with stronger passwords. Via a strong password, it will be harder for the attacker to steal the user's coins. Note that setting up the wallet for the first time allows the user to get the secret key in the form of a QR code. This code can be scanned by an authenticator app and should be printed out and kept in a safe place, with the aim to help the user to access his/her account if the smartphone is lost or stolen. • The user should have direct control of the private keys that correspond to the bitcoins. Without the private keys, there is no control on the bitcoins. • Storing multiple backups of the Bitcoin wallet <p>Wallet backup protects the user's bitcoins against computer or software failures and also retrieves user's funds if his device is stolen or lost. There are three popular methods that are currently used to back up a Bitcoin wallet.</p>

	<p>1- Wallet.dat</p> <p>This method aims at backing up a wallet.dat file that includes private keys, public keys and the corresponding addresses, transactions information, and other metadata.</p> <p>2- BIP 32</p> <p>This method is responsible for determining the key pairs from a single seed. This seed is used to create a master private key (xprv) that can be imported to any BIP 32 compliant HD wallet.</p> <p>3- BIP 39</p> <p>BIP 39 is a method of generating a mnemonic sentence, that is a group of easy-to-remember words. BIP 39 represents a seed that determines wallet addresses (e.g. BIP 32). Upon creating a wallet for the first time, a list of 12 to 24 seemingly random words will appear. More words ensure better security. On the other hand, some wallets like Trezor allow the user to select the entropy himself.</p>
--	--

Activity Template	
Number	1
Title	Explain why it is computationally infeasible for anyone to generate a Bitcoin transaction that references its own output as an input.
Type	<ul style="list-style-type: none"> • Research
Aim	1- Understand Bitcoin and its usage.
Description	This activity shows the students understandings of Bitcoin and its technologies
Timeline	1 week
Assessment	Each student is required to submit a one-page report. The report will be assessed on the basis of completeness, correctness and overall quality.

Activity Template	
Number	2
Title	How would social media (like Facebook) be impacted by blockchain? What do you mean specifically by 'impact' in this context?
Type	<ul style="list-style-type: none"> • Reflection
Aim	4- Illustrate the blockchain technology
Description	Currently, Facebook knows everything about its users and can sell advertisers much targeted information for advertising purposes. With blockchain-based social media, individuals will have much more control over their personal information and will be able to use cryptocurrencies to facilitate micro-transactions for selling bits of their information. This model is fundamentally a distributed model, which is the opposite of the current Facebook-style social media.
Timeline	3 hours
Assessment	Each student is required to submit a one-page report. The report will be assessed on the basis of completeness, correctness and overall quality.

Think Template (MCQs)	
Number	1
Title	5.3. Bitcoin
Type	Choose the correct answer
Question	Who created Bitcoin?
Answers	a Notosohi Sakamoto b Gavin Andresen c Satoshi Nakamoto d Dorian Nakamoto e Paul Krugman

Think Template (MCQs)	
Number	2
Title	5.3. Bitcoin
Type	Choose the correct answer
Question	How many bitcoins will ever be created?
Answers	a Unlimited b 77,340,109 c 21 million but it can be adjusted by the Bitcoin Foundation by majority vote. d 21,000,000 e The square root of 2^2

Think Template (MCQs)	
Number	3
Title	5.2 Blockchain
Type	Choose the correct answer
Question	What is the name of the general ledger that tracks all Bitcoin transactions?
Answers	a The Gox-Chain b The Block-link c The Block-chain d Ledger-Link e Satoshi-square

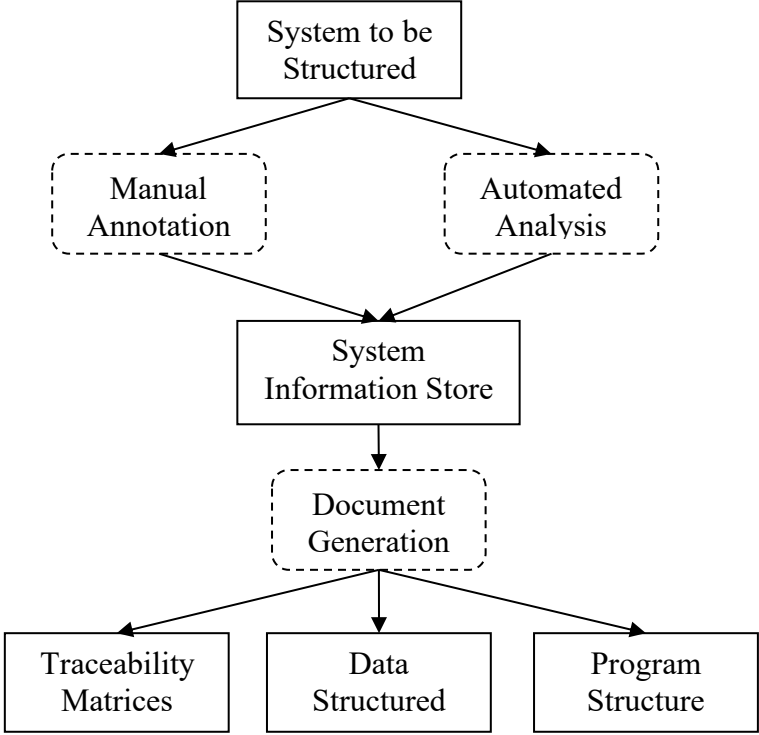
Extra Template	
Number	5.1
Title	Research paper that includes references to most of the literature on Bitcoin
Topic	Section 5.1, 5.2
Type	J. Bonneau, A. Miller, J. Clark, A. Narayanan, J. A. Kroll and E. W. Felten, "SoK: Research Perspectives and Challenges for Bitcoin and Cryptocurrencies," <i>2015 IEEE Symposium on Security and Privacy</i> , San Jose, CA, 2015, pp. 104-121. doi: 10.1109/SP.2015.14

6. Digital Forensics Reverse Engineering Fundamentals

Scope Template	
Number	6
Title	Digital Forensics Reverse Engineering Fundamentals
Introduction	<p>This chapter entitled "Digital Forensics Reverse Engineering Fundamentals ", is an introduction chapter about using reverse engineering in the field of digital forensics. Section one defines Reverse Engineering terms and their general use. The section 6.2 explains the vulnerability analysis using reverse engineering by reviewing some reverse engineering methods and approaches that can be used to find systems vulnerability. An example of using reverse engineering to analyze internet protocols is discussed in section 6.3. Moreover, section 6.3 contains a list of reverse engineering tools and approaches that can be used to analyze internet protocols. Section 6.4 explains how to use reverse engineering techniques and tools to analyze malware. In the last section, the reader can find a brief summary of some reverse engineering tools that can be used to analyze malware.</p>
Outcomes	<ol style="list-style-type: none"> 4. Gain basic knowledge about using reverse engineering tools and techniques in the field of digital forensics. 5. The ability to analyze network protocols using reverse engineering tools and techniques. 6. The ability to analyze system vulnerability using reverse engineering tools and techniques. 7. The ability to detect and analyze malware using reverse engineering tools and techniques to detect and analyze malware.
Topics	<ol style="list-style-type: none"> 1. Introduction to Reverse Engineering 2. Vulnerability Analysis Using Reverse Engineering 3. Protocol Reverse Engineering 4. Malware Analysis Using Reverse Engineering Techniques 5. Reverse Engineering Tools for Malware Analysis
Study Guide	<p>Instructions on how to study this unit.</p> <ul style="list-style-type: none"> • Required study time: 13 Hours. • Unit comprehensive reading. • Refer to external resources for more details such as the references appeared in the text • You are required to have a PC or laptop to install a virtual machine and some other tools and software to be able to try the examples and do the activities. <ul style="list-style-type: none"> ✓ Required software: <ul style="list-style-type: none"> ▪ A virtual machine. ▪ Disassembler, debugger, and decompiler. ▪ Penetration testing tools such as Namp, Hping, p0f, Httprint, Xprobe2, and Super Scan.

Content Template	
Section Number	6.1
Section Title	Introduction to Reverse Engineering
Introduction	This section is dedicated to explaining the definition of reverse engineering and how we can use reverse engineering and for what purposes. Besides, it briefly discusses the fields and its applicability to apply reverse engineering tools and procedure.
Content	<p>Reverse engineering can be defined as an evaluation process of a product, system or an object to extract knowledge from them by analyzing its structure, functions, and operations. By analyzing the extracted knowledge of an object, the analyst can define the weakness and strengthens of the object, and then he can use this knowledge to redesign the object to improve its functions and operations. Reverse engineering can also be used to copy an existing object or some of its parts, or recover a damaged object or some of its parts.</p> <p>Reverse engineering applies to variety of fields such as software engineering, mechanical engineering, and chemical engineering. Depending on the nature of the field, the reverse engineering procedure should be adapted and should use the appropriate tools to be applied to the selected field. For example, we apply reverse engineering in software engineering to get the source code of a program, improve program performances, fix bugs and identify malicious content in a program by using many tools such as system monitoring, disassemblers, debuggers, and decompilers.</p> <p>In contrast to the forward engineering, reverse engineering starts from the target (final product, system or object) to get the requirements by going through design and analysis phases. Reverse engineering processes in the software field can be described in abstract with the following activities: collecting information about the target; examining and analyzing information; recording functionalities; recording dataflow and control flows; reviewing recovered designs.</p> <p>In conclusion, reverse engineering is a systematic procedure to study an object in detail to construct a basic knowledge of the object that is studied. Then, reverse engineering uses the basic knowledge about the object in several tasks such as copying the object, redesigning the object to get an enhanced version, and more.</p>

Content Template	
Section Number	6.2
Section Title	Vulnerability Analysis Using Reverse Engineering
Introduction	This section reviews some reverse engineering methods and approaches that can be used to find systems vulnerabilities.
Content	<p>Security is one of the critical issues of any information system. There are many security assessment methods to deal with the information systems security such as proof of correctness methods, legend designs and software engineering environments.</p> <p>The primary goal of penetration testing is to identify the vulnerabilities of the information system by simulating an attack against the target system. Based on the information available and known about the system, penetration testing uses one of three strategies: white box, black box, and gray box. Based on its objective, penetration testing can be carried out in two ways: external and internal testing.</p> <p>Penetration testing can be performed manually or by using automated tools. There are many free penetrating testing tools such as Namp, Hping, p0f, Httprint, Xprobe2, and Super Scan.</p> <p>"Reverse engineering and Vulnerability Analysis is also a part of Cyber security." (Kumar and Alka, 2017). Reverse engineering can be used to track attackers. This strategy provides very useful information that can be analyzed to find the vulnerability and the weaknesses of the system.</p> <p>As introduced in section one, the reverse engineering process is adaptable to be compatible with the field on which it is applied. (Kumar and Alka, 2017) proposed a reverse engineering process to be used in cyber security, as illustrated in Figure 1.</p>

Content Template	
Section Number	6.3
Section Title	Vulnerability Analysis Using Reverse Engineering
Introduction	This section explains the analysis of vulnerabilities by using reverse engineering. It reviews some reverse engineering methods and approaches that can be used to find systems vulnerabilities.
Content	 <pre> graph TD A[System to be Structured] --> B[Manual Annotation] A --> C[Automated Analysis] B --> D[System Information Store] C --> D D --> E[Document Generation] E --> F[Traceability Matrices] E --> G[Data Structured] E --> H[Program Structure] </pre> <p>Figure 7: Process of Reverse Engineering (Kumar and Alka , 2017, p. 951)</p> <p>The system to be structured is the system that needs to be studied and examined to find its vulnerability issues. Reverse engineering tools and techniques can be used to examine a system to get its specific structure. After structuring the system, the system should be examined to find any illegal activity or unwanted operation. This operation can be performed manually or automatically. System information store contains information about the illegal activities and the unwanted operations to be used later in the attack investigation. The investigation of attacks generates a document that describes the illegal activities and operations. The final step of the proposed process by (Kumar and Alka, 2017) which include program structure, data structure and traceability matrices is used to track the attack to find the source of the attack.</p>

Content Template	
Section Number	6.4
Section Title	Protocol Reverse Engineering
Introduction	An example of using reverse engineering to analyze internet protocols is discussed in this section.
Content	<p>A network protocol is a set of rules and standards that describe and control the communication between two or more devices over a network (Forouzan and Fegan, 2007). For example, to initiate and accomplish basic data communication, the software and the hardware involved in the communication have to use some protocols such as TCP/IP and HTTP. HTTPS, SSL, and SFTP are examples of protocols that control data communication with some security level.</p> <p>Each protocol has its parameters, format, semantics and some other specifications that are described in detail in a formal specification document that is published by the protocol developer or owner. However, sometimes we cannot find the protocol specification document for many reasons (e.g., the protocol developer or owner wants to keep it secret). Because the protocol is an essential part of any communication, knowing its specifications is required in digital forensics and network investigation and security. For example, the knowledge of the used protocols and their specifications provide very useful information that can be used in network testing software, intrusion detection, fingerprint generation, and detecting services running on non-standard ports. In the case of unavailable protocol specifications, protocol reverse engineering can be used to extract the protocol parameters, formats, semantics, and some other specifications. (Sija et al., 2018)</p> <p>Protocol reverse engineering can be done manually, but it requires a long time, and it may give incorrect output. Therefore, several automatic reverse engineering tools and methods were developed to overcome the weakness of the manual methods.</p> <p>In conclusion, reverse engineering can be used in the network environment to infer the used protocol's specifications. Then, exploit the inferred specification in several actions such as networks testing and monitoring, and network digital forensics.</p>

Content Template	
Section Number	6.5
Section Title	Malware Analysis Using Reverse Engineering Techniques
Introduction	This section is dedicated to discussing the issues about malicious software and its types. Moreover, this section explains how to use reverse engineering techniques and tools to analyze malware.
Content	<p>Malicious software (Malware) is a general term used to describe a program that runs on a computer to perform specific malicious actions, such as making the infected machine be controlled by unauthorized users or encrypt the data of the system). Malware includes computer Viruses, Worms, and Trojans. Computer Viruses are set of instructions that inject themselves into another program or file. The virus cannot infect the computer unless the user runs or opens the infected object. Unlike Computer Virus, Worms are standalone malicious software, i.e. they do not need to be injected into other programs. They can replicate themselves many times and spread themselves without the need for user interaction. Other examples of malware are Trojans, attacks which that look like legitimate programs with some useful functions, but that hide some harmful functions that can be started when a user activates them. (Sija et al., 2018)</p> <p>Several malware detection methods were proposed and developed. These detection methods belong to one of the following categories: anomaly-based methods, specifications-based methods, and signature-based methods. Anomaly-based methods are based on monitoring the system and program activities to look for anomalous activities. Anomaly-based methods are carried out in two phases: the training phase, and the detection phase. During the training phase, the methods learn the standard behaviors of the system and record them as legal activities. At the detection phase, the methods look for any activity that is not listed in the normal activities list that is learned during the first phase. Anomaly-based methods feature two detection approaches: static and dynamic. In the static approach, the methods read and monitor the structure of the program to detect the malicious code and do not require running the program on the machine. On the contrary, the dynamic anomaly-based detection monitors the program execution to detect malicious code and behaviors. (Idika and Mathur, 2007)</p> <p>"Signature-based detection attempts to model the malicious behavior of malware and uses this model in the detection of malware. The collection of all of these models represents signature-based detection's knowledge. This model of malicious behavior is often referred to as the signature." (Idika and Mathur, 2007)</p>

Content Template	
Section Number	6.6
Section Title	Malware Analysis Using Reverse Engineering Techniques
Introduction	This section is dedicated to discussing the issues about malicious software and its types. Moreover, this section explains how to use reverse engineering techniques and tools to analyze malware dynamically.
Content	<p>Reverse engineering can be used to analyze malware. It can be used to observe the behavior of malware by running it on isolated systems and observe how it interacts with environment elements such as the file system, registry, and the network. Another way in which reverse engineering can be used to analyze malware is by regenerating its source code and learning its structure and capabilities.</p> <p>In the following, you can find the general steps of malware dynamic analysis:</p> <ul style="list-style-type: none"> • Dedicate an isolated host for analysis purpose. Restore the host if anything suspicious occurs. • Take a snapshot of the state of the machine's file system and registry. • Run and interact with the malware. • Stop the malware. • Take another snapshot of the state of the machine's file system and registry and compare it with the snapshot that was taken before to find what changes have occurred on the system. Regshot is a free and very useful tool that can be used to take registry snapshots and compare between them, you can visit the following link for more information an Regshot download link: (https://sourceforge.net/projects/regshot/).

Content Template	
Section Number	6.7
Section Title	Reverse Engineering Tools for Malware Analysis
Introduction	This section contains a brief summary of some reverse engineering tools that can be used to analyze malware such as disassemblers, decompilers, virtual machines, and sandboxes.
Content	<p>As introduced in the previous section, malware dynamic analysis techniques require studying the structure of the malware sample without executing it on a machine. To do so, the analyst can disassemble or decompile the malware and go through malware source code to do the analysis. Malware dynamic analysis techniques require running the malware first, then observing and monitoring its behavior by recording the changes and activities that occurred on the operating system. Therefore, disassemblers, decompilers, debuggers, sandboxes, and network traffic analysis tools can be used to analyze software to detect malware. IDA in an interactive disassembler that can be used to generate assembly language source code from the machine executable code. By using the HEXRays decompiler add-on, IDA enables us to convert assembly language into pseudo code, which allows us to understand the functionality of the code more quickly than reading the assembly code. There are three versions of this tool: demo version which is limited edition for evaluation, IDA free version for non-commercial use, and the commercial version.</p> <p>Debuggers like Immunity Debugger, GNU Project Debugger, WinDbg, Wind River and PE32 are great tools that can be used in malware static analysis, as they allow the user to view and change the running state of a program that is debugged.</p> <p>In addition to debuggers, disassemblers and decompilers, other tools are required to monitor the activities of the system and of the network during malware analysis, such as Regshot(which we already described in the previous section), Wireshark [https://www.wireshark.org/download.html] and TCPDumps [https://www.tcpcdump.org/].</p> <p>Sandboxes and Virtual Machines are very useful tools that can be used in malware analysis tasks; they enable us to run malware in isolated environments. In addition to isolating malware while running, some sandbox software has additional features and tools that enable us to perform some types of analysis by just clicking on an option and get very rich readable report. For example, Malwr has static analysis, behavioral analysis and network analysis while running malware or any application. Cuckoo, Hybrid Analysis, Virus Total are other examples of sandboxes that can be used as reverse engineering tools for malware analysis. The sandboxes listed in this section are available for free to use.</p>

Activity Template	
Number	6.1
Title	Malware Analysis Using Reverse Engineering Techniques
Type	Practice and training of using some reverse engineering tools.
Aim	After completing the activity, the student will be able to use decompiler to learn the structure and the capabilities of malware.
Description	In this activity, the student is required to select one of the available decompilers and use it to regenerate the source code of a malware. Then, the student should write a short report describing the structure and the capabilities of the selected malware sample.
Timeline	<ul style="list-style-type: none"> • Required time: 5 – 7 hours. • The student has to download and install one of the available decompilers. • The student has to find and download a malware sample. Several malware repositories are available to provide malware samples to researchers such as https://zeltser.com/malware-sample-sources/ • The student has to generate the source code of the selected malware. • Write a short report describing the structure and the capabilities of the selected malware sample.
Assessment	<p>This activity will be assessed based on:</p> <ul style="list-style-type: none"> • The ability to use the tools. • The ability to find the critical behaviors of the malware. • The completeness. • The correctness. • The overall quality. • The followed process.

Activity Template	
Number	6.2
Title	Protocol Reverse Engineering
Type	Practice and training of using some reverse engineering tools.
Aim	After completing the activity, the student will be able to use some reverse engineering tools to analyze internet/network protocols.
Description	In this activity, the student is required to select one of the protocol analysis methods and tools and use it to analyze and extract the specification of one of the internet/network protocols.
Timeline	<ul style="list-style-type: none"> • Required time: 4 – 6 hours. • The student has to select one of the methods of analyzing protocols that has been discussed in the third section of this chapter. • The student has to select one protocol and try to find its specifications using the method selected in the previous step.
Assessment	<p>This activity will be assessed based on:</p> <ul style="list-style-type: none"> • The ability to apply the protocol analysis methods and tools. • The ability to find the protocol specifications. • The completeness. • The correctness. • The overall quality. • The followed process.

Activity Template	
Number	6.3
Title	Reverse Engineering Tools for Malware Analysis
Type	Practice and training of using some reverse engineering tools.
Aim	After completing the activity, the student will be able to use some reverse engineering tools to analyze software and decide whether it is malicious software or not.
Description	In this activity, the student is required to find some reverse engineering tools and use them to detect malware activities and analyze its behaviors.
Timeline	<ul style="list-style-type: none"> • Required time: 4 – 6 hours. • The student has to download and install one of the tools that have been discussed in the last section of the chapter. For example, the student can install Malwr, which is a sandbox with some analysis features. • The student has to find and download a malware sample. Several malware repositories are available to provide malware samples to researchers such as https://zeltser.com/malware-sample-sources/ • The student has to run the malware on the sandbox and generate a report to describe the malware behaviors.
Assessment	<p>This activity will be assessed based on:</p> <ul style="list-style-type: none"> • The ability to use the tools. • The ability to find the critical behaviors of the malware. • The completeness. • The correctness. • The overall quality. • The followed process.

Think Template (MCQs)	
Number	6.1
Title	Introduction to Reverse Engineering
Type	Choose the correct answer
Question	What is the purpose of reverse engineering?
Answers	<p>E. To take something apart to better understand it.</p> <p>F. To put something together to make it work faster.</p> <p>G. To develop new ways of building software programs.</p> <p>H. A+B.</p>

Think Template (MCQs)	
Number	6.2
Title	Malware Analysis Using Reverse Engineering Techniques
Type	True / False
Question	Dynamic analysis of malicious software can be accomplished by examining the code of malicious software without running it.
Answers	C. True. D. False.

Think Template (MCQs)	
Number	6.3
Title	Malware Analysis Using Reverse Engineering Techniques
Type	True / False
Question	Static analysis of malicious software can be accomplished by examining the code of malicious software without running it.
Answers	C. True. D. False.

Think Template (MCQs)	
Number	6.4
Title	Reverse Engineering Tools for Malware Analysis
Type	Choose the correct answer
Question	Registry analysis software is used to:
Answers	<p>E. Retrieve activity log.</p> <p>F. Recover lost files.</p> <p>G. Detect if and how a system has been hacked.</p> <p>H. B + C.</p>

Extra Template	
Number	6.1
Title	Reverse Engineering and Vulnerability Analysis in Cyber Security
Topic	Vulnerability Analysis Using Reverse Engineering
Type	Journal Article: Kumar, M., Alka, A., 2017. Reverse Engineering and Vulnerability Analysis in Cyber Security. Int. J. Adv. Res. Comput. Sci. 8.

Extra Template	
Number	6.2
Title	Data Communications and Networking
Topic	Protocol Reverse Engineering
Type	Book: Forouzan, B.A., Fegan, S.C., 2007. Data Communications and Networking. Huga Media.

Extra Template	
Number	6.3
Title	A Survey of Automatic Protocol Reverse Engineering Approaches, Methods, and Tools on the Inputs and Outputs View
Topic	Protocol Reverse Engineering
Type	Journal Article: Sija, B.D., Goo, Y.-H., Shim, K.-S., Hasanova, H., Kim, M.-S., 2018. A Survey of Automatic Protocol Reverse Engineering Approaches, Methods, and Tools on the Inputs and Outputs View. Secur. Commun. Netw. 2018.

Extra Template	
Number	6.4
Title	Discoverer: Automatic Protocol Reverse Engineering from Network Traces.
Topic	Protocol Reverse Engineering
Type	Conference Paper: Cui, W., Kannan, J., Wang, H.J., 2007. Discoverer: Automatic Protocol Reverse Engineering from Network Traces., in: USENIX Security Symposium. pp. 1–14.

Extra Template	
Number	6.5
Title	Tupni: Automatic reverse engineering of input formats
Topic	Protocol Reverse Engineering
Type	Conference Paper: Cui, W., Peinado, M., Chen, K., Wang, H.J., Irun-Briz, L., 2008. Tupni: Automatic reverse engineering of input formats, in: Proceedings of the 15th ACM Conference on Computer and Communications Security. ACM, pp. 391–402.

Extra Template	
Number	6.6
Title	Polyglot: Automatic extraction of protocol message format using dynamic binary analysis
Topic	Protocol Reverse Engineering
Type	Conference Paper: Caballero, J., Yin, H., Liang, Z., Song, D., 2007. Polyglot: Automatic extraction of protocol message format using dynamic binary analysis, in: Proceedings of the 14th ACM Conference on Computer and Communications Security. ACM, pp. 317–329.

Extra Template	
Number	6.7
Title	Automatic Protocol Format Reverse Engineering through Context-Aware Monitored Execution.
Topic	Protocol Reverse Engineering
Type	Conference Paper: Lin, Z., Jiang, X., Xu, D., Zhang, X., 2008. Automatic Protocol Format Reverse Engineering through Context-Aware Monitored Execution., in: NDSS. pp. 1–15.

Extra Template	
Number	6.8
Title	Prospex: Protocol specification extraction
Topic	Protocol Reverse Engineering
Type	Conference Paper: Comparetti, P.M., Wondracek, G., Kruegel, C., Kirda, E., 2009. Prospex: Protocol specification extraction, in: Security and Privacy, 2009 30th IEEE Symposium On. IEEE, pp. 110–125.

Extra Template	
Number	6.9
Title	Exploiting intra-packet dependency for fine-grained protocol format inference
Topic	Protocol Reverse Engineering
Type	Conference Paper: Huang, Q., Lee, P.P., Zhang, Z., 2015. Exploiting intra-packet dependency for fine-grained protocol format inference, in: 2015 IFIP Networking Conference (IFIP Networking). IEEE, pp. 1–9.

Extra Template	
Number	6.10
Title	Dissecting customized protocols: automatic analysis for customized protocols based on IEEE 802.15. 4
Topic	Protocol Reverse Engineering
Type	Conference Paper: Choi, K., Son, Y., Noh, J., Shin, H., Choi, J., and Kim, Y., 2016. Dissecting customized protocols: automatic analysis for customized protocols based on IEEE 802.15. 4. In: Proceedings of the 9th ACM Conference on Security & Privacy in Wireless and Mobile Networks. ACM, 183–193.

Extra Template	
Number	6.11
Title	A reverse engineering tool for extracting protocols of networked applications
Topic	Protocol Reverse Engineering
Type	Conference Paper: Shevertalov, M. and Mancoridis, S., 2007. A reverse engineering tool for extracting protocols of networked applications. In: Reverse Engineering, 2007. WCRE 2007. 14th Working Conference on. IEEE, 229–238.

Extra Template	
Number	6.12
Title	Reverse engineering of protocols from network traces
Topic	Protocol Reverse Engineering
Type	Conference Paper: Antunes, J., Neves, N., and Verissimo, P., 2011. Reverse engineering of protocols from network traces. In: Reverse Engineering (WCRE), 2011 18th Working Conference on. IEEE, 169–178.

Extra Template	
Number	6.13
Title	A Survey of Automatic Protocol Reverse Engineering Approaches, Methods, and Tools on the Inputs and Outputs View
Topic	Malware Analysis Using Reverse Engineering Techniques
Type	Journal Article: Sija, B.D., Goo, Y.-H., Shim, K.-S., Hasanova, H., Kim, M.-S., 2018. A Survey of Automatic Protocol Reverse Engineering Approaches, Methods, and Tools on the Inputs and Outputs View. Secur. Commun. Netw. 2018.

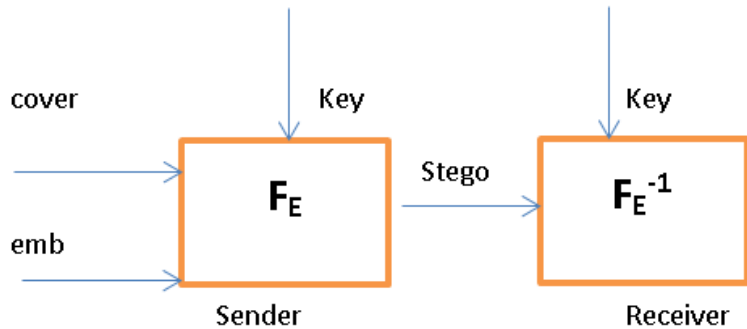
Extra Template	
Number	6.14
Title	A survey of malware detection techniques
Topic	Malware Analysis Using Reverse Engineering Techniques
Type	Journal Article: Idika, N., Mathur, A.P., 2007. A survey of malware detection techniques. Purdue Univ. 48.

7. Steganography

Scope Template															
Number	7														
Title	Steganography														
Introduction	This chapter introduces Steganography and stego-system. It gives a brief overview of steganography history, steganography categories, classifications and steganography in information system. This chapter also shows the usage of steganography in different fields; such as medical industries, music, movies and in Terrorism. Additionally, This chapter presents some tools used in steganography.														
Outcomes	<p>After reading this chapter, you will learn:</p> <ul style="list-style-type: none"> • what is steganography • the importance of steganography in information systems • the categories and classifications of steganography • techniques for detecting steganography • Popular tools that may be used for steganography. 														
Topics	<ul style="list-style-type: none"> • Introduction • History of steganography • Stegosystem Model • Usage of steganography • Classification of Steganography • Steganography versus Cryptography • Watermarking • Issues should be considered in steganography • Detecting Steganography • Stego-Forensics • Technical Tools 														
Study Guide	<table border="1"> <thead> <tr> <th>Task</th><th>Time</th></tr> </thead> <tbody> <tr> <td>Preparation</td><td>3 hr</td></tr> <tr> <td>Textbook Content</td><td>2 hr</td></tr> <tr> <td>Thinking</td><td>2 hr</td></tr> <tr> <td>Tutorial Work:</td><td>3 hr</td></tr> <tr> <td>Self-Reading extra resources</td><td>1 hr</td></tr> <tr> <td>Total</td><td>11 hours</td></tr> </tbody> </table> <ul style="list-style-type: none"> • Required study time: 11 hours • Required hardware/software: <ol style="list-style-type: none"> 1. Preferred Virtual machine that runs windows 7. 	Task	Time	Preparation	3 hr	Textbook Content	2 hr	Thinking	2 hr	Tutorial Work:	3 hr	Self-Reading extra resources	1 hr	Total	11 hours
Task	Time														
Preparation	3 hr														
Textbook Content	2 hr														
Thinking	2 hr														
Tutorial Work:	3 hr														
Self-Reading extra resources	1 hr														
Total	11 hours														

Content Template	
Section Number	7.1
Section Title	Introduction
Introduction	This section will introduce Steganography. We will give a brief overview of steganography in information system. Also this section will define cryptography and its usages.
Content	<p>Steganography is the science that is used to hide information. In computer domain steganography gives users ways to hide information within mediums such as files. Most current digital Steganography techniques and tools hide and embed data inside media such as system file, image, video, and audio files.</p> <p>In digital security there is another popular term called cryptography that is used to protect data, but it is different from Steganography. Steganography is used to hide information and make it unseen, while cryptography used to encrypt data and make it unreadable but it can be seen. Steganography and Cryptography can co-exist, and thus a file which contains stego-information can be also encrypted.</p>

Content Template	
Section Number	7.2
Section Title	History of Steganography
Introduction	In this section we will discuss a historical example of Steganography.
Content	<p>There are a lot of historical examples that provides situations where information should be hiding to traverse enemy territory until it reaches its destination undetected. For example In Ancient Greece, If they want to deliver a hidden message they choose a messenger to carry out the message, by shaving his head and write the text message in his bold head using tattoo, then wait until his hairs grow again, after that he will delivers the message to its destination. This method has many drawbacks such as the waiting time till the hairs grow again, and if they make a mistake in the message it cannot be erased since tattoo is not erasable.</p>

Content Template	
Section Number	7.3
Section Title	Stegosystem Model
Introduction	This section will discuss the stegosystem. The objectives for this section are to discuss the components of this system and to show the mechanism of how the stegosystem works.
Content	<p>A stegosystem is the way that is used to perform steganography. Each stegosystem consists from the following components:</p> <ul style="list-style-type: none"> • Embedded message: The secret message that we want to hide in a medium. • Cover medium: Is the medium, such as image that is used to hide the embedded message. • Stego-key: The secret key that is used in message encryption and decryption. • Stego-medium: consists of the embedded message and the cover medium. <p>(Figure 7.1) explains how steganography system works.</p>  <pre> graph LR cover --> FE[FE] emb --> FE Key --> FE FE -- Stego --> FE_inv[FE^-1] Key --> FE_inv FE_inv --> emb_out[emb] subgraph Sender FE end subgraph Receiver FE_inv end </pre> <p>f_E: steganographic function - embedding</p> <p>f_E^{-1}: steganographic function - extracting</p> <p>emb: message to be covered</p> <p>cover: coverdata for emb message</p> <p>key: parameter for f_E</p> <p>stego: coverdata + embedded message</p> <p>Figure 7.1: Steganography Mechanism. (Image redrawn from Computer Forensics Investigating Data & Image Files book, chapter 1, section 1:2 figure 1).</p>

Content Template	
Section Number	7.4
Section Title	Usage of steganography
Introduction	We will show in this section the usage of steganography in different fields; such as medical industries, music, movies and in Terrorism.
Content	<p>Steganography applications are not necessarily bad or illegal. Let's discuss the following scenarios:-</p> <ul style="list-style-type: none"> • Medical records: using Steganography in medical records can be helpful by avoiding mixing up these records. Since every patient has an EPR (Electronic Patient Record), which has a medical records inside it and other medical information's. • Digital music: Steganography is also used to protect music from being copied by introducing subtle changes into a music file that act as a digital signature. • The movie industry: Many companies use Steganography in DVD and VCDs for copyright protection. • Terrorism: Dark websites is known to use Steganography to hide messages inside images for secret communication between terrorists all over the world.

Content Template	
Section Number	7.5
Section Title	Classification of Steganography
Introduction	In this section we will show the different categories of steganography. Additionally, we will show the various techniques used to hide information. Also, this section discusses some examples for how to use these techniques practically.
Content	<p>Three major categories for Steganography which are:</p> <ul style="list-style-type: none"> • Technical steganography • Linguistic steganography • Digital steganography <p>7.5.1. Technical Steganography In technical steganography many techniques may be used, in this sub section we will discuss the following techniques:</p> <ul style="list-style-type: none"> • Invisible inks: includes Messages that are written and cannot be seen without performing specific action. For example, if we use the mix of onion juice and milk to write a message, the resulted message cannot be seen without heating. • Microdots: This technique applies image size shrinking within a page. <p>7.5.2. Linguistic Steganography Many ways Linguistic steganography used to hide messages. In this sub section we will discuss two main methods, which includes semagrams and open codes.</p> <p>A. Semagrams Semagrams hides information through the use of signs or symbols embedded in the message. Two Semagrams types exist which are Visual and Textual:</p> <ul style="list-style-type: none"> • Visual Semagrams: In this method, paint (image) may be used to hide information. For Instance, the location of an image on a web page used to give a special message. • Text Semagrams: In this technique, a message is hidden by changing in the followed style in a paragraph or a page. For example changing the font size for a word in a letter. <p>B. Open Codes Open codes make use of openly readable text. This text contains words or sentences that can be hidden in a reversed or vertical order, (figure 7.2) is an example of open codes hidden message. Open codes have two types which are:-</p> <ul style="list-style-type: none"> • Jargon codes: in this type, a group of people define a specific language, and only this group can understand the exact meaning for words in this language. For instance, if a group of people define a word test as a home 123 at address xyz, whenever this word appears in a shared medium (such as a website), only this group will match the exact meaning of this word. • Covered ciphers: In this type a message is embedded in a carrier medium that is visible to everyone, but it can be read by specific peoples who understand the carrier medium. For example using hidden field in HTML form to hide a message.

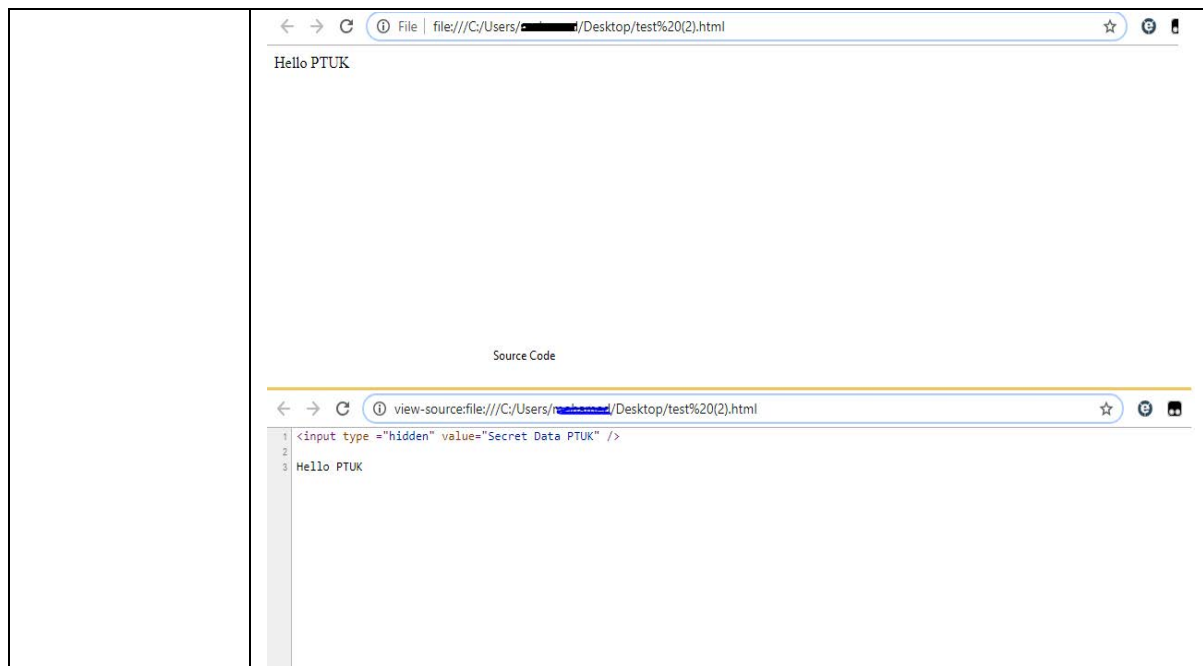


Figure 7.2: The source file can reveal an injected message.(This image was taken during an experiment in PTUK university)

7.5.3. Digital Steganography

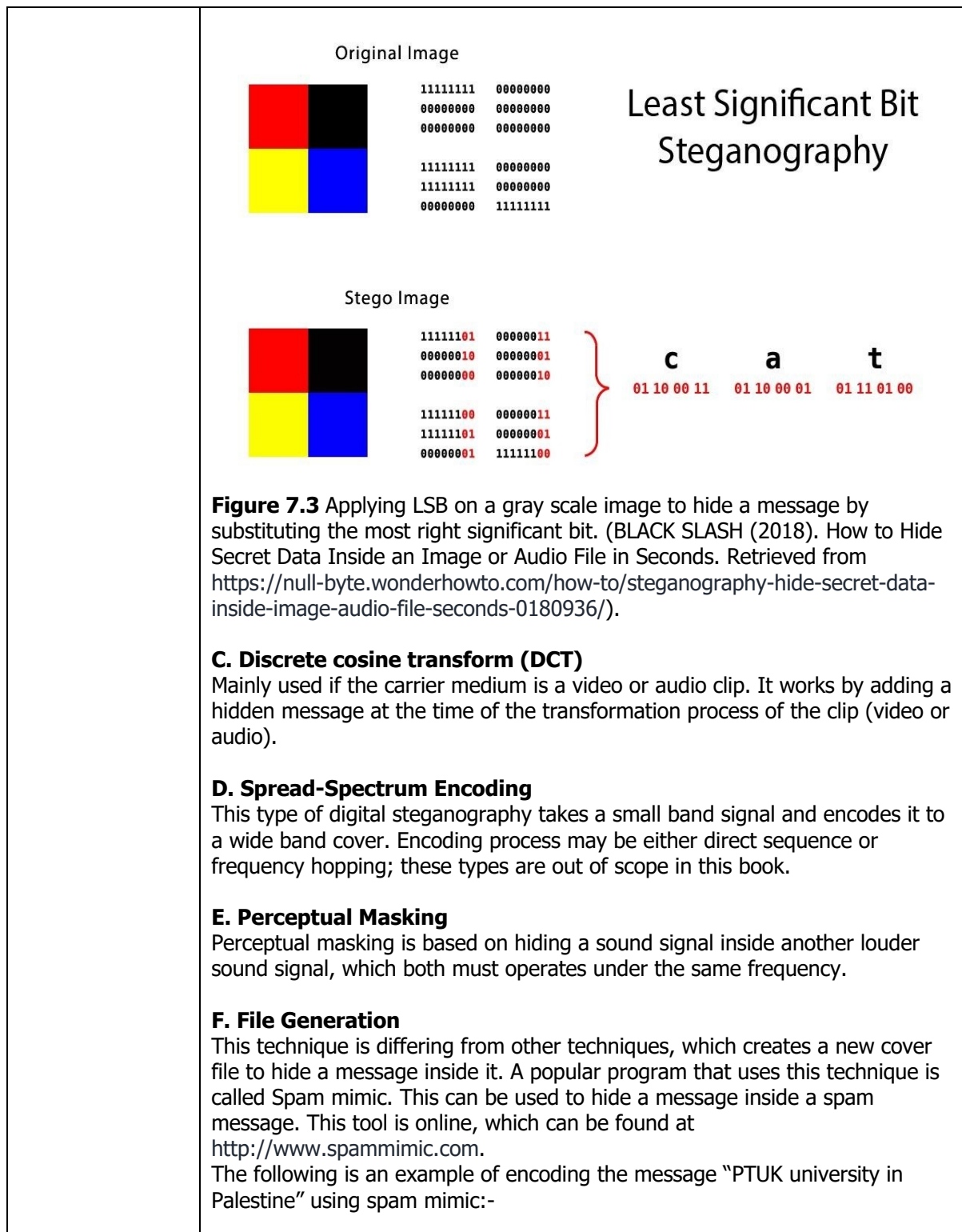
Many digital steganography techniques can be used to hide a message, in this sub section we will discuss the following techniques:-

A. Injection

In this type of digital steganography technique, the host file holds the secret information, which may be a picture, sound file, or even a video clip. The disadvantage for this method that it will increase the size of a host file, thus making it easy to be detected. But we can avoid this problem by deleting the host file once the message is created.

B. Least Significant Bit (LSB)

LSB technique is based on the principle that the rightmost bit on a media is the least impact on the binary data. In other words, the rightmost bit in a media will be replaced with a bit from the embedded message without affecting the media itself. (Figure 7.3) shows an example of applying LSB techniques on a gray scale image to hide a message. So, you should consider that more bits modification meaning more information to hide, hence more distortions to the carrier medium, making it more detectable.



Your message **PTUK university in Palestine** gets encoded into spam as:

Dear Friend ; You made the right decision when you signed up for our mailing list . If you no longer wish to receive our publications simply reply with a Subject: of "REMOVE" and you will immediately be removed from our directory ! This mail is being sent in compliance with Senate bill 2616 , Title 7 , Section 301 . This is not multi-level marketing . Why work for somebody else when you can become rich within 42 days . Have you ever noticed people love convenience and people love convenience . Well, now is your chance to capitalize on this . WE will help YOU SELL MORE and turn your business into an E-BUSINESS ! You are guaranteed to succeed because we take all the risk ! But don't believe us . Ms Jones of Kentucky tried us and says "I've been poor and I've been rich - rich is better" . We are a BBB member in good standing ! We IMPLORE you - act now ! Sign up a friend and you'll get a discount of 30% ! Thanks ! Dear Friend ; Your email address has been submitted to us indicating your interest in our newsletter ! We will comply with all removal requests

Figure 7.4: Using Spam mimic to hide a message. .(This image was taken during an experiment in PTUK university)

G. Digital File Types

In this subsection we will introduce techniques which best fit each file type. Four digital file types, which are images, audios, videos, and system files.

1. Image files

Three commonly image file formats used in steganography, which are Graphics Interchange Format (GIF), Joint Photographic Experts Group (JPEG), and Tagged Image File Format (TIFF). In GIF format the image file is compressed. JPEG format is used for images that are small in size. TIFF image format used as standard image format for file exchange.

The Least-significant-bit (LSB) insertion and Masking and filtering are techniques which may be used to hide a message inside an image file.

2. Audio

Human ear cannot mask quiet sound from louder sound, especially if the two sounds use the same frequency. So technique such as Spread Spectrum will be best to hide a message inside audio file. Additionally LSB technique may be used in this type of files.

3. Video Files

Best technique to hide a message inside a video file will be the DCT technique.

	4. Steganographic File System Hiding information inside a file system is very common; for example, you can embed and hide huge amount of information in an NTFS windows file system metadata without making the file suspicious and detectable. Techniques such as file generation may be used in these types of files.
--	---

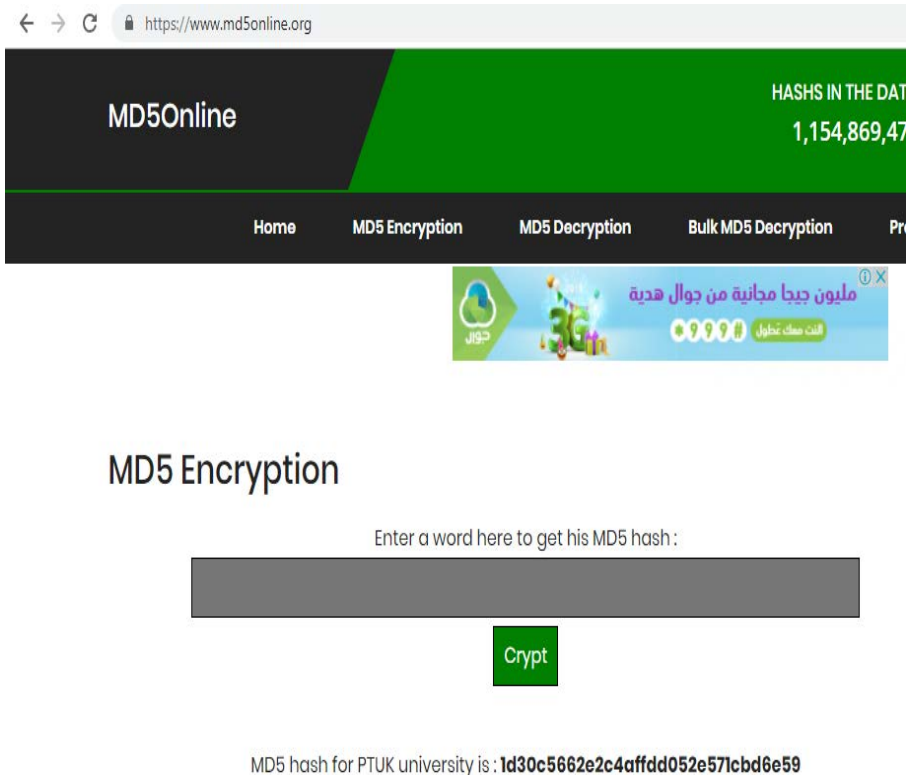
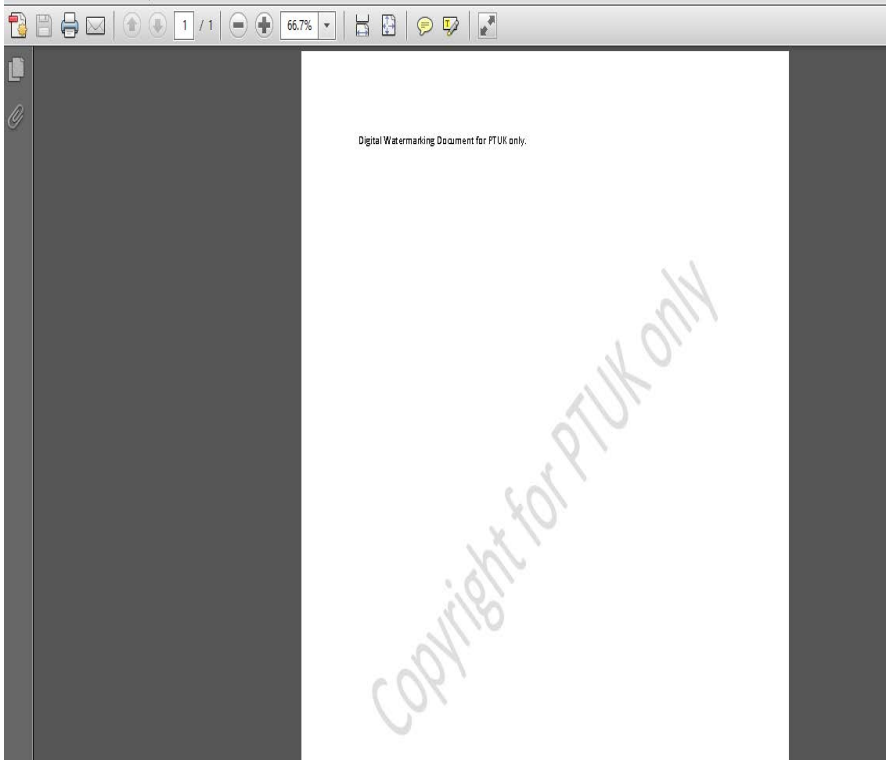
Content Template	
Section Number	7.6
Section Title	Steganography versus Cryptography
Introduction	This section explains differences between steganography and cryptography.
Content	<p>As mentioned earlier in this chapter, Steganography is used to hide information within a medium; this medium may be a system file or image or audio or video file. On the other hand, Cryptography is basically applying an encryption algorithm to specific data, so that it can't be read by a person who has no permission to read it. Encryption algorithm may be one way or two ways, it depends on the data that you want to safe. An example of one way encryption algorithm is mds algorithm. (Figure 7.5) illustrates encrypting the data "PTUK university" using md5 algorithm.</p>  <p>MD5 hash for PTUK university is: 1d30c5662e2c4affdd052e571cbd6e59</p>

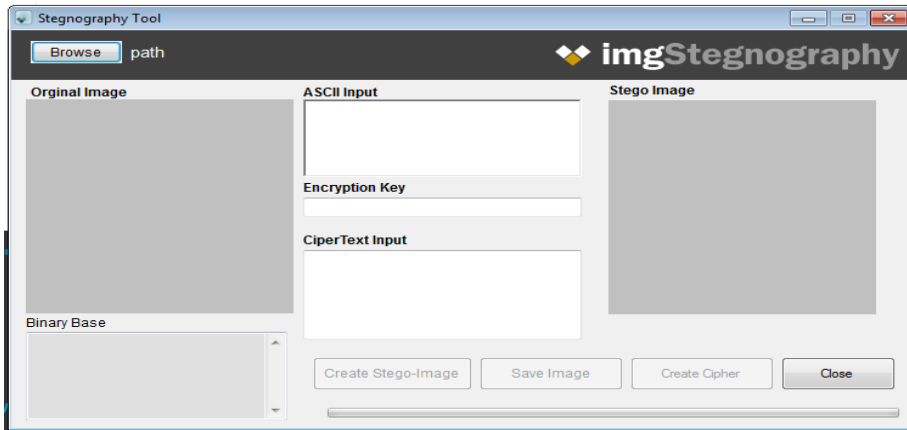
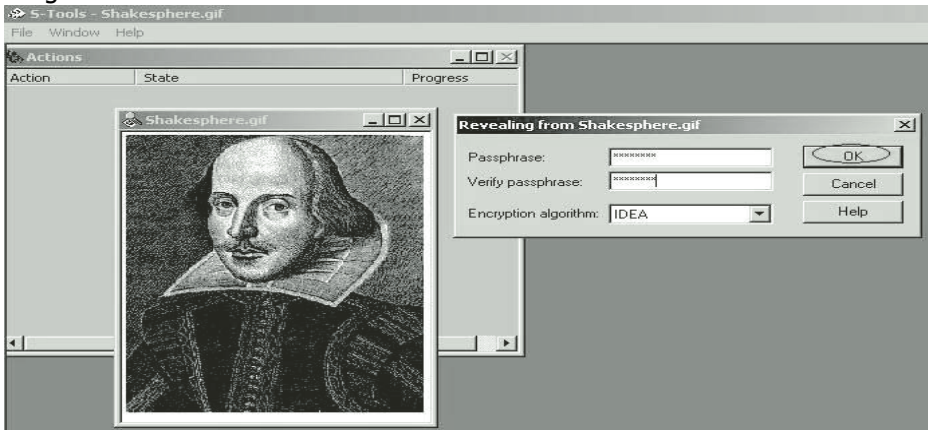
Figure 7.5: Encrypting "PTUK university" data using md5 algorithm. .(This image was taken during an experiment in PTUK university)

Content Template	
Section Number	7.7
Section Title	Watermarking
Introduction	This section describes watermarking. It shows where this technology can be used. Additionally, this section shows how watermarking is different from steganography.
Content	<p>Digital watermarking is used to prohibit copying or modifying the copyright data without permissions. On other words, watermarking is concerned for copyright. Also it is not necessarily that the watermark should be hidden as steganography. For example, in the following PDF document (Figure 7.6), a watermark was created in the middle of each page to ensure the copyright permission in this document is for PTUK university.</p>  <p>Figure 7.6: Watermark for a PDF document. .(This image was taken during an experiment in PTUK university)</p>

Content Template	
Section Number	7.8
Section Title	Issues which should be considered in steganography
Introduction	There are three issues that should be considered while hiding information. This section describes these issues in detail.
Content	There are three issues that should be considered while hiding information, which are 1) the level of visibility, 2) robustness versus payload, and 3) the file format dependence. Level of visibility, means how tangible is the way of insertion of the hidden message in the carrier medium. For example if we use an image as a medium and we use LSB technique to hide a message, and we notice a big and noticeable change in the carrier image. Robustness versus payload, meaning we should choose a technique that makes unnoticeable or least changes to the carrier medium. File format dependence, meaning the compression of a file that has lossless information to a file with lossy information can erase the hidden message that exists in the cover of the carrier medium. Such as compression GIF image file, this is known extension as a lossless compressions extension.

Content Template	
Section Number	7.9
Section Title	Detecting Steganography
Introduction	There are several tools and techniques for detecting steganography. This section will shows some of these tools and techniques used to reveal information that is hidden within a specific media.
Content	<p>Many techniques may be used to detect steganography, note that these techniques may be automated using different tools as we will see later in this chapter. In this section we will discuss the following techniques:-</p> <ul style="list-style-type: none"> • Statistical test: This technique works by examining the statistical information of the original image (any media can be used) and compare it to the suspicious image. • Stegdetect: is a tool that is used to reveals hidden information from an image. • Steganalysis: This is a reverse engineering process for steganography. It works by examining differences in bit patterns and unusual large size of a file. <p>There are many other techniques and methods, such as Appended spaces and invisible characters, Steganalysis Methods/Attacks on Steganography, ...etc, which is out of scope in this book.</p>

Content Template	
Section Number	7.10
Section Title	Stego-Forensics
Introduction	This section will define Stego-Forensics term to the student.
Content	Stego-forensics is a part of forensic science that applies steganography methods and tools to find the cause of a crime. For instance, discover the communication channels between terrorists around the world.

Content Template	
Section Number	7.11
Section Title	Technical Tools
Introduction	In this section we will discuss tools used in steganography, and tools that a digital forensic investigator may use to discover steganography in different media.
Content	<p>7.11.1 Steganography Tool</p> <p>Steganography is security software used to secure folders, encrypt files, and encrypt, split, and transmit data. It contains other useful security tools for the user (Figure 7.7). This tool is helpful in steganography and stego analysis.</p>  <p>Figure 7.7: Steganography can be used to encrypt messages into image files.</p> <p>7.11.2 S-Tool</p> <p>S-Tool is a GUI tool that can be used to hide multiple files within a single carrier medium. (Figure 7.8) shows how to hide multiple files inside a JPEG image file.</p>  <p>Figure 7.8: Hiding multiple files inside a JPEG image file.</p> <p>7.11.3 Masker</p> <p>Masker is commercial software that provides It encrypts files and hides messages behind image. (Figure 7.9)</p>

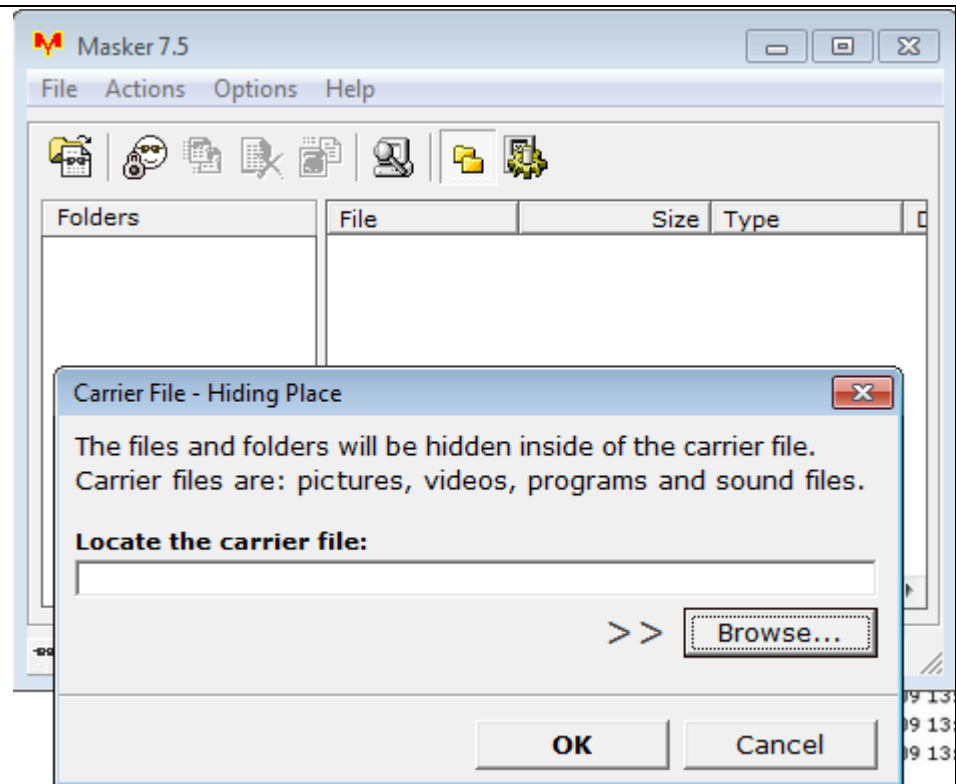


Figure 7.9: Masker software.

Content Template	
Section Number	7.12
Section Title	Chapter Summary
Introduction	
Content	<ul style="list-style-type: none"> • Steganography is the science of hidden information. • Cryptography and Watermarking are totally different from steganography. • Many techniques and methods are used in steganography, such as Injection, LSB, Spread-Spectrum Encoding, Spread-Spectrum Encoding etc. • There are three issues that should be considered while hiding information, which are the Level of visibility, robustness versus payload, and file format dependence. • There are many tools that can be used for steganography, such as Masker. Additionally, there are many tools that might be used to discover steganography in a file such as Steganography tool.

Activity Template	
Number	7.1
Title	Case study
Type	Research
Aim	The aim of this activity is to put the student in a real forensic problem, to measure his/her ability for solving real life scenario.
Description	A disgruntled employee using his/her work email to send a picture of the children to a friend but actually they are shipping out your most commercially sensitive information?
Timeline	1 week
Assessment	The document will be assessed based on Logic, correctness and overall quality.

Activity Template	
Number	7.2
Title	Practical Exercise
Type	Practical
Aim	The aim of this activity is to allow the students to practice on using steganographic tools.
Description	Use one of the mentioned steganography tool to hide a message inside a media, After that use one the mentioned stego analysis tool to discover the hidden message that you created earlier.
Timeline	2 Days
Assessment	Each student is required to submit a working tool; additionally, the student must submit a document of features for the selected tools. The report will be assessed based on completeness, correctness and overall quality.

Think Template (MCQs)	
Number	7.1
Title	Introduction
Type	Fill in the blanks
Question	_____ is the art and science of hiding information in plain sight. By ensuring that data is hidden from casual observers
Answers	Steganography

Think Template (MCQs)	
Number	7.2
Title	Steganography versus Watermarking
Type	Fill in the blanks
Question	_____ is used to protect data from distortion by others
Answers	watermarking

Think Template (MCQs)	
Number	7.3
Title	Detecting Text, Image, Audio, and Video Steganography
Type	Fill in the blanks
Question	_____ and _____, are used to discover Steganography in an audio files
Answers	<ul style="list-style-type: none"> • Scanning information for inaudible frequencies • Determining odd distortions

Think Template (MCQs)	
Number	7.4
Title	Steganographic File System
Type	Fill in the blanks
Question	There are _____ possible methods used to embed messages in a text file
Answers	two

Think Template (MCQs)	
Number	7.5
Title	Steganography versus Cryptography
Type	Fill in the blanks
Question	_____ requires the same key or two different keys for encryption and decryption.
Answers	Encrypted message

Extra Template	
Number	8.1
Title	Steganography
Topic	1
Type	URL: https://www.webopedia.com/TERM/S/steganography.html

Extra Template	
Number	8.2
Title	Computer Forensics Investigating Data & Image Files
Topic	8.2,8.3,8.4,8.8,8.9
Type	Book, Chapter 1: ISBN- 13: 978-1-4354-8351-4 [Book title and reference?]

Extra Template	
Number	8.3
Title	Steganography and Its Applications in Security
Topic	8.5 [is the numbering of topics correct?]
Type	URL: http://www.ijmer.com/papers/Vol2_Issue6/EN2646344638.pdf

Extra Template	
Number	8.4
Title	Hide Secret Data Inside an Image or Audio File in Seconds
Topic	8.5
Type	URL: https://null-byte.wonderhowto.com/how-to/steganography-hide-secret-data-inside-image-audio-file-seconds-0180936/

Extra Template	
Number	8.5
Title	MD5Online
Topic	8.6
Type	URL: https://www.md5online.org/

Extra Template	
Number	8.6
Title	Digital Watermarking
Topic	8.7
Type	URL: https://en.wikipedia.org/wiki/Digital_watermarking

Extra Template	
Number	8.7
Title	Steganography Tool
Topic	8.11
Type	URL: https://sourceforge.net/projects/stegtool/reviews

Extra Template	
Number	8.9
Title	Steganography Tool
Topic	8.11
Type	URL: http://www.softpuls.com/masker/