PRIVACY AND SECURITY CONFERENCE 2019

PRIVACYANDSECURITYCONFERENCE.PT

Proceedings of the Digital Privacy and Security Conference 2019

16 January 2019 *Porto, Portugal*

Editors
Carla Cordeiro and Hugo Barbosa

UNIVERSIDADE LUSÓFONA DO PORTO

Copyright

1st Edition 2019

Issue EOI:10.11228/dpsc.01.01

Editors: Carla Cordeiro and Hugo Barbosa

Proceedings Design: Hugo Barbosa

Graphical Design/Website: Hugo Barbosa

E-mail: hugo.barbosa@ulp.pt

Conference Website: http://privacyandsecurityconference.pt

Conference EOI :10.11228/dpsc

DPSC2019 Proceedings EOI: http://eoi.citefactor.org/10.11228/dpsc

Universidade Lusófona do Porto
Rua Augusto Rosa, nº 24
4000-098 Porto – Portugal
Telephone: +351 222 073 230

FOREWORD

ORGANIZING AND SCIENTIFIC COMMITTEES

Digital Privacy and Security Conference 2019 organization and Scientific Committees welcome you to the second conference. The main goal of a scientific event is to discuss, disseminate and create knowledge. Organizing this conference proved to be a challenging opportunity for us to achieve this goal.

Currently, we are living in a digital world, where we share all information consciously and subconsciously about our life with any one. This situation puts the people in a worrying situation. Our commitment and hard work have as aims to contribute for all participants to acquire tools to better protect themselves. This area is in constant evolution and need we improved our knowledge.

The young students that devote themselves to research deserve our praise for their efforts in the search of new knowledge and better intellectual and technical skills. Persistence and strong motivation constitute the driving force which stimulates students of Supplementary Networking course, Informatics Engineering degree, from the Lusofona University of Porto (ULP), to the creation of scientific papers related to this field of study, to the promotion of research, and to the knowledgeable discussion and practical demonstration on a variety of issues addressed, particularly in the context of computer science and computer networks. The grouping of this information, which takes the shape of a book, is the natural result of these principles put into practice.

We would like to thank all those authors whose participation in this endeavor contributed to its success, hoping it will promote a better understanding of the issues that were addressed.

Thanks to all the sponsors who made the conference possible, as well as all those who contributed to the success of DPSC2019.

Porto, January 2019

Carla Cordeiro and Hugo Barbosa

Conference Committees

ORGANIZING COMMITTEE

Carla Moreira Cordeiro – Universidade Lusófona do Porto, Portugal
Lusofona University of Porto, Portugal

Hugo Azevedo Barbosa – Universidade Lusófona do Porto, Portugal
Lusofona University of Porto, Portugal

SCIENTIFIC COMMITTEES

Hugo Azevedo Barbosa – Universidade Lusófona do Porto, Portugal
Lusofona University of Porto, Portugal

Óscar Ferreira Ribeiro – Universidade Lusófona do Porto, Portugal
Lusofona University of Porto, Portugal

José Lobinho Gomes – Universidade Lusófona do Porto, Portugal
Lusofona University of Porto, Portugal

João Paulo Magalhães – Escola Superior de Tecnologia e Gestão | Politécnico do Porto
School of Technology and Management | Polytechnic of Porto, Portugal

Nuno Santos – Instituto Superior Técnico | Universidade de Lisboa, Portugal
Higher Technical Institute | University of Lisbon, Portugal

João Ulisses – Universidade de Vigo, Espanha
University of Vigo, Spain

Kiavash Satvat – University of Illinois at Chicago, United States
University of Illinois at Chicago, United States

Contents

SESSION 1 - Privacy and Security Issues

SESSION 2 - Preventive Measures in the Era of Digital

SESSION 1
PRIVACY AND SECURITY ISSUES

Data Security in Modern Cars

Pedro Martins

Lusofona University of Porto, Portugal
pedro.miguel.martins9@gmail.com

Abstract. As the automotive industry progresses there's been a growing investment in connected cars. With it the need for ensuring secure communications inside of, from and to said cars has also become a growing focus. This paper will talk about the known threats, how they are identified and their impact on systems, ways to avoid them, as well as methods to contain and resolve them in order to raise interest and awareness over these matters, using as means of reference, a case study from the industry.

Keywords: Communication, Vehicle, Automobile, Data Security, Car.

1 Introduction

With advancements in technology, daily use objects have steadily and increasingly become more connected in networks in order to give them more flexibility and to adapt to the users' needs. Vehicles have also followed this trend, and although it brought very useful features, it also brings with it some risks.

Data security and mainly personal data security have been subjects of extreme importance with the development of digital technology. We live in a time where information can be easily shared and extracted, as such, the need to keep a person's personal data private has become the focus with the recent push to develop security measures and methods of prevention.

After the introduction of the Electronic Control Unit (ECU) to the commercial car, the access to vehicular telemetry has become more common place, firstly used to monitor the engine and to control variables such as the injector on time in order to improve efficiency [1], it soon expanded to many more areas of the vehicle and quickly more types of ECUs followed and made the vehicle's information easily accessible.

Nowadays, after the introduction of parking assist and other technologies, information such as the steering angle, braking and displays can now be tampered with [2]. This means that someone with physical access to a vehicle, can use the ECU's functions to gain control of the vehicle.

With the advance of current communications technologies however there's been a growing push to connect the vehicles to a network, sharing information between vehicles to avoid common problems such as vehicle blind spots and information on what's coming ahead.

With some vehicles having the need for software updates using wireless networks, some of which being more important to the vehicle's safety than others, the importance for secure communications becomes evident as well as the risks associated.

This paper will be structured in two main parts, in-vehicle connections and mobile connections. In the first topic the in-vehicle connections will be separated in two kinds, the wired and the wireless, after a brief introduction, some common attacks will be described followed by some mitigation and prevention methods. The second topic will be about mobile connections, describing first how the vehicle uses such connections, followed by the results of an interview in the field.

2 In-Vehicle Connections

Starting with the last line of attack on a connected vehicle, we have ECUs, these devices are capable of processing signals and are usually coupled with sensors, other ECUs or even interfaces connecting to the users or even outside systems [3]. They create a network of data, responsible for tracking the vehicle's components' status and process them in order to maximize efficiency, safety and even to aid and inform the driver and its passengers.

It's clear why attackers might want to get access to the ECUs' data. Gaining access to one or more interfaces would be a gateway to the vehicle's control and information on its users. As such, security in these interfaces is critical for the whole system's security.

To understand how data flows in a network of ECUs, we must first investigate how the various nodes are connected and how the communication is established between them.

Below, some of the more common possible attacks on the in-vehicle network will be described, all of which require either physical access or very close proximity to the target vehicle.

2.1 Attacks over wired networks

In a vehicle there are one or more serial buses and gateways connecting the various nodes together and ensuring they can send and receive information. One such system is called Controller Area Network (CAN) [4]. The CAN is a broadcast protocol as such, all nodes receive a sender's message, but only the receiver will accept them [5].

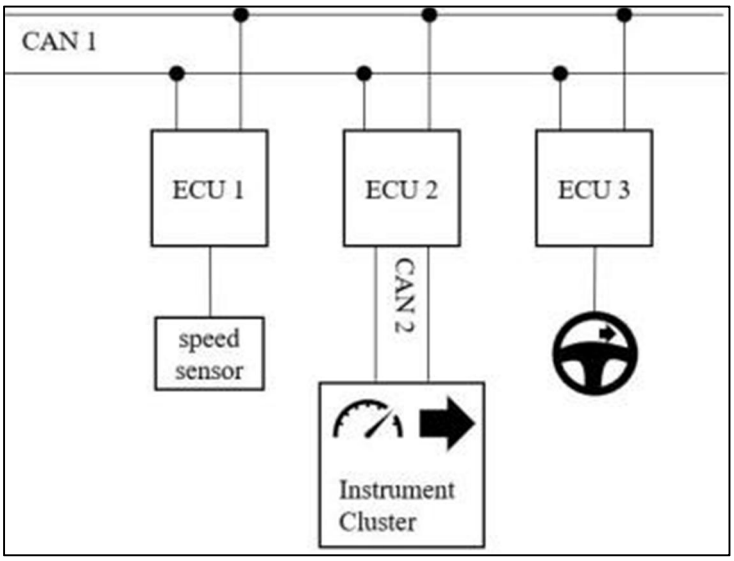

Fig. 1. CAN Architecture example [4]

2.1.1 Denial of service

The first and most basic of attacks across networks is the Denial of Service (DoS). This attack consists of overloading the bus so that the communication between nodes becomes obstructed and no packets can be delivered. The effects and results of the attack however vary between networks and ECUs, in some cases even disabling power steering and limiting the wheel's range of motion. [2]

A DoS can be achieved in several ways, ranging from sending false Request to Send (RTS) packets to repeating packets it sniffs in the CAN bus [6].

The damage of the attack can't however be considered the same for every vehicle, while some vehicles are older and make less use of the processing capabilities of said networks, newer vehicles are filled with sensors and ECUs but are also more developed in terms of data security as the technology got more time to mature.

Some ways to mitigate the damage caused by a DoS can be:

- Checking for excessive packets during a time frame, this can be achieved by programming the ECU to drop requests it has already responded from the same source for a limited time frame [6], since the frequency of normal packets is predictable [2];
- Ensuring minimal safety when one or more ECUs lose control avoiding cases such as the reduced steering range [2].

2.1.2 Packet Sniffing and Data Injection

Data injection is a more complex attack as it involves the attacker sending packets impersonating nodes in the bus and sending potentially malicious information requiring knowledge of the network structure which in most if not all cases implies that the attacker needs to reverse engineer the target vehicle's CANs [2].

To aid with this process the attacker would need to analyze packets in the buses (sniffing), this however is similar across vehicles' CANs and there's already a wide variety of tools for this purpose easily accessible to most. To achieve this, some ECUs allow you to do diagnostic operations requiring an authentication key. While in some cases it might be necessary to extract its firmware and reverse engineer the key to gain access to the mentioned mode, in some cases, depending on the system you can get away with brute force since some use a fixed seed while authenticating. [2]

Although sniffing by itself already poses a privacy concern, when used for the purpose of packet injection it poses a much bigger threat, in some cases even making it possible to gain control of the vehicle's steering by resending packets used in park assist used to control the steering wheel. Some other attacks range from sending misleading information to the vehicle's dashboard which includes the speedometer and door signals causing confusion and misleading the driver. [2]

Suggested methods to avoid these attacks, some of which already used in commercially available vehicles include:

- Random seeds for authentication purposes, lowers the chance of a successful authentication by brute forcing as described previously [2];
- Physical network separation, although most systems use CAN buses, some manufacturers avoid possible attacks on other systems such as the cruise control by connecting them directly to the Powertrain Control Module (PCM) where the accelerator is also connected [2];
- Message obfuscation, it tries to maintain messages known only to authorized parties, although it doesn't stop all attacks it adds an extra layer of difficulty, forcing the attacker to reverse engineer in order to decode proprietary messages [7]

2.2 Attacks over in-vehicle wireless networks

With the increase of sensors in the car, wired connections across sensors directly affect the weight and the wire complexity of these networks. To avoid this, wireless networks are thought to become increasingly used across the vehicle and although vehicle-tovehicle and vehicle-to-infrastructure connections have been getting much attention, the first wireless network widely installed in every new vehicle has been the Tire Pressure Monitoring System (TPMS). [8]

This system consists of one sensor per tire that periodically broadcasts the tire's variables such as temperature and pressure over to a receiver fitted to an ECU which after filtering the received packets processes the sensors' data [8].

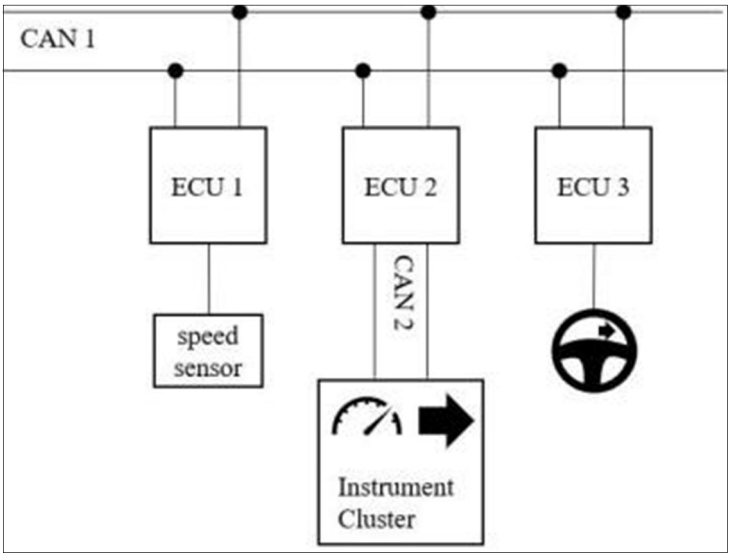

Fig. 2. TPMS Architecture with four antennas [8]

The TPMS's lack of security however doesn't get much attention since the car's metal body works as a shield shortening the maximum connection distance, and because a third--party getting tire pressure seems harmless [8].

Furthermore, having most modern vehicles pushing for handsfree interaction with the vehicle using Bluetooth [9] which is nowadays also common across smartphones and even fitness devices, a whole new vector of attack is possible.

Although some of the previous attacks over wired connections still apply over invehicle wireless networks, attacks unique to these connections will be given more attention.

2.2.1 Indirect attacks (Trojans)

For an attacker it's much harder to physically pair a device to the vehicle, instead, using an app the target will likely have installed in one of his devices paired connected over Bluetooth, the attacker can send harmful information to the vehicle's system over the target's phone (Trojan) [9].

These sorts of attacks where harmful payloads are sent to the vehicle's ECU responsible for the Bluetooth communication could potentially exploit security flaws only available over paired devices [9].

Some possible ways to avoid these attacks start with:

- Increasing safety measures in phone application stores, this will only get you so far however, since there are exploits that can compromise devices through malicious Web sites [9];

- User awareness, education on methods of prevention and the risks of visiting rigged websites [10] or installing harmful software;

2.2.2 Tracking

Even though the TPMS doesn't seem to contain much information on its own since its purpose is to be used to track tire pressure and due to its lacking security measures, the sensor's IDs can be used to identify a vehicle [8].

Coupled with the reported possibility of eavesdropping the system over ranges of over 10 meters using a strategically placed receiver make it a viable way of locating a car. Although some similar systems exist, such as the automatic license plate identification this system and the toll tags, this method is can prove to be more reliable against the usage of fake license plates since the system is hard for drivers to deactivate [8] and doesn't need the vehicle to be equipped with an extra equipment.

This method doesn't make it possible to track a vehicle with much precision due to the high cost of placing receivers along a stretch of road. However, by well thought out placements the receivers could track the frequency of highway entrances and exits' usage [8].

It is however possible to avoid this by:

- Adding cryptographic measures to the wireless communications, making it more difficult to decipher and obtain information such as the sensors' IDs in the TPMS;

- Better insulation of radio frequencies, avoiding the possibility of sniffing on the wireless networks, lowering the maximum distance at which the attack can occur.

3 Mobile Connections

With the European Union making eSIM based system eCall mandatory, modern vehicles are now always connected via mobile networks.

Although by itself the system doesn't pose security risks, many manufacturers are now pushing to give more functionality to the vehicle over mobile networks, enabling intractability with the vehicle wirelessly over mobile networks.

The most common and used mobile communication system today is the Global System for Mobile Communications (GSM), first used mainly for voice communications it is now the most

widely used wireless technology worldwide and although technologies like 4G LTE (Long Term Evolution) have gained popularity, GSM still handles a great amount of voice calls as systems without access to Voice over LTE (VoLTE) usually fall back to it. [11]

The main problem with mobile communications however is the fact that they use an unprotected medium such as air. To mitigate this problem, current technologies such as 2.5G now using cryptographic techniques, and 3G having new authentication systems. [12]

3.1 Attacks over Mobile Networks

Since GSM is the most common, has some similarities over the other alternatives and is the one that all systems will roll back to when needed, it's going to be described in greater detail.

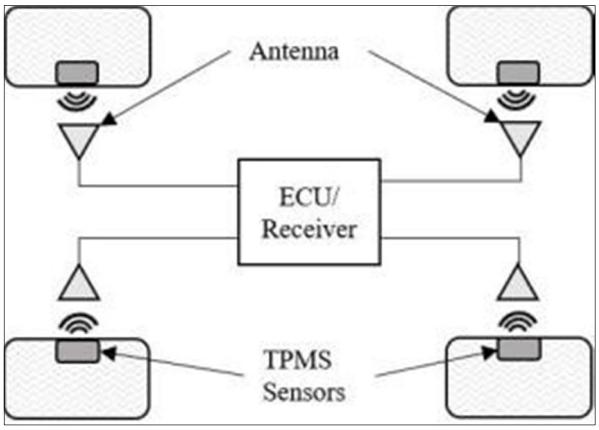

Fig. 3. GSM Network Architecture Simplified [14]

This technology is structured in three major subsystems:

- The Mobile Station (MS), which is all the equipment and/or software required for the subscriber to communicate with the network. It includes the mobile equipment, such as a mobile phone, and a Subscriber Identity Module known more commonly as a SIM card;
- The Networking and Switching Subsystem (NSS), responsible for connecting calls from one end to another, to manage mobility of the subscribers and the communication with the telephone network and the internet;
- The Base Station Subsystem (BSS), that manages the radio transmission between MSs and other subsystems. [13]

3.1.1 Man in the middle

The man in the middle attack as the name suggests, is performed by being an intermediary in the communication of two systems. In the case of GSM this attack can be achieved by the attacker impersonating a Base Station to the Mobile Statin and impersonating the Mobile Station to the Base Station. [14]

Although these kinds of attacks can be used in the systems above, in the case of mobile networks they appear to get more of a focus, since multiple protocols such as GSM and UMTS operate simultaneously.

These attacks over GSM and UMTS connections target encryption keys which would grant the attacker access of the content of the message as it is, this could mean text messages such as SMS could be read, the attacker could also send messages impersonating the target and even manipulate them. [14]

Attacks like this can and have been mitigated using:

- Better authentication mechanisms, as has been proved to be critical in resolving some threats such as the A5/2 attack on GSM [14];
- Use of Certificate Authorities, even if the message is read and deciphered, another cypher will be behind, this as will be described ahead, is a technique already in use.

4 Security Measures in use

As an interview revealed, some manufacturers use mobile networks in conjunction with APIs to provide functions to the user via smartphones and other smart devices. These APIs limit and control the access an outside system has over the vehicle, avoiding the risk of an attacker controlling a function outside of the API.

Although the eSIM could be used for tracking, it can be disabled, leaving only the functionality of emergency calls available.

A system can be further secured from possible man in the middle attacks, by means of certificates where asymmetrical keys are used to keep the communication private and that, besides communications, all stored data is also encrypted in the systems, keeping the users' information private.

Furthermore, it has also been revealed that to avoid indirect attacks such as trojans mentioned above, all workers in the working place take part in mandatory training, to alert them over risks, such as the use of unknown USB devices, and possible threats of human engineering.

5 Conclusion

We see the world and the vehicular industry growing each day with newer developments, bringing with it much more flexibility over how we interact with our daily vehicles. A great amount of data is now used for our comfort in cars, going to the extent of using wireless methods to avoid cable complexity. With the modern car being so connected to the end user, some security risks surface, namely some privacy concerns and health hazards.

The industry and governments however seem to be steadily adapting to this new age of technology where so much information is connected between devices and has so much control over our daily life, with laws like the European Union's General Data Protection Regulation (GDPR) and some businesses investing in cybersecurity awareness of their workers.

It could be seen from this work that threat awareness is also vital, given that one form of attack, the trojan, can be mitigated by properly educating the population over possible risks caused by some common actions we deem harmless, can lead to serious security flaws.

From the results we've seen it's clear that the technology we use is in a constant wrestle between security methods and new attacks, which leads me to believe that a problem with the industry in the future could be the end of life support of the software in these vehicles.

References

1. Honda Motor Co Ltd: Engine control unit, (2003).
2. Chris Valasek, Charlie Miller.: Adventures in Automotive Networks and Control Units, IOActive, (2014).
3. Siemens Aktiengesellschaft: Electronic control unit for a motor vehicle, (1999).
4. Robert Bosch GmbH: Bosch Automotive Electrics and Automotive Electronics. Springer, (2014).
5. Lin, C., Sangiovanni-Vincentelli, A.: Cyber-Security for the Controller Area Network (CAN) Communication Protocol. International Conference on Cyber Security. IEEE (2012).
6. Mukherjee, S., Shirazi, H., Ray, I., Daily, J., Gamble, R.: Practical DoS Attacks on Embedded Networks in Commercial Vehicles. 12th International Conference, ICISS (2016).
7. Zhang, T., Antunes, H., Aggarwal, S.: Defending Connected Vehicles Against Malware: Challenges and a Solution Framework. IEEE Internet of Things Journal. 1, (2014).
8. Rouf, I., Miller, R., Mustafa, H., Taylor, T., Oh, S., Xu, W., Gruteser, M., Trappe, W., Seskar, I.: Security and privacy vulnerabilities of in-car wireless networks: a tire pressure monitoring system case study. 19th USENIX conference on Security (2010).
9. Checkoway, S., McCoy, D., Kantor, B., Anderson, D., Shacham, H., Savage, S., Koscher, K., Czeskis, A., Roesner, F., Kohno, T., Comprehensive Experimental Analyses of Automotive Attack Surfaces. SEC'11 Proceedings of the 20th USENIX conference on Security (2011).
10. Luo, X., Liao, Q.: Awareness Education as the Key to Ransomware Prevention. Information Systems Security. 16, 195-202 (2007).
11. Sauter, M.: From GSM to LTE-advanced pro and 5G. Wiley (2017).
12. Corallo, A., Cremonini, M., Damiani, E., Vimercati, S., Elia, G., Samarati, P.: Security, Privacy, and Trust in Mobile Systems (2015).
13. Ochang, P., Irving, P.: Evolutionary Analysis of GSM, UMTS and LTE Mobile Network Architectures (2016).
14. Meyer, U., Wetzel, S.: On the Impact of GSM Encryption and Man-in-the-Middle Attacks on the Security of Interoperating GSM/UMTS Networks. IEEE 15th International Symposium on Personal, Indoor and Mobile Radio Communications (2004).

The Benefits and Risks of Privacy and Security in the Era of Digital Health: Health National Service (SNS) of Portugal

Ana Moreira
Lusofona University of Porto
Porto, Portugal
a21602146@mso365.ulp.pt

Hugo Barbosa
Lusofona University of Porto
Porto, Portugal
hugo.barbosa@ulp.pt

Abstract. This paper intends to address privacy and security issues according to new digital health times, considering all devices available or integrated into the system. The possible risks and benefits of this new era should be known for better risk prevention and more effective use of benefits. As cyberattacks are increasing and increasingly customized, health cybersecurity has become a fundamental issue, but it is a complex process because it involves a large amount of information to be managed by various types of users. To ensure data privacy, integrity and security, protection mechanisms must be implemented, keeping in mind certain precautions and encourage users to modify certain behaviours. Digital Health is continually being improved, especially regarding the use of the current technologies. This paper includes an analysis to the applications of Health National Service (SNS) of Portugal.

Keywords: Digital Health, Privacy, Security, Risks, Benefits, Prevention, Cybersecurity, Cyberattacks, Confidentially.

1 Introduction

The desire to improve health care through timely access to information and decisionsupport aids, the need for simultaneous access to records by doctors, nurses and administrators is accelerating the move to electronic patient records. [1] [2] We have a strong expectation that our patient records will be used only in the context of providing effective care, and otherwise, the information will be kept secret. [2] Our medical records contain much mundane information about us but may contain some of the most sensitive information (about emotional problems, HIV status, substance abuse, genetic predisposition to disease and others). The access to patient records must be controlled because this can affect the life of the patient, for example, causing social prejudice, affecting insurability and the ability to get a job. [2] [3]

The healthcare security allows preserving the confidentiality, integrity and availability of information. We all are responsible for the information security and have the responsibility of protecting your data. [4] The human factor, that is fundamental to healing patients, is critical to healthcare security, the hospital leaders have to recognize it and otherwise all the security could be compromised. [5] The encryption of medical records is essential to keep protected in the event of a breach, that could result in the loss of patient trust if their data are exposed. All the records that is created, stored or transmitted must be encrypted and we must understand how data flows in the organization. According to [6], only 63% of healthcare organizations encrypt patient health information on their work devices.

In that sense and after the study of several data, such as concepts and numbers associated with the privacy and security in the era of digital health process, the analysis of the existing resources allowed the knowledge of the current spectrum of existing systems which in conjunction with the survey conducted in this paper have helped to understand the shortages. Thus, based on this acquired knowledge, some risks and benefits were made.

The rest of the paper is structured as follows. Section 2 presents an introduction to privacy and security in digital health including terminology and concepts. Section 3 introduces the motivation of era of digital health environment. Section 4 introduces some applications that exist in Health National Service (SNS) of Portugal. Finally, in Section 5, conclusions are drawn and referred future work.

2 Privacy and Security in Digital Health

First of all, we must distinguish three concepts involved in protecting health care information: privacy, confidentiality and security. [2] • Privacy, the right of a person to control the disclosure of personal information; • Confidentiality, the controlled release of personal health information to a care provider; • Security, procedures that help maintain the integrity and availability of information systems and control access to their contents. Confidentiality, integrity and availability are widely accepted as the information Security Triad, describing the three core objectives of information security. Availability refers to the guarantee of consistent access to the information by authorized people and is a crucial objective in the healthcare institutions. [7]

Fig. 1. Schema of Information Security Triad. [7]

The most important threats to patient information confidentiality are: [8]

- From inside the patient care institution (accidental disclosures, insider curiosity and insider subornation);
- From within secondary user settings (uncontrolled secondary usage – those who have access rights to patient information for a purpose in support of primary care);
- Outsider intrusion into medical information systems (unauthorized access).

Cybersecurity will be the main challenge for healthcare institutions, as the number of successful threats at these institutions has grown significantly in recent months. It is known that in the United States, more than 110 million individual health records have been targeted by cyber-attacks in hospital institutions, which represents a 3 times superior number of affected individuals when compared to the total threats occurred between 2009 and 2014. The institutions attacked invested a value five times higher than the one applied in the acquisition of hardware, software and services related to Information Security. [9]

3 Era of Digital Health

In this section we will talk about benefits, risks and forms of prevention in this era of Digital Health.

3.1 Benefits

The health records have the function of continuation of health care, documentation of their processes and communication between health care professionals. [10] As more healthcare facilities, from hospitals to private practices, move from paper charts to electronic medical records, the benefits will increase to physicians and patients: electronic health records can be accessed on demand, and can potentially save lives, time and money. [11]

One major and obvious benefit is legibility, historically, illegible handwriting has been a prime source of medication errors, in one source, more than 60% of medication errors in hospitals were traced to poor handwriting. [12] The use of an electronic medical records helps to reduce medical errors by utilizing computerized prescription entry, predicting drug interactions and displaying a warning for the health-care provider. [13] Some software is designed to integrate with bar code scanning technology, if a nurse scans the wrong medication, an alert pops up alerting to a problem. [12]

The paper charts generate large volumes of documentation that require a lot of storage space, it is difficult to organize and properly store so much information. The search for a paper chart will always be more difficult and time consuming and there is a strong possibility that it is recorded on sheets of paper stored in different compartments. This makes it impossible to locate records quickly and does not allow access to complete health information. Information may become difficult to recover from missing parts or even damages that render handwritten information illegible due to the paper's fragility when stored for a long time. [10]

Another major plus of electronic health records is that patients can be seen sequentially by different providers with up-to-date information immediately available. Other advantages include the fact that the system allows the patient to access their own medical information easily anywhere. [12]

Communication between physicians can be greatly improved with this, allowing each party full access to a patient’s medical history. This access allows for a more indepth evaluation and enables doctors to reach an accurate diagnosis more quickly. In addition, electronic health records can make it easier for doctors to follow up with patients and track continuing care, both under their supervision and that of the patient’s other doctors. The doctors can quickly and

easily pull up test results of their patients and also can verify when they had past exams or procedures. This can save time during a doctor's office visit, and in case of emergency, these records can provide critical and life-saving information to emergency care providers. [11] This can reduce the number of duplicate tests and exams too. Some software can flag each critical value of exams results for clinical staff and can also help physicians determine when to repeat an exam. [12]

Catastrophic events have demonstrated that patients in these situations are often confused and frightened, making it easy to forget personal medical details. Every second counts during an emergency, so having access to a patient's medical history, blood type and allergy information, when the patient is unable to communicate can be the difference between life and death. [11]

The electronic health records can reduce costs through decreased paperwork by reducing the number of employees to do this work and by reducing waste and redundant tests. [14] Facilities can also use these data in more expansive ways, for example, clinical researchers to improve health. Electronic health records have many more benefits than disadvantages, so implementing them is well worth the investment. [12]

3.2 Risks

There is a temptation to use patient health information for business purposes than those initially authorized, for example, to manage risk in insurance or to guide marketing of medical products. According to [2] there are 206 reported cases of direct discrimination from unauthorized use of genetic tests information. At least three companies in health information services industry are members of the "terabyte club" that are organizations with large data warehouses used to collect and analyse data for business applications.

According to [15] a healthcare record that is rich with personal, medical and financial detail, may be 50 times more valuable to a cybercriminal than a stolen credit card information.

Cyberattacks can occur in two forms, one is used to attack data and other is focused on control systems. The first type attempts to steal or corrupt data and deny services, the second type attempt to disable or take power over operations used to maintain physical infrastructure. The clear majority of internet attacks have fallen into the first category, such as important data stealing. [16] [17] According to [18], in 2013 and 2014 healthcare companies saw a 70% increase in cyber-attacks, it's important that everyone acknowledges the situation and understands the risks and impact, an investment in security of the organization will bring down the risk but some risk will always exist.

A cyber-physical attack on building equipment pales in contrast to the damage a determined hacker can do if he gains access to active medical devices. [19] A hospital network controls the diagnostic, treatment and life support equipment on which patient lives depend and these attacks can do a lot more than get information, they can really disrupt the day-to-day operations of your facilities. [20] The active medical devices require wireless communication and internet connectivity for software-based control of therapies and network-based transmission of patients stored medical information, this combination makes active medical devices more vulnerable to cyberphysical attack. [21] [22]

For example, an infusion pump could be controlled remotely by a malicious hacker to cause the machine to dump an entire vial of medication into a patient and the hospital staff member wouldn't notice even if they keep an eye on the pump from a centralized monitoring station. We have read how hackers struck hospitals with ransomware that prevented staff from accessing patient records or scheduling appointments, now that hackers realize the value of the medical networks to the hospital staff, they have seen that hospitals are more likely to pay up without fight because they cannot afford the downtime of restoring patient files from backups. [22] [23]

A terrorist hacker, on the other hand will not ask for money, he will simply hack active medical devices to deliver fatal doses of drugs or he will change pharmacological databases so drug to drug interactions go undetected until it is too late for hospital staff to notice. Unless hospitals take the steps necessary to secure their active medical devices they will be targeted for cyber-physical attack and possibly with lifethreatening consequences. [22]

According to [24] there are some statistics from the year of 2017:

- 1 in 13 Web requests lead to malware Up 3% from 2016;
- Percentage spam rate 55% in 2017 in email;
- 5.4B WannaCry attacks blocked;
- 600% increase in attacks against Internet of things devices;
- 24,000 Average number of malicious mobile apps blocked each day.

This numbers remind us that is very important prevent our devices from being attacked. Many users continue to make life of the attackers easier, by continuing to use older operating system versions. [24]

Ransomware is a type of malware that takes control over a computer system by encrypting all the data on the drive, the data is then held at ransom until a predetermined cost is paid. Due to the use of cryptocurrencies for payment it is difficult to track those demanding the ransom making it tough to prosecute. Ransomware can be transmitted by emails as legitimate business or tempting links, trojans acting as update requests, gaining access by exploiting known network or security software vulnerabilities. [25]

A rapid increase in the computerization of health care organizations, many without the capacity to keep up to date with the extensive privacy and security measures required, has made them targets for cyber-criminals. Health care organizations may be perceived as more vulnerable targets by cyber-criminals due to a potentially smaller IT staff and older set of IT infrastructure. [20]

3.3 Prevention

To prevent and secure your system from threats you need to know exactly what happens to patient health information after it enters your environment. It is your job to ensure it is stored, transmitted or destroyed in the most secure way possible. To do this you must record all the hardware, software, devices, systems and data storage locations that can access information, you

should find gaps in your security and then properly encrypt all the information when it enters in your environment. [6]

There are three common data handling processes often confused: masking, hashing, and encrypting.

- Masking is hiding part of the data from view, it is still there in clear text, you just can't see all of it on the screen. [6]

- Hashing is running the data through a mathematic algorithm to change it into something indecipherable. The best hashing algorithms are designed so that it's impossible to turn a hash back into its original string.

- Encrypting turns data into a series of unreadable characters, that aren't of a fixed length. The key difference between encryption and hashing is that encrypted strings can be reversed back into their original decrypted form if you have the right key. [26]

Over 100 organizations since 2009, according to [6], have had patient health information stolen because of inadequate email encryption, due to the nature of this mean of communication and the struggles to properly secure it through encryption, consider avoiding the transmission of health information via email whenever possible. The use of patient portals is preferred for sending information to patients and it's a way to reduce risks.

According to a study by [27] 36% of healthcare organizations and 55% of business associates that have been breached point to unintentional actions by their employees as the cause. This needs to change and for this they need introductory training sessions for new employees and partners to regular updates and refreshes. Julian M. Goldman says that physicians respond best when they understand why something is important, what the outcome could be and what the risks are. Then they become partners in the solution. [5] Security and privacy of health care information is a "people" problem to, if it is known that the system will record the identities, times and circumstances of all users accessing information and that these records are reviewed regularly, ethical users will think twice about abusing their privileges. Of course technology can help to make sure that only personal authorized access information and that information gets from one place to another accurately and securely, but technology can do very little to ensure that the person receiving the information will handle it according to confidentiality standards. It is unthinkable that we would impose system security constraints so tight that they would prevent an emergency room doctor from accessing the records. [2]

According to [20] there are some primary prevention methods like employee security training and awareness, confirm that backup routines are actively deployed and can be effectively restored, anti-virus programs, firewalls and network access control.

4 Applications of the Health National Service (SNS)

In this section we will talk about the applications that exist in Health National Service (SNS) of Portugal and some measures that helps protect the data of the users.

According to [28] the security and protection of their user data is taken seriously. They seek to comply with good international safety practice as well as with the "EU guidelines on assessment of the reliability of mobile health applications" and the code of conduct and privacy for mobile health applications. They have a page dedicated to cybersecurity but there are still many topics that are under construction. [28]

Mobile health applications have a strong potential to bring important benefits to citizens and society, however, data concerning health is highly privacy sensitive. Therefore, mobile health apps must be planned in such a way that the privacy of the users is protected. [29]

According to [29] you must obtain informed consent as users install the app to process their data for the purposes you've described to them. You must carefully consider what data is strictly necessary for your app and don´t collect or process more data for a longer duration than strictly necessary. The developers should ensure the confidentiality, integrity and availability of the personal data processed via their apps. Is required to implement appropriate measures to protect personal data against destruction, loss, alteration, disclosure, access and other unlawful forms of processing. Processing of personal data must be compatible with the purposes for which you first collected the personal data, as communicated to the users of your app. Secondary processing of the data for scientific research purposes is however still considered as compatible with the original purposes if it is done in accordance with EU level rules adopted. [29]

4.1 MySNS

The app MySNS allows you to consult news, your health information, provide a list and a map of health institutions, evaluation of the quality and satisfaction of the SNS by the citizen, consultation of the SNS Contact Center 24 and you will also receive notifications such as heat alerts and others associated with your location. [30]

Fig.2. Screenshot of the app MySNS. [30]

4.2 MySNS Tempos

Mobile application that allows you the consultation of the average waiting time in the hospitals of the National Health Service. The user can consult, per institution, the average time of attendance in the emergency room. The application also allows to obtain more data about the

hospital institution, such as address, telephone contacts and geographical location, through the use of the GPS of the mobile device. [30]

Fig. 3. Screenshot of the app MySNS Tempos. [30]

4.3 MySNS Carteira

Through the SNS user number and validated with information in the National Register of Users, the electronic health portfolio allows the citizen to associate specific "cards" with informative components of interest. [30]

Fig. 4. Screenshot of the app MySNS Carteira. [30]

4.4 MyADSE

This application allows you to save the ADSE beneficiary card digitally, receive messages when a refund is made, apply for the European sickness insurance card, allow you to update beneficiary and family members data, also allow you to make refund simulations and know the values and applicable rules. [30]

Fig. 5. Screenshot of the app MyADSE. [30]

4.5 Dador CHVNG

With the app of the Blood Service of the Hospital Center of Vila Nova de Gaia/Espinho you will be able to access your personal page and obtain all the essential information as donor, namely the records of previous harvests, analysis results, personal notifications, among many others. [30]

Fig. 6. Screenshot of the app Dador CHVNG. [30]

4.6 Dador S João

This app is identical to the Dador CHVNG but it is destined to Hospital S. João. [30]

Fig. 7. Screenshot of the app Dador S João. [30]

4.7 Dador.pt

The application is intended to promote blood donation, the user will have real-time access to information on where and when to give blood. [30]

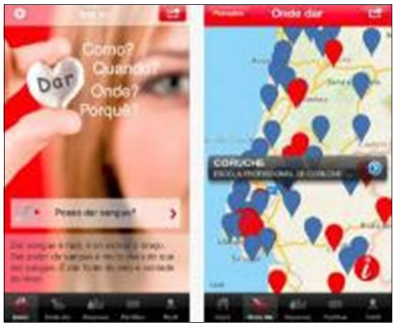

Fig. 8. Screenshot of the app Dador.pt. [30]

4.8 eMed.pt

INFARMED provides a mobile application that facilitates users access to cheaper drug prices. This tool allows that, during the moment of the choice or the acquisition of a certain drug, the user identifies another drugs equivalent, but economically cheaper. The application also provides information about the information flyer and allows you to create a drug consumption plan. Lastly, the application provides the location of the pharmacies closest to the point where the user is. [30]

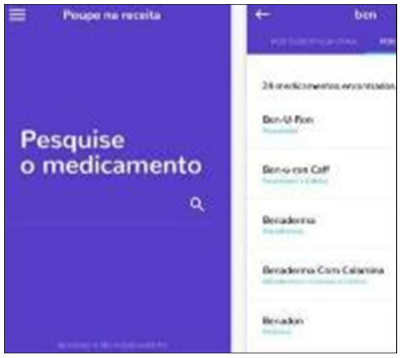

Fig. 9. Screenshot of the app eMed.pt. [30]

5 Conclusion

Today's society is turning into a digital society, with the growing influence of the technology in our daily lives. It is present in so many places, sometimes without people realizing it, that being already frequent is considered natural. This society increasingly directs its attention to new technologies, fostering their development and reaching a closer approximation of knowledge, making it increasingly accessible to all.

There are numerous advantages of using technologies in healthcare institutions, nevertheless it is necessary to reflect that there are vulnerabilities. There is no system totally safe, however attacks can be prevented and damage reduced, considering all the prevention measures as mentioned in this paper.

References

1. IOM: “The computer-based patient record: An essential technology for health care”, Institute of Medicine, National Academy Press, Washington, DC, (1991).
2. Rindfleisch, T. C.: “Privacy, information technology, and health care”, pp.1-10, (1997).
3. Rothfeder, J.: “Privacy for Sale”, New York, Simon and Schuster, (1992).
4. Portuguese Republic - Ministry of Health: “República Portuguesa - Ministério da Saúde, SNS, SPMS: A segurança da informação”, (2017).
5. AT&T: “Physicians & Cybersecurity Risk: A Cybersecurity Handbook for Healthcare CEOs”, pp.3-14, (2018).
6. SecurityMetrics: “Medical data encryption 101”, (2015).
7. OpenText, https://www.opentext.com/products-and-solutions/business-needs/informationgovernance/ensure-compliance/information-security-and-privacy, last accessed 2018/12/20.
8. NRC: “For the Record: Protecting Electronic Health Information”, National Research Council, National Academy of Sciences, (1997).
9. Information Security, http://spms.min-saude.pt/seguranca-da-informacao/, last accessed 2018/12/10.
10. Leal, M.: “Avaliação da Qualidade do Registo Clínico Eletrónico”, University of Minho, Master's Dissertation, (2013).
11. USF Health, https://www.usfhealthonline.com/resources/healthcare/benefits-of-ehr/, last accessed 2018/12/10.
12. Hoover, R.: “Benefits of using an electronic health record”, Nursing2016, (2016).
13. Alpert, J.: “The electronic medical record in 2016: Advantages and disadvantages”, Digital Medicine, (2016).
14. HealthIT, https://www.healthit.gov/faq/what-are-advantages-electronic-health-records, last accessed 2018/12/10.
15. Medscape: “Stolen EHR Charts Sell for $50 Each on Black Market”, (2014).
16. Bogdanoski, M.; Petreski, D.: “Cyber terrorism - global security threat”, (2014).
17. CCRC, http://www.crime-research.org/library/Robert1.htm, last accessed 2018/11/29.
18. Gomes, R.: “Health sector Cybersecurity Strategic Plan”, pp.10-21, (2016).
19. Reel, M.; Roberson, J.: “It’s Way Too Easy to Hack the Hospital”, Bloomberg, (2015).
20. Johnson, C.; Zapotosky, M.: “Under pressure to digitize everything, hospitals are hackers’ biggest new target”, Washington Post, (2016).
21. Kramer, D.; Baker, M.; Ransford, B.; Molina-Markham, A.; Stewart, Q.; Fu, K.; Reynolds, M.: “Security and Privacy Qualities of Medical Devices: An Analysis of FDA Postmarket Surveillance” PLOS One, (2012).
22. Ayala, L.: “Cybersecurity for Hospitals and Healthcare Facilities: A Guide to Detection and Prevention”, Apress, New York, (2016).
23. Leetaru, K.: “Hacking Hospitals and Holding Hostages: Cybersecurity in 2016” Forbes, (2016).
24. Symantec: “ISTR Internet Security Threat Report”, vol.23, pp.7-50, (2018).
25. HITEQ: “Ransomware Guidance For Health Centers”, (2017).
26. Security Innovation Europe, https://www.securityinnovationeurope.com/blog/page/whatsthe-difference-between-hashing-and-encrypting, last accessed 2018/12/04.
27. Ponemon Institute: “Sixth Annual Benchmark Study on Privacy and Security of Healthcare Data”, (2016).
28. SPMS, https://comunidade.mysns.pt/seguranca.html, last accessed 2018/12/15.
29. European Commission, “Draft Code of Conduct on privacy for mobile health applications”, (2016).
30. SPMS, https://www.sns.gov.pt/home/apps-da-saude/, last accessed 2018/12/14.
31. Capelão, F.; Barbosa, H.: “Cybersecurity in Healthcare: Risk Analysis in Health Institution in Portugal”, International Journal for Research & Development in Technology (IJRDT), vol. 9:3, pp. 25 -31, (2018).

Perception of Risk and Precautionary Behavior in CyberSecurity: Hints for Future Research

Eliza Oliveira[1[0000-0002-3518-3447]] and Vania Baldi[1[0000-0002-7663-3328]]

[1] CIC-DIGITAL/Digimedia, Department of Communication and Arts, University of Aveiro
Aveiro, Portugal
elizaoliveira@ua.pt, vbaldi@ua.pt

Abstract. People are using the internet each day more and more, being exposed to the risk of harms in the cyberspace. Thus, it's necessary to identify how individuals perceive those risks and what are the safety behaviors they take to avoid them. However, while risk perception is not a recent area of study and a plethora of cyberthreats frequently emerge, little attention has been given to these kinds of risks in the academy field. The goal of this paper is to present a literature review of the studies concerning risk perception and precautionary behavior associated with the use of digital technologies. To accomplish this, a survey in the multidisciplinary data basis of Science Direct and Web of Science was conducted, focusing in publications after 2016. Seven articles have been analyzed. The small number of studies proves that risk perception concerning cyber-security is still underexplored so that only a few authors commonly have published in this area. Methodological limitations will be noted, as well as other issues, regarding the conducted experiments and data analysis. Additionally, significant gaps to be fulfilled in the next investigations will be pointed out, to provide hints for future researches. Studies in this area provide a dataset for policy-makers, directing educational efforts and predicting public responses to technologies, making this subject-matter highly significant.

Keywords: Cyber-security, risk perception, precautionary behavior, survey.

1 Introduction

In recent years, the whole world has testified an incredible evolution of the digital technologies, highly associated with the feasibility of the global communication and the power of penetration of the information in everyone's life. In this direction, humanity has been living an age of great success concerning to technological advances. In this context, people communicate through the cyberspace, being able to use different kinds of technological devices and being able to connect anytime and anywhere they wish [1]. With this fast development and the lower cost in Information and Communication Technologies (TICs), people are using technological solutions each day more [2]. While this progress has provided countless advantages for internet users, digital technologies bring a series of hazards related to the use of TICs, such as risks related to information-sharing security and privacy in the network [3]. Moreover, despite its advantages, the fact of data and information be easily moved, shared, copied and stored in digital media, cyber threats are frequently forged, being the cyberspace continuously a dangerous environment for sharing personal information [4][2]. These include identity thief, virus, spyware, user surveillance and cyberbullying [5].

While everyone is exposed to risks, young people tend to be more vulnerable to cyber-attacks. Simultaneously, recent studies have verified that university students are lax about the use of technology, more specifically in terms of mobile devices [2].

According to [2], risk scenario regarding the use of TICs leads to the obligation of the person to be aware of risks to which they are exposed online to protect its personal information on the internet [2]. For [6], those risks immediately brings up the subject of awareness, being risk perception directly related to this issue, as well as the precaution of individuals when facing risks.

Risk perception can be defined as an intuitive risk judgement by individuals [7]. Risk perception is an interdisciplinary field and has been widely studied by different areas, including geography, anthropology, social sciences and psychology. Some studies were conducted to find out how people perceive risks related to different kind of hazards such as nuclear waste, chemical risks and automobile safety defects [8]. Also, some of the major challenges in risk science has been the attempt to identify user's risk perception and security behavior regarding the use of technologies. However, as different hazards frequently appear with the use of digital technologies, recently little attention has been given to these kinds of risks in the academy field.

The main goal of this paper is to present a literature review based in recent research publications that show studies related to risk perception and precautionary behavior in the cyber-security domain. Other objectives are identifying gaps to be fulfilled in the next researches and analyze methodological and procedure possible issues.

This article is organized as follows. Next section (2) will present the state of art regarding the meanings of risk, risk perception and precautionary behavior. Also, relevant studies will be highlighted. The methodology associated to the survey is presented in section 3. Results will be described in section 4, as well as the discussion. Finally, section 5 provides the most relevant conclusions and considerations for future works.

2 Background Theory

Throughout the history of the humankind, people always suffered and survived to the inherent risks of each period. Although risks always exist, its meanings have suffered changes over time, as well as the way people perceive and deal with it. In spite studies of risk have begun during the Renaissance, it remains a lack of consensus about the etymology and the meaning of risk in literature nowadays [9][10]. According to [10], some historicist agree that the term came from the Arabic word *risq* which refers to acquisition of wealth and good fortune, while some others believe that "risk" derives from the Latin, *risco*, and it was used primarily by the sailors when entering in uncharted waters [10]. Moreover, the author Peter Bernstein [9] points out that the word risk derives from the earlier Italian, *risicare*, and it means "to dare", defending that risk is a choice rather than a fate. The author also says that "The actions we dare to take, which depend on how free we are to make choices, are what the story of risk is all about" (Bernstein, 1998, p. 29).

Concerning risk perception, Paul Slovic [7], relates the term to intuitive daily risk judgments in which the majority of citizens rely on to evaluate typical and catastrophic hazards. For the author, "The ability to sense and avoid harmful environmental conditions is necessary for the survival of all living organisms. Survival is also aided by an ability to codify and learn from past experience. Humans have an additional capability that allows them to alter their

environment as well as respond to it. This capacity both creates and reduces risk." (Slovic, 2000, p. 220). In this text, Paul Slovic [7] correlate the human capacity of identifying and perceiving risks to the survival, defending that individuals have the unique quality of learning with the environment, facing it and changing it according to their own benefits. This significative adaptation capacity can be notice in everyday life, in which people develop personal safety techniques as a habitual daily activity [10]. In this sense, the perception of risk, as well as the way people will respond to the hazards, is strictly related to personal characteristics, such as the individual's experience, cultural and societal background and the subject interpretation of the risks [11]. Consequently, risk is culturally and socially dependent, varying the way different people perceive it and behave to avoid it [10].

Perception of risk, as well as the risks itself, has been suffered changes throughout the ages, presenting different contours formatted according to social construction and technological development of the historical epochs. In Pre-Enlightenment, Christian understandings of human existence typically perceived the risks and dangers as fate, as an unpleasant but inevitable aspect of life [10]. In the Post- Enlightenment era, dominant religious methods for explaining risk have given rise to technical and scientific rationality [10]. Ulrich Beck presents the science as a driving force behind the definition of risk and claim that the transition from religion to technical and scientific rationality was reflected by the emergence of a distinct institutional form, which should protect citizens from potential dangers of the time. In this period, novel methods of risk assessment linked to mathematics and probability theory have been developed, rising what's called calculation of risk [10]. Therefore, the emergence of risk assessment led to a different mode of perceiving risks, from the lenses of logic and probability. For example, in this period of history, actuarial insurance systems have emerged to determine the likelihood of accidents occurring [10].

In contemporary western society the manufactured threats increase, creating "serial risks", which reproduce to such an extent that risk management mechanisms become insufficient. These "serial risks" bring with them peculiar features: neither institutions nor even scientists hold the knowledge about all the risks associated with new technologies or the repercussions of their dangers to the population. Therefore, the very use of technology itself represents a risk for people, who begins to distrust politicians and regulated organizations [11]. Lash [12] presents that in contemporary times there is a significant intensification in the perceived risks by population, which comes from the dissemination of information through media and the increase of interpersonal communication using digital technologies. For the author, contemporary society introduces an overwhelming flow of judgements, since images, sounds and narratives are brought under the lens of media. Concerning this state of affairs, Wildavsky [13] commented as follows:

"How extraordinary! The richest, longest lived, best protected, most resourceful civilization, with the highest degree of insight into its own technology, is on its way to becoming the most frightened. Is it our environment or ourselves that have changed? Would people like us have had this sort of concern in the past?… Has there ever been, one wonders, a society that produced more uncertainty more often about everyday life? Today, there are risks from numerous small dams far exceeding those from nuclear reactors. Why is the one feared and not the other? Is it just that we are used to the old or are some of us looking differently at essentially the same sorts of experience?" (Wildavsky, 1979, p, 32).

During the last decades, few investigations have been an attempt to answer these questions by identifying the opinion of individuals about hazardous activities, substances and technologies, developing technics and methods for assessing the opinions about several kinds of risks [7]. For [7] the basic assumption underlying risks studies is those promoted and regulated by health and safety needs, to understand the way people think and respond to risk, improving communication between politicians and the public, directing efforts to new risk management strategies (e.g warning labels, regulations, substitute products).

In the next subsection, major significant works carried out in the risk perception and precautionary behavior area will be presented.

2.1 Relevant Studies

Historically, the bulk of empirical studies related to risk research has been conducted in the United States [7]. In this context, the majority of the earlier researches in risk perception utilized the psychometric paradigm as a methodological approach, seeking out to probe behavioral intentions in hypothetical situations and access the decision making progress [10]. Previous study has proposed several risk dimensions as a predictor of perceived risks [14][15]. Starr [15] underlines two categories: those in which the individual is involved on a voluntary basis and those in which the participation is involuntary, imposed by society. His findings show that voluntary activities are perceived as less risky. Following this perspective, Slovic [14] presented nine dimensions which influence risk perceptions, such as immediacy of the effects, the severity of consequences and knowledge about risk. The empirical study evaluated 30 different activities (e.g smoking, alcoholic behaviors) and technologies (e.g X-ray, nuclear power). The outcomes of the indicated dimensions proved to be effective predictors of the perception of risk, risk acceptance and perceived benefits. Other studies conducted by the author proved that, in general, people feel themselves to be immunity from risks related to familiar activities, individuals overestimate the risk presented by atypical but remarkable events, whilst underestimate typical daily risks [16] and that immediacy has a positive impact in risk perception, since they are more likely to provoke anxiety [17].

Kahneman [18] discovered that the more voluntary people believe exposure to the risks associated with phishing is, the less risky they perceive phishing to be. In addition, [18] presented that when positive consequences are immediate and negative repercussion of an activity are not, perceived risk is reduced. Other relevant findings include the reduction of risk perception when individuals feel they are in control, understand about the risks and when they are experts [19].

Despite the fact that the majority of the works related to cyber-security focus on protection and security of the systems, and that studies related to individual's awareness and behavior regarding the use of information systems are limited [2], some works are worth to be mentioned. In this sense, [19] and [20] analyzed a set of hazard (21 and 15, respectively) on internet in terms of risk perception and other risk dimensions. The authors concluded that severity of consequences, accident history, voluntariness, duration of impacts, understanding and possibility of exposure were significant positive predictors of risk [19]. In [20] outcomes showed that voluntariness, knowledge to science, controllability, newness, dread and severity were significant predictors of risk.

Concerning precautionary behavior, [21] proposed a new model to assess user's computer security behavior specifically, susceptibility perception, severity perception and cues to action, together with the component of effective risk management, suggesting that individual's security behavior is strictly related to the perception of threat and evaluation of the behavior to resolve the threat. The authors state that security behavior demands the user to take additional steps towards avoiding security accidents (e.g the use of strong passwords and conduct regular backups) and conclude, that, practices and behaviors related to cyber-security could be considered as protective approaches and as preventive actions. After presenting the grounded theory, the next section details the methodology used in this survey, encompassing all the steps taken in the process of finding and choosing the articles in the data basis.

3 Methodology

Literature review has become an increasingly significant aspect of research for gathering advanced knowledge regarding a specific area and has been detailed previously [22]. The first stage towards conducting the survey was the definition of the appropriate databases to find research articles associated with risk perception and precautionary behavior related to cyber-security. In this direction, as perception of risk and precautionary behavior is an interdisciplinary area, the databases chosen for the research of the articles were the Web of Science and Science Direct. These data aggregate systems are a multidisciplinary research tool that grants easy access to a wide range scientific literature since they encompass publications in physical sciences and engineering, biological sciences, health sciences and social and human sciences. Following this stage, the correspondent keywords must be defined. In this scope, three keywords were designed to find articles correlated with this paper's subject: risk perception, cyber-security and precautionary behavior. Papers and proceedings published before 2016 have been excluded, as well as papers related to workshops, courses, book and book chapter. The year-based exclusion criterion took into account the fact that there is one relevant review published in 2015 [23]. The inclusion criteria were: being original research articles, being published after 2016 and present empirical work respecting the previously mentioned keywords. It is important to emphasize that papers which present studies related to only one of the principal topics (i.e works that provide studies related to perception of risk or precautionary behavior) were also accepted. However, all papers should approach cyber-security issues.

The steps for the selection of the papers used in this survey were the following ones: first, the titles were analyzed. At this point, it has been excluded those papers which do not fit the inclusion criteria of this study. After, a pre-selection was done through the reading of the abstract. Next, all the selected papers were fully read to select the documents that effectively illustrate relevant research that should be considered in this literature review.

As a consequence of the search in the previously mentioned databases, 13 articles were found in Science Direct and only one, repeated article, has been found in Web of Science. Thus, 13 research articles were considered as possible to use in this survey, although only seven were selected as they comply with the inclusion criteria. The next following section presents these selected works.

4 Results

As previously said, seven articles were selected for this survey as they fulfil the inclusion criteria of the study. According to the analysis, the first four main regions dealing with cyber-security problems include the United Kingdom, The Netherlands, The USA and Ireland. The following Table 1 lists the authors and year of publications, the study conducted, and principal contributions of the approaches contemplated in each document. Considering the year of publication, two papers were published in 2019, two in 2018, two in 2017 and only one in 2016. Also, among the seven selected works, four were written by the same author, showing that only few researches have focused on cyber-security studies. Issues related to cyber-security were similar across studies, emphasizing the aspect related to the protection of personal data by individuals, namely regulation of the user's information-sharing and only one measured security and privacy regarding social media.

In respect to the methodological approaches, all works carried out quantitative empirical study with internet users, conducting an online survey. In addition, participants were chosen through different methods, but all through online recruitment such as through Mechanical Turk [24], an invitation sent by Toluna [25], using recruitment services of online panels [26][5], recruitments by e-mail [3][27] and selection as a sample [2].

Table 1. General information of the studies contemplated for this review.

Authors / Year of Publication	Conducted Study	Main Result
Bavel et. al (2019) [25]	Online experiment with a sample of 2024 internet users from several countries. Explored the effects of notifications (message advised and threat appeal) on security behavior when purchase in a mock e-commerce store.	Both of kinds of notifications increased security behavior, but coping message more so.
Jansen & Schaik (2019) [26]	Online experiment with 786 Dutch internet users. Examined the impact of fear appeal messages on user cognitions, attitudes and security behaviors against phishing attacks.	Positive effects on cognition, attitudes and security behaviors.
Cain et. al (2018) [24]	Online experiment with 268 internet users from different countries. Explore the user's knowledge and behavior regarding cyber hygiene, such as the use of antivirus, firewall and providing name in social media.	Gender and age are determinants in user's behavior and knowledge regarding cyber hygiene.
Schaik et. al (2018) [5]	Online study with 201 UK non-student internet users. Examined risk perception, other risk dimensions and precautionary behavior related to set of hazards that correspond with security and privacy settings of the Facebook.	Perception of risk was highest for cyberbullying and information sharing.
Schaik et. al (2017) [3]	Online study with 436 UK and US college students. Examined risk perception and precautionary behavior of a set of internet security hazards, including phishing, identity thief, keylogger and cyberbullying.	Risk Perception was higher for identity thief, keylogger and cyberbullying.
Jeske & Schaik (2017) [27]	Online study with 323 college students of US and UK. Examined the familiarity of users about 16 online threats, internet attitudes and security behavior, including phishing, identity thief, keylogger and cyberbullying.	Three different clusters of knowledge were labeled, which influences in security behavior and internet attitudes.
Ögütçü et. al (2016) [2]	Online study with a total of 881 (169 academic, 317 administrative and 395 college students) individuals from Turkey Examined levels of awareness toward information security in terms of perception and behavioral aspects, using four elaborated scales.	Results show significant differences within samples and habits of internet usage

Measurements data were collected using the Psychometric Paradigm by the administration of Likert scales [2][27][3][5][26][24] in six studies. Therefore, only one research article presented a distinct strategy for obtaining outcomes, determined by the decisions made by participants during the experiment with the mock online purchase [25].

In addition, the numbers of participants were different in all research studies, with a minimum of 201 individuals in [5] and a maximum of 2024 parties in [25]. Crosscultural works was conducted by [3][25][27]. The sample was different in all studies: In [25] the mean age was 40.8 years with 50.3% of females with 40.84% with upper secondary education. In [26] the age range was 19-76 years with 50.6% females and 52.5% with high education. Participants of paper [24] presented age range of 18-55+ years with 142 females, being 38.06% with a 4-year degree, while the sample of [5] have an average age of 42 years with 92 females and 55% with an undergraduate degree. In [3] the mean age of participants was 23 with 336 females. Finally, participants of paper [27] presented an age range equal to 18-60 years, with 74% female and [2] had a mean age of 28.1, being 70% undergraduate. Additionally, all participants in papers [3] and [27] were US or UK college students.

4.1 Discussion

Given the characterization of the selected articles, some observations can be made. Thus, this subsection will provide a general discussion encompassing procedure concerns, methodological limitations, measurement methods and data analysis. Results will not be detailed in this section since the goal is to point out opinions and comments to improve future research in the cyber-security field. However, considerations will be made concerning variables, such as age and multinational approach always where such data are important to be mentioned.

Since the use of internet in countries of Europe and in the US are quite intense, studies in cyber-security shows to be highly significant. Thus, in order to grant a safe use of technologies, citizen's security behavior and awareness must be considered.

Two papers selected in this review utilized theory models that are currently used in health interventions. In this sense, both papers, [26] and [25], used the Protection Motivation Theory in order to accomplish their goals. These two studies presented similar objectives related to the use of copying and threat appeal messages to measure improvements in security behavior. Thus, these were the only works which intervened with participants. Other studies have no intention to intervene with individuals. They only investigate risk perception and security behavior. This ups on to new future works possibilities, since it is important, not only to determine people's awareness regarding the use of technologies but also to improve it.

In regard to methodological approaches used by the researchers, only one work did not use the Psychometric Paradigm to collect data [25]. Therefore, this is one of the methodological limitations of the analyzed works, which utilizes quantitative methods to achieve the purpose of the study. Therefore, there is a lack of qualitative data concerning perception and precautionary behavior in cyber-security domain. This is highly significant, since the perception of risk and precautionary behavior are both dependent on subjective and cultural factors. In addition, all the works were conducted through the online environment, which makes difficult to reach a qualitative personal opinion of the participants with respect to cyber-security issues. Also, the

recurrence of the same authors in the papers makes the theme centralized and little explored. The centralization of publications leads to the recurrent use of only one methodological strategy. The results and contributions are thus limited. Mythen [10] presents a critique of the use of the psychometric paradigm in research on risk perception by social sciences, arguing that risk perception and Internet security behavior are culturally and subjectively constructed. In this sense, should be analyzed through qualitative measures which include interviews and focus group.

Since the whole world has been suffering a significant transformation in human's habits, changing the physical environment to the virtual one, where the construction of online interpersonal relations become natural and all the daily activities seem to migrate too, many contributions can be achieved by conducting a cross-cultural study. [2] obtained interesting outcomes by people of different countries, being Polish and Spanish people's behavior the most insecure among the other countries (UK, German, Sweden). Other findings highlight the way that individual decision-making strategies differ across countries and the author suggests the need to adopt a multi-national approach in studies such as his, especially if the aim is to produce policy options that are generalizable across contexts.

The age is another point to be discussed. Works with a wide range of age found that security behavior tends to be higher by the older individuals. One of the findings of [2] is that students (between the ages of 18 and 30) seem to be the most at-risk group. The reason for this may be the high use of the Internet, especially the social networks and social media. Another reason may be the low possibility that the youth had a negative experience with Internet/information security or technology and that they are less controlled by parents. In article [25] the relationship between age and cyber-security is highly complex. For example, older adults are more vulnerable than younger adults to certain types of phishing attack but less vulnerable to others. Both younger and older adults are likely to modify their security behaviors following a warning of some kind, but older adults are particularly affected by trust violations. Also, Cain et. al (2018) present that, although it is commonly believed that age has an impact on cyber hygiene behaviors, older users tended to behave more securely than younger users. According to the author, this finding was counterintuitive because younger people are believed to have the most know-how about technology. Is important to highlight that studies which encompass a wide range of ages are more likely to identify disparities in the behaviors of people who present themselves in different age groups. Thus, future works must provide information among individuals, in order to provide behavioral comparisons between different ages.

In conducting studies that address individuals' awareness of risk perception and precaution, the authors agree that data security and privacy in the online environment is primarily the responsibility of individuals. Paper [2] points out that the overall success of both software and hardware security mechanisms are based on the effective behavior of users of the IS specification. Also, [10] highlights that in contemporaneity, with the technological improvements, people are more enlightened about the risks they are exposed to and how to prevent or to behave more securely. However, since the risk exposure are each day bigger and the cyber threats are more and more prominent, it is important to conduct training concerning the secure use of the internet and other TICs.

Regarding the perception of risk, in the two works published by Schaik [3] and [5] the particular hazards/activities that were judged to be most risky were cyberbullying, sharing telephone

number, sharing e-mail address and failing to receive login notifications about the Facebook user. Additionally, the same author, in [3], identified that keyloggers, identity theft and cyberbullying have higher perceptions of risk.

Finally, the works address issues regarding, predominantly, with the security of information-sharing on the internet, while topics related to online aggression are not presented. This gap provides hints for future works since cyberbullying threats has been widely studying in literature and the incident has been each day more intense in the entire world. Is also important to highlight the importance of studies regarding risk perception and precautionary behavior for providing a dataset for policy-makers, directing educational efforts and predicting public responses to technologies, making this subject-matter highly significant. This will give significative insights for the production of future commercial product's release in regard cyber-security, such as better protective software and hardware.

5 Conclusion and Future Work

This paper presents a literature review regarding risk perception and precautionary behavior concerning cyber-security threats. Additionally, with the analysis of the papers is possible to conclude that work's majority focus is on cyber-security related to security of personal information, being others significant subjects like cyberbullying and online harassment not addressed by the researchers. Finally, future works should address alternative subject-matters, such as precautionary behavior and risk perception regarding online aggression. In addition, intervention with users in order to educate them for better use of TICs should be considered to be conducted.

Future works should take into account that cross-cultural studies provide contributions to the literature since they enable comparisons and analysis between different countries and cultures. Upcoming researches should be made using alternative methodologies, in order to cover qualitative data of cultural and subjective determinants for perception of risk and security behavior. Information about these subjects will provide great information concerning how the risks are perceived be cultural and societal individuals.

Finally, as previously said, Studies in this area provide a dataset for policy-makers, directing educational efforts and predicting public responses to technologies, making this subject-matter highly significant.

References

1. Baldi, V., Oliveira, E.: A queda do império Facebook: uma análise sobre os motivos que levam ao afastamento da rede social. In: Educación y comunicación mediada por las tecnologías. pp. 41–60. EGREGIUS Ediciones, Sevilha (2018).
2. Öıütçü, G., Testik, Ö.M., Chouseinoglou, O.: Analysis of personal information security behavior and awareness. Comput. Secur. 56, 83–93 (2016).
3. Van Schaik, P., Jeske, D., Onibokun, J., Coventry, L., Jansen, J., Kusev, P.: Risk perceptions of cyber-security and precautionary behaviour. Comput. Human Behav. 75, 547–559 (2017).

4. Loon, J. van: Risk and Technological Culture: Towards a Sociology of Virulence. Routledge, London and New York (2003).
5. Van Schaik, P., Jansen, J., Onibokun, J., Camp, J., Kusev, P.: Security and privacy in online social networking: Risk perceptions and precautionary behaviour. Comput. Human Behav. 78, 283–297 (2018).
6. Assailly, J.-P.: The psychology of risk. Nova Science Publisher, INC, New York (2010).
7. Slovic, P.: Perceptio of Risk. In: The Perception of Risk. p. 511. Taylor & Francis Group; Routledge, Nova York (2000).
8. Kahan, D.M.: Handbook of Risk Theory. (2011).
9. Bernstein, P.: Against-the-Gods-The-Remarkable-Story-of-Risk. John Wiley & Sons, INC, New York (1996).
10. Mythen, G.: A critical introduction into the risk society. Pluto Press, Londres (2004).
11. Adam, B., Beck, U., Loon, J.: The Risk Society and Beyond: Critical Issues for Social Theory. Crit. Issues Soc. Theory. 232 (2000).
12. Lash, S.: Risk Culture. In: The Risk Society and Beyond: Critical Issues for Social Theory. pp. 47–63. SAGE Publications Inc, Londres (2000).
13. Wildavsky, A.: No Risk Is the Highest Risk of All. Am. Sci. 67, 32–37 (1979).
14. Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S., Combs, B.: How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits. Policy Sci. 9, 127–152 (1978).
15. Starr, C.: Social benefit versus technological risk. What is our Society Willing to Pay for Safety. Science (80-.). 1232–1238 (1969).
16. Slovic, P., Fischhoff, B., Lichtenstein, S., Roe, F.J.C.: Perceived Risk: Psychological Factors and Social Implications. Proc. R. Soc. A Math. Phys. Eng. Sci. 376, 17–34 (1981).
17. Slovic, P.: Perception of Risk: Reflections on the Psychometric Paradigm. In: Krimsky and D. Golding (ed.) Social Theories of Risk Westport. pp. 117–52. Praeger, New York (1990).
18. Kahneman, D.: Thinking, Fast and Slow. Routledge, New York (2017).
19. Huang, D.L., Rau, P.L.P., Salvendy, G.: Perception of information security. Behav. Inf. Technol. 29, 221–232 (2010).
20. Garg, V., Camp, J.: End User Perception of Online Risk Under Uncertainty. In: Proceedings of 45th Hawaii International Conference on System Sciences. pp. 3278– 87. , Manoa (2012).
21. Ng, B.Y., Kankanhalli, A., Xu, Y. (Calvin): Studying users' computer security behavior: A health belief perspective. Decis. Support Syst. 46, 815–825 (2009).
22. Gough, D., Oliver, S., Thomas, J.: An Introduction to Systematic Reviews. SAGE Publications Inc, London (2017).
23. Quigley, K., Burns, C., Stallard, K.: "Cyber Gurus": A rhetorical analysis of the language of cybersecurity specialists and the implications for security policy and critical infrastructure protection. Gov. Inf. Q. 32, 108–117 (2015).
24. Cain, A.A., Edwards, M.E., Still, J.D.: An exploratory study of cyber hygiene behaviors and knowledge. J. Inf. Secur. Appl. 42, 36–45 (2018).
25. Van Bavel, R., Rodríguez-Priego, N., Vila, J., Briggs, P.: Using protection motivation theory in the design of nudges to improve online security behavior. Int. J. Hum. Comput. Stud. 123, 29–39 (2019).
26. Jansen, J., Van Schaik, P.: The design and evaluation of a theory-based intervention to promote security behaviour against phishing. Int. J. Hum. Comput. Stud. 123, 40–55 (2019).
27. Jeske, D., Schaik, P. Van: Familiarity with Internet threat: Beyond a wareness. 66, 129– 141 (2017).

Cybersecurity and Cybercrimes in Portugal

Justino Silva

Lusofona University of Porto, Portugal
jdps96@gmail.com

Abstract. With the development of technologies, not everything that originated from them is good. This paper intends to share knowledge regarding cybersecurity, as well as cybercrimes. Basic security principles of cybersecurity, as well as types of cybercrime will be addressed. Most users are inexperienced and susceptible to cyber-attacks, so it is important to warn them to protect themselves to avoid being cybercrime victims. This paper also addresses the state of cybersecurity in Portugal, and one of the most recent cybercrimes that damaged the world.

Keywords: Cybercrimes, Cybersecurity, Cyber Threats, Cyber-attacks, WannaCry.

1 Introduction

The security of an organization is only as strong as its weakest component. [1]

Most people don't realize that the Internet isn't safe and use it daily in their devices like phones and computers, unaware of the dangers it has. The need for cybersecurity is increasing, because with the Internet of Things being developed, the drawback is that our everyday objects now feature an IP (Internet Protocol) address for Internet connectivity, which an intruder can access its control, stealing personal data or causing catastrophic damage. [2]

Computer crime can be summarized as a criminal activity which involves an unauthorized access, illegal interception, data interference, abusage of devices, forgery, blackmail, fraud, and others, to a system in order to steal confidential data or cause damage. Cybercrime can cause harm to any organization. As the use of technology is increasing day-by-day, the crime is also increasing gradually. [3]

In this paper, we start by defining the meaning of security, complementing with the basic security principles that define a secure system.

Next, we define what are cybercrimes, splitting them in different types, and show the most common types of cybercrime. In this chapter we also present a cybercrime that occurred in 2017, forcing all governments to invest in cybersecurity education and implementation.

After this chapter, we focus in Portugal specifically, showing some research results, that give us and idea of the state of cybersecurity in this country. The main reason of the creation of this paper was to inform about this last chapter referring to Portugal, in order to show that security is

a very important topic that is often forgotten until it is too late, and the message should be shared with everyone, so people get the information that they are missing.

2 Cybersecurity and Cybercrimes

2.1 Cybersecurity

The author [4] defines cybersecurity as the practice of protecting systems, networks, and programs from digital attacks. These attacks are usually aimed at accessing, changing, or destroying sensitive information; extorting money from users; or interrupting normal business processes.

Implementing effective cybersecurity measures is particularly challenging nowadays because there are more devices than people, and attackers are becoming more innovative.

According to [5], there are four basic security principles:

- Access - Using physical and software controls to protect your hardware or data from intrusion.
    - For hardware, access limits usually mean physical access limits.
    - For software, access limits usually mean both physical and virtual means.
- Authentication - provides a means to identify a person or entity. Setting up all authentication features such as a password system in your platform operating systems to verify that users are who they say they are.
    - Authentication provides varying degrees of security through measures such as badges and passwords. For example, ensure that personnel use employee badges properly to enter a computer room.
- Authorization - defines what an authenticated user or entity can do. Use authorization to ensure company personnel can only work with hardware and software that they are trained and qualified to use.

  For example, set up a system of read/write/execute permissions to control user access to commands, disk space, devices, and applications.
- Accounting - Customer IT personnel can use software and hardware features to monitor login activity and maintain hardware inventories.
    - Use system logs to monitor user logins. Track system administrator and service accounts through system logs because these accounts can access powerful commands.

10.11228/dpsc.01.01.004

- Periodically retire or archive log files when they exceed a reasonable size, in accordance with the customer company policy. Log files can become very large over time, so it is essential to maintain them.
- Use component serial numbers to track system assets for inventory purposes. Oracle part numbers are electronically recorded on all cards, modules, and motherboards.

2.2 Cybercrimes

Cybercrime consists of a criminal act that is committed online by using electronic communications networks and information systems. It is a border less problem that can be classified in three broad definitions: [6]

- Crimes specific to the Internet, such as attacks against information systems or phishing (e.g. fake bank websites to solicit passwords enabling access to victims' bank accounts).
- Online fraud and forgery. Large-scale fraud can be committed online through instruments such as identity theft, phishing, spam and malicious code.
- Illegal online content, including child sexual abuse material, incitement to racial hatred, incitement to terrorist acts and glorification of violence, terrorism, racism and xenophobia.

The most common types of cybercrime are:

- Malware – Malware is malicious software that exploits a network´s vulnerability and accesses it in order to install something like spyware, ransomware, virus, worms. Once inside the system, this software can block access from the user(ransomware), obtain information(spyware), render the system inoperable [7]
- DDOS (Distributed Denial of Service) – A Denial of Service works by overwhelming an online service, in order to overload the system and make it shut down or reboot. A DDoS is a DoS on a larger scale, the attacker uses various IP addresses so there are less chances of the attack failing, and tracking the perpetrator is harder. [8]
- Botnets - Botnets are hacked networks where the computers connected to them are being remotely controlled by hackers. The hackers then use those computers to access the network's database or to perform crimes like a DDoS or steal data. [9]

- Phishing – phishing attacks work by either making the user click a fraudulent link that appears to be from an official source and it automatically installs malware, or the user deliberately types personal information on a fraudulent website. [10]

- Virus – a virus is a malware that infects a computer, and it's usually disguised as a creditable source. As a biological virus, it spreads itself on the computer and to other computers, causing hardware and software problems. [9]

- Ransomware – ransomware is a malicious software that can for example be planted on a fraudulent link or email, that once clicked infects the computer and encrypts parts of the data or even all of it, blocking the user from accessing it. The software then displays a pop-up message where it informs the victim that his machine is encrypted and asks for payment, usually in Bitcoin. Once the ransom is payed, ideally a decryption key is shown on the screen. [11]

- Sextortion – sextortion is blackmail using sexual information in return of sexual favours or money from the victim. Either the criminal contacts the victims using an online social network for adults, by assuming the identity of an attractive man or woman, or he hacks them, then proceeds to record and save nude pictures and videos from the victims, which then threatens to release online unless they either pay a ransom or appeal to the blackmailers demands. [12]

A cybercrime can happen to anyone, or anything, that is online. The two main reasons are either for money, where the hackers ask for a ransom, or just to destroy data or even the system itself.

According to [13], 53% of the cyber-attacks resulted in damages of 500,000 dollars or more. As referred in the same study, here's a graphic that shows the top 10 malicious file extensions in emails, in the year of 2017.

Fig. 1. Top 10 malicious file extensions [13]

- 38% refer to Office formats such as Word, PowerPoint or Excel
- 37% refer to Archive files, such as compresses files in .rar or .zip format
- 14% refer to PDF files

The user, when seeing these types of files on an email that looks legit, will open them without questioning because it looks official.

2.3 A recent and dangerous Cybercrime that failed

In 2017, the world was attacked with a ransomware named WannaCry. It was a worm that spread by exploiting vulnerabilities in the Microsoft's Windows Operating System, most precisely those that weren't up to date as of March 14, 2017, and once installed, encrypted files and demanded a payment in exchange of the decryption key. It had two primary components, a module for self-propagation and a ransom module to handle the extortion process. The ransom could be between 300 dollars and 600 dollars. Each file was encrypted using separate AES encryption key, and each key was separately encrypted using 2048-bit RSA encryption. [14]

The first ransomware ever appeared in 1989, the AIDS Trojan, it spread in floppy disks and to pay the ransom, people had to send 189$ to a post office in Panama. [15]

WannaCry resulted in over 200,000 organizations affected around the world, spread over 150 countries.

One of the most affected services was the National Health Service in the United Kingdom, infecting the computer system of 47 hospitals, forcing operations and doctor's appointments to be cancelled for a lot of patients. In the end, the hackers only managed to profit less than 70,000 dollars converted from bitcoin. [16, 17]

3 A focus in Portugal

As mentioned in [18], according to Microsoft, as of 2017, there was found malware in 8.3% of the analysed computers in Portugal, 7.0% of those were Trojans.

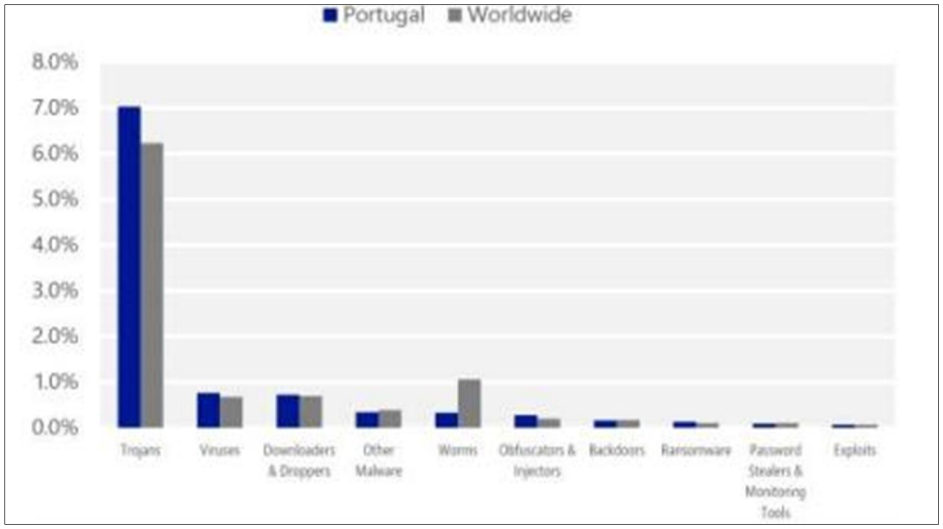

Fig. 2. Percentage of malware in the amount of analyzed Computers in Portugal as of March 2017 [18]

Source: Website Builder Expert (2017)

In the same document, we can see a study from Website Builder Expert, also in 2017, that shows the most vulnerable countries to Cyber Threats, and the countries with the greatest number of cybercrime victims.

EU COUNTRY	CYBERCRIME VULNERABILITY SCORE				
1. MALTA	42%	11. SLOVENIA	38%	21. SWEDEN	32%
2. GREECE	41%	12. CROATIA	37%	22. ITALY	31%
3. ROMANIA	41%	13. DENMARK	36%	23. FRANCE	31%
4. SLOVAKIA	40%	14. LATVIA	35%	24. UK	31%
5. SPAIN	40%	15. CZECH REP	35%	25. NETHERLANDS	30%
6. LITHUANIA	39%	16. POLAND	34%	26. GERMANY	30%
7. CYPRUS	39%	17. IRELAND	33%	27. ESTONIA	30%
8. PORTUGAL	39%	18. LUXEMBOURG	32%	28. FINLAND	29%
9. HUNGARY	39%	19. AUSTRIA	32%		
10. BULGARIA	38%	20. BELGIUM	32%		

Fig. 3. UE's countries in greatest danger to cybercrimes [18]

Source: Website Builder Expert

In the previous table, we can see that Portugal occupies the eight position on the most vulnerable countries to cybercrimes.

The same study also shows that on the top five countries with the most percentage of cybercrime victims, Portugal is on the third place.

	% OF POPULATION WHO HAVE EXPERIENCED CYBERCRIME	ANNUAL AVERAGE MALWARE ENCOUNTER RATE	CYBERCRIME VICTIMHOOD RATING
1. ROMANIA	18%	28%	23%
2. NETHERLANDS	27%	14%	21%
3. PORTUGAL	15%	24%	20%
4. POLAND	16%	23%	20%
5. ITALY	17%	21%	19%

Fig. 4. UE's countries with the most victims of cybercrimes [18]

Source: Website Builder Expert

Portugal's National Cybersecurity Centre was created in 2014, formally assuming the powers of national authority over cybersecurity subjects. It also deals with national management and coordination of response to cyber incidents, also ensuring international cooperation in this subject. [19]

The official website to CNCS informs the public about cyber threats, displaying alerts of vulnerability in compromised computer programs and smartphone apps. It also shows news about cyber conferences, or job vacancies in the area of cybersecurity, or even cybersecurity learning courses. There's also an option to notify the organization about a criminal cyber act suffered, but it should also be notified to the police.

CNCS counts with the cooperation of different enterprises like NATO, ENISA, the European Commission, and others. It also partners with the project "No More Ransom", that vouches to stop criminal activities connected to Ransomware.

Cybersecurity can be a pretty rewarding job, with salaries reaching as high as 62,000€, as an IT Security Engineer in a bank for example. [20]

As of an ENISA study shown in [21], here's a table that represents the top 15 cyber threats in the last four years prior to 2018. The top three greatest threats occupy the same positions in every year, being them Malware, Web-based Attacks and Web application Attacks. This clearly shows that these are well built and are hard to cypher and fight.

	2014	2015	2016	2017
1	Malicious code: Worms / Trojans	Malware	Malware	Malware
2	Web-based attacks	Web based attacks	Web based attacks	Web based attacks
3	Web application / Injection attacks	Web application attacks	Web application attacks	Web application attacks
4	Botnets	Botnets	Denial of service	Phishing
5	Denial of service	Denial of service	Botnets	Spam
6	Spam	Physical damage / theft / loss	Phishing	Denial of servisse
7	Phishing	Insider threat accidental)	Spam	Ransomware
8	Exploit kits	Phishing	Ransomware	Botnets
9	Data breaches	Spam	Insider threat	Insider threat
10	Physical damage / theft / loss	Exploit kits	Physical manipulation / damage / theft / loss	Physical manipulation / damage / theft / loss
11	Insider threat	Data breaches	Exploit kits	Data breaches
12	Information leakage	Identity theft	Data breaches	Identity theft
13	Identity theft / fraud	Information leakage	Identity theft	Information leakage
14	Cyber espionage	Ransomware	Information leakage	Exploit kits
15	Ransomware / Rogueware / Scareware	Cyber espionage	Cyber espionage	Cyber espionage

Fig. 5. Greatest Cyber Threats in Portugal [21]

4 Conclusion

It is very easy to stay outdated in technologies, so it is very important we try to refine our knowledge every day in order to keep up with the times. Most times, systems get hacked only because they are outdated. Often businesses keep the same hardware and software unchanged for decades, thinking they are saving money instead of buying new equipment because the old works, but they end up paying even more when they get hacked, updating all the equipment, plus the security system and to recover lost data. Portuguese schools and hospitals or health centers, for example, are using outdated technology. Health facilities with outdated technology, are always at a higher risk of machine malfunctions, which puts human lives at risk, and schools that nowadays don't have functioning computers can't provide its students with means of learning computer sciences and other subjects alike.

People should start worrying a lot more about their safety online, mainly when they just type their personal data anywhere, and when they upload their photos online. It makes it easier for hackers to get the information they need about someone to hack them.

This paper was hard to conceive because cybersecurity is still a very recent term, and Portugal is only starting now to invest in it, so the information available is either scarce and in small detail, or it requires to be purchased. Still, I believe there's a lot to learn and uncover about cybersecurity. Maybe it's because most of cyber criminals are self-taught, and most of the information about this subject is hidden to the public because it would help the authorities to reverse engineer solutions to cybercrimes faster.

It is good news that Portugal is creating cybersecurity-oriented degrees and post High School technical courses, in order to create cybersecurity jobs and improve the country's national security. But, in order to take the next step, I believe that basic security should be taught in schools as early as elementary school, because nowadays, kids begin using devices at a very young age, so it is fundamental that they should know how to stay protected online.

References

1. Conteh, N.; Royer, M.; "The Rise in Cybercrime and the Dynamics of Exploiting the Human Vulnerability Factor", International Journal of Computer, Volume 20, Number 1 (2016)
2. Bruijin, H.; Janssen, M.; "Building Cybersecurity awareness: The need for evidence-based framing strategies", Government Information Quarterly, January (2017)
3. Tiwari, S.; Bhalla, A.; Rawat, R.; "Cyber Crime and Security", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 6, Issue 4, April (2016)
4. "What is Cybersecurity", Cisco Security, Retrieved From: "https://www.cisco.com/c/en/us/products/security/what-is-cybersecurity.html", Last Access: 10 December 2018.
5. "Basic Security Principles", Oracle Basic Security, Retrieved From: "https://docs.oracle.com/cd/E79568_01/html/E79571/glymd.html", Last Access: 10 December 2018.
6. "What is Cybercrime?", European Commission Migration and Home Affairs, Retrieved From: "https://ec.europa.eu/home-affairs/what-we-do/policies/organized-crime-andhuman-trafficking/cybercrime_en", Last Access: 10 December 2018.
7. "What are the Most Common Cyberattacks?", Cisco Security, Retrieved From: "https://www.cisco.com/c/en/us/products/security/common-cyberattacks.html", Last Access: 18 December 2018.
8. "Types of Cybercrime", Panda Security Mediacenter, Retrieved From: "https://www.pandasecurity.com/mediacenter/panda-security/types-of-cybercrime/", Last Access: 18 December 2018.
9. "Top 5 most prominent forms of cybercrime", CBR Online, Retrieved From: "https://www.cbronline.com/list/top-forms-cybercrime", Last Access: 18 December 2018.
10. "Common Online Threats and How to Protect Yourself", LastPass, Retrieved From: "https://blog.lastpass.com/2013/06/common-online-threats-and-how-to-protectyourself.html/", Last Access: 18 December 2018.
11. "Ransomware", Enisa Ransomware, Retrieved From: "https://www.enisa.europa.eu/topics/csirts-in-europe/glossary/ransomware", Last Access: 18 December 2018.
12. "Online Safety", Interpol Cybercrime, Retrieved From: "https://www.interpol.int/en/Crime-areas/Cybercrime/Online-safety/Sextortion", Last Access: 18 December 2018.
13. "The attack landscape", Cisco Annual Report 2018, Retrieved From: "https://bit.ly/2sMN4Fi", Last Access: 18 December 2018.
14. "Ramson Wannacry", Symantec Security Center, Retrieved From: "https://www.symantec.com/security-center/writeup/2017-051310-3522-99", Last Access: 18 December 2018.
15. Mohurle, S.; Patil, M.; "A brief study of Wannacry Threat: Ransomware Attack 2017". International Journal of Advanced Research in Computer Science, Volume 8, No.5, May-June 2017, Department of Computer Science MITACSC, India, (2017).
16. "NHS cyberattack: Everything you need to know about 'biggest ransomware' offensive in history", The Telegraph News, Retrieved From: "https://www.telegraph.co.uk/news/2017/05/13/nhs-cyber-attack-everything-need-knowbiggest-ransomware-offensive/", Last Access: 20 December 2018
17. "What is WannaCry and how does ransomware work?", The Telegraph Technology Intelligence, Retrieved From: "https://www.telegraph.co.uk/technology/0/ransomware-doeswork/", Last Access: 20 December 2018.
18. Barros, G.; "A Cibersegurança em Portugal". Temas Económicos, Number 56, August 2018, Gabinete de Estratégia e Estudos, Ministério da Economia, Lisboa, Portugal, (2018).
19. "Centro Nacional de Cibersegurança", Fundação para a Ciência e a Tecnologia, Retrieved From: "https://www.fccn.pt/inteligent-transitions-in-ux-design/", Last Access: 20 December 2018.
20. "Security Salaries in Portugal", Glassdoor, Retrieved From: "https://www.glassdoor.com/Salaries/portugal-security-salarySRCH_IL.0,8_IN195_KO9,17_SDAS.htm", Last Access: 20 December 2018.
21. Barros, G.; "A Cibersegurança em Portugal". Temas Económicos, Number 54, August 2018, Gabinete de Estratégia e Estudos, Ministério da Economia, Lisboa, Portugal, (2018)

Digital Investigation of a Cybercrime: Sextortion as a Case Study

ALJI Mohamed and CHOUGDALI Khalid

Electronics and Telecommunication Systems Research Group, National School of Applied Sciences,
Ibn Tofail University, Kenitra, Morocco
mohamed.alji@uit.ac.ma and chougdali@gmail.com

Abstract. Sextortion is a way of blackmailing a victim for money, or favors, in which a sexual content in a digital format is being used for the extortion. In order to apprehend the criminal and present him to the court of justice, law enforcement agencies need to identify the suspect and link its computer to the happening of the crime. The present article is a case study of a digital investigation on a sextortion crime. The aim is to present a way of solving such a crime based on the digital artifacts that may be found on the suspect computer. We also held an experiment to demonstrate the utility and the reproducibility of our approach and concluded by a discussion on some unique indicators of committing the crime using the suspect computer.

Keywords: Digital investigation, Forensic analysis, Software Forensic artifacts.

1 Introduction

1.1 Cybercrime

The internet can be reached by more than 46% of the world population according to the Internet Live Stats project [1], that's more than 3.9 billion people all over the world that can use internet up to now. Imagine if a small fraction of those people know a malicious way to exploit the endless possibilities of the interconnected devices, that may engender the happening of a lot of digital crimes or a so-called cybercrime.

The term 'cybercrime' is used to express a crime that a computer device is somehow involved in it. That means not only the illegal hacking of an online banking system is considered a cybercrime, the offense of stealing a notebook is also a cybercrime. In fact, according to [2], cybercrime means two things either it is the crime that has migrated from the real world to the cyberspace, for instance : harassment can either be on schools during recreation time or on online blogging platform, or it is the offenses that are specifically designed for the cyberspace, such as distributed denial of service, spam or website defacing.

One of the offenses that migrated from the real world to the cyberspace and got an amplified effect because of the easiness of multimedia sharing content through social networks and online platforms is Sextortion.

1.2 Sextortion

According to the definition available on the official Interpol website[3], sextortion is a blackmail in which a sexual information or images are used to extort sexual favors and/or money from the victim. A sexual content at the disposal of a malicious person can lead to an attempt of a favors extortion from the victim. The question in our context is "how an internet user can commit such

a cybercrime ?". For illustration purposes, we describe one scenario of a typical happening of the crime as follow: the criminal assumes the identity of an attractive woman and uses a suggestive video broadcasted as a live video to an instant messaging application (such as IMO or Skype). The victim believes that the conversation being held is real and instantaneous, thus, answering the suggestive solicitations with similar ones. The criminal takes advantage of the situation and proceed to the recording footage of the victim in the nude or performing a sexual act. After that, the criminal blackmails the victim and threatens to diffuse the recorded video to an online video sharing platform such as Youtube or to share it with the victim's Facebook friends, unless a certain amount of money is paid [4].

1.3 Fight back : Digital forensics efforts

Some victims pay the money taking no guarantee of the none reiteration of the extortion. Other victims choose to deposit a complaint to the police, the investigation yields most of the time to the identity of the suspect. The local police proceed to the seizing of the suspect computer. At that moment, it is up to the computer forensic expert to assert whether the suspect is the searched for extortionist or not, through analyzing the digital forensics artifacts present on the digital evidence device that may link the computer of the suspect to the happening of the crime. Knowing the digital forensics artifacts and how to figure them out for a typical scenario of such cybercrime will easy the computer forensic expert task. The main aim of this article is to bridge such gaps.

The present article is structured as follow: In section 2, we will describe the adopted research methodology and the way the experiment has been designed. In the second section 3, we depict the results of the research. Lastly, in section 4, we discuss those results in details and we conclude.

2 Research Methodology

2.1 Overview of the Computer Forensic Examination Process

Understanding the computer forensic examination process is from two different standpoints: the theory and the practice. At first, the computer forensic examination is the practice of collecting, analyzing and reporting on digital data, as visualized in the following diagram:

The collection part consists of the gathering of the digital evidence following the most current best practices [5]. The analyzing step consists of finding out what happened in the digital environment based on the available digital forensic artifacts. The reporting phase consists of answering the issues raised to solve the case [6].

Fig. 1. The process of computer forensic examination

In practice and in general, the steps of a computer forensic examination proceeded by a law enforcement agency are as follow [7]:

During the collection phase, the police department needs to obtain the authorization from the court of justice to search and seize the digital devices that may contain digital evidence. After that, they should secure the area where the devices are found (the crime scene). From the seizing of the devices, a document of the chain of custody is established to ensure the traceability. The devices are uniquely identified and securely transported to the digital forensic laboratory. In some cases, the digital evidence acquisition is to be done on the scene on a running live system. A forensic image is created from the devices using forensically sound methods and tools.

During the analysis step, and while securing the original material in a safe location, the digital forensic examiner designs a review strategy of the digital evidence including, for instance, a list of keywords and search terms. Then proceed to the examination and the analysis of the forensic images according to that strategy.

During the reporting phase, the digital forensics examiner interprets and draw inferences based on facts gathered from the digital evidence. Then, he writes down in a standardized format report the description of its findings. He may also give testimony under oath in a deposition or in a courtroom as a witness expert.

In what follows, we will simulate an extortionist computer and study its digital forensic artifacts. So we need to have a computer-like running system with a sextortion predefined scenario set-up. We then acquire the forensic image, and we consider the bitstream image as the input. The simulated bitstream image is the forensically sound acquired image from the suspect computer that has been used to commit the cybercrime.

2.2 Setting up the computer suspect Virtual Machine

In order to experiment with the possibility to go back on time, we made the choice of using Virtual Machines. We chose Oracle VM VirtualBox as a Virtual Machines manager on a Linux machine for stability reasons. We downloaded the operating system Windows 10 on ISO format available from the following link [8]. Following the recommendations of the Oracle VM VirtualBox official manual [9], we created a Virtual Machine. We installed the guest OS (Windows 10) on the Virtual Machine after configuring the virtual disk as dynamically allocated and relatively small starting size. We installed the VirtualBox Guest Additions on the guest os. Finally, We executed the steps of the predefined scenario of sextortion as described in the following paragraphs. As stated before, we are supposed to create a suspect machine with the cybercrime already committed using it. One way of committing a sextortion is visualized on figure 2.

10.11228/dpsc.01.01.005

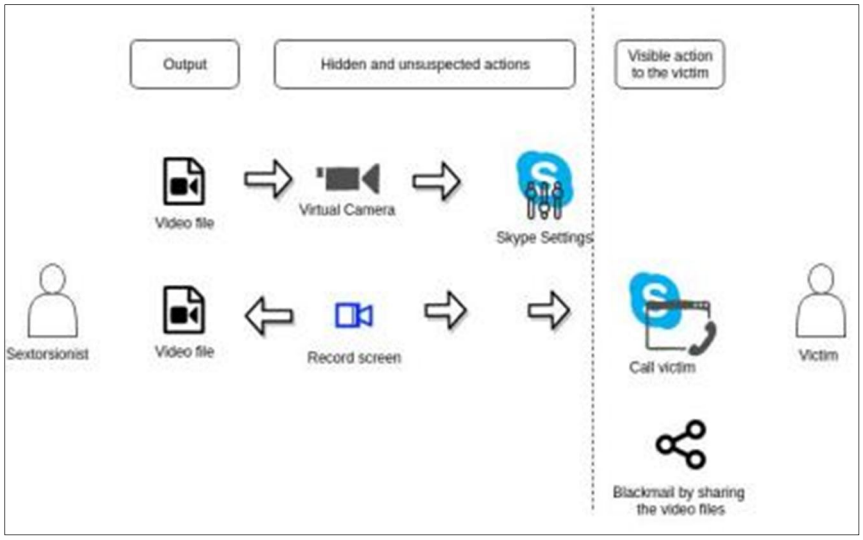

Fig.2. A simplified illustration of a way of commiting a sextortion crime

We describe the steps to make the virtual machine of the suspect computer commit the sextortion crime as the following predefined scenario :

- We installed the instant messaging software (Skype) and video camera effects manipulator (SplitCam) and configured them in a way to set in Skype settings the SplitCam Virtual Camera as a default camera for Skype video conversation.
- We created one account for the suspect (Sarah-suspect) and one account for the victim (victim-one) on Skype. We made a text conversation and a call to the victim account.
- We recorded the footage using one of the available screen recorder and save the recording result to a file.
- We shared the recorded file on a private channel on Youtube, and then proceeded to the extortion discussion through Skype.
- We finally uninstalled the video camera effects manipulator (SplitCam).

2.3 Acquisition of forensic images for analysis

From the VirtualBox manager on the Linux machine, we proceed to the snapshots needed while performing the steps of the previously predefined scenario. For instance, a snapshot is done after the installation of the instant messaging software and the video camera effects manipulator. Then, we create a clone of the virtual disk using the following command:

```
VBoxManage clonemedium snapshot_i.vdi snapshot_i.raw --format RAW
```

The i corresponds to each snapshot required. Then, we are performing a bitstream copy and a conversion from the virtual disk format (.vdi) to a RAW format because the digital forensic analysis platform (Autopsy) does not support a data source in .vdi format.

The acquired images are forensic images according to [10] and will be useful for any postmortem analysis. The following flowchart 3 details the design of the experiment retained to establish the necessary forensic images for further analysis.

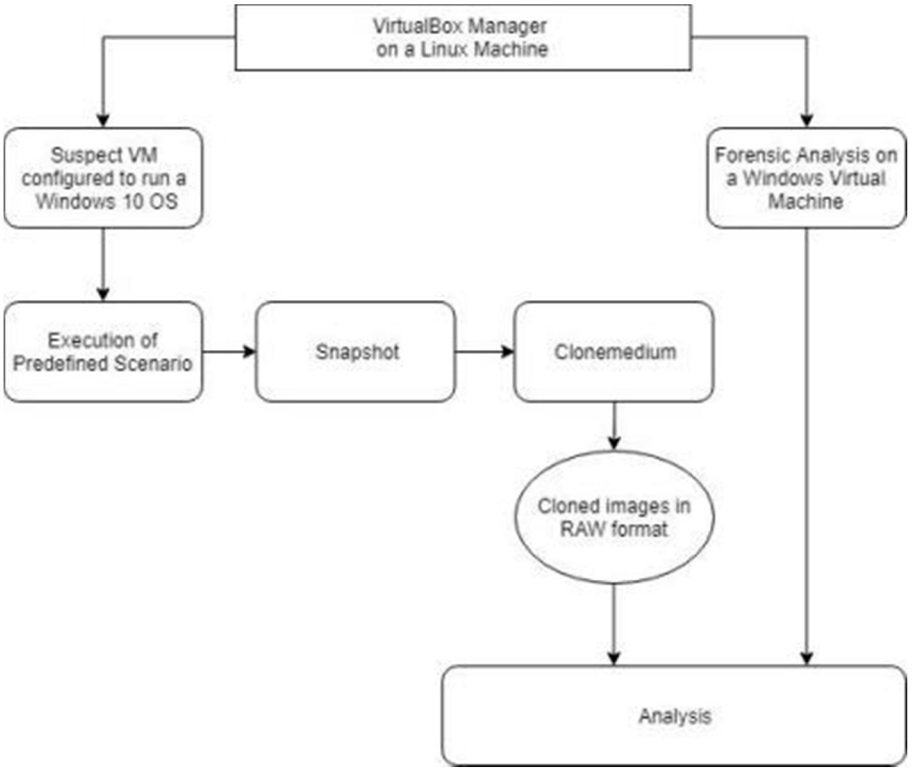

Fig.3. Experiment design flowchart for the acquisition of forensic images

2.4 Tools used for the Analysis of the suspect virtual machine

Since we are playing with a simulated version of a suspect machine, we can use multiple tools to analyze multiple aspects of digital forensics artifacts. We will do a postmortem analysis of the content of the filesystem using the digital forensic analysis platform: Autopsy. The registry content will be frozen before and after each step and a comparison is established using: RegShot. Registry Viewer will allow to view the registry content and parse some registry values. The hexadecimal viewer: HxD will allow the representation of a hexadecimal data and a search within unstructured values. The file recovery: Photorec will allow the recovery of deleted or lost files independently from the filesystem.

Oracle VM VirtualBox v5.2.18	create forensic images
Autopsy v4.8.0 on Windows os	parse the file system content
RegShot v1.9.0	save state of the registry for analysis
Registry Viewer v2.0.0	view registry keys and hives
RegLookup v1.0.1	recover registry keys
HxD v1.7.7 on Windows oS	search unstructred data & view hexadecimal values
PhotoRec 7.0 on linux	recovery program of lost files

Table 1. A list of used tools

3 Results Analysis

3.1 The recovery of SplitCam log file

The video camera effects manipulator SplitCam is installed on the following directory %ProgramFiles(x86)%\SplitCam knowing of course that it is provided only as a 32-bit program. Some data specific to SplitCam are store don the following directory C:\Users\%USERPROFILE%\AppData\Roaming\SplitCam. An interesting text file named "SplitCam.log" is created in this directory after the first run of the application. The log file has an important forensic value since it records all the uses of SplitCam with their timestamps. For instance, we can observe the line "> Source ". It tells the source of content used at a specific time for SplitCam. It may be a Camera, a file or a folder of a list of multimedia content. Other digital artifacts are a desktop shortcut that points to splitcam.exe, a quick launch shortcut, and a start menu folder. The main objective of using the file recovery PhotoRec is to recover the SplitCam log file as a remnant digital forensic artifact. In order to narrow the process, we limited the recovery to txt and tx? files format. And for further optimization, we used our own custom signature for the SplitCam log file which corresponds to:

extension name : log

offset of the signature : 0

signature or magic value :"==== ["

footer: 0d-0a 0d 0a hex

We managed to figure it out as described in following documentation web page [11]. But before doing so, we checked successfully that PhotoRec can recover the SplitCam log by using the beta version of PhotoRec online checker. The recovered log file contains a line that starts > Source: that indicates the source of the broadcasted video content. In our case, we found a video file in Windows Media Video format (WMV) entitled "Wildlife.wmv".

3.2 Recovery of meaningful registry subkeys

Windows registry keys and subkeys left by software can contain a valuable forensic information [12]. In our case, the main registry key artifact left by Splitcam can be found on the following location: HKEY_CURRENT_USER\Software\ SplitCam\SplitCam. On the file system, the file "NTUSER.dat" comprises those pieces of information is located in the user main directory "Documents and Settings". In the previously mentioned registry key path, there are some registry subkeys, named "LastDevice" and "LastVideo", that have a forensic value. "LastDevice" indicates which device was lastly used by SplitCam as a source of content. For instance, if the subkey contains the value "SuperVision HD". That means the last device use don SplitCam that added effects toitsaninstantaneous stream is the suspect machine camera "SuperVision HD". After uninstalling SplitCam, such registry key may disappear. Recovering from windows registry using reglookup-recover tool [13] & [14] will allow to uncover the value of the registry subkeys "LastVideo" and "LastDevice". If this recovery does not succeed, an approximate search of the unstructured data within the file "NTUSER.DAT" using a hexadecimal viewer such as HxD, yields its values.

3.3 Exploring the content in Windows filesystem

There is another artifact, that allows as to check whether the suspect had set the SplitCam virtual video camera as a default source camera for the instant messaging application Skype. We found that the settings file "settings.dat" located at %USERPROFILE%\AppData\Local\Packages\Microsoft.SkypeApp_ID\Settings contains a parameter that shows "SplitCam Video Camera" set as prefered camera. The digital forensic analysis platform Autopsy allows the parsing of the filesystem content. So that, a simple search for some files based on their types will allow the finding of Videos and Images that may have been used to commit the cybercrime. A search for the video file that was set as a source in SplitCam log can ease the task. In addition, reading and parsing the skype discussions content logged in the suspect machine may reveal the extortionist discussion. While this part will not be detailed since Skype forensic artifacts are widely explored by the digital forensic research community. Sometimes, the parsing of the suspect's browser history can disclose the presence of links to an online video sharing platform such as Youtube. And those pointed out videos may have been removed from the platform because of their sexual content.

4 Discussion

The main objective of the digital investigation is the findings of unique incriminating indicators of the happening of the cybercrime (Sextortion). In our case, the presence of digital forensics artifacts of the typical following software such as a Camera Effects Manipulator, an Instant Messaging Application and even a Screen Recorder can indicate that the suspect may have used this trio of software to commit the sextortion crime. But that is still insufficient, since the suspect may have used that software for other common reasons. Setting a virtual camera from a video effect manipulator that points to a suggestive video file as a default source video for Skype will reveal the criminal intention of a suspect. In addition, the recovered registry keys may have disclosed that the registry key "Last Video" points to a suggestive video file. The presence of incriminating content of files such as video files that may be used to lure, like

suggestive videos of girls or the screen recorded scenes of unknown people will help the judge decision. Further more, the extraction of artifacts of Instant Messaging such as the blackmail discussion and the log of made calls to the victim can assert the happening of the sextortion. Any further extracted web history of the accessed links and shared links of uploaded videos related to the matter will help incriminate the suspect.

5 Conclusion

In this paper, we described a typical way to solve a Sextortion cybercrime and we performed an experiment to assess the reproducibility of the results. We used virtualization capabilities to simulate a suspect machine. We examined some digital forensics artifacts that lead to the incriminating of the suspect. Still, there are much more ways to commit the Sextortion crime. Indeed, the sextortionists can innovate and use other instant messaging applications rather then Skype. They can record the screen in so many ways, for instance using a lightweight standalone application, or an installed one, they can use old version of the tools. They can run the portable version from a USB thumb. It is difficult to size all the malintentioned possibilities. In a future work, we will try to size those cases in order to provide a more integrated way to deal with such a crime while aiming at assisting the justice.

References

1. InternetLiveStats.com. Real time statistics project. internet live statistics. 2016. [Online; accessed 28-July-2018].
2. SusanW.Brenner. Cybercrime:Re-thinkingcrimecontrolstrategies. Crimeonline, pages 12–28, 2007.
3. Interpol. Online Safety. https://www.interpol.int/Crimeareas/Cybercrime/Online-safety/Sextortion, 2018. [Online; accessed 07-Avril2018].
4. Michael Joyce. Video chat extortion and sexual abuse. 2012.
5. Tom Killalea and Dominique Brezinski. Guidelines for evidence collection and archiving, 2002.
6. J. Kävrestad. Guide to Digital Forensics. Springer Briefs in Computer Science, 2017.
7. Reynaldo Anzaldua Linda Volonino. Steps to take in a computer forensics investigation. http://www.dummies.com/computers/pcs/computer-security/steps-to-takein-a-computer-forensics-investigation/, 2008. [Online; accessed 15-Juin-2018].
8. Microsoft Corporation. https://www.microsoft.com/fr-fr/softwaredownload/windows10ISO, 2018. [Online; accessed 20-Avril-2018].
9. Oracle Corporation. Oracle vm virtualbox manual. https://www.virtualbox.org/manual/, 2004-2018. [Online; accessed 13-Sept2018].
10. Manish Hirwani, Yin Pan, Bill Stackpole, and Daryl Johnson. Forensic acquisition and analysis of vmware virtual hard disks. The 2012 International Conference on Security and Management, 2012.
11. CGSecurity. Add your own extension to photorec. https://www.cgsecurity.org/wiki/Add_your_own_extension_to_PhotoRec, 2016. [Online; last accessed 14-Sept-2018].
12. Harlan Carvey. The windows registry as a forensic resource. Digital Investigation, 2005.
13. Timothy Morgan. Reglookup. http://projects.sentinelchicken.org/reglookup/, 2011. [Online; last accessed 14-Sept-2018].
14. Timothy Morgan. Recovering deleted data from the windows registry. The Digital Forensic Research Conference, 2008.

Security Issues in Serious Games Web Environments

Nuno Pontes

Lusofona University of Porto Porto, Portugal
nunopontes04@gmail.com

Abstract. The purpose of this paper is to have an overview about the concept of Serious Games, to try and understand the various types of games that exist in that realm, as the contextualization on where they sit through where "Serious" is concerned. Adding to that an array of examples as well as the contribute to real professional jobs, the importance these games bring to the professional world, the message that Serious Games transmit. The possible solutions to solving security problems, the importance that serious games can bring to cyber security issues and demonstration case study.

Keywords: Security, Educational Games, Digital games, Security Issues, Serious Games, Games, Game Web Environments

1 Introduction

When we reference games, we often associate the term with leisure and entertaining, however, serious games have another objective other than pure entertainment. The games which it refers to are more focused on to educational areas, publicity, military training, and medicine. Their objective is mostly to train professionals, workers, students and in some cases, via simulation when it comes to practical fields corresponding to the military. Similarly, to simulating military games, there's also Advert games, which pertains to publicity games, where the focus promotes a brand, a product or a service. Ultimately, whenever serious games are brought up, their objective will pertain to educational and learning purposes, either to help new businesses arise assist diversified professionals getting prepared to start their new work. Additionally, backtracking into the origin of the term, Serious Games, the idea of using games aimed at dealing with serious matters is an old concept. According to [1] "Americas Army game was the first successful and well-executed serious game that gained total public awareness." Even though a flush of Serious Games appears to begin in 2002, quite a few were simply designed for serious purposes before this date. By current day definitions, any digital game that was designed for a purpose going beyond entertainment can be considered a Serious Game. The very first game to ever be created, regarding popular culture, was Pong, for Atari console in 1972. Was effectively the first video game to embrace commercial success. Among the video games created before Pong, some titles were not designed for pure entertainment, but rather for serious purposes, to illustrate a scientific research study, to train professionals and to broadcast a message according to [1]. According to [1] "We are concerned with Serious Games in the sense that these games have an explicit and carefully thought out process and are not intended to be played primarily for amusement." The definition stated on [1] in the 1970s paper entitled "Serious Games", conveys to us, that it is a rule nowadays whenever the word game in brought up in a sentence it is strictly regarding the entertainment and leisure, which is far from the truth.

2 Serious Games about Web Environments

During the cold war, the army invested a sum of money into the research field, as well as some investments to fuel a few projects from this particular time-frame, which led to technology that is now prevalent in our daily lives, such as computers and the internet. Despite that fact, many of the first computers were initially conceived to operate under military purposes, ranging from ballistics, computations, resource management, to simulations. Officers around the globe used war games in place of training in addition to educational assistance, and such thing had been the spark to fuel the idea to create computer-influenced war games in research departments [2]. According to [2] "Hutspiel" was a strategy game created in 1955 and worked on this fashion by allowing two human players to experiment with the impact of nuclear weapons on a global battlefield. The game was highly detailed, as it simulated ammunition and fuel supply for each unit controlled by two players.

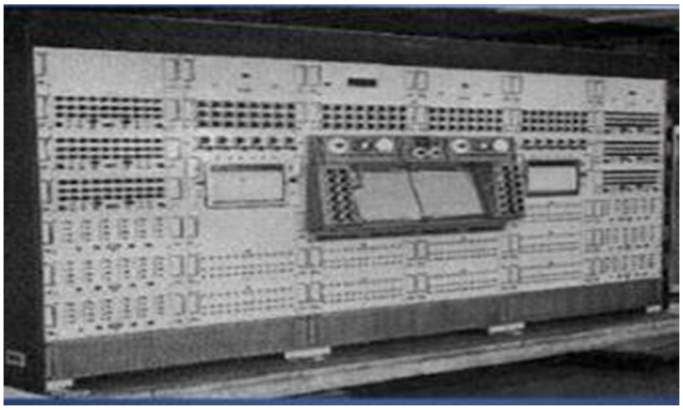

Fig.1. Hutspiel

Several other games were given the birth "post-Hutspiel" project, namely “T.E.M.P.E.R”, a cold war simulation game created in 1961, and “ARPA-AGILE COIN GAME”, another game which simulates an internal revolutionary conflict in a country. These strategical military games represent a step too much more complex simulation games [2]. Most of these games were not available for public use, and the little information found nowadays about them is mainly treated as unclassified military documents. None of these games were available for general public, the little information we could find about them comes from unclassified military documents. Which makes the tracking down for reference images rather complicated, most of them do not exist online for the public to see. However, we can classify the games mentioned previously on this paper as being the ancestors, so to speak, pf the simulation video games that appeared on personal computers in the 80s, when it comes to military topics games such as “Dunningan, 1992. The old projects were created to help to boost this field and served as good source for the current serious games available nowadays. Additionally, next will be approached their importance and impact to professional and educational fields.

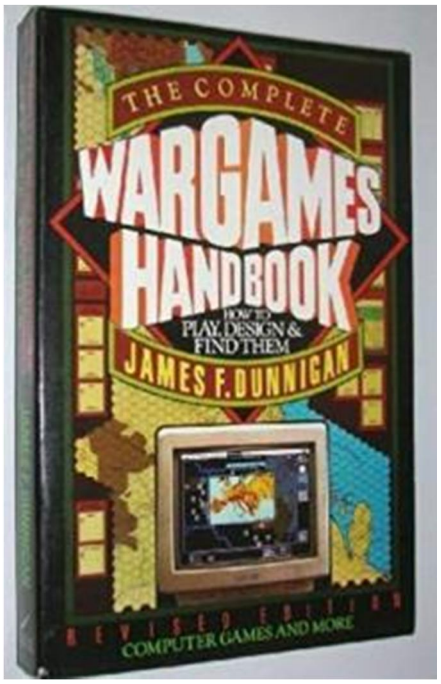

Fig.2. Dunningan

3 Importance of Serious Games

Serious Games carry the weight and responsibility to transmit messages in a more meaningful, interactive and interesting way. It is certainly a viable addition to aid professional work, in order to educate interns/workers on their fields of work, by making the learning process into an appealing experience.

There have been given a few different purposes for Serious Games over the course of years, such as personal development, medical health, publicity, product promoting, military, social science and corporate purposes. A clear sign of the importance over the last decade, the UK and the US carried out a cyber war game to test the resilience of the financial sector in the City of London and on Wall Street. The attack orchestrated in London banks was designed to test out the City's defenses against online saboteurs. The exercise was useful but the real challenge lays in co-coordinating across the industry to make sure a crisis scenario is never reached [3].

Serious Games work because they appeal to nearly all ages, certainly to adults, making teaching bearable and transferring knowledge, a much more enjoyable and intuitive experience. The feeling of being connected to Serious Games is often due the social aspect of our human nature, allowing ourselves to be connected to others like us, like a ranking system commonly seen in many kinds of games, rather scoring high on the ranking charts allowing individual games to reference other players and create a feeling of kinship.

According to a study [4] carried out by the Pew Research Center, proving that 97% of teenagers play some sort of digitalized game, which comes to enable Serious Games to be highly relatable to a huge portion of the population, as players tend to feel in control, or rather autonomy. Having met that feeling of autonomy, they learn and manipulate outcomes within the game

itself, while learning and educating themselves as they play. Efficiency talks volumes, because if we make a mistake in a game we do not die, for example when it concerns piloting training. Much the same is relatable with other professional roles in society, we will not lose a real case if we are lawyer roleplaying. And that is the beauty of games, they are in a different reality as the one we live in. They guide us, providing enough ammo and fuel, to transfer enough confidence to make mistakes, while actively learning, and not being afraid of the consequences. The more a game is played the less frequently the mistakes are made, growing out our efficiency [4].

Serious Games are useful in that regard because they have this gist in turning what could be taxing and boring, into potentially transfiguring a subject into a more entertaining, powerful way of introducing new concepts. However, even if things seem quite promising and revolutionary when it comes to new learning processes, the threats to that web environment are not free of problems, which is something that will be addressed next, the issues concerning web environment.

4 Cyber Security Problems

In the midst of education process from online services, which provide coverage for professional games, there might be certain complications, adapting the form of viruses, trojan, hacking, amidst others. Most of us enjoy playing, either casually, socially or professionally, however, online threats exist and more often than not they are issues to pay attention to. For instance, there are few specific and well-known threats usually can cause complications. Teslacrypt was designed to encrypt game-play data for dozens of video games, prompting the user to pay a ramson to decrypt those files. Targeting some well-known games including Call of Duty and Minecraft, it basically blocks the access to saved game files, configuration files or game files [5]. There are other problems which are alarming when it comes to security issues. Password stealing is known practice among cyber-gaming, either professionally or casual play. In this case, there are various types of spyware called keyloggers, which capture keyboard events and try to steal access credentials, there are also pieces of malicious code that attempt to steal access credentials for online games or platforms [5]. One of the most popular scam plots is when a player receives a chat message from another player offering to join his team, then the unknow player is usually super friendly and praises the victim for his skills, telling him that he should join this team of great players. The deceit resides when the victim is encouraged to download and install an application, it can adopt the form of a voice communication program. The attacker will be insisting on the fact that the victim cannot be part of the group, if the software is not installed on his system. Least as we expect the executable is definitely not something to help the victim joining the group, but rather a malicious software that is capable of stealing account credentials [5]. One of the most problematic cases are those of an android trojan which hides himself among the games in google play, which allows the attacker to control devices remotely, thanks to its backdoor capabilities. By imitating games and e-learning apps [5]. Another case remotes to a fake Minecraft app that installs scareware, which was downloaded by more than 600,000 android users. Which consists on the following, after showing the victims sign of fake viruses or threats on their device, it then tried to convince the users to subscribe to a premium SMS service in order to remove the fake threats [5].

Usually the TeslaCrypt attacks as follow: infects a machine, scans all drivers for files, encrypts the files with an algorithm, and replaces the desktop with some sort of ransom message [6]. There might be much more ways to disrupt the peaceful web gaming or just web browsing, but there are very common and effective in way, because they can be easily used all that's needed is for the victim to be unaware of what their circumstances currently are and they will fall for the deceit. One particularly dangerous threat to computer users is phishing. This is a type of attack, in which victims get invited by scam emails to visit fraudulent websites. The attacker creates a fraudulent website which has the look and feel of the legitimate website. The users are invited by sending scam emails to access the fraudulent website and steal their money. [7]

Another thing to note are fake apps. Nowadays apps are our day-to-day download or rather we use them quite often if we compare it to a few years ago, such is the evolution of technology [8] Apps tend to be very efficient in many of our daily tasks, school, cooking, car service, grocery shopping and of course gaming. And the apps can also be used in many electronic devices, such as laptops, tablets, and smartphones. Therefore, being careful is an overstatement, because the bottom line is that there are fake apps masquerading as official games, and even updates. In the theme which I am writing this article for it is important to note that there are many apps for professional learning, being games or simply any kind of apps to serve as support to digital learning.

Showcasing solutions to the main issue on the next chapter, when it comes to fight off all the threats referred. Usually when we mention Cyber Security it is often concerned to applying measures to ensure confidentiality, integrity and availability of digital data that is either stored, sent or received. The security defenses can range from protection software to risk, to training about awareness, all of them do prevent security breaches, such as loss of data, theft or damage to our computing system.

According to a recent report released by [9] on the threat overview during the period from October'1st to December 31st, 2017 "fourth quarter," most of the threats are targeting government entities (48%), energy and telecommunication sectors (15% and 11%), which reflects the threat actors' intent in impacting the national economy, while the number of the Threat Alerts was slightly higher (7%) as compared to the third quarter of 2017. As explained on the graphic bellow:

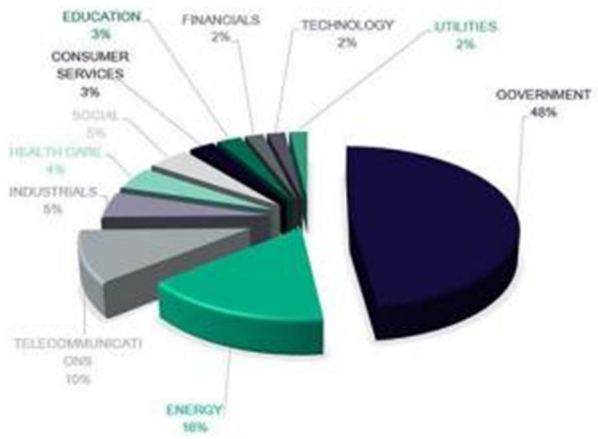

Fig.3. Graphic for the amount of security violations

There are currently some practical curses that help enhancing cybersecurity skills two practical, project-oriented, courses on cyber-attacks and defense. They focus specially on adversary thinking, a crucial skill for cybersecurity experts who must be able to think like an attacker in order to set up effective countermeasures. While this skill can be exercised in Capture the Flag games, challenges, and competitions, their courses introduce an innovative approach. The learners are guided to create a serious security game deployed at the KYPO cyber range, which allows emulating real threats and attacks in a controlled environment [10].

In order to boost the cybersecurity knowledge amongst ourselves the Dutch government issued a serious game, naming it ThreatBattle. A game that basically helps boosting the cybersecurity awareness to avoid eventual collateral damage done by security threats.

In this game a concept called "transreality" is used, on which its meant that the game takes security elements from the players real-life situations and incorporates them into the game [11]. Another serious game exists named Qbit in which the participants are challenged to show their true colors in a game situation on the prison island of Alcatraz. Basically, the game helps identifying the success factors and points to the attention regarding security awareness, management or project group, providing an insight into behavioral and practical working methods [12]. Although the game doesn't focus mainly on computing aspect, in other words, the technological part of it, it does, however still boost up and improves the sense of responsibility and risk awareness, and a better understanding and commitment regarding cyber security.

According to [13] which launched its first training product in December after over 20 years in the online gaming industry, named Ares. The benefit of cyber ranges [13] says, is that "in a virtual environment, if we break the network, we just respawn it." They approximate a real-life scenario without the real-life risk. The project was useful on the way that it does not rely on instructors and is intuitive enough. Their new product Ares introduced game theory and artificial intelligence to cyber security training [13].

It is important to note that serious games nowadays play such an important role when it comes to give pointers to educational fields, professional fields as well, serious Games, can be a great asset to society, research, and business, because lately society focuses much more in digital world, and with the upcoming of internet, the smartphone, and bitcoin [14]. In general, adults and even children, live inside a system that already required them to adapt with modern technology and be ready to move rapidly to more advanced technologies. It is quite logical that people grown to adapt to this situation and learn how to make use of it. In other words, modern media will become their learning objective [14].

The positive aspect of this is that the cybersecurity professionals behind the project itself are also gamers, and thus it makes this simulation-based gaming platform the potential solution for cybersecurity training.

Another study that was issued in 2016 namely '2016 Data Security Incident Response Report' by [15], the top reason that caused 37% of all data security incidents last year was Phishing/hacking/malware. The second largest number of incidents, 24%, can be tracked back to employee actions or mistakes. External theft of devices (17%), vendors (14%), internal theft (8%), and lost or improper disposal (6%) complete the figure bellow.

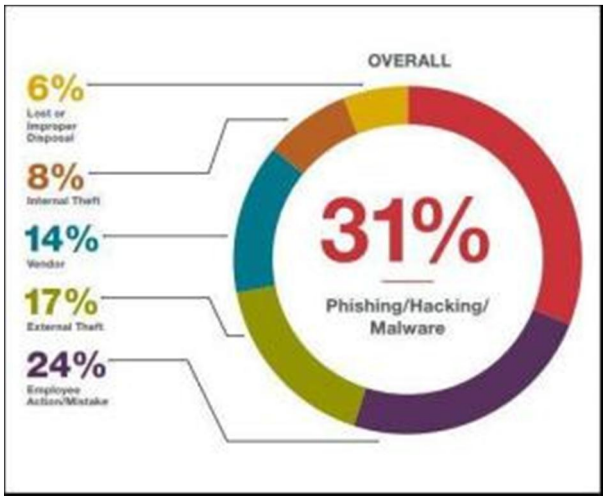

Fig.4. Study about the data security incident

Furthermore, on the table below there was a study done by [15] where it displays a few games that concern to cyber security, while also having details that describe the type of game, topics and target audience. Most are free to play, and their main focus is online security. The study is as follows according to [15] where they say that the number of studies into cyber security is rising exponentially, and nearly all studies focus on efforts to train and raise awareness within the general public.

Game Name	***Game Type***	***Methodology***	***Results***
TiER	Interactive role-play	EEG and Eye tracking	Unclear
Anti-Phishing Phil	Mobile application training safety of link URLs	Think aloud, pre-test & post-test experimental vs. control, SUS usability questionnaire	Positive impact on learning, awareness and phishing susceptibility
Security games by Next Generation Security (NGSEC)	Web-based	Comparing on-task performance	Significant improvement in game
CyberCIEGE	3D virtual world (sims style)	Unclear Experiment & self-assessment Theoretical review of cognitive principles	Sufficiently flexible to illustrate a wide range of topics and positive early indication Positive Unlcear, but there is a need to create a science of games
PicoCTF	Web-based	Survey	Positive educational experience according to students & instructors

Table 1. CyberSecurity Games

5 Solution for Security Problems

All these security exploits can be daunting to deal with if all we want is merely to make use of technology to learn our job craft, or simply to indulge ourselves in casual web browsing or playing games. There are, however, various ways to also defend ourselves against these security breaches to our web environment.

The principles for our daily technologic devices, such as computers, smartphones and tablets, is much the same. We have to be cautious about opening unknow files attached to email messages or instant messages, as well as checking the authenticity and security of downloaded files and new software.

The use of an antivirus, malware software and antispyware software is a must, as well as using a firewall if the case is to use a computer device. Identity and also a backup to our personal data and creating and using testing passwords do play an important role when it comes to security. The protective software cannot do all the work, this must come from each of us, the sensitivity to avoid certain things that might slip through our protective systems. [16]

There are also when it comes to professional gaming environments, security practices, that can help boost our experience risk-free of rather problematic events. The Administrator Mode risk, which consists in some games requiring the use of administrator mode. It's important to make sure the game vendor is reputable and download the game from a site we believe we can trust. Usually free download of games sometimes conceals malicious software. Including "plug-ins" sometimes required to run games. Bottom line is by operating on administrator mode, we open ourselves to the risk that an attacker could gain complete access to our personal computer. [16]

Playing the game at the game site can often make the experience risk-free. Because when we play an online game usually on a web browsing, we can use the administrator mode, and when we are done playing the game, we can just switch back to a user account and simple browse the web. This will reduce the risk of ending up on a malicious web site as well [16].

6 Conclusion

In this paper about security on web game environments, the theme was approached with special emphasis to security, in order to have it safe it's important to know the risks. First of all, as I talked about in this paper, a Serious Games, is a game which is more than just entertainment, it is often aimed to education, or helping people, supporting them to change their behavior or simply a cause in the world. In media we have educational films, documentaries, and the rise of serious games comes from the acknowledgement, the interactivity. The interactive games are known a powerful media, in order to make a message stick deeper and also being more convincing.

From a learning point of view this method of teaching is more effective, its more about active learning because there is no choice in order than participate to progress on a video game, thinking about what we are doing as opposed to be sitting on a living room watching the documentaries or simply watching videos to learn, this would be under a very passive experience. The great benefit from using serious games is to change specific patterns in our

behavior, changing them for the best, to stray away from regular mistakes and being aware of certain details we would do normally. For instance, on the medical field, serious games, can potentially save a few lives, in terms of the results, it is aiming to get the great thing about games, since they are digital we can measure everything as they are really analytical.

When it comes to cyber defense, cyber security, I've noticed the absence of these types of games in the casual entertainment of game market, and in my opinion, this is relevant because if we've got serious games in the casual game market we can reach a larger audience and that is relevant special in education if we want to reach a larger audience and fight more efficiently against cybercrime. The ideal scenario is to reach a broader number of people not just a specific group of people of X company or school, the idea should be to reach everyone or as much as possible.

This paper served to educate me in this field further, as well as helping me understand how important the serious games are in this current century, even more so since we live in a digital era. This way of learning is perfect to boost our confidence, to help us know what is dangerous, and to cyber security specially, it's a way for experts and non-experts raise awareness for what can be dangerous when dealing with digital threats, whether they reside on a serious game web environment or simply when browsing the web.

References

1. Samari N.; "Brain Feasts, Longer Reads", (2015).
2. Minhua M.; Andreas O.; Lakhmi C. J.; "Serious Games and Edutainment Applications", (2011).
3. Scuffham M., Franklin J.: "Cyber-attack 'War Game Tests", (2013).
4. Lenhart A.; Kahne J.; Middaugh E.; Rankin M. A.; Evans C.; Vitak J.; "Teens, Video Games, and Civics", (2008).
5. Porolli M.; "Online Threats and How to Avoid Them", (2016).
6. Wang S.; "Analysis of the TeslaCrypt Family", (2017).
7. Vayansky I.; Kumar S. A.; "Phishing - Challenges and Solutions", (2018).
8. Yan L. and P.; "Fake Apps: Feigning Legitimancy".
9. NCSC; "2017-Q4 Threats and Rist Report", (2017).
10. Svabensky V.; Vykopal J.; Lastovicka M.; "Enhancing Cybersecurity Skills by creating Serious Games", (2018).
11. Grevelink J.; "Serious Games for Cybersecurity", (2015).
12. QBIT; "A Serious Game", <https://www.qbit.nl/awareness-training/serious-games/>.
13. Oesch T; "Serious Games for Serious Topics: Training Cybersecurity Professionals Us-ing AI-Powered Games", (2017).
14. Mans B.; "Serious Games", (2017).
15. Bartl P.; "Enterprise Gamification & Serious Games for Cybersecurity – The Human Factor", (2016).
16. Hayes J. E.; "Playing it Safe: Avoiding Online Gaming Risks", (2008).

A Zero Trust Approach to Network Security

Pedro Assunção

University Lusofona of Porto Porto, Portugal
pedroj_9@hotmail.com

Abstract. In the last years, we have seen an increase in the use of wireless networks due to new forms of communication. The online security has become a hotly debated topic in the community. People want to have access to all of your applications and resources anywhere, anytime. With the increase in the use of Cloud computing and IoT, the number of connected devices increases that consequently also increase the targets of cybercrime. A simple change of mentality can help protect data and the entire network. This paper describes what a Zero Trust Network is and show some concepts behind this architecture/philosophy. Zero Trust is an architecture that has a principle that everything inside or outside the network is not reliable until verified.

Keywords: Cybersecurity, Business Security, Zero Trust Network, Google BeyondCorp.

1 Introduction

Cybersecurity is a hotly debated topic today because of the breach large amounts of sensitive information from large companies. Which has caused a big question on who we can trust our data.

So, the IT Landscape has changed, and the use of networks substantially have risen. The users want to access applications everywhere, every time and with this, we have a great amount of sensitive data on these networks/applications that cybercriminals want to be able to profit with. So, the security models have to support this evolution to keep user data safe.

According to the authors [1], the conventional model normally used has in mind build a wall between trusted and untrusted resources, local network and the internet for example. And according to the Computer Security Institute (CSI), approximately 60 to 80 percent of network misuse incident is originated inside the network [2].

With these needs come a security architecture called “Zero Trust” that was developed by John Kindervag of Forrester Research [1].

Zero Trust, unlike the conventional model, has as its principle "never trust, always verify," where both internal and external networks cannot be trusted. This principle is the basis for reducing the risk of attacks not only external but also internal.

This model brings new concepts on how to design a corporate network such as segmentation gateway, which allows increasing the micro-segmentation of a network with the aim of having more visibility over all traffic by inspecting all types of users and devices that connect in the

network. BeyondCorp is an example of a zero trust architecture designed by Google that allows employees to work more securely in any location without the need for a traditional VPN.

2 Network Landscape

Initially, with the appearance of the first computer systems, companies would be more "isolated" which mitigated the number of attacks focusing on their efforts to restrict access only within the company by hierarchical levels. Since these days the safety models developed have focused on separating the "trusted resources" from the "untrusted resources" using layers of protection to build digital perimeters. The according to the author [3] the traditional perimeter security depends to the firewalls, VPN's and web gateways who has to deal with employee's skill shortages, overloaded, and an ever-expanding number of cloud apps and mobile devices which leads to an increase in the attack surface of cybercriminals. The growing of cloud computing and the internet of things have caused these perimeters to be eliminated. What can we say that the conventional model is no longer functional. With that model as much as we invest in cybersecurity of our company, new and more sophisticated attacks are launched against our defenses, so we must look at cybersecurity not as an investment but as a necessity over time. According [4] it is estimated that cybercrime targets increase considerably due to the existence of an ever-growing universe of connected people (figure 1).

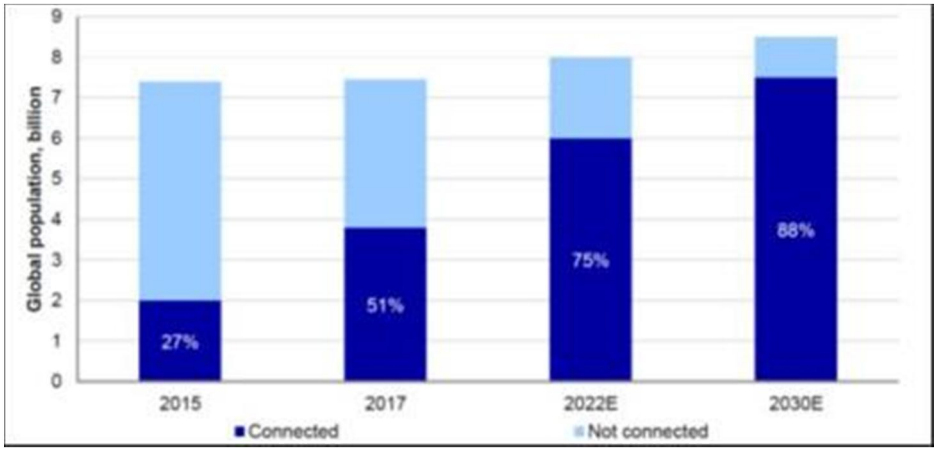

Fig.1. Growing of connected people [4]

2.1 Impact of Cybersecurity

With the increase in the number of devices connected to the Internet and consequently more attack area for cybercriminals, the monetary values involved in cybersecurity has been increasing [4].

Cybersecurity is about protecting information and cyber threat systems. Threats can be used as malware (malware, ransomware, phishing, worms) these weapons are increasingly sophisticated and automated and can be purchased at low cost.

With this, companies have a "punctual products" approach to combat these threats, which we will see later in this article that this is not viable because this model is expensive and complex.

The cybersecurity market is expected to reach 170 billion by 2020 [5].

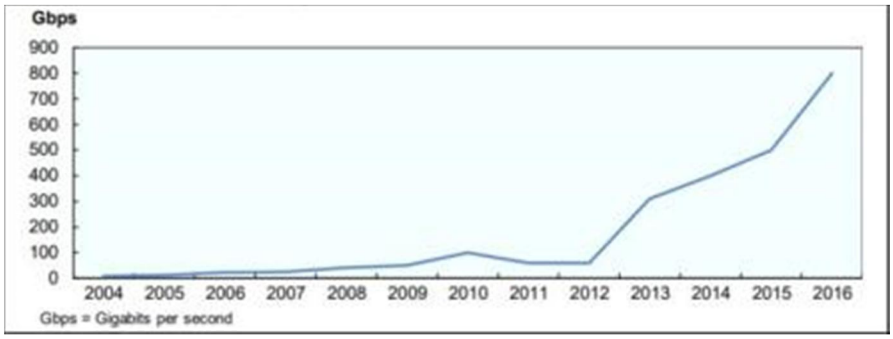

Fig.2. Evolution of bandwidth used for major DDoS [4]

3 Zero Trust Fundamentals

As the conventional model is not more functional, we cannot let the security of our organization be relied upon purely on a firewall or intrusion prevention system.

The Zero Trust is a security model developed by John Kindervag at Forrester research which has the principle of "never trust, always verify,", This architecture is designed to mitigate threats within the network-specific data or assets so that more granular rules can be applied.

According to John Kindervag [6], Zero Trust is not making networks, clouds, or endpoints more reliable; is to eliminate the concept of trust from digital systems, Trust is binary, it's on or off which is different from the real world interacting with people.

So Zero trust is all about how you think and there is no single formula for implementing this type of architecture. When building a network with zero trust DNA you need to keep in mind the following topics: [6]

- Ensure all data are securely accessed based on user and location
- The use of access control is strongly advised/required
- Inspect de log's of all traffic

This is important in a world where mobility is more dominant with tablets, smartphones, laptops, and IoT devices accessing the internet. These devices need to access these resources in a secure way.

3.1 Zero Trust Architecture

With this we can say that no longer exist a trusted interface in our devices, no longer exist trusted network, no longer exist trusted users and this is an important concept when we need to move packets from one place to another. If we look at the traditional model (figure 3) we have several layers of the network that we put several layers of security devices. This eventually making the network heavy, unmanageable, hard to keep safe and always invest in new devices over time.

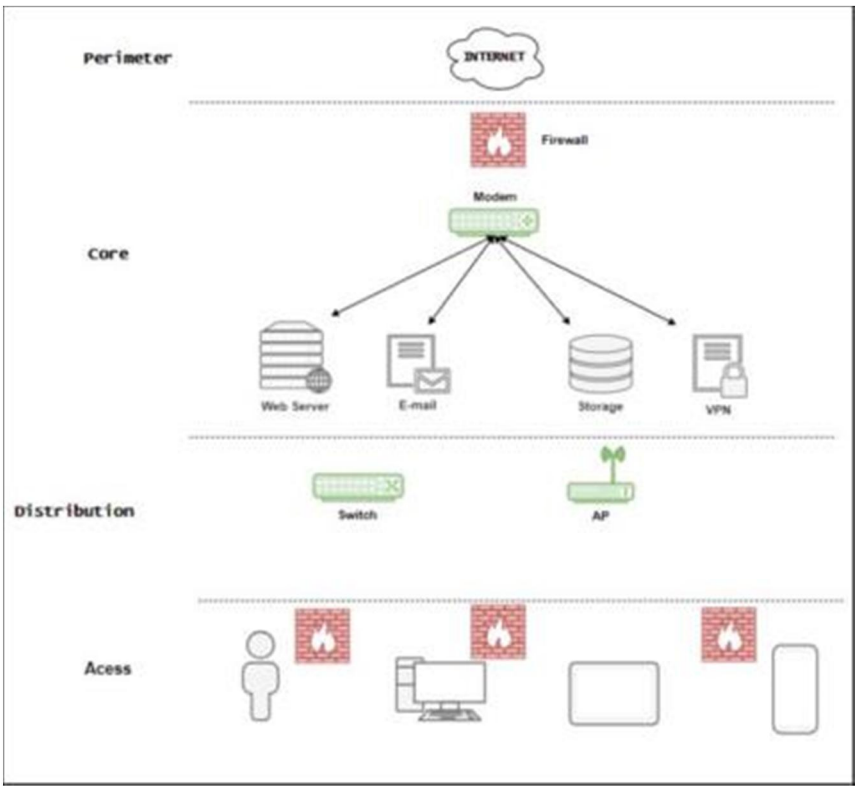

Fig.3. Traditional Architecture base on [6]

So zero trust model redraws the network and creates a new idea of segmentation gateway. This concept is built to concentrate all the resources that are used in a modern network like, content filtering, access control, firewall, cryptographic engines, package forwarding.

This type of segmentation is used in a modular way, is scalable and can adapt to any type of business, without having to restructure the entire network structure. But if a company has to structure the network, it has to do from the inside out, what will allow having a network that adapts and evolves with a security ADN where all packets can be delivered in a secure way.

This segmentation gateway model is considered by [7] as a next-generation firewall, since it is developed to increase the micro-segmentation of the networks, becoming quite versatile in terms of being scalable, adapting to all types of business and virtualization-friendly.

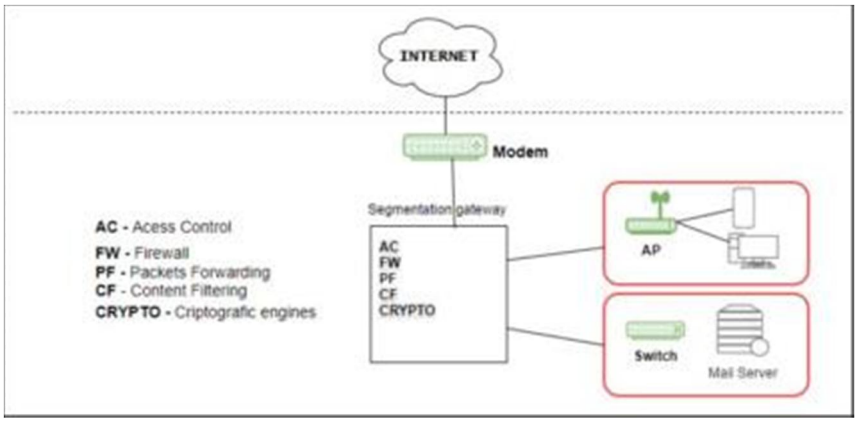

Fig.4. Zero Trust Architecture based on [6]

In the figure above it is represented a segmentation gateway in a basic way making the separation in micro-segmentation (MCAP) so that it is easier to inspect all the traffic of the network.

So, with the segmentation and next-generation firewalls, we can control who, what, where and when someone gets connected in the network. After a user is authenticated, the privileges must be tightly managed.

The thinking behind this is to prevent lateral movement inside the network after being compromised this reduces the range of damage caused. The term firewall, in this case, cannot be confused with the fact that we want to place near the perimeter of the network because the segmentation gateway must be placed in the center of the network.

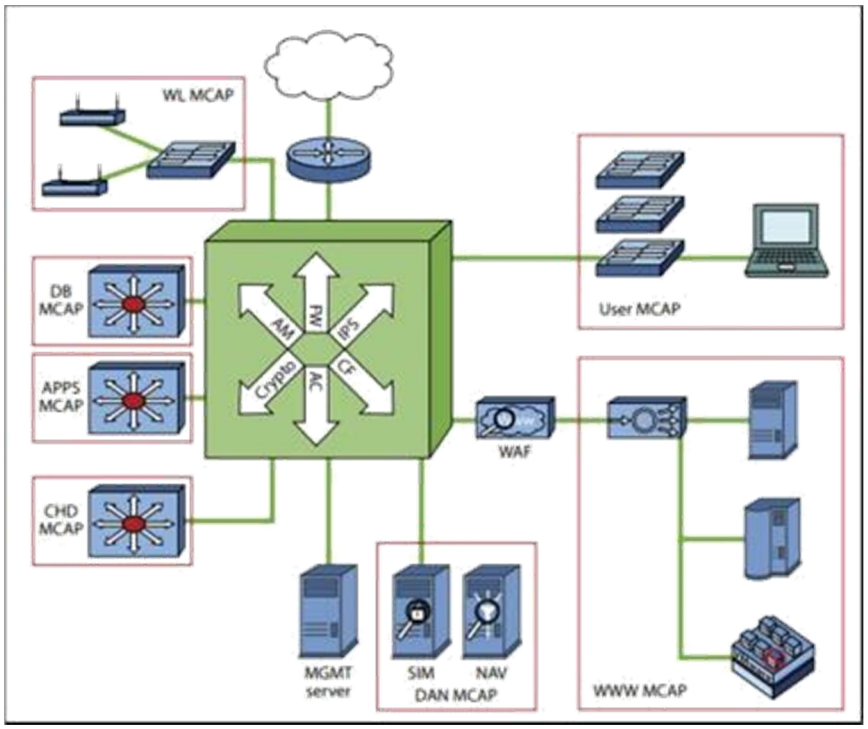

Fig.5. Extensible network architecture [6]

This is an example of a network that is scalable and can be augmented in any way we want or need.

3.2 Security benefits of Zero Trust

Zero trust is a philosophy that brings innumerable advantages in several levels to the company. So, I believe that the future of cybersecurity goes through this model.

According to [8] Zero trust delivers security and impressive business results. This model delivers a considerable business value like greater enterprise visibility protecting your customer data and business. In a “front-end” perspective Zero trust avoid financial costs in security audits, maintaining a good reputation towards other companies. In a “back-end” perspective we have reduced time to breach detection and get visibility into all your corporate traffic by inspecting the user request, devices, and data. Reducing the complexity of the security stack is a great support for the network maintenance team to deliver security and excellent end-user experience.

With this, we can be completely rested? Of course not. Some attacks against Zero Trust networks are well mitigated, while others we can only detect the attack. No model is perfect and 100% effective but we can reduce the impacts caused by any type of attack.

4 BeyondCorp by Google

BeyondCorp is a business security model of building zero trust networks in Google. By changing network perimeter access controls for individual devices and users, BeyondCorp allows employees to work more securely in any location without the need for a traditional VPN.

Basically, according to [9] Google BeyondCorp is a new model that excuses privileges in the corporate network, instead, access only depends on the credentials of the user and the device. whether it is in a home network, a hotel or a coffee shop. All access to enterprise resources is fully authenticated, fully authorized, and fully encrypted based upon device state and user credentials. BeyondCorp can enforce fine-grained access to different parts of enterprise resources

BeyondCorp consists of many cooperating components to ensure that only appropriately authenticated devices and users are authorized to access the requisite enterprise applications (figure 7).

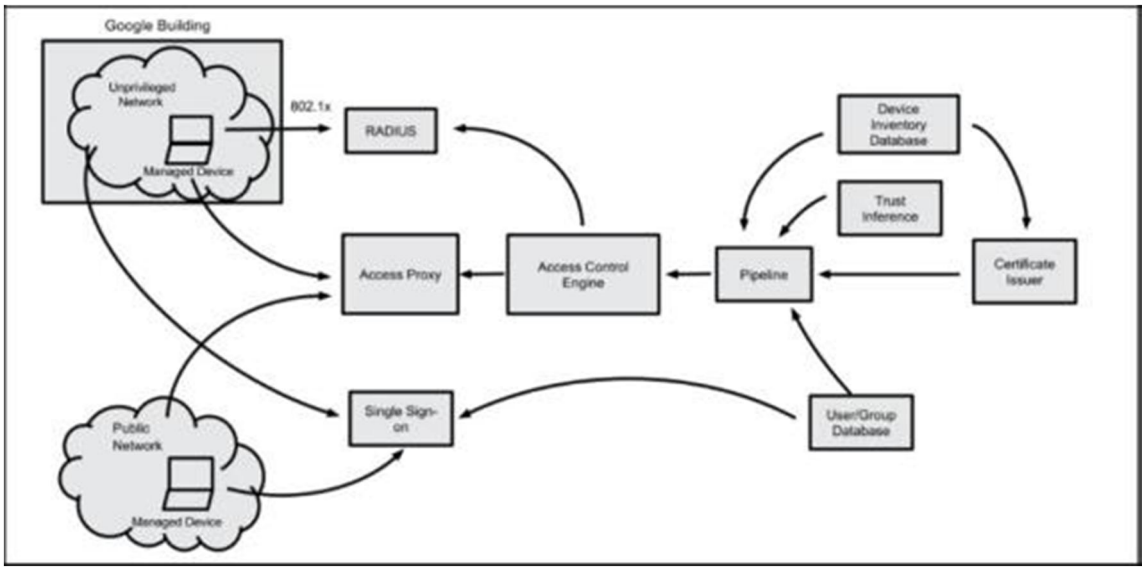

Fig.6. BeyondCorp components and Accessflow

In this model, all the devices normally used by the user are stored in a device inventory database in order to have better visibility of all the devices that are used in the network. All such information collected about devices and users is needed to understand what applications are used internally and what security policies should be applied in each application. It is necessary to understand the roles of the position and decide who has access to specific services in order to have effective access control.

With this, BeyondCorp brings many benefits to corporate network security by keeping the business consistent.

BeyondCorp benefits

- Keep devices up-to-date with the latest software
- Maintain an inventory of employee devices
- Monitor all endpoints & log all traffic
- Only communicate over fully encrypted channels
- Incorporate multi-factor AUTH
- Eliminate Static credentials

But not all models are 100% efficient there are also some difficulties that may appear in the implementation of this type of model. These difficulties may vary depending on the area of activity of the company itself. Depending on the area of the company, some vendors need network access inside the enterprise in order to maintain installed products or provide direct services, making access management difficult.

5 Conclusion

Cybersecurity is an area of emerging security that has been felt in recent decades as an exponential concern since the number of devices connected to the Internet is increasing dramatically, with almost 90% of the world population expected to be connected to the Internet by 2030.

And with this, traditional security models are becoming increasingly impractical due to the increase in the sophistication of the attacks and the elimination of the perimeters of computer networks.

With this scenario, there was a need to have another type of mentality and approach on data protection.

Zero trust arises from the need to simplify data security. In the absence of a simple implementation formula, this philosophy is based on "never trust, always verify". This is the beginning of the change of the mentality eliminating the trust of our network. Zero Trust segmentation platform is the basis of any Zero Trust initiative that allows us to break the network into micro-segmentation, giving us the ability to adapt our needs without restructuring our entire network.

Nowadays we have already seen several companies following this type of model because they bring many advantages. One of these cases is Google BeyondCorp.

BeyondCorp is the Google-designed business solution that allows users to work anywhere without VPN.

References

1. Gilman E., Barth D.: Zero Trust Networks, O'Reilly, (2017)
2. Sivaraman R.: "Zero Trust Security Model". S3tel Inc, White Paper (2015)
3. Williams C.: Zero Trust Security, Centrify Special Edition. John Wiley & Sons, Inc., Hoboken, New Jersey (2019)
4. Osorio de Barros G.: "A Economia da Cibersegurança", Gabinete de Estratégia e Estudos, Ministério da Economia(2018)
5. Morgan S.: "Cybersecurity Market Reaches $75 Billion In 2015; Expected To Reach $170 Billion By 2020", Forbes (2015)
6. Kindervag J.: Build Security Into Your Network's DNA: The Zero Trust Network Architecture, Forrester (2010)
7. Kindervag J.: Clarifying What Zero Trust Is and Is Not (2018)
8. Akamai: "The 6 Business and Security Benefits of Zero Trust." White Paper (2018)
9. Ward R., Beyer B,: "BeyondCorp A New Approach to Enterprise Security". Usenix, vol. 39:6 (2014)

Internet of Things: Privacy and Security Implications

Roberto Ferreira

Lusofona University of Porto, Portugal
rober@live.com.pt

Abstract. The Internet of Things or IoT has become one of the major concerns relative to security. IoT has a big impact on our lives. Certain aspects need to be reviewed in order to understand their features and their flaws. And knowing their weaknesses, we can prepare and develop better solutions that allow us to use these devices in a safe and responsible way. This paper gives an overview of what is Internet of Things and implications in security and privacy of the IoT which will be divided in Attacks and Threats, Security and Privacy in IoT and a Demonstration of Security breach in a CCTV System.

Keywords: Internet of Things, IoT, Security, Privacy.

1 Introduction

Internet of Things is becoming an increasingly growing topic. We can find IoT devices pretty much everywhere like home appliances, home energy management, home security and safety, health and fitness, information and entertainment. IoT devices like refrigerators, coffee machines, thermostats, smart bands, smart watches, intelligent ovens, smart bulbs etc. It has attracted strong interest from both academia and industry. But what is IoT after all? IoT are all devices that are capable to identify and communicate data between each other. These devices can collect a vast amount of information about the environment and how they are used, with or without the active involvement of the human being [1].

IoT is a network of physical objects. The Internet is not only a network of computers, but it has evolved into a network of devices of all type and sizes, vehicles, smart phones, homes appliances, toys, cameras, medical instruments and industrial systems, animals, people, buildings, all connected, all communicating and sharing information based on stipulated protocols in order to achieve smart reorganizations, positioning, tracing, safety, control and even personal real time online monitoring, upgrade, control and administration [2].

Enterprises and government agencies are embracing a digital transformation that is reinventing business models to better serve customers and drive new growth. The rapid adoption of new technologies and innovations is driving a global rethinking of traditional business processes and creating new ways to generate better business outcomes and quality-of-life improvements. Over a million new IoT devices are connected to the internet dally and that process is accelerating. Experts predict that between 25 and 50 billion new IP-enabled IoT devices will be deployed and online by 2020.

This is being called the Fourth Industrial Revolution-a period of explosive productivity improvements driven by innovation and the combination of technologies that unlock new business models [3].

This rapid innovation and adoption of this Internet of Things devices bring privacy and security challenges. Along in this paper we will get to know more IoT security and privacy problems and try to solve them.

2 Attacks and Threats

Internet of things devices are rapidly becoming ubiquitous while IoT services are becoming pervasive. Their success has not gone unnoticed and the number of threats and attacks against IoT devices and services are on the increase as well. Cyber-attacks are not new to IoT, but as IoT will be deeply interwoven in our lives and societies, it is becoming necessary to step up and take cyber defense seriously. With the necessity of secure IoT has resulted in a need to comprehensively understand the threats and attacks on IoT infrastructure [4].

According to [5], in the future, maybe the year 2020 with IPv6 and 5G network, millions of heterogeneous things will be the major factor of concern at that time. The IoT can be viewed in different dimensions by the different sections of academia and industry. Whatever the viewpoint, the IoT has no yet reached maturity and is vulnerable to all sort of threats and attacks.

Security issues are divided in three dimensions, based on phase, architecture and components. The IoT devices requires five phases, from data collection to data delivery to the end users.

Phase 1, Data collection, acquisition or perception, foremost step is to collect or acquire data from devices or things. Based on the characteristic of the thing, different types of data collectors are used. Thing may be static body (body sensors or RFID tags) or dynamic vehicle (sensors and chips).

Phase 2, Storage, the data collected in phase 1 should be stored. If the thing has its own local memory, data can be stored. Generally, IoT components are installed with low memory and low processing capabilities. The cloud takes over the responsibility for storing the data in the case of stateless devices.

Phase 3, Intelligent processing, the IoT analyzes the data stored in the Cloud Data Centers and provides intelligent services for work and life in hard real time. As well as analyzing and responding to queries, the IoT also controls things. There is no discrimination between a boot and a bot, the IoT offers intelligent processing and control services to all things equally.

Examples of such IoT systems are pervasive healthcare, advanced building management systems, smart city services, public surveillance and data acquisition, or participatory sensing application [8].

Phase 4, Data Transmission, the data transmission occurs in all phases, from sensors, RFID tags or chips to Data Centers, from Data Centers to processing units and from processors to controllers, devices or end users.

Phase 5, Delivery, delivers the processed data to things on time without errors or alteration, is a sensitive task that must always be carried out.

In all five phases occurs a variety of attacks. Data leakage, sovereignty, breach and authentication are the major concerns in the data perception phase. The image demonstrates the variety of attacks in every single phase.

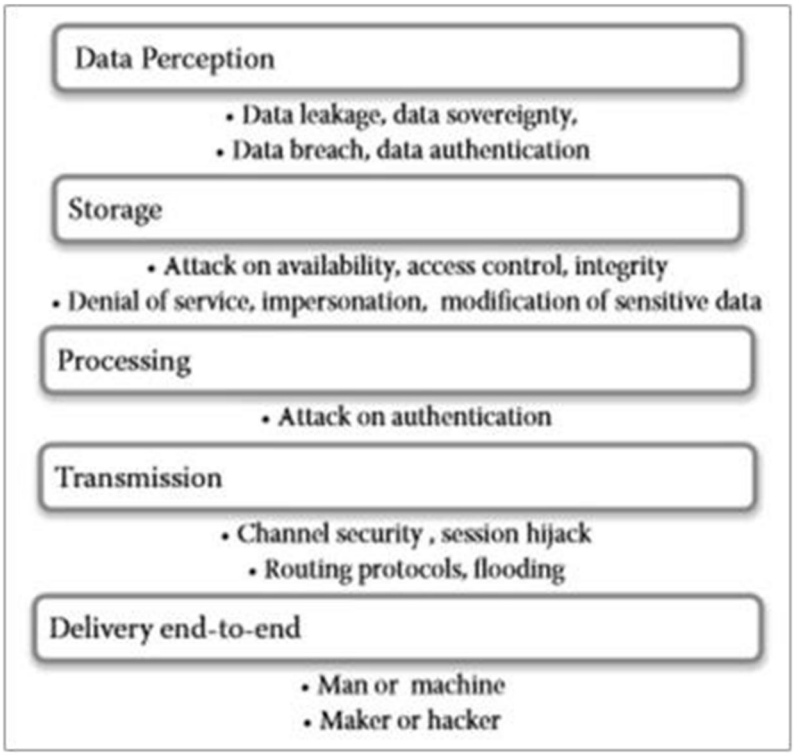

Fig.1. Attacks based on Phase

Relative to the architecture, the IoT has not yet been confined to a specific one. Different vendors and applications adopt their own layers. The attacks based on architecture, its assumed to have four layers. Sensing/Perception layer where External Attack, which attack, or worm hole and sewage pool are the most common to happen. The next layer, Network layer are routing protocol and address compromise. In the Transport Layer are Denial of Service and Man-in-the-Middle. And finally, the Application layer, which are more common attacks like, revealing sensitive data, data destruction, user authentication or intellectual property.

According to [6], the most common cyber attacks in IoT are Man-in-the-Middle, Data & Identity Theft and Denial of Service, between others.

The Man-in-the-Middle concept is where an attacker or hacker is looking to interrupt and breach communications between two separate systems. It can be a dangerous attack, because it is one where the attackers secretly intercept and transmits messages between two parties when they are under the belief that they are communicating directly with each other. As the attacker has the original message, they can trick the recipient into thinking they are still getting a legitimate message. Many cases have already been reported within this threat area, cases of hacked vehicles and hacked “smart refrigerators”.

In Data & Identity theft, careless safekeeping of Internet connected devices like, mobile phones, iPad´s, smartwatch´s are playing into the hands of malicious thieves and opportunistic finders. The more details that can be found about a user, by the IoT devices, the easier and the more sophisticated a target attack aimed at identity theft can be.

Denial of Service (DoS) attack happens when a service that would usually work is now unavailable. There can be many reasons for unavailability, but it usually refers to an infrastructure that cannot cope due to capacity overload. In DDoS, Distributed Denial of Service attacks, many systems attack one target. This often done through a botnet, where many devices are programmed (often unbeknownst to the owner) to request a service at the same time to a system.

The attacks based on components, by [5], the IoT connects “everything” through the internet. These things are heterogenous in nature communicating sensitive data over a distance. Apart from attenuation, theft, loss, breach and disaster, data can also be fabricated and modified by compromised sensors. Verification of the end user at the entry level is mandatory, distinguishing between humans and machines is extremely important.

3 Security and Privacy Preservation in IoT

For the use of the IoT to be safe and private, there must be a series of concepts that must be understood.

The Internet of Things is the interconnection of billions of smart things around us, with the ability to collect, store, process and communicate information about themselves and their physical environment. IoT systems will deliver advanced services of a whole new kind base on an increasingly fine-grained data acquisition in an environment densely populated with smart things. Privacy is a very broad and diverse notion for which literature offers many definitions and perspectives. From a historic view, the notion of privacy shifted between media, territorial, communication and bodily privacy to an increasing use and efficiency of electronic data processing information, privacy has become the predominant issue today. So according to [8], the definition of privacy of IoT is the guarantee of:

- Awareness of privacy risks imposed by smart things and services surrounding the data subject
- Individual control over the collection and processing of personal information by the surrounding smart things
- Awareness and control of subsequent use and dissemination of personal information by those entities to any entity outside the subjects personal control sphere

About security, based on the security issues discussed before there are necessary countermeasures so that authenticity, confidentiality, integrality, privacy and availability be

preserved. According to [7], some of the countermeasures are certification, access control, data encryption and security in cloud computing.

Relative to certification, it's a secure way of confirming the identity of both parties which communicate with each other. This can be achieved by using a Public Key Infrastructure or through a third party like a certificate authority that facilitates interactions between the users to assure the properties of data exchange [7].

Access Control is another mechanism which gives secure environment of IoT by limiting the access control for machines, objects or people which are illegal to access the resources. Access control can be implemented by using encrypted passwords, confidential directories or files, configurating and update rights etc.

Nowadays, there is a myriad of access control models that are applied to different Internet of Things scenarios in which security is required. The most popular models are Mandatory Access Control (MAC), Discretionary Access Control (DAC), RoleBased Access Control (RBAC) which are traditional access control models that do not consider additional parameters such as resources information and dynamic information (such as time, location). In order to provide a more flexible mechanism, the Attribute-Based Access (ABAC) was proposed, in which authorization decisions are based on attributes that the users must prove (eg: age, location, roles, etc.). One of the main advantages of ABAC is requests do not have to be known a priori by targets, providing a higher level of flexibility for open environments, compared to RBAC models. Nevertheless, in ABAC everyone must agree on a set of attributes and their meaning when using ABAC, which is not easy to accomplish. IoT scenarios imposes significant restrictions on privacy and access control, tradition access control approaches solution was not designed with these aspects [10].

Insufficient Authentication/Authorization regards to access control, can result in a data loss or corruption, lack of accountability or denial of access and can lead to a complete compromise of the device and/or user accounts. Attackers uses weak passwords, insecure password recovery mechanisms, poorly protected credential or lack of granular access control to access to an interface. Attack could come from Internal or external users. So, to make it secure according to [9]:

- Ensuring that the strong passwords are required
- Ensuring granular access control is in place when necessary
- Ensuring credential are properly protected
- Implement two factor authentications where possible
- Ensuring that password recovery mechanisms are secure
- Ensuring re-authorization is required for sensitive features
- Ensuring options are available for configuration password controls

Data Encryption technique is used to maintain confidentiality as well as integrity of the information. This way data that is intercepted by an attacker, encryption prevents that data from being deciphered [7].

According to [11], current cryptographic models and security schemes are based on widely adopted encryption algorithms, and privacy standards. Confidentiality is ensured in most of the cases with Advanced Encryption Standard (AES). The asymmetric algorithm RSA serves for asymmetric algorithm encryption, digital signatures, as well as for key management. SHA standards are used as secure hash functions. Alternatively, Diffie-Hellman (DH) and Elliptic Curve Cryptography (ECC) supplement the privacy schemes, basically in asymmetric cryptography. The applied suites have been designed for general purpose uses and their functionality is based on significant processing power, good memory resources and power availability. Since the applicability of these cryptographic models and security schemes is a bit unclear, detailed analysis is needed, in order to be ensured, that they can be implemented in the specified resources of IoT. Especially in the case, of minimizes capabilities of hand-held and portable devices [11].

Lack of encryption can result in data loss and depending on the data exposed, could lead to a complete compromise of the device or user accounts. There for according to [9], enough encryption requires:

- Ensuring data is encrypted using protocol such SSL and TLS while transiting networks
- Ensuring other industry standard encryption techniques are utilized to protected data during transport
- Ensuring only accepted encryption standards are used and avoid using proprietary encryption protocols

Relative to Cloud Computing, cloud is just a name for a huge data storage capacity with high performance and affordable low cost. In the essential, working with IoT large number of sensors nodes collects and analyses huge amount of data, storing and processing of data where cloud computing can be used very effectively. So, cloud computing security is very important to prevent attacks to that data [7]. So, according to [9], An insecure cloud interface could lead to compromise of user data and control over the device, enough cloud computing security requires:

- Default passwords and ideally default usernames to be changed during initial setup
- Ensuring user accounts cannot be enumerated using functionality such as password reset mechanisms
- Ensuring account lockout after 3-5 failed login attempts
- Ensuring the cloud-based web interface is not susceptible to XSS, SQLi or CSRF
- Ensuring credentials are not exposed over the internet

- Implement two factor authentications if possible

4 Demonstration of Security breach in a CCTV System

Security cameras are meant to monitor homes, public spaces, stores among other locations, however their main purpose is to monitor robberies or any suspicious behavior. The goal of security cameras is to protect those who use them and not the contrary. So, I tried to test the security of these CCTV systems, for academic reasons, that are connected to the internet and I did have access to a few. The one´s that I had access to, simply had default usernames and passwords, like username=admin and password=12345, like the Fig.2.

Fig.2. Login online to a CCTV system

But those that did not have the default credentials, it is possible to enter multiple times different combinations of users and passwords that can be obtained throw lists of millions of different records online. That is, I automatically can test various combinations of usernames and passwords until eventually find the correct one and get access to the CCTV system like Fig.3, through brute force cracking tools like Hydra. In this case it was easy to access to the CCTV system, because the user didn't change the default username and password. And this is more common than we think, especially in IoT devices that have weak security features. A solution to this situation would be the manufactors of the CCTV systems to force the users to change at least the password to a different and stronger one and have a maximum of login try's, like 6, until the device block and notice the user. Some companies have already done this in their devices, but others don't. Once I have access to de CCTV system, I can see what the cameras can see. But I'm not suppose to. The same camera system that is supposed to protect the user property, is a privacy and possible security breach. I can monitor when the owner arrives at the property and when he leaves, comprehending is routine and then take advantage of it by for example to steal the house. Use it as botnet to DDOs attack. Or even have access to the local network through the IoT device.

Fig.3. Access and control of the CCTV system

5 Conclusions

This paper has as main objective to understand what IoT is, how it works and what its vulnerabilities are regarding its security and privacy.

First, I have provided a basis for better understand how these devices are used in quotidian and how these have great potential in the future. Then I discuss the possible attacks according to their operation and definition of the most common attacks. Next, regarding to preservation of IoT security and privacy, I outline the main concepts that should be present in these devices so that they can be considered safe and a brief notion of privacy regarding to IoT devices.

In the demonstration of security breach, it was possible to demonstrate that is not hard to access to this IoT devices, because most of them don't offer any resistance at hall. Some IoT devices simply did not reach their maturity and therefore constitute a possible threat for those who use them.

IoT devices are capable of sensing and storing data from the environment around them and how they are used, which can jeopardize the safety of those using them or even be used as botnets in DDoS attacks.

So, the IoT devices are fantastic but, the safety levels are not adequate unless the companies responsible for their production have security and privacy in mind when developing them.

References

1. Keyur, k .P.; Sunil, M. P.; Internet of Things-IoT: Definition, Characteristics, Architecture, Enabling Technologies, Application & Future Challenges . vol. 6, Issue No.5, (May 2016).
2. CNCS,Homepage,https://www.cncs.gov.pt/a-internet-das-coisas-iot-internet-of-things/, last access 2018/12/21
3. Fortinet: Understanding the IoT explosion and its impact on enterprise security, (2017)
4. Mohamed A.; Geir M. K.; Cyber Security and the Internet of Things: Vulnerabilities, Threats, Intruders and Attackers, Norway, (2015)

5. Fei Hu.: Security and Privacy in Internet of Things (IoTs) Models, Algorithms, and Implementations, pp.27-39, Taylor & Francis Group, (2016)
6. GlobalSign, https://www.globalsign.com/en/blog/five-common-cyber-attacks-in-the-iot/, last access 2018/12/21
7. Mayuri, A. B.; Sudhir, T. B.; Internet of Things: Architecture, Security Issues and Countermeasures, vol.125, No.14, (September 2015).
8. Jan, H. Z.; Oscar G. M.; Klaus, W.; Privacy in the Internet of Things: Threats and Challenges, (28 May 2015)
9. OWASP, Internet of Things Top Ten, (2014)
10. Y. Andaloussi; M. D. El Ouadghiri; Y. Maurel; J. M. Bonnin; H.Chaoui; Access control in IoT environments: Feasible scenarios, (2018)
11. Nicolas, S.; I. D. Zaharakis; Cryptography and Security in Internet of Things (IoTs): Models, Schemes, and Implementations, (November 2016)

About Security in Internet of Things

José Adame

Lusofona University of Porto, Portugal
joseangelromeroadame@gmail.com

Abstract. In this work, we start with an introduction to know what the IoT (internet of things) is, what are the Some examples. Then its main problems in security and privacy are explained. As for the solution, it is determined according to the purpose for which it is used device and not the device itself, this is also explained in two final examples. And finally a small conclusion for the future about security and privacy in IoT.

Keywords: Internet of things, IoT, Security, Security Services, Privacy, Personal Data Protection.

1 Introduction

Nowadays, information architecture based on The Internet allows the exchange of goods and services among all elements, equipment and objects connected to the network. The IoT refers to the network interconnection of all objects everyday, who are often equipped with some kind of intelligence. In this context, The Internet can also be a platform for devices that communicate electronically and share specific information and data with the world around them. Thus, the IoT can be seen as a true evolution of what we know as the Internet adding more interconnectivity extensive, a better perception of information and more complete intelligent services. In its For the most part, the Internet was used for protocols aimed at connecting applications such as HTTP (Hypertext Transfer Protocol) and SMTP (Simple Mail Transfer Protocol). However, today a large number of intelligent devices communicate with each other and with other control systems. This concept is known as M2M (machine-to-machine communications).

IoT (Internet of Things) is an emerging architecture based on the Global Internet that facilitates the exchange of goods and services between networks of the chain of supply and that has a major impact on the safety and privacy of the actors involved [1].

Some highlights in the history of the IoT are the following:

- The term: Internet of Things was first used by Kevin Ashton in 1999 that I was working in the field of network RFID technology (identification by radiofrequency) and emerging detection technologies.

- However, the IoT "was born" sometime between 2008 and 2009 [2].

- In 2010, the number of everyday physical objects and devices connected to the Internet was around 12.5 billion. Currently there are about 25 billion devices connected to the IoT. More or less an intelligent device per person [2].

- It is expected that the number of smart devices or "things" connected to the IoT will be more than 50 billion in 2020.

IoT introduces a radical change in the quality of life of people, offering a great number of new opportunities for access to data, specific services in education, in safety, health care or transport, among others fields. On the other hand, it will be the key to increasing the productivity of companies, offering a wide distribution of the network, smart local networks of devices smart and new services that can be customized according to the needs of the client. The IoT brings benefits to improve the management and monitoring of assets and products, increases the amount of information data and allows the optimization of equipment and use of resources that can translate into cost savings. In addition, it offers the opportunity to create new intelligent interconnected devices and explore new models of deal.

Next we will see what makes up any "thing" exposing at the end of that point concrete examples as the case of Amazon echo that is so controversial nowadays. Finally I am making a step to the possible solution and specifying in it with a good level of details ending with two simple examples that make it easier to understand and visualize it as a summary.

And finally a brief and concise personal conclusion about all this.

2 Development and demonstration of results

IoT can be seen as a combination of sensors and actuators that are capable of provide and receive information digitizes and place it in bidirectional networks capable of transmit all the data to be used by a lot of different services and end users [3].

Multiple sensors can be attached to an object or device to measure a wide range of physical variables or phenomena and then transmit all the data to the cloud. Detection can be understood as a service model. Nowadays, the state-of-the-art devices, such as conventional items of households such as refrigerators or televisions, include communication and detection. These capabilities will be increasing with the incorporation of communication smarter and detection tools.

The architecture of IoT systems can be divided into four layers: detection layer objects, the data exchange layer, information integration layer, and the layer of application services [4].

Smart devices can already be connected through the traditional Internet. Without However, the IoT incorporates the detection layer that reduces the capacity requirements of those devices and allows their interconnection. Data consuming sensors communicate with sensors or owners thereof through the integration layer of information that is responsible for all communication and transactions. Meanwhile, new requirements and challenges for data exchange, filtering and integration of the information, the definition of new services for users, as well as an increase in

complexity of the network architecture. On the other hand, the use of cloud technologies is growing exponentially New infrastructure platforms and applications software are offered within the framework of the IoT. Some of the main advantages and benefits of the IoT will be the creation of innovative services with better performance and solutions added value, together with the reduction of data acquisition costs of services existing and the opportunity to create new sources of income in a context of a model of sustainable business These applications can be oriented to consumers, businesses, commercial activities, and survey activities, to the industrial and scientific community through use of application developers.

The number of applications and services they can provide is practically unlimited and it can adapt to many fields of human activity, facilitating and improving its quality of life in multiple forms.

As for specific devices, it is worth highlighting the case of Google home and Amazon echo, According to a study made by Mozilla these products (and others) can track the user and share your data with third parties for no apparent reason. The Mozilla report adds that, in the case of Google, its Smart speaker can listen to conversations but, as it is difficult to know what else you can do with that information. As for Amazon Echo, this device is in a position similar to that of the device Google. According to the document, the intelligent speaker of Amazon does not eliminate the stored data, it has a complex privacy policy and share data with other companies. In addition, the application is able to access to the user's camera and microphone and trace their location.

2.1 Security and Privacy

Security presents a significant challenge for IoT implementations due to the lack of a common standard and architecture for IoT security.

In heterogeneous networks, as in the case of the IoT, it is not easy to guarantee security and privacy of users. The main functionality of the IoT is based on the exchange of information among the billions or even billions of objects with an Internet connection. One of the security problems in the IoT that has not been considered in the standards is the distribution of the keys between devices.

On the other hand, privacy issues and profile access operations between the IoT devices without interference are extremely critical. Still, ensure the Exchange of data is necessary to avoid losing or compromising privacy. Increasing of the number of intelligent things that surround us with sensitive data requires management of transparent and easy access control so that, for example, a provider only can read the data, while another is allowed to control the device. In this meaning, some solutions have been proposed such as the grouping of devices integrated into virtual networks and only those devices desired within each virtual network. Another approach is to maintain an access control in the application layer in function of each seller.

In order to have a widespread adoption of any identification system for objects, there is a need to have a technically sound solution to guarantee the privacy and security of the clients. Although in many cases, security has been done As an added bonus, the feeling is that public acceptance of the IoT will occur only when Apply sound security and privacy solutions. In particular, the attacks must be intercepted, authenticated data, controlled access and customer privacy (natural and legal persons) guaranteed. It could be hybrid security mechanisms that

combine, for example, hardware security with key diversification for offer superior security that makes attacks significantly more difficult or even impossible. The selection of security features and mechanisms will continue being determined by the impact on business processes; and the counterparts will be between chip size, cost, functionality, interoperability, security and Privacy. Security and privacy issues should be addressed by future standards that must define the different security functions to provide services of confidentiality, integrity or availability.

There are also a number of issues related to the identity of people. These should be addressed in policy and legislation, being of vital importance for the efficient public administrations of the future.

As to how this situation must be faced [5], first you must understand that you can offer different levels of security depending on the service that the device needs. The Recommendation X.800 defines the following security services [6]:

- Authentication: to identify the communicating entity and the data source.
- Access control: to prevent the unauthorized use of resources.
- Confidentiality of the data: to protect them against unauthorized disclosure.
- Integrity of the data: to ensure that they have not been altered or destroyed an unauthorized way.
- Non-repudiation: to give proof of the origin of the data or delivery thereof.
- Availability: to guarantee the continuity of accessibility and use by the authorized entities.

These services are provided through security mechanisms alone or in combination, such as: encryption, digital signature, mechanisms for access control, data integrity mechanisms, authentication exchange, traffic filler, control of routing and notarization.

Low power consumption and light processing protocols are used to make the most of WSN resources. In this sense, you also have to prevent attacks whose objective is not to violate privacy but to overload the nodes and cause consumption extra [7].

The levels of security and privacy on the elements that must be protected depend on the imperatives imposed by the framework legal to which the final service is subject.

At European level, the Council of the European Union Convention n. 108, Strasbourg 28/1/1981, 5 ratifications 1/10/1985, established common data protection criteria for all members of the CE [8], coordinated by Directive 95/46 / CE of the European Parliament of 24/10/1995 [9].

In its report on the Data Protection Directive of 24/02/2004, the EU recognizes the legislative heterogeneity of its member countries and emphasizes the need for European states and institutions to adopt a equivalent level of protection of the rights of individuals [10]. Highlights

that this heterogeneity of national legislations on data protection hinders the development of the European internal market. As a result of the lines of action established by the European Parliament in the Communication of the European Economic and Social Committee [11], Europe is moving towards a common regulatory framework with the Proposal for a general regulation on data protection [12]. Once approved will be of direct application in two years for the entire European Union. This Regulation will affect those who process data of a personal nature and have an establishment in a member state, even if the information is processed performed outside the European Union. Companies not established in Europe will be affected if they process personal data for provide goods and services to residents of the EU.

Working Party Working Group 29 (WP29) has approved the first joint opinion on internet of the things, the Opinion 8/2014 on the new developments in the Internet of the things of 09/16/2014 [13], whose preparation has been led by the Agency Spanish Data Protection Agency (AEPD) together with the French authority Commission nationale de l'informatique et des libertés (CNIL).

The General Data Protection Regulation [14] will homogenize the legislation European Union on the protection of personal data, although in a global market It will still be necessary to live with the legislative heterogeneity.

In the technological field there are numerous works such as Dener, Fatema and Brad, Maw et al. , Kumari and Shukla, Shukla and Kumari, Malik, Kuthadi, Rajendra and Rajalakshmi, Karlof and Wagner and others [15-20], who have elaborated efficient security mechanisms for WSNs, and to avoid putting at risk the quality of the service for excessive consumption of resources. Companies and suppliers of equipment and networks They are very active in devising services for society and increasing their commercial catalog.

Among the proposals that exist, one gathers the knowledge about security and privacy generated for internet of things by the legal, technological and business areas, in a computer system able to channel collaboration between these areas. The purpose is to select automatic security and privacy policies that should be applied to new products and services. The collaboration between these three areas, would allow the issuance of certifications of trust for stakeholders and eliminate possible barriers of distrust. This proposal proposes the following collaborative environment:

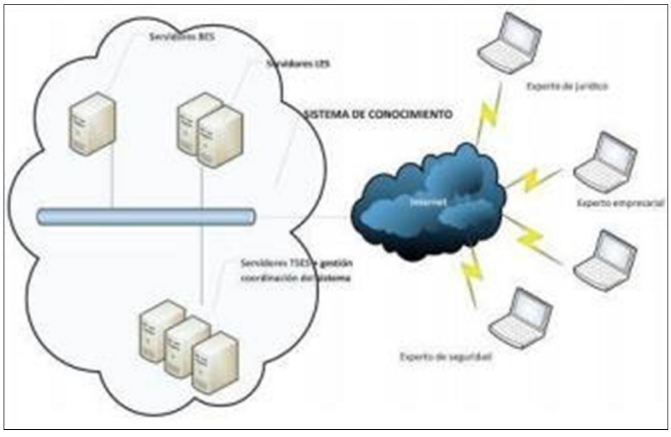

Fig.1. Collaborative environment among experts from the three knowledge areas BES (business expert system); LES (legal expert system); TSES (technological solutions expert system).

Companies dedicated to the design of new products and services can perform simulations virtual events prior to making decisions about the real market. In the legislative field, the impact on society and the market of possible modifications and new legislations in this matter, being able to know how and to what extent they would affect the products and services existing and their future developments. Technological areas can quickly know the critical aspects that need new research and innovations. To carry out complete tests, a model has been designed to test the possibilities of automation from the decision of the security policy to the configuration of the operation of the elements of the WSN. These tests have been carried out in the GRyS team (Group of Next Generation Networks and Services) of the Citsem (Center for Research in Software Technologies and Multi- media Systems for Sustainability) of the UPM (University Polytechnic of Madrid).

The design and test model (figure 2) consists of a WSN, a middleware platform oriented to services, Aware Project [21], and PDPS-IOT expert system (personal data protection system - internet of things) [22].

The PDPS-IOT expert system decides the security and privacy policy to be applied to the service (object of this article), which is communicated to the Aware platform, which knows the WSN to which it connects (its technology, the way in which it must dialogue with them, its possibilities, the security mechanisms they support) managing their configuration possibilities. Generate the commands and actions to configure them remotely using their middleware. PDPS-IOT knows the set of security mechanisms that Aware is able to manage, with that a homogeneous level of security can be established for all WSN networks.

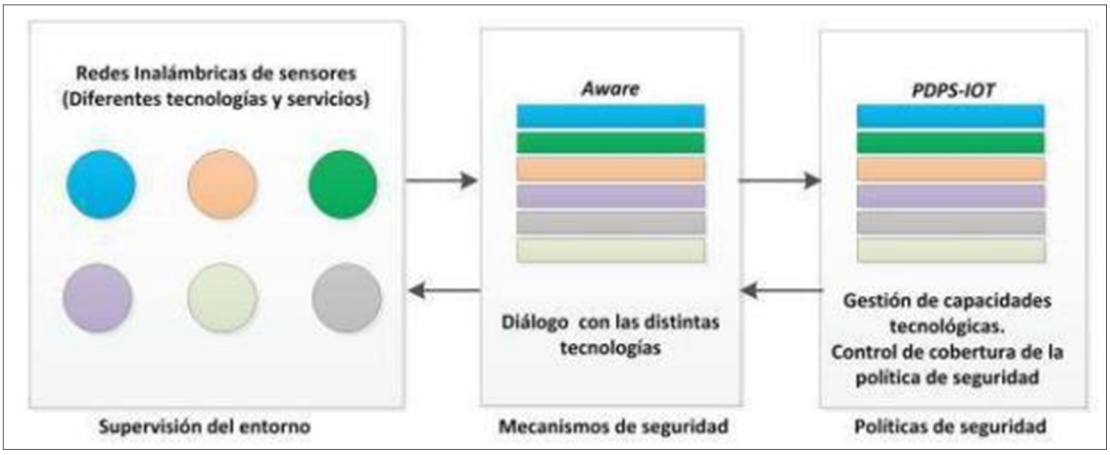

Fig.2. The design model presented.

The final service used for testing monitors health by collecting pulse data and body temperature. These data are processed and sent to the base station to inform a medical equipment. This final service involves health data of people, protected in Spain by Royal Decree 1720/2007, of 21/12/2007, on the protection of personal data, and in Europe by general regulation of data protection. Through the PDPS-IOT expert system (figure 3), which has the service specifications as input, the policy is obtained as output of security and privacy (security level) that the service should have.

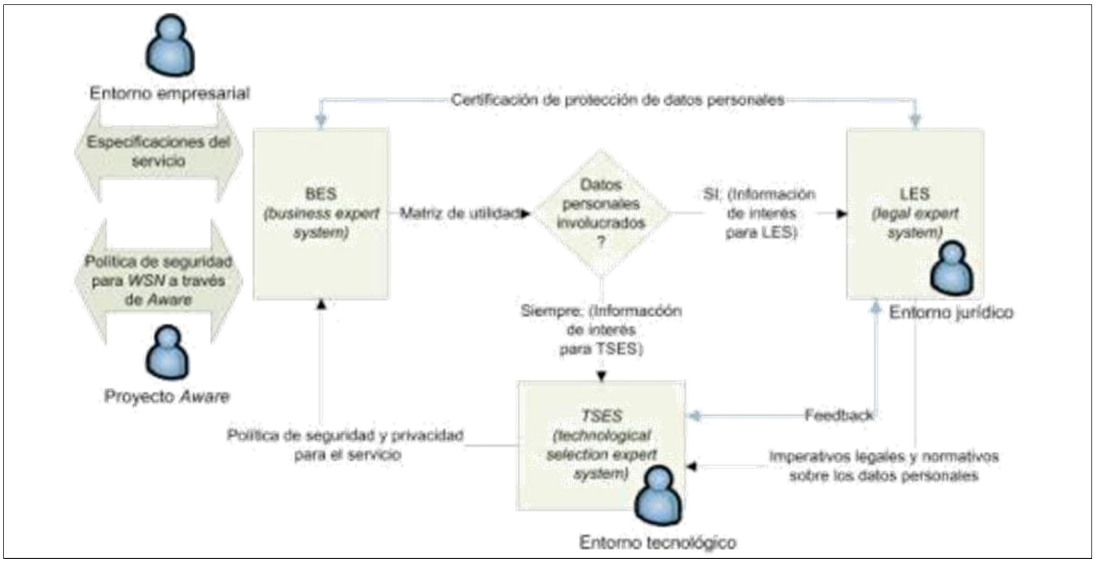

Fig.3. Personal data protection system for internet of things services (PSPS-IOT). BES (business expert system); LES (legal expert system); TSES (technological solutions expert system).

This is possible thanks to the fact that the relevant information has been formalized in what we call the matrix of utility, from which the legislative framework that affects the service is obtained. What you are looking for is to obtain the security and privacy imperatives that must act on the elements of information that should be protected. These imperatives will be transformed into a set of services and security and privacy mechanisms. In the event of a change in the use or reuse of a WSN to provide another service, or the same service in another area, after communicating it to the managers and duly update the utility matrix generates the process that culminates with the Remote reconfiguration of security in the WSN, if necessary.

As an example of the different behaviors to be had, two examples of cases of uses in which the IoT is the same, a cardiac belt:

In the case of a farm, own service to control the health of livestock. Does not involve data personal The information collected is not useful for anyone else. It is not a critical service (it does not replace food controls) and has no special continuity requirements. Access to the system does by user and password without additional mechanism.

In the case of a football team, own service to control the health of the players. There are data personal (health) protected by law before disclosure / disclosure (hide everything, or separate identity of the person of his measurements). This data must be managed and managed only by staff authorized: veracity of the actors and authorization of access. Content is required, according to the law, personal data must be true (we assume that the sensors are calibrated). It is a service that is not critical and without special continuity requirements.

3 Conclusion

In short, everything depends on the control of the bodies that regulate and can create new laws for the security and privacy in IoT and that these measures evolve at the same time as these devices and do not are left behind, this way we can have that confidence in the products that are certified by these entities (in principle) and little by little in the future will go through much safer terrain. All this does not mean that there are always some non-certified or simply new devices that have certain innovative data treatment that endangers privacy and security as it is impossible guarantee a perfect security and that there are not things that always escape us but ultimately for a standard user of these devices (like me) is worth using these devices even with that possible danger of security and privacy.

References

1. R. H. Weber, (2010). "Internet of Things - New Security and Privacy Challenges". Computer Law & Security Review 26: 23-30. Consulted in 2018.
2. Dave Evans. (2011). How the Next Evolution of the Internet Is Changing Everything. Cisco Internet of Things White Paper. Consulted in 2018.
3. Charith Perera et. al. (2014). Sensing as a Service Model for Smart Cities Supported by Internet of Things. Transactions on Emerging Telecommunications Technology 25 (1): 81– 93. Consulted in 2018.
4. Ma HD. (2011). “Internet of things: Objectives and scientific challenges”. Journal of computer science and technology 26 (6): 919-924. Consulted in 2018.
5. Al-Ameen, Moshaddique; Liu, Jingwei; Kwak, Kyungsup (2012). “Security and privacy issues in wireless sensor net- works for healthcare applications”. Journal of medical systems, v. 36, n. 1, pp. 93-101. Consulted in 2018.
6. Unión Internacional de Telecomunicaciones (1991). Recomendación X.800. Arquitectura de seguridad de la interconexión de sistemas abiertos para aplicaciones del CITT. Consulted in 2018.
7. Comisión Europea (2012). Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos (Reglamento general de protección de datos), n. 2012/0011/ COD, 25/01/2012. Consulted in 2018.
8. Council of Europe Treaty Office (1981). Convention for the protection of individuals with regard to automatic processing of personal data, n. CETS 108, Strasbourg, 28/1/1981, pp. 110. Consulted in 2018.
9. Unión Europea (1995). “Directiva 95/46/CE del Parlamento Europeo y del Consejo, 24 de octubre de 1995, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos”. Diario oficial, n. L 282 de 23/11/1995, pp. 0031-0050. Consulted in 2018.
10. Parlamento Europeo (2004). “Resolución del Parlamento Europeo sobre el primer informe sobre la aplicación de la Directiva sobre protección de datos (95/46/CE)”. Diario oficial n. C 102E de 28.4.2004, pp. 147-153. Consulted in 2018.
11. Comité Económico y Social Europeo (2009). “Dictamen sobre la Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones - Internet de los objetos - Un plan de acción para Europa [COM(2009) 278 final]”. Diario oficial, n. C 255, 22/09/2010, pp. 116-120. Consulted in 2018.
12. Comisión Europea (2012). Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos (Reglamento general de protección de datos), n. 2012/0011/ COD, 25/01/2012. Consulted in 2018.
13. European Commission (2014). “Opinion 8/2014 on the on recent developments on the internet of things”, n. 14/EN WP 223, adopted on 16 September 2014, pp. 1-24. Consulted in 2018.
14. Palafox-Maestre, Luis E.; García-Macías, J. Antonio (2008). “Security in wireless sensor networks”. En: Yan Zhang; Miao, Ma. Handbook of research on wireless security. Hershey, PA: IGI Global, pp. 547-564. ISBN: 978 1 59 904899 4 . Consulted in 2018.
15. Dener, Murat (2014). “Security analysis in wireless sensor networks”. International journal of distributed sensor networks, v. 2014, pp. 1-9. Consulted in 2018.
16. Fatema, Nusrat; Brad, Remus (2014). “Attacks and counterattacks on wireless sensor networks”. International journal of ad hoc, sensor and ubiquitous computing, v. 4, n. 6, pp. 1-15. Consulted in 2018.

17. Karlof, Chris; Wagner, David (2003). “Secure routing in wireless sensor networks: Attacks and countermeasures”. Ad hoc networks, v. 1, n. 2-3, pp. 293-315. Consulted in 2018.
18. Kumari, Babli; Shukla, Jyoti (2013) “Secure routing in wireless sensor networks”. Ijarcsse, v. 3, n. 8, pp. 746-751. Consulted in 2018.
19. Kuthadi, Venu-Madhav; Rajendra, C; Rajalakshmi, Selvaraj (2010). “A study of security challenges in wireless sensor networks”. Journal of theoretical and applied information technology, v. 20, n. 1, pp. 39-44. Consulted in 2018.
20. Malik, M. Yasir (2012). “An outline of security in wireless sensor networks: Threats, countermeasures and implementations”. En: Zaman, Noor; Ragab, Khaled; Abdullah, Azween. Wireless sensor networks and energy efficiency: protocols, routing and management. Hershey, PA: IGI Global, pp. 507-527. ISBN: 978 146 660102 4.Consulted in 2018.
21. Santos-Familiar, Miguel; Martínez-Ortega, José-Fernán; López, Lourdes (2012). “Pervasive smart spaces and envi- ronments: A service-oriented middleware architecture for wireless ad hoc and sensor networks”. Architecture for wire- less ad hoc and sensor networks, v. 2012, pp. 1-11. Consulted in 2018.
22. Sánchez-Alcón, José-Antonio; López, Lourdes; Martínez-Orte- ga, José-Fernán; Castillejo, Pedro (2013). “Automated deter mination of security services to ensure personal data protection in the internet of things applications”. En: 3rd Intl conf on innovative computing technology (Intech), August, pp. 71-76. Consulted in 2018.

SESSION 2
PREVENTIVE MEASURES IN THE ERA OF DIGITAL

A Review on Cyber Attacks and Its Preventive Measures
Valdemar Sousa

A Review on Cyber Attacks and Its Preventive Measures
Pedro Teixeira

Malicious URL Detection using Machine Learning Algorithms
Marcelo Ferreira

Address Resolution Protocol (ARP) Spoofing: Attacks and Defenses
Bruno Duarte

Creating GDPR Compliant Interpretable Models
Pedro Strecht

A Data Encryption Application: Development Proposal
Diogo Vilas Boas

A Review on Cyber Attacks and Its Preventive Measures

Valdemar Sousa

Lusofona University of Porto, Portugal
valde_sousa@hotmail.com

Abstract. Today and on the current Internet, cyberattacks are a headache for all users of the World-wide network, businesses and digital security experts fight every day and try to find solutions so that data or information is not affected for these attacks. With this there is no "medicine" that solves all cyberattacks, but preventive measures that try to avoid major damage to those who suffer an attack. This paper focuses on summarizing a bit of what are cyberattacks and realizing better than they are capable, and also preventive measures that can ensure greater security for Internet users.

Keywords: CyberAttacks, Preventive, Measures, Internet, Crimeware, CyberPrevention, CyberDetection

1 Introduction

We are currently witnessing great transformations in our society, more specifically at the technological level. In the world where we live the Internet has a fundamental role in creating a new space, where we do not need physical presence to commit something illicit.

All data stored on a computer or network represents a three-dimensional model in which the virtual user can move. [1]

In this paper, we will address one of the events that in recent years has taken over the Internet, cyberattacks.

In fact, the network society promotes that there is a development and innovation with regard to the Internet, but it is also the promoter of various dangers, including cybercrime.

There are several sectors that constantly suffer from this type of tendency, including governmental institutions, large companies and also security forces, because in these areas there is information that would be very advantageous for the Could get caught.

Cyberattack is an attempt to damage or disrupt a computer system, or obtain information stored in a system through hacking. [2]

In this paper we will address a little more thoroughly this trend of cyberattack, showing which cyberattacks most used, also some preventive measures where we will demonstrate and try to dispose possible solutions in which the common user this great and complex network can "defend".

2 Cybercrime and Cyber Attack

2.1 Cybercrime

The internet is not an impartial field, it is a space in which viruses, worms, crackers in which they pierce firewalls and access confidential data in which this information can be worth many millions in monetary terms. Computer security specialists work hard every day to keep our privacy intact and to increase the difficulty of a cyberattack on our machines.

Cybercrime is the name given to cybercrimes that encompass any activity or practices in a network, are computer crimes practiced through computers, against them or through them. In another perspective, cybercrimes can be assessed as a conduct of unauthorized access to computer systems, destructive actions, data modification, communication interception, child pornography dissemination, copyright infringement, Terrorism, among others [3].

The term cybercrime had its first appearance in a G-8 meeting (group composed of the 7 richest countries in the world and Russia) close to the end of the years 90. This meeting had the purpose of addressing the methods used to combat the Elicits practices that occurred on the Internet.

The transactional predominance is one of the strong characteristics of cybercrime, which hinders investigations and the clearance of evidence against the accused. Another feature of cybercrime is the increase of personal computers, which facilitates anyone in the world to perform illicit practices, anywhere on the planet and without leaving home.

The practice of cybercrime is so common, that, according to Norton, a company specializing in digital security, about 65% of the cybercriminals have been victims of some kind of cybercrime, where the greatest difficulty in fighting, is faced with the lack of efficient laws and punishments In several countries in the fight against criminals.

In a summarized way, cybercrime uses technology as a way to divert information online, whether it's a computer or smartphone. The development of this theme, was not an instantaneous act, was having evolutionary changes [4].

Some curiosities about cybercrimes:

- Cybercrime has already surpassed illegal drug trafficking as a criminal money maker.
- Every 3 seconds, an identity is stolen.
- Without sophisticated security measures, a PC can be infected within 4 minutes after connecting to the Internet.

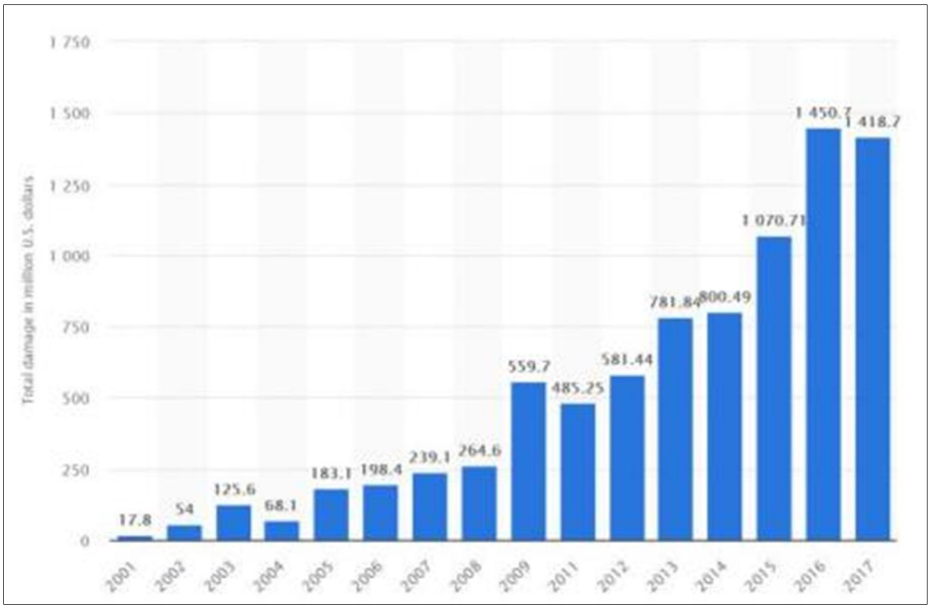

Fig.1. Amount of monetary damage caused by reported cybercrime to the IC3 from 2001 to 2017 (in million U.S. dollars) [5]

2.2 Cyberattack

Cyberattacks are an increasingly common reality in the current Internet, it is a reality that does not yet have a global and effective response in order to combat and avoid them. There are several terminologies used to designate a crime that is practiced by a computer connected to the Internet, such as cybercrimes, digital, computer, computer fraud, among others.

Cyberattack is the deliberate exploration of computer systems and technology-dependent networks. In other words, it is an unlawful, non-ethical or unauthorized conduct involving the automatic processing of data or transmission of data [6], is one that is performed by people with the objective of obtaining information or benefits using technological tools connected to the Internet and disturbing the functioning or structure of those being attacked.

Cyberattacks use malicious code to change your computer's coding, logic, or data, resulting in consequences that can compromise your data and lead to cybercrime, such as information and identity theft, a cyberattack is also Known as an attack from a computer network.

Cyberattacks have numerous consequences:

- Identity theft, fraud and extortion
- Malware, pharming, phishing, spamming, spoofing, spyware, trojans e virus
- System infiltration
- Unauthorized access

- Password sniffing
- Abuse of Instant Messaging
- Among others

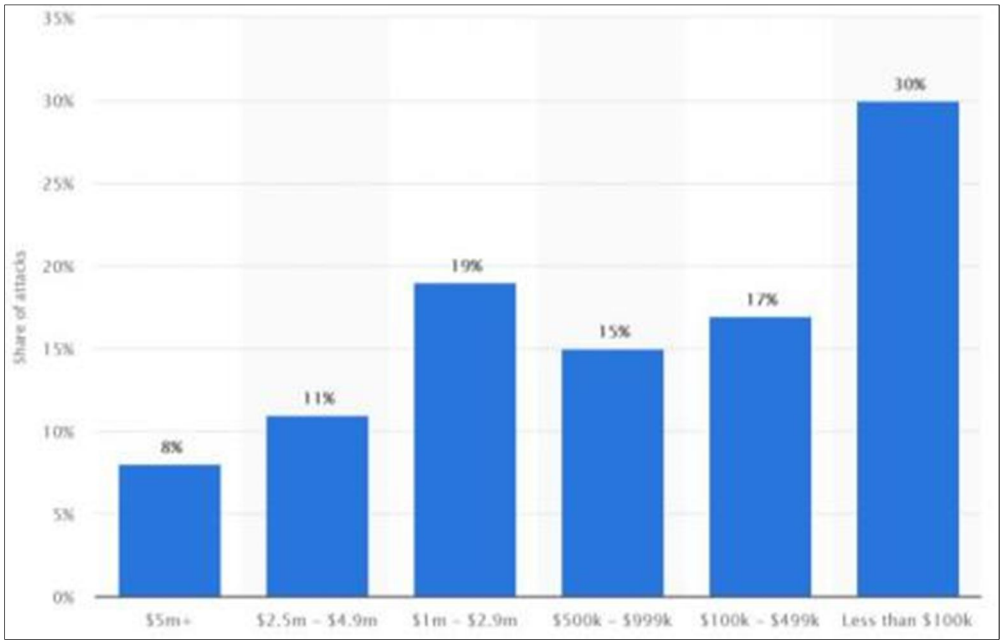

Fig.2. Average financial damages of cyberattacks caused to businesses worldwide as of April 2018 (in U.S. dollars) [7]

According to the "Convention on Cybercrime" adopted by the Council of Europe in 2001, it can be highlighted as cybercrimes:

- Infringements against the Confidencialidade, integrity and availability of data and computer systems
    - Illegal access to a computer system;
    - Unlawful intercetation of data or telematic communications;
    - Attack on data integrity (own conduct of a hacker subgroup, known as a cracker);
    - Attack on the integrity of a system;
    - Production, marketing, obtaining or possession of applications or access codes that allow the practice of the above-mentioned crimes.
- Computer infractions

- Falsification of data

- Infractions relating to the content:
  - Child pornography
  - Racism and xenophobia

- The attack on intellectual property and the rights associated with it
  - Public display of movie without permission of the Rightsholder

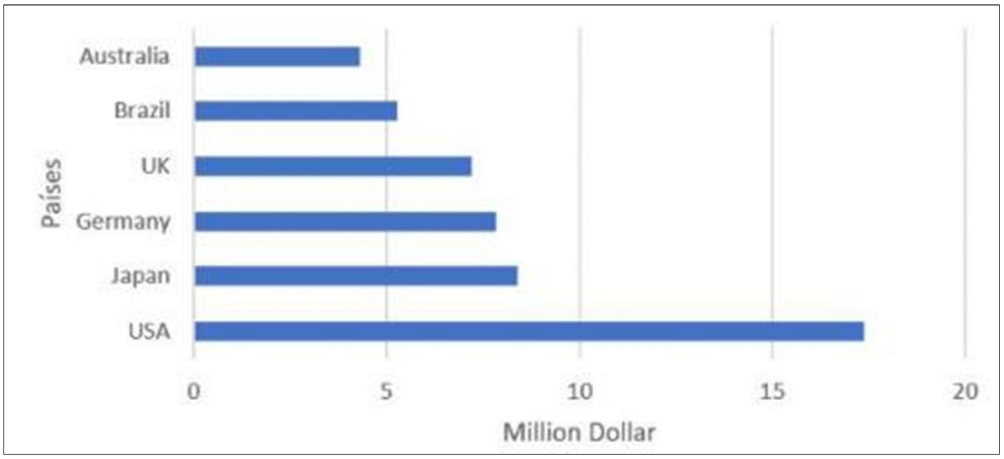

Fig.3. List of countries which have the average cost of cybercrime in the world

2.2.1 The Most Common Types of Cyberattack

With the frequency of use and sophistication of the Internet, service providers have as one of their main missions control denial of service (DoS) attacks and distributed denial of service (distributed denial) attacks of Service-DDoS), which are increasing increasingly. The internet is part of a worldwide infrastructure, which has a unique feature, that of not having boundaries that defend it from the attacks. Some of these attacks are quite harmful and can cause ample blackouts on the internet.

The DOS/DDoS attack is an attempt to make an overload happen on a device or server, so that the resources are unavailable to users, in other words, it is an attack that aims to disconnect from the Internet an equipment that is connected to it. For the attack to be performed, the attacker uses a range of techniques, using software, sending several packages to the target, for the purpose that is overloaded and unable to respond to any package requests, so put, users do not Access to the data to which the server was attacked.

The target can be a server, a router, or other equipment at times with the physical destruction of the hardware that was attacked.

DoS, it is the attack in which a computer with an Internet connection is used to flood a server with packets, the purpose of this attack is denial of service and the bandwidth overhead.

DDoS, it's practically similar to a DoS attack, but with very different results, because instead of using a computer and a connection, they use multiple computers and so many other connections. The computers behind this type of attack are part of the botnet. The main difference between DOS and DDoS is the target server, with DDoS, the server will be overloaded by thousands of requests when compared to DOS that the overhead is significantly lower.

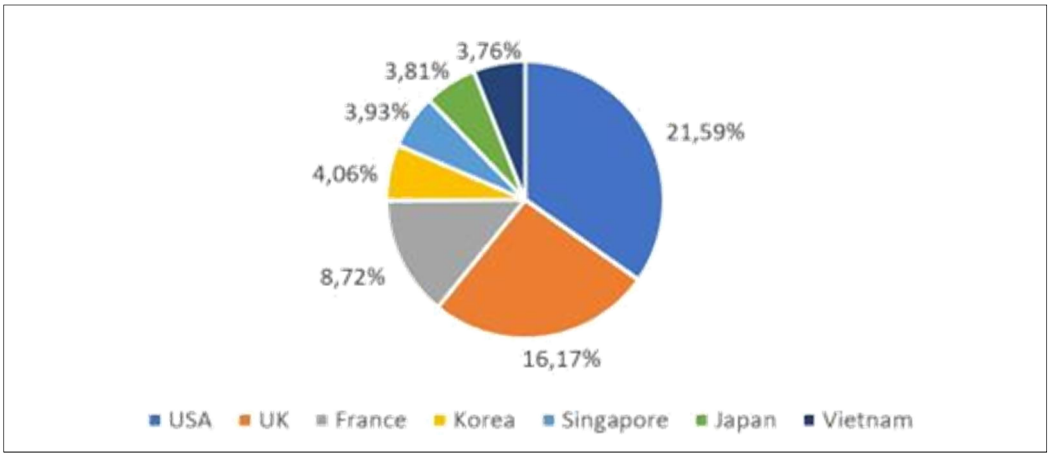

Fig.4. The list of countries from which highest percentage of Global Denial of Service Attacks (DDoS)

2.3 Tools Used in Cyberattack

In the current Internet, hackers help themselves in software to commit their actions, these help themselves in software used to practice cybercrime, called Crimeware. Crimeware is a software that is used to commit a criminal act, and is generally considered with a non-desirable software.

The software tools used in cybercrime are sometimes referred to as crimeware. Like cybercrime itself, the crimeware termination achieves a wide range of different malwares, or potentially malicious software. Crimeware is a class of malware designed specifically to automate cybercrime [8].

David Jevans created the term "Crimeware" in February 2005 in a response from the Anti-Phishing Working Group to the FDIC article "ending the identity theft of account Invasions", published on December 14, 2004.

2.3.1 Crimeware: Bots

The term "bot" is approximated by robot, is one of the most sophisticated software types of crimeware that the internet currently faces.

Bots are like Trojans and worms, perform a variety of automated tasks on behalf of hackers, who are safely located. Bots can perform various tasks, from sending spam to the detonation of Internet sites.

Bots can invade computers in many ways, sometimes they are scattered over the internet, looking for vulnerable and unprotected computers to infect [9]. When you reach your goal and find an exposed computer, quickly infect the machine found and report to your "master". Your main goal is to stay hidden until you receive some kind of task transmitted by your master.

The service provider informs you that your computer is sending spam to other Internet users [10]. Bots are part of a network of infected computers called "botnet", these networks are created by attackers repeatedly infecting victims ' computers.

2.3.2 Crimeware: Trojans and Spyware

One of the most popular methods among hackers is Trojan horses and spywares.

Trojan Horse Software presents itself to the computer as a useful program, while in fact wreaking havoc and damage. Trojan horses are the first part of an attack, and your goal is to remain hidden during a download. These can not be scattered alone, such as a virus or worm [11].

Spyware is the term used for programs that secretly observe and monitor the activity of a computer, storing information that may be sold to entities that benefit from this information. Spyware can be installed in several ways, most of the time it is improperly installed and bundled with other software's that install purposely [12].

3 The Concept of Cyber Prevention and Detection

The internet is a large network where there are multiple computers interconnected by other smaller networks, and they communicate with each other. The computers communicate with each other through an IP address, where a series of information is exchanged and there arises the biggest problem, because there is a large amount of personal information that crosses this network, being the disposition of thousands of people who Have access to the Internet, and when they are not available by the user themselves, are sought by other users and then committing the so-called cybercrimes [13].

Cybernetic Prevention is the act of restricting, controlling, removing or preventing the occurrence of cyberattacks in computer systems. On the other hand, Cyber Prevention is responsible for detecting irregularities in the activities of the Internet user.

In current times, cyberattacks can capture a lot of information and even keep it out of the control of its owners, to protect this information it is very important to prevent and act against possible cyberattacks that may arise [14].

A good security engineer always has to in mind that "security is not a product, but rather a process", more than designing a system that solves some network gap, a strong encryption in a system, is to be designing a model in which all strategies and Med Safety, so that systems are more and more protected from threats.

On a computer or IT system, the security status goes through three processes: threats prevention, detection and response [15].

These three processes are based on various system policies and components, such as access and cryptographic controls, firewalls that play a very important role in preventing systems, intrusion detection systems (IDSs), and the response of Systems to the attacks.

Access and encryption controls can protect data from systems.

The firewall is by far the most common prevention system, with regard to network security, since, properly configured, provides essential safeguards and protects access to internal network services and to catch certain types of attacks, due to its packet filtering capability.

"Answer" is necessarily defined by the security requirements assessed in an individual system and can cover from simple update protections to notification of legal authorities, counterattack.

3.1 Prevention Tips

Keep your computer current with the latest patches and updates. The updates provided are useful for the computer to be protected and prevents attackers from leveraging software failures (vulnerabilities) that they could use to invade the system [16]. Using the "Automatic Updates" feature is a great start to keep seguro on-line.

- Make sure your computer is configured securely

A configuration Efficient Of the most popular applications on the Internet, such as the browser or email software, is one of the most important areas, and where the utmost attention to detail can help prevent an attack.

- Choose strong passwords and keep them safe

The passwords are a fact of life on the Internet, are for our information what keys are to the door of our house, they protect our property or our data, for the rest of the world. In the current generation of internet, "discovering" the password of another is a crime, because we can pass by someone else or acquire private data. As such, act with some seriousness at the time of electing one, as well as not sharing. Keep the password private and a way to ensure its effectiveness.

- Protect your computer with security software

Protecting your Pc with specific security software is essential to a basic online security, these software's include firewall and antivirus programs that can assist in the protection of possible cyberattacks. There are several software packages that encompass numerous combinations of protection, such as firewall, antivirus, anti-spyware Among other features such as antispam and parental controls that assist in the sudden counterattacks protection.

- Protect your personal information

On the Internet, to take advantage of many services, we inevitably provide personal information. As the disclosure of personal information is rarely possible, the following list contains some advice on how to share personal information securely online:

- Keep an eye on fake email messages.

- Do not respond to email messages requesting personal information.
- Avoid Fraudulent websites used to steal personal information.
- Attention to privity policies on websites and software.
- Save Email address, and not disclose unnecessarily.

4 Example of Cyberattack and Preventive Measures

4.1 Man on the middle

Regardless of the type of business, the goal of hackers is to steal information from Internet users, whether through individual or large-scale attacks. Almost always, offenders start by trying to introduce some kind of virus into the target computer, if for some reason it goes wrong, one of the most popular attacks is called "Man in the Middle", as its own name implies, the attacker sits between the client and the Server that is available to you.

During a "man in the middle" attack the communication between a and B is intercepted by the attacker, and relayed in an unlimited way, the attacker can retransmit the data without change, but also with change or block part of the data.

One of the most common strands of the attack is the attacker taking over a Wi-Fi router, interconnecting conversations with them. The practice of free Wi-Fi is the most common one for which an attacker can perform the attack, because these networks can have multiple users and network security is not so restricted.

Fig.5. “man in the middle” - normal communication and communication intercepted

4.1.1 Example

Communication between "A" and "B" where "C" is the attacker:

1 - "A" sends a message to "B" that is intercepted by "C". A “Hello B, send me the key” → C B

2 - "C" sends message to "B" (all points that was sent by "A") “C” → “Hello B, send me the key” → B

3 - “B” Responds with the key “C” ← [Key] ← B

4 - “C” Replaces the "B" key with yours and passes it to "a" A ← [Key “C”] ← C

5 - “A” Encrypts a message and with what believes that the key is “B” “A” → “Message” [Key “C”] → “C”

6 - The message was encrypted with the "C" key So "C" sees the content and modifies it if you want to “C” → "Message Changed" ["B" key] → “B”

4.1.2 Defenses

Although they cause great damage and are dangerous, there are different ways to defend themselves from the man in the middle attack, and most of them should be installed on the router and servers. A technique widely used, and that can help combat the attack is the encryption between client and server, in this case the server identifies the client due to its digital certificate and establishes the cryptographic connection, where the information is completely secure.

Another caution that the user should take into account is Internet browsing and choose to browse websites that use the HTTPS protocol, because this protocol is encrypted that prevents the hacker from accessing the information.

Access to free Wi-Fi networks is the oasis for hackers, because there are inexperienced users at the level of digital security, in these networks is where there are the greatest number of attacks man in the middle. Whenever the user uses Wi-Fi networks that are not encrypted, one of the forms of defense is to create a VPN service. This VPN service ensures that data circulating on the network is not exposed in these unsafe environments. The VPN encrypts the connection, and in which hackers will not have access to the IP.

5 Conclusion

The new technologies, more specifically the Internet, have lots of positive points, such as globalization making computer equipment indispensable in our daily life, but with this information circulation so fast and the learning it offers, the Internet also has its less good points, such as cybercrime and cyberattack. With the ease of access that the Internet offers and the increase in the number of users, many of them without any knowledge at the computer level, the most serious problems begin to appear.

However, because of this increase in users, opportunities for digital crime practices also grow, where more experienced users are taking on less experienced users to commit crimes, in which attackers take their Benefit.

Therefore, to combat these attacks, we need preventive measures and good practices in the use of this great global network, the Internet.

In this paper, the concept of cybercrime and cyberattack were studied, and we discussed some tools that attackers use to take possession of the information circulating in the "arteries" of this great global network, the Internet.

In the end, we conclude and discuss the tools and preventive measures that, adopted by users, hinder the work of attackers and thus protecting private information.

References

1. “all of the data stored in a large computer or network represented as a three-dimensional model through which a virtual-reality user can move” Collins English Dictionary – Complete & Unabridged 2012 Digital Edition.
2. Collins English Dictionary, Corpyrigth HarperCollins Publishers.
3. Pinheiro, P. P. Direito Digital. 4. ed. São Paulo: Saraiva, (2010).
4. DONNER et. al. Low self-control and cybercrime: Exploring the utility of the general theory of crime beyond digital piracy. Journal Computers in Human Behavior, (2014).
5. https://www.statista.com/statistics/267132/total-damage-caused-by-by-cyber-crime-in-the-us/
6. Neto, M. F. and Guimar˜aes, J. A. C., Crimes na internet: elementos para uma reflexão sobre a ética informacional. Revista CEJ, 7(20). ISSN 2179-9857, 2003.
7. https://www.statista.com/statistics/881158/average-financial-damages-via-cyber-attacks/
8. Markus Jakobsson, Zulfikar Ramzan, Crimeware: Understanding New Attacks and Defenses, 2008
9. Shuchi Juyal and Ruchika Prabhakar, “A comprehensive study of DDoS attacks and defense mechanism”, Journal of Information and operation management, 2012.
10. Soumya Tiwari , Anshika Bhalla and Ritu Rawat, “Cyber Crime and Security”, International Journal of Advanced Research in Computer Science and Software Engineering, 2016.
11. Wadhwa, A. and Garg, A., Studying and Analyzing Virtualization While Transition from Classical to Virtualized Data Center. International Journal of Computer Applications, 2015.
12. Seema Vijay Rane and Pankaj Anil Choudhary, “Cyber Crime and Cyber Law in India”, Cyber Times International Journal of Technology and Management, September 2012.
13. INELLAS, G. C. Z., Crimes na Internet. São Paulo: Editora Juarez de Oliveira, 2004.
14. Enrique García, www.larepublica.co, 20 de julho de 2018
15. Wadhwa, Amit. "Comprehensive Analysis of Security Issues and Solutions While Migrating to Cloud Environment." International Journal of New Innovations in Engineering and Technology 4.4 (2016)
16. Wadhwa, Amit. "Comprehensive Analysis of Security Issues and Solutions While Migrating to Cloud Environment." International Journal of New Innovations in Engineering and Technology 4.4, (2016)

A Review on Cyber Attacks and its Preventive Measures

Pedro Teixeira

Lusofona University of Porto, Portugal
pedroalv_23@hotmail.com

Abstract. The aim of this paper is to be able to approach the questions of privacy and security on a review about cybernetic attacks and its preventive measures. It will be discussed the impact of these attacks on modern society and how they change the way of how we need to view things in order to improve the way we work, play and entertain ourselves. This paper will also include a study of the potential attack we can suffer and the explanation of why it is one of the most dangerous in my opinion. It will also show the capabilities of the attack and how we can try to prevent the attack.

Keywords: Inside Job, Social Engineering, Blackmail, Ransomware, WannaCry, Encryption.

1 Introduction

In today's standards Cybercrime is a well-structured business in which big teams made of highly trained hackers make a profit by just attacking vulnerable companies or even common individuals when they are exposed to these hackers by not taking the necessary measures to prevent the attacks. [1]

As the years progress not only does the technology advance, but the criminal environment in the cybernetic world also does, which means that almost everything is connected to the internet in some way or another, meaning, that the number of doors from where hackers can attack is endless. Starting with simple pranksters joking on the internet sending some spam into mail-boxes, to professionally organized people attacking bank accounts and breaching into big companies to able to steal their personal data, therefore we live in a world where cyber threats and attacks are recognized as political and commercial challenges with levels of financial and reputational consequences. [2]

Knowing all this, we need to stay up to date in every way possible in order to keep up with what we can be struck with, so it is more than important to know what cyber-attacks are, how to spot a potential one and protect ourselves. [3]

The first section of this paper is going to explain how we can describe a cyber attack as well as the common scenarios that they occur on, thereafter, preventive measures will be detailed so we can have a better understanding how on to avoid certain risks.

In section three we are going to look at the current situation about cyber-crime around the world and a brief topic on how things are developing in Portugal.

Lastly, in the final section of this research paper, an exposition about one of the potential attacks based on encryption will be explained as well as detailing how we can prevent getting infected and how to proceed if we were already attacked by the virus.

2 Cyber Attacks

A cyber-attack is generally known by the common user as someone trying to steal our information or compromising our own device. Attacks can vary in terms of goals, one can be aimed to disable or turn offline a desired computer or electronic device, another goal is to gain access to said computer or electronic device and steal its data and gain admin privileges on it. [4]

As said before, cyber-attacks can have behind them a diverse spectrum of intentions, attacks on the general public or corporate organization can be or will be carried out through the spread of malicious viruses, fake websites, denial of services etc. The amount of ways today on how we can breach security for personal gain is enormous. [24]

2.1 Common Cyber Attack Scenarios

In terms of organizational attacks, we can trail a vast number of scenarios where they can occur.

Inside jobs are one of the most common ways of attacking a company or organization, most of the time when your company gets attacked, that threat is coming from the inside, where an employee working at said organization may exploit his role in the company to have access to confidential information by hacking the computer's network that is available to him. One of the most dangerous aspects of inside jobs is that the access to these personal information's are coming from trusted systems, and because of this most of these threats will pass undetected mostly because the attacker can erase the evidence since he has all the control.

The 2017 Cyber Security Intelligence Index by IBM concluded that 60% of the attacks organized against corporations were made by insiders, the aim of these attacks mostly involved malicious intent. [5]

Social Engineering is considered the technique used by attackers to lure users into sending them their confidential data. Social Engineers take advantage of human behavior to accomplish their illegal goals, one of the methods to capture the personal data of the victims is either by impersonating, for example, an IT support employee and trick the victim into divulging passwords and confidential data or simply by using the method of "phishing", were the victim opens a link to a fake website where they are prompted to insert their personal information, what they don't realize is that the website where the data is being inputted is fake and the information is, therefore, going directly into the hacker's hands. [5][6]

Blackmailing and extortion have been around for a long time, it has been used long before the internet was even a "thing", for ages now people around the world try to gain advantages in

terms of money, property, promotions or simple revenge with the use of blackmailing another person by threatening to expose private information. [5]

Nowadays it is even more common these types of crimes because a hacker can for have to access private pictures of someone and then ask them for money, so they won't expose the photos. Everyone now has a phone that is connected to the internet making it relatively easy for some expert to gain access to it, the sheer amount of cases of blackmail and extortion around the world is exponentially higher each year, Fig.1 shows the statistics of the number of recorded offenses in England and Wales from 2002/03 to 2017/18. [5][7]

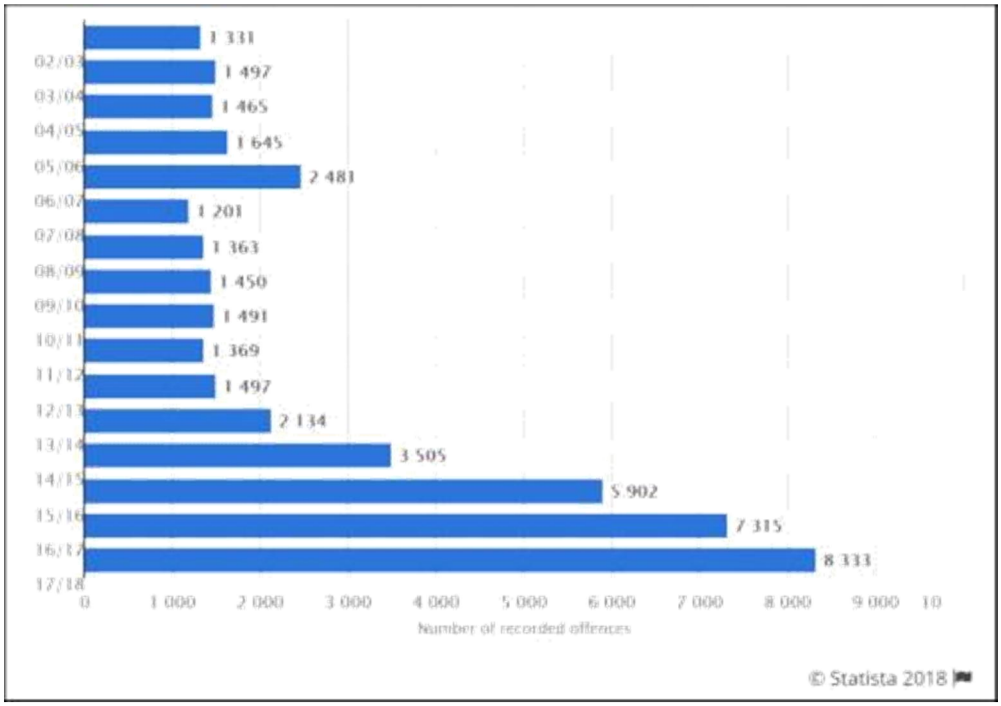

Fig.1. England and Wales Offences Recorded

2.2 Common Cyber Attack Preventive Measures

According to multiple articles, 2017 was a year where many companies suffered from security breaches and largely because of the lack of preventive measures. 2018 on the other hand, is proving to be a year where most of these companies are investing more and more money on security, this said we can follow some of the steps to prevent some of the described examples of cyber-attacks:

- Inside jobs

To try and prevent these types of attacks we can conduct background checks before proceeding to employ any candidate to the company, doing a review on the criminal, financial and commercial background will only benefit the company by making sure the employee is always telling the truth, since there is also an estimate that 40% of resumes might contain false information, so making sure we are employing the right person is crucial to prevent an inside job. [8][9]

Other common ways to prevent this issue is by having more than one trusted person overseeing the financial system of the company as well as restricting the number of employees who have access to back-end information, this way we can create an efficient and controlled environment of the crucial information the company has, preventing the risk of an inside job. [8][9]

- Social Engineering

According to most opinions the best way to combat social engineering is no more than educating yourself against these threats, you cannot protect yourself if you don't know where the threat is coming nor how the threat is coming, so it is essential for oneself to study and understand how threats are conducted so we can have a bigger understanding on the precautions to take and possible educate others into not making the mistakes that can lead to a possible act of social engineering. [10]

These types of attacks are also common and dangerous are in video game clients, where people can impersonate admins by asking passwords of un-educated and naïve young gamers, these people will often ask for personal information in order to steal either the account itself that has value money wise thanks to the games that the account contains or credit card information. Yet again, educating oneself and being aware of messages like the ones in Fig.2 and Fig.3 where there will always be a message saying that no admin will ever ask for your password of personal information. [10]

Common sense in general is the biggest weapon, being aware of the dangers and always suspect of the most obvious questions and simple forms on the internet, always checking if the website we are navigating on has the SSL (Secure Socket Layer) enabled, so we know there is no danger in clicking or providing our information. [10]

- Extortion and Blackmail

Extortion and blackmail are always a sensitive theme to describe since most of these attacks are often about intimidating us directly through threats and money in exchange of deleting any private information, they allegedly have acquired from using viruses. To be clear most of these threats normally are fake and only exist to intimidate us into giving the culprit money in form of a bitcoin, so when we normally receive an e-mail like the one in Fig.2, we should just ignore it, or simply never open an e-mail from an unknown source before checking it with an antivirus, this is just a scam so we should remain calm and collective the sender did not install anything on our computer nor did he use a keylogger to get access to our credential and private photos and videos of us. [11]

GMX - (Part num your Hacked phone. +XX XXXXXX0837) Seite 1 von 1

GMX ProMail

(Part num your Hacked phone. +XX XXXXXX0837)

Von: [redacted]
An: [redacted]
Datum: 10.08.2018 17:00:12

It seems that, +XX XXXXXX0837, is your phone. You may not know me and you are probably wondering why you are getting this e mail, right?

actually, I setup a malware on the adult vids (porno) web-site and guess what, you visited this site to have fun (you know what I mean). While you were watching videos, your internet browser started out functioning as a RDP (Remote Desktop) having a keylogger which gave me accessibility to your screen and web cam. after that, my software program obtained all of your contacts from your Messenger, FB, as well as email.

What did I do?

I backuped phone. All photo, video and contacts.
I created a double-screen video. 1st part shows the video you were watching (you've got a good taste haha . . .), and 2nd part shows the recording of your web cam.

exactly what should you do?

Well, in my opinion, $1000 is a fair price for our little secret. You'll make the payment by Bitcoin (if you do not know this, search "how to buy bitcoin" in Google).

BTC Address:

1GYNGZLEUGkkQjHo19dHDnGE87WsAiGLL8

(It is cAsE sensitive, so copy and paste it)

Important:
You have 48 hour in order to make the payment. (I've a unique pixel in this e mail, and at this moment I know that you have read through this email message). If I do not get the BitCoins, I will certainly send out your video recording to all of your contacts including relatives, coworkers, and so on. Having said that, if I receive the payment, I'll destroy the video immidiately. If you need evidence, reply with "Yes!" and I will certainly send out your video recording to your 6 contacts. It is a non-negotiable offer, that being said don't waste my personal time and yours by responding to this message.

Fig.2. E-mail sent by the attacker [11]

Nonetheless we should change our passwords regardless if we get these types of e-mails or not, a good practice is changing the passwords often and not having the same password for different types of platforms, but more importantly, than that is having a two-step verification in order to diminish the risk even more. [11]

A powerful and easy tool to use to see if our credentials have been leaked or stolen, we should use this website called 'https://haveibeenpwned.com/', this useful tool will actually tell us which platform had the credentials leaks and a Pastebin with a conjunction of thousands of other e-mails that have also been affected. [11]

Fig.3. "Have I been PWNED" results [25]

When we type our e-mail, it will show us the number of breaches our account has suffered has shown on Fig.3, in this example a game, a video sharing platform, and some other data breaches from other websites our e-mail was registered in. This tool is very handy has it also shows the actual data that was compromised and released to the public as well as the year that it occurred.

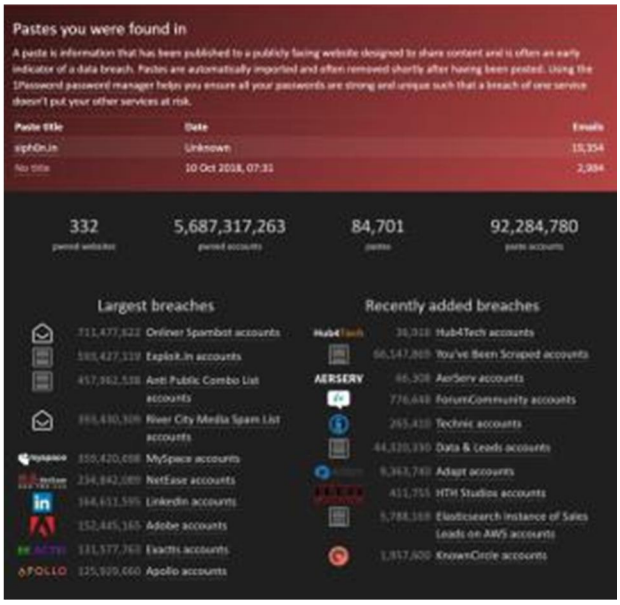

Fig.4. "Have I been PWNED" Results [25]

Fig. 4 shows us the Pastebin that can be consulted, there we can see the actual e-mail that was stolen and released to the public alongside ours, below that we have some stats of the websites that were breached, the number of 'pwned' accounts the numbers of pastes and paste accounts. Even so, we can see some highly regarded websites like Adobe and LinkedIn were also attacked.

3 Brief Overview of Cybercrime Around the World

Cybercrime is a non-stop topic around the world, global cybercrime costs are estimated to grow annually and will reach a historical number of 6 trillion dollars by 2021 compares to its half counterpart of 3 trillion dollars in 2015 [12]. These numbers consequently projections based on historical cybercrime figures, these costs will include the destruction of data, stolen money, theft of personal and financial data, fraud and many other vulnerable types of data. [13]

Converging all these topics around attacks and threats the United Nations specialized agency for information and communication technologies measured that countries around the world are committed to preventing and improving their priorities around this matter, even though less developed countries like African regions still lack the means necessary to combat these threats remain mostly dependent of other countries support. The main areas being the focus today are:

- Safeguarding digital business models
- Utilizing Intel and info-sharing programs
- Security of the Internet of Things
- Managing geopolitical cyberthreats

Reports from multinational professional services firms believe based on surveys conducted to CIOs and CISOs concluded that organizations appear to be moving in the right direction in terms of adapting to the current situation of attacks, meaning they can now prevent and foresee these threats resulting in a recovery time that may vary between minutes and within a certain amount of hours for 44% of the companies. [13] [14]

Ransomware, banking malware and mobile malware continues to be the most significant threat at this moment in time because to date there is no system immune to the infection. [13] [14]

3.1 A Brief Look at Portugal

As shown cybercrime grows exponentially each year, and stats show how well and how much countries around the world are evolving to prevent the attacks and actually help others with these problems, therefore in 2016 a unit focused on cybercrime and cyberterrorism was created in Portugal the law was approved and enabled the National Police to have this unit based on a model adopted by EC3 (European Cybercrime Center) from EUROPOL. Thanks to this unit Portugal can and will be able to respond in a fast and efficient way to cybercrime. [15]

However, creating this unit will not solve all the problems since according to APAV (Associação Portuguesa de Apoio à vítima) estimated that 78% of people in Portugal that surf the web are not well informed about the risks and threats. [16]

Attorney Pedro Verdelho alerts that cybercrime has gone up, which is a fact and common occurrence around the globe, but on the other hand, there are not credible statistics about it. The attorney states that no one in Europe has such rigorous numbers, this is due to the stats not being covered in the right way, meaning that the crimes are being recorded yes, but not how the actual crime was convicted, the same way a crime occurs but we don't have a record if the crime was online or not. [17]

Cybercrime fighting unit from the National Police responsible Carlos Cabreiro, admits also that cybercrime is going up stating extortion and identity theft being the one that is growing the most. Exposing our personal data on the internet is the catalyst for this growth since gathering all our details is easy which leads to fake accounts on social media and impersonation for all other kinds of activities. [18]

4 Ransomware

Ransomware is one of the various types of malware that exists today, these true purposes of these viruses is to block the access to computers and mobile devices or encrypting the data existing in a system, usually, the attacker demands money for the restoration of the device and/or data that he managed to gain access to. [19]

Many times, the logo of INTERPOL or a law enforcement agency is displayed to the victim making them believe the authorities are involved in any activity, even though this is not the case because INTERPOL or any other agency will ever block or encrypt data of any user. [19]

4.1 "WannaCry" – The Attack

WannaCrypt also known as WannaCry on May 12th of 2017 started spreading along multimer networks infecting thousands of computers across the globe, 24 hours after the first attack the actual number of infections reached a number as high as 185.000 machines in over 100 countries. [20]

This attack was aimed at mostly hospitals, telecommunication companies, and gas plants. One of the companies that were damaged the most was the National Health Service (NHS) located in the United Kingdom. [20]

Ransomware is normally used to infect small to large businesses across the world making its way to systems and networks via attachments on e-mails, browsers or third-party exploits, ransomware like WannaCry automated the exploitation of a vulnerability present in most versions of Windows. This potent Ransomware is one of the most dangerous nowadays thanks to the ability that allows the attacker to run code on the vulnerable computer enabling him to plant ransomware without human and local action. This behavior of an attack was never seen

before allowing it to be the perfect attack against specific environments or infrastructures like servers running vulnerable versions of the Server Message Block (SMB protocol). [19] [20]

The attacker solicited a rescue of 275€ to 550€ which makes it believe that the attackers aimed at smaller businesses, so they amassed attacks based in quantity and not in quality. [21]

4.2 Preventing WannaCry

The correct way to prevent them from being affected by WannaCry is using a a copy of Microsoft Windows that is not up to date:

- Windows 10 (1507,1511,1607)
- Windows 8/8.1
- Windows 7
- Windows Vista
- Windows XP
- Windows Server 2008, 2008 R2, 2012, 2012 R2

Therefore, the first step into prevention is always to keep the device up to date with all security patches installed. [22]

But what if your machine is already affected?

Fig.5. "WannaCry" Infected Warning

As shown in Fig.5 this is what will appear if your computer is already infected with the ransomware, you will be prompted with a countdown that tells us that, if we don't pay the demanding money, our files will get deleted.

There is no current fix for the virus if one's computer is already infected even though experts are working towards finding ways to decrypt the files. The advice that should be taken is reformatting the computer, restoring a previous version or installing a new version of windows and restoring the data files that have been backed up. Although in recent months project "No More Ransom" is being developed by the National High-Tech Crime Unit of the Netherlands police, Europol's European Cybercrime Centre and McAfee. [19] [22] [23]

In no way or form should we ever adhere to their demands for the money they ask as reported by The Guardian via Europol. [22]

In general, we can be avoiding ransomware by updating and installing software and security patches, this is key to avoid infection as well as never jailbreaking or rooting any device. [22]

5 Conclusion

Human behavior is and will keep being the reason why most crimes happen, taking advantage of one's ingenuity or lack of knowledge on the subject, will result in someone exploiting a service, blackmail someone, extort another or steal their identity. The trend for cybercrime is to actually keep increasing as well as the diversity of the threats we may face in a near future, ransomware like WannaCry that surged in a way no one was expecting, creating such an impact that even today experts don't have an official way of preventing or recuperating the encrypted data, are signs that these methods of exploitation will keep evolving and developing, we can come to terms that if someone wants to break in they will, it is just a matter of time and resources until someone breaches whatever their target is.

Spreading the word as much as possible to keep preventing and minimize damage is crucial. Reducing the number of violations of privacy and theft must be a priority, seeing the numbers and statistics means that if nothing is done, more and more crimes will be committed.

Nonetheless, not everything is bad, as we can clearly see giants like INTERPOL and EUROPOL keep moving masses and creating ways to give tools to start preventing most of the crimes as well as national police in Portugal being allowed to have a unit to combat these threats and be able to respond faster and more efficiently.

Taking in consideration most of the aspects reviewed in this research we can conclude that the wave of crime is indeed increasing at a steady rate either by the creation of new exploits, taking advantage of misinformed people, poorly configured systems, outdated software and the most common of all human behavior in which leads to other types of threats like extortions and blackmailing, we can also rest assured that the big companies and associations worldwide are aware of this problem and are creating ways and the means necessary to swim against this current that seems too heavy to be taken care of. Countermeasures must evolve as fast, if not faster than exploits, no one wants to be in the bitter end of the deal, so its everyone's duty to

make the cyberspace a cleaner and healthier place, overall every single individual is responsible for alerting and contribute to better environment online and offline.

References

1. "Anatomy of a Cyber Attack the Lifecycle of a Security Breach", Oracle Linux, Retrieved from: "http://www.oracle.com/us/technologies/linux/anatomy-of-cyber-attacks-wp-4124673.pdf", Last Access: 20 December 2018.
2. "The Evolution of Cyber Attacks", ITBusinessEdge, Venafi (2013) Retrieved from: "https://www.itbusinessedge.com/slideshows/the-evolution-of-cyber-attacks.html", Last Access: 20 December 2018.
3. Ledford, J.; "Could a Cyber Attack Knock Out Your Computer?", Lifewire (2018).
4. Fruhlinger, J.; "What is a Cyber Attack Recent examples show disturbing trends", CSO from IDG (2018).
5. Farhat, V.; McCarthy, B.; Raysman, R.; "Cyber Attacks: Prevention and Proactive Responses", Holland & Knight LLP, Practical Law Company (2011).
6. Hulme, G.; Goodchild, J.; "What is Social Engineering? How criminals take advantage of human behavior", CSO from IDG (2017).
7. "The Crimes of Blackmail and Extortion", Law Shelf Educational Media, A project of National Paregal College, Retrieved from: "https://lawshelf.com/videos/entry/the-crimes-of-blackmail-and-extortion ", Last Access: 20 December 2018.
8. Doyle, A.; "What Employers Look for in Background and Credit Checks", Background Checks for Employment, the balance careers (2018).
9. Mueller, K.; "Deter the Inside Job. 5 Ways to Avert Employee Theft and Fraud.", Entrepreneur Europe (2018).
10. Olavsrud, T.; "9 Best Defenses Against Social Engineering Attacks", eSecurity Planet (2010).
11. Jareth.; "How to deal with cyber blackmail?", EMISOFT, Security Essentials (2018).
12. Periman, K.; "How to Prevent the Bank Robbery No One Can See", 2017 Midyear Cybersecurity, Report Cisco Blogs, Financial Services (2017).
13. Steve, M.; "2017 Cybercrime Report", Cybersecurity Ventures, Herjavec Group (2017).
14. "2017 Global Enterprise Security Survey", Fortinet (NASDAQ: FTNT) (2017).
15. EC3Europol.; 2015, 7 October 2016.; Retrieved from: "https://twitter.com/EC3Europol", Last Accessed: 18 December 2018.
16. "A Realidade do Cibercrime", APAV, Retrieved from: "http://apav.pt/cibercrime/ ", Last Access: 20 December 2018.
17. Verdelho, P.; No stats from cybercrimes, Retrieved from: "https://rr.sapo.pt/noticia/46824/cibercrime-temaumentado-mas-nao-ha-estatisticas", Last Access: 20 December 2018.
18. Cabreiro, C.; Identity theft is growing, Retrieved from: "https://rr.sapo.pt/noticia/46824/cibercrime-temaumentado-mas-nao-ha-estatisticas", Last Access: 20 December 2018.
19. "Online Safety", International Criminal Police Organization (INTERPOL) Retrieved from: "https://www.interpol.int/en/Crime-areas/Cybercrime/Online-safety/Ransomware ", Last Access: 20 December 2018.
20. "Bitdefender next-generation machine-learning and memory introspection technologies ensure that Enterprises worldwide have always been safe from the WannaCry ransomware mega-attack and the underlying Eternal Blue zero-day exploit", Bitdefender, Retrieved from: https://www.bitdefender.com/business/usecases/wannacry.html, Last Access: 20 December 2018.
21. Séneca, H.; "WannaCry: 12 mil computadores infetados em Portugal", Exame Infomática (2017).
22. Justin.; "How to Prevent and Fix WannaCry Ransomware", My Private Network (2018).
23. "No More Ransom", Retrieved from: "https://www.nomoreransom.org/en/index.html", Last Access: 20 December 2018.
24. "What constitutes a cyber-attack?", NEC, Retrieved from: "https://www.nec.com/en/global/solutions/safety/info_management/cyberattack.html", Last Access: 21 December 2018.
25. "; -- have I been pwned?", Retrieved from: "https://haveibeenpwned.com/", Last Access: 22 December 2018.

Malicious URL Detection using Machine Learning Algorithms

Marcelo Ferreira

Lusofona University of Porto, Portugal
ferreira_marcelo@outlook.pt

Abstract. Malicious URLs are a dangerous threat to cyber security, these types of attacks can lead to scams, where people lose money, their information and accounts. It is important to be able to detect and act against these threats, the most conventional way is the use of blacklists, but this technique has many difficulties in acting against new URLs, so we are increasingly focused on machine learning algorithms, and that is precisely the focus of this paper. In this paper, we'll cover the most common and dangerous attacks through malicious URLs, how they work, and how to prevent them. We will focus on and analyze more concretely a technique to detect, which uses algorithms of Machine Learning.

Keywords: Malicious URLs, Machine Learning, Detection, Algorithms.

1 Introduction

With the rise of the internet and the evolution of technologies, we have moved away from the society of industries and we have moved on to the information society. We are in a society where we use various forms and technologies of communication and information, such as online purchases, online banking and betting sites. But with the increase of the use of these technologies, also comes the increase of risks, because third parties try to steal and take advantage of our information, networks and systems for their own benefit. [1]

People using, for example, home computers are targets that are very vulnerable to external attacks and threats because in most cases they are not aware of the various types of attacks and how they should be protected. They are naiver and fall into the simplest attacks. Even people who use computers at work and are alerted and aware of the attacks and that they must be careful and protect the company and themselves fall into the simplest attacks. [1]

These attacks can be reconnaissance only, which means that attackers attempt to map systems and networks in order to find flaws, can be access attacks, where attackers gain access to systems and networks, gain access to services, databases, among others. Another form of attack, is to make a machine or network resource unavailable or make it so slow that it is practically impossible to use it. [2]

Going to the numbers and the facts, 24,000 malicious applications are blocked every day, and information that most applications release is 63% mobile phone numbers and 37% device location. From 2016 to 2017 the percentage of cybersecurity costs increased by 22.7%, with malware attack costing companies 2.4 million dollars on average. The applications with most

security problems are those of lifestyle and music, 411 million accounts of dating sites were stolen in 2017. [3] [4] Cybersecurity has emerged as a set of tools, policies, concepts, protocols, actions, techniques, and best practices that help protect organizations and individuals in general from attacks, helping to maintain confidentiality, integrity, authenticity and nonrepudiation of information. [5]

2 Universal Resource Locator

URL's, also known as Universal Resource Locator, as the name indicates are used to find a particular resource on the Internet, they are also known as web address. A URL finds by providing, to the browser for example, an abstract of the location of the resource, when this resource is found the system can execute a great diversity of operations. [6]

A URL is formed by the protocol used to access the resource, the location of the server to be accessed, which may be on the form of the domain or on the form of the IP address and the path where the resource is located. [7]

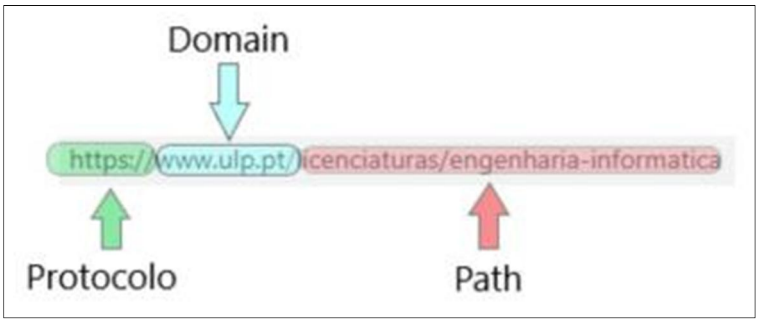

Fig.1. This figure, represents an URL and consequently all the parts that constitute an URL.

The following protocols are the most commonly used today: [8]

- FTP, File Transfer Protocol, allows the transfer of files between two machines. The client connects to the server to obtain some file and the server receives the request and supplies the file.
- SFTP, Simple File Transfer Protocol, does the same as the FTP protocol, but uses a different technology that allows you to authenticate and secure the connection between the client and the server, ensuring greater file security.
- POP3, Post Office Protocol, works like a mailbox for emails, when the client accesses the server, has access to all the emails that he received.

- SMTP, Simple Mail Transfer Protocol, this protocol is aimed at sending emails, this protocol is very effective, but only allows the sending of text.

- IMAP, Internet Message Access Protocol, this protocol joins the best of the two worlds, allowing the user to access and manage their messages directly on the server, in a counterpart you have to always be connected to the network and it has storage limit.

- HTTP, Hypertext Transfer Protocol, it is used for site navigation. The protocol also works with a connection between the client and the server, where the client is the browser that the user uses, and the server is the site that is intended to be accessed. The browser sends a request to access a page and the server sends an access permission response. With it, come the files that form the page that the user wants to access.

- HTTPS, Hypertext Transfer Secure, works like HTTP but has a layer of protection because it uses SSL, the SSL protocol ensures privacy and data integrity between two applications that communicate over the internet. This is done through the authentication of the parties involved and the encryption of the data transmitted between the parties.

When a URL directs the browser to a file that it can open, such as images or PDF's, the browser displays content without having to download the file, but many other types of files require a download. Because of all the complexity and diversity of functions, URLs can also be made to do evil and attack the user. [9]

3 Attacks using URL

One of the most common attacks is called phishing, the main purpose of the attack is to trick the user into giving them their data, usually login data. The attack consists of getting the target to click on a link that takes you to a page similar to what the target wanted, but when they login, the attackers get their login data. [10]There are several techniques for phishing, one of them is to buy a domain similar to the site that we want to copy for example facebok.com instead of facebook.com, because the human brain is accustomed to when reading words with almost imperceptible errors, it processes as if the error did not exist, then the user does not even notice that he is not on the real site and logs in giving access to his data to the attackers. In an experiment at Carleton University, giving a list of links to multiple users to identify if it is an attack attempt or the original site, users used various techniques to verify that the site was true or false by analyzing the URL, testing the features of the sites, looking at google if the URL of the original site and the one provided for the experience were the same, checking if the site uses SSL, among others, but outside of this experiment, people do not do this for all sites that enter , only if they notice that some is suspect. The problem is that they try hard to go unnoticed. [11] Another type of attack is the Driven-by download, which consists in unintentionally downloading malicious code, this attack does not require the user to click something or do

something. The purpose of these codes is to communicate with another computer to gain access to the device. These malicious files are very common in corrupted websites. [12]

There are attacks that are based more on the psychological of the person, such as social engineering, that the first step of this attack is to know your target, such as sites they visit the most, their concern about security, their location, among other data. This attack is very common because it is easier to identify the vulnerabilities of people than of systems and networks. [13] Some attacks using this technique are [14]:

- Bating, these attacks try to wake greed and curiosity to the victim, they can be physical like leaving a PEN in a strategic place so that when the victim finds it, it enters in your computer and installs malware in your system, or through ads that lead to malicious sites that lead to malware downloads.

- Scareware, victims receive false alarms and threats leading them to believe that their system is infected with malware. Usually are websites that say these messages and almost always have a button to download a tool that will correct all the evils of the system, but ironically this tool will infect the system.

- Honey trap, gaining the confidence of the target, until you are in a kind of online relationship, with the confidence of the target, you can send a malicious URL that the person will most likely trust you and provide their data.

3.1 Malicious URL Detection using Backlists

To combat these malicious URLs, we must detect them because of their ease of deceiving us, so we created forms to detect them, one of the most common being the use of Blacklists. Blacklists are databases where the data of the URLs that have already been confirmed as malicious are saved, and more URLs are added to the list over time. Whenever a new URL is visited, a database lookup is performed. If the URL is present in the blacklist, it is considered to be malicious and then a warning will be generated; else it is assumed to be benign. Although it seems a safe and effective method, backlists are slow, because they cannot keep up with the growing number of URLs, meaning that these databases will never be able to have all the malicious URLs that exist because new ones are created every day and new ways to get around blacklists. To combat this problem and find a new way to detect malicious URLs, scientists have, in recent years, sought a solution in Machine Learning algorithms. [15]

4 Malicious URL Detection using Machine Learning.

Machine learning is a category of algorithm that allows software applications to become more accurate in predicting outcomes without being explicitly programmed. The basic premise of machine learning is to build algorithms that can receive input data and use statistical analysis to predict an output while updating outputs as new data becomes available. Machine Learning

approaches, use a set of URLs as training data, which sometimes can be Blacklists, and based on the statistical properties, learn a prediction function to classify a URL as malicious or benign, which gives them the ability to generalize to new URLs unlike blacklisting methods. [16]

4.1 Extraction of Features

The first requirement for a Machine Learning model is a "training date," which corresponds to a large number of URLs. These algorithms can be supervised or unsupervised, meaning that the algorithm knows whether URLs are malicious / benign or do not know. There are also semi-supervised algorithms which mean that they know the classification and a part of the training URLs. The next step is to extract information from the URL and transform it so that it can be rendered by the template. [17]

These features can be: [18]

- Lexical, are features obtained through the properties of the URL, because through the aspect of the URL it should be possible to identify if it is malicious, because these URLs try to pretend to be benign URLs, changing its aspect a bit. The most commonly used features are the length of the URL, the length of each URL component (Domains, subdomains, path), number of special characters, and each character sequence separated by a special character ("/", ". ") Is considered a feature. Based on all the words in all the URLs, a dictionary was built. If the word was present in the URL, the value of the resource would be 1 and 0 otherwise. This is also known as the bag-of-words model. The whole bag-of-word resource approach can be seen as a form of blacklist compatible with Machine Learning, instead of focusing on the entire URL string, analyze the URL based on smaller components. Not to be detected by Blacklists hackers use algorithms to generate URLs that are not in them but it will be harder to go through the bag-of-words approach because the template will detect that the URL has words that are not in the dictionary. • Host-based, are obtained through the host properties of the URL. This allows us to know the location of the host, its identification and some other features. URL lifetime is one of the most important features since malicious users typically have much less lifetime than benign ones. It is possible to obtain information about when and who registered the domain as well as the location of the IP. Due to the difficulty of obtaining new IPs these features are very important to detect malicious URLs.

- Content Based, these features are obtained by downloading the contents of the URL page, thus being the most dangerous type of feature to obtain but can greatly help prediction accuracy. We can get the HTML code of the page and analyze the number of words, average words per line, distinct words, links to remote scripts and invisible objects. Often, malicious code is encrypted in HTML and this is directly linked with a larger number of words with great length and with the use of concatenation. Other

features we can get are related to JavaScript as they are used by hackers to encrypt malicious code or execute without permission. We analyze features such as the number of long strings, number of events, number of strings, suspicious objects and tags, and the use of functions such as exec (), link (), eval (), escape (), search () because they are used to distribute malware. [19]

- Other, there has currently been a growth of platforms that allow you to shorten a URL to a totally different and smaller string to allow sharing on social networks, such as Twitter that, per tweet, only allows 140 characters. But these platforms can be used for evil and spread malicious URLs, although they try not to produce short URLs for the malicious, they have little effectiveness since they use Blacklists as a basis to block the malicious ones. So, to bridge this, we can get information about the shortened link itself and the devices that accessed the platform to shorten them as well as the devices that shared them, also counting the number of shares on social networks. Other features that may help in detection are the measures of popularity, one of these measure is the Link Popularity which is scored based on incoming links from other webpages. Also used are the number of popups and the behavior of plugins. The number of times that the URL is accessed can also be an indicator, because malicious ones, are less accessed. [20]

The next step is to transform and convert all the features extracted from the URLs into a numerical vector so that they can be powered by machine learning algorithms. In addition, some data normalizations can usually be used to deal with the scale problem. [17]

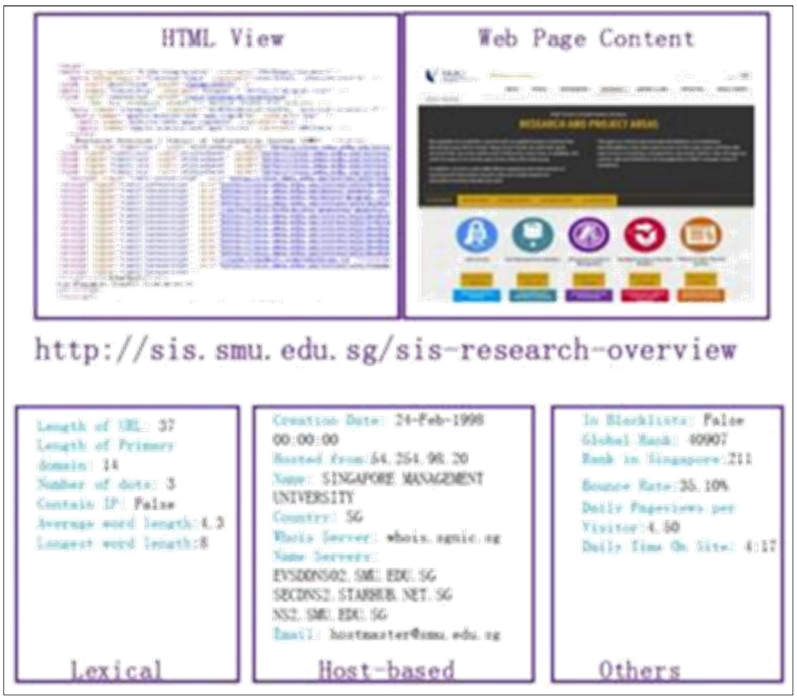

Fig.2. This figure represents the features that can be obtain from an URL.

4.2 Machine Learning Algorithms for Malicious URL Detection

After converting the features to vectors, there are many algorithms that can be applied to predict if an URLs is malicious or not, so we are going to study some of them. We are going to study some algorithms from the Batch Learning family and others from the Online Learning family. The Batch Learning algorithms, work with the assumption that all the training data is available after the training task, some of the Batch Algorithms are:

- Naive Bayes, this algorithm generatively classifies the URLs and it's a "naive" algorithm, in the sense that it considers that all the features (x) are independents from each other and, for each feature, it calculates the conditional probability, described in the next equation, are y=1 stands for URL malicious and y=0 for URL benign. After making this calculation for all features it decides if its malicious or benign. [22]

$$P(y=1|\mathbf{x}) = \frac{P(\mathbf{x}|y=1)}{P(\mathbf{x}|y=1) + P(\mathbf{x}|y=0)}.$$

- Support Vector Machine (SVM) are well-known for binary classification of great dimensions of data. This algorithm uses a single rule that is expressed by kernel function K(x, x ′) that computes the similarity between two feature vectors and non-negative coefficients {αi} n i=1 that indicate which training examples lie close to the decision boundary. SVMs classify new examples by computing their (signed) distance to the decision boundary. Up to a constant, this distance is given by

$$h(\mathbf{x}) = \sum_{i=1}^{n} \alpha_i (2y_i - 1) K(\mathbf{x}_i, \mathbf{x}),$$

  The sign of this distance indicates the side of the decision boundary on which the example lies. In practice, the value of h(x) is to predict a binary label for the feature vector x. [21]

Online Learning are algorithms that treat the data as an instance flow, thus they are Learning from the training data and predicting the real URLs almost at the same time. So, from the Online Learning family, here are some of them:

- First order, are algorithms that learn by updating a vector with the labels (benign or malicious) using only first order features and the training data. [23]

- Second order, these algorithms, instead of using first order features, tries to boost the learning efficiency, by exploring second order features like statistical features. [23]

- Online Active Learning, supervised algorithms assume that the classifications of the training data are received by them with no cost, but that's not true, in real systems this

process can be expensive and slow, to overcome this issue Online Active Learning tries to reduce this cost. These algorithms aim to only consult if an URL was already classified (malicious or benign) if it low transmits confidence levels. [24]

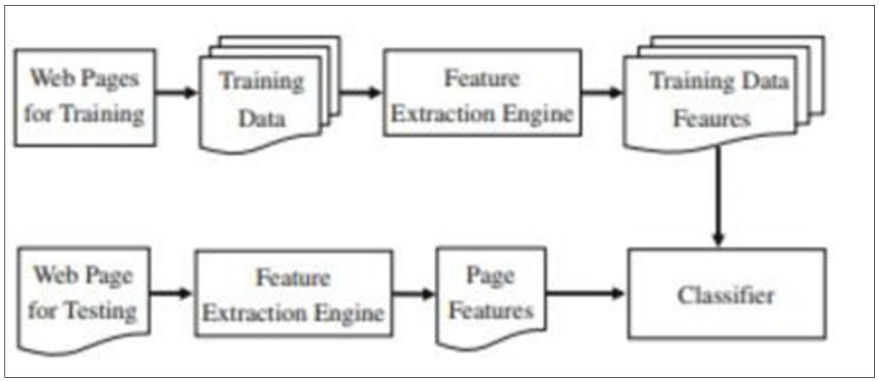

Fig.3. This figure represents all malicious URL Detectors, based on Machine Learning.

5 Conclusion

In this paper we learn that through some techniques, hackers can obtain our personal and private information and that they can bypass several of our ways to detected them and protect from them. One of these serious treats and the main focus of this paper was the malicious URL, hackers have several techniques and algorithms to obfuscate their URLs to bypass out defenses. To overtake them in this race to protect our data, one promising choice, that we aboard on the paper, are the Machine Algorithms projected to Detect or classify URLs as benign or malicious. Although they are a good way to improve security they are expensive and hard to adapt to some applications, like browsers, but because of their potential, there is a need to wager in them and to study more, so that they can grow and be more present in people's life.

References

1. Wang, W.; Lu, Z.: Cyber security in the Smart Grid: Survey and Challenges. Computer Networks, vol. 57:5, pp. 1344-1371. ISSN 1389-1286, https://doi.org/10.1016/j.comnet.2012.12.017. (2013)
2. Uma M., Padmavathi G.: A Survey on Various Cyber Attacks and Their Classification, International Journal of Network Security, vol.15:5, PP.390-396. (2013)
3. Internet Security Threat Report, Vol.23, Symantec.
4. Richards, K.; LaSalle, R.; INSIGHTS ON THE SECURITY INVESTMENTS THAT MAKE A DIFFERENCE, COST OF CYBER CRIME STUDY, Accenture. (2017)
5. Rossouw, S.; Niekerk, J.;From information security to cyber security, Computers & Security, vol. 38, pp. 97-102, ISSN 0167-4048, https://doi.org/10.1016/j.cose.2013.04.004.(2013)
6. What is URL?, https://searchnetworking.techtarget.com/definition/URL, last accessed 22/11/2018.
7. What is URL? Definition from Techopedia, https://www.techopedia.com/definition/1352/uniform-resource-locator-url, last accessed 22/11/2018.
8. Protocolos de Rede, https://www.weblink.com.br/blog/tecnologia/conheca-os-principaisprotocolos-de-internet/ last accessed 24/11/2018.
9. What's the difference between a site and a URL?, https://www.lifewire.com/what-is-a-url2626035, last accessed 24/11/2018.

10. Ataque Homografico, https://www.techtudo.com.br/noticias/2017/11/ataque-homograficotruque-na-url-engana-usuarios-com-paginas-falsas.ghtml, last acessed 30/11/2018.
11. Alsharnouby, M.; Alaca, F.; Chiasson, S.; Why phishing still works: User strategies for combating phishing attacks, International Journal of Human-Computer Studies, Volume 82, pp. 69-82, ISSN 1071-5819, https://doi.org/10.1016/j.ijhcs.2015.05.005.(2015)
12. What is a Drive-by Download, https://www.kaspersky.com/resource-center/definitions/drive-by-download, last acessed 30/11/2018.
13. What is Social Engineering?, https://searchsecurity.techtarget.com/definition/social-engineering, last acessed 30/11/2018.
14. What is Social Engineering?, https://www.incapsula.com/web-application-security/socialengineering-attack.html, last acessed 1/12/2018.
15. Sahoo, D.; Liu, C.; Steven C. H.; Malicious URL Detection using Machine Learning: A Survey, eprint arXiv:1701.07179. (2017)
16. What is machine learning?,https://searchenterpriseai.techtarget.com/definition/machinelearning-ML, last accessed 12/12/2018.
17. Detecting Malicious URLs with Machine Learning, https://ritcsec.wordpress.com/2017/12/07/detecting-malicious-urls-with-machine-learning/, last accessed 11/12/2018.
18. Utilizando Aprendizado de Máquina para Detecção Automática de URLs Maliciosas Brasileiras, Lucas Dantas Gama Ayres, Salvador – BA. (2018)
19. Y.-T. Hou, Y. Chang, T. Chen, C.-S. Laih, and C.-M. Chen, "Malicious web content detection by machine learning," Expert Systems with Applications, vol. 37:1, pp. 55–60. (2010)
20. Sahoo, Doyen; Liu, Chenghao; Hoi, Steven C. H, "Malicious URL Detection using Machine Learning: A Survey", arXiv e-prints, January 2017.1
21. Kolari, P.; Finin, T.; Joshi, A.; "Svms for the blogosphere: Blog identification and splog detection," in AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs, pp. 92–99. (2006)
22. Aggarwal, A.; Rajadesingan, A.; Kumaraguru, P.; "Phishari: automatic realtime phishing detection on twitter," in eCrime Researchers Summit (eCrime), IEEE, pp. 1–12. (2012)
23. Ma, J.; Saul, L.; Savage, S.; Voelker, G. M.; "Identifying suspicious urls: an application of large-scale online learning," in Proceedings of the 26th Annual International Conference on Machine Learning. ACM, pp. 681–688. (2009)
24. Zhao, P.; Hoi, S. C.; "Cost-sensitive online active learning with application to malicious url detection," in Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, pp. 919–927. (2013)

Address Resolution Protocol (ARP) Spoofing: Attacks and Defenses

Bruno Duarte

Lusofona University of Porto, Portugal
a21802084@mso365.ulp.pt

Abstract. Nowadays, address resolution protocol spoofing attacks are becoming more and more, so it is necessary to know how to defend ourselves from it. With this work will be addressed several topics within ARP Spoofing, firstly will be started by explaining what the address resolution protocol (ARP) is as well as the constituents of it. Next, will be explained the types of ARP (ARP caching, Inverse Arp, Reverse Arp, among others). In a second part, will focus my work on ARP Spoofing explaining how it is well to say the types of attacks of the same, last presented the defenses that the user can use to be able to prevent these same attacks.

Keywords: ARP, Spoofing, Attacks, Defenses.

1 Introduction

With the desire to improve the security of the users, the address resolution protocol (ARP) Spoofing comes increasingly to be topic of conversation by the worse reasons since this is one of the ways of the people being attacked by hackers, being possible that the same ones obtain steal data, steal passwords using them to perform larger crimes, even get DoS attack to deploy servers below with the successive sending of multiple packets to the network.

ARP operates by sending out “ARP request” packets. An ARP request asks the question, “Is your IP address x.x.x.x? If so, send your MAC back to me.” These packets are broadcast to all computers on the LAN, even on a switched network. Each computer examines the ARP request, checks if it is currently assigned the specified IP, and sends an ARP reply containing its MAC address.[1]

2 Address Resolution Protocol (ARP)

The address resolution protocol (ARP) is a low-level network protocol for translating network layer addresses into link layer addresses. [2] ARP finds the hardware address, also known as Media Access Control (MAC) address, of a host from its known IP address. ARP maintains a cache (table) in which MAC addresses are mapping to IP addresses. [3]

IP addressing occurs at Layer 2 (data link) and Layer 3 (network) of the Open System Interconnection (OSI) reference model Layer 2, addresses are used for local transmissions between devices that are directly connected. The use of Layer 3 addresses is for indirectly

connected devices in an internetwork environment. Each network uses addressing to identify and group devices so that transmissions can be sent and received. [3]

For devices to be able to communicate with each other when they aren´t part of the same network, the 48-bit MAC address must be mapping to an IP address. Some of the Layer 3 protocols used to perform the mapping is:

- Address Resolution Protocol (ARP);
- Reverse ARP (RARP);
- Serial Line (SLARP);
- Inverse ARP

The ARP was developed to enable communications on an internetwork and is defined by RFC 826. Layer 3 devices need ARP to map IP network addresses to MAC hardware addresses so that IP packets can be sending across Network. [3]

Before a device sends a datagram to another device, it looks in its ARP cache to see if there are a MAC address and corresponding IP address for the destination device. If there is no entry, the source device sends a broadcast message to every device on the network. Each device compares the IP address to its own. Only the device with the matching IP address replies to the sending device with a packet containing the MAC address for the device (except in the case of "proxy ARP"). The source device adds the destination device MAC address to its ARP table for future reference, creates a datalink header and trailer that encapsulates the packet, and proceeds to transfer the data. [3]

The ARP request message has the following fields:

- HLN, Hardware address length. Specifies how long the hardware addresses are in the message. For IEEE 802 MAC addresses (Ethernet) the value is 6.
- PLN, Protocol address length. Specifies how long the protocol (Layer 3) addresses are in the message. For IPv4, the value is 4.
- OP, Opcode specifies the nature of the message by code:
    - ARP request
    - ARP reply
    - through 9 -- RARP and Inverse ARP requests and replies
- SHA, Sender hardware address specifies the Layer 2 hardware address of the device sending the message.

- SPA, Sender protocol address specifies the IP address of the sending device.
- THA, Target hardware address specifies the Layer 2 hardware address of the receiving device.
- TPA, Target protocol address specifies the IP address of the receiving device.

3 Types of ARP

3.1 Arp Caching

All operating systems maintain ARP caches that are check before sending an ARP request message. Each time a host needs to send a packet to another host on the LAN, it first checks its ARP cache for the correct IP address and matching MAC address. The addresses will stay in cache for a couple of minutes. You can display ARP entries in Windows by using the arp -a command. [4]

3.2 Static and dynamic entries in the ARP Cache

The ARP cache takes the form of a table containing matched sets of hardware and IP addresses. Each device on the network manages its own ARP cache table. There are two different ways that cache entries can put into the ARP cache [5]:

- Static ARP Cache Entries: These are addressing resolutions that are manually added to the cache table for a device and are keeping in the cache on a permanent basis. Static entries are typically managed using a tool such as the arp software utility.
- Dynamic ARP Cache Entries: These are hardware/IP address pairs that are added to the cache by the software itself resulting in successfully-completed past ARP resolutions. They are kept in the cache only for a period and are then removing.

A device's ARP cache can contain both static and dynamic entries, each of which has advantages and disadvantages. However, dynamic entries are used most often because they are automatic and don't require administrator intervention.

3.3 Inverse ARP (InARP)

The inverse Arp is the opposite of ARP. Instead of using Layer-3 address (IP address) to find MAC address, Inverse ARP uses MAC address to find the IP address. This same method is only used for device configuration and is enabled by default in ATM (Asynchronous Transfer Mode) networks. Inverse ARP is used to find Layer-3 address from Layer-2 address (DLCI in frame relay) dynamically mapping local DLCIs to remote IP addresses when you configure Frame Relay. When using inverse ARP, we know the DLCI of a remote router but do not know its IP

address sending a request to obtain that IP address and map it to the Layer 2 frame-relay DLCI. [6]

3.4 Reverse ARP (RARP)

Reverse ARP is a networking protocol used by a client machine in a local area network to request its Internet Protocol address (IPv4) from the gateway-routers ARP table. The network administrator creates a table in gateway-router, which is used to map the MAC address to corresponding IP address. When a new machine is configuring or any engine which doesn't have memory to store the IP address, needs an IP address for own use[7].

3.5 Proxy ARP

Proxy ARP is the technique in which one host, usually a router, answers ARP requests intended for another machine. By "faking" its identity, the router accepts responsibility for routing packets to the "real" destination. Proxy ARP can help machines on a subnet reach remote subnets without the need to configure routing or default gateway. [8]

3.6 Serial Line Address Resolution Protocol (SLARP)

Serial Line Address Resolution Protocol (SLARP) requests. AutoInstall will use the first available method (DHCP, BOOTP, RARP, or SLARP) for configuration. If all LAN interface configuration options fail, AutoInstall will attempt to configure an available serial interface using SLARP. Though DHCP is the preferred method for AutoInstall over LAN interfaces, these other options remain enabled to ensure backward compatibility with older network topologies. SLARP is an extension of Cisco HDLC, where if the remote router of the serial link does not yet have a config saved in NVRAM, the router will SLARP on the connected serial link to obtain a valid IP address. This feature benefits the Cisco autoinstall feature where an admin can connect a recent router to the network and have limited connectivity with little effort.[8]

3.7 Authorized ARP

Authorized ARP addresses a requirement of explicitly knowing when a user has logged off, either voluntarily or due to a failure of a network device. It is implemented for Public wireless LANs (WLANs) and DHCP.[3]

4 ARP Spoofing

Address Resolution Protocol (ARP) spoofing attack is a type of network attack where an attacker sends fake Address Resolution Protocol (ARP) messages inside a Local Area Network (LAN), with an aim to deviate and intercept network traffic.

In normal Address Resolution Protocol (ARP) operation, when a network device sends an ARP request (as broadcast) to find a MAC address corresponding to an IPv4 address, ARP reply comes from the legitimate network device which is configuring with the IPv4 address which matches the ARP request.

The ARP reply is caching by the requesting device in its ARP table. A network attacker can abuse Address Resolution Protocol (ARP) operation by responding ARP request, posing that it has the requested IPv4 address.

Once the attacker's MAC address is mapping to an authentic legitimate IPv4 address, the attacker will begin receiving any data that is intended for that legitimate IPv4 address.

Now the attacker can launch a man-in-the-middle attack can start capturing the network traffic for any sensitive user data. [9]

Arp Spoofing is different from Spoofing because the spoofing is fraudulent or malicious practice in which communication is sending from an unknown source disguised as a source known to the receiver. [10]

4.1 ARP Spoofing - Attacks

A man-in-the-middle attack requires three players. There's the victim, the entity with which the victim is trying to communicate, and the "man in the middle," who are intercepting the victim's communications. Critical to the scenario is that the victim isn't aware of the man in the middle.

How does this play out? Let's say you received an email that appeared to be from your bank, asking you to log in to your account to confirm your contact information. You click on a link in the email and are taking to what appears to be your bank's website, where you log in and perform the requested task.

In such a scenario, the man in the middle (MITM) sent you the email, making it appear to be legitimate. The attacker also created a website that looks just like your bank's website, so you wouldn't hesitate to enter your login credentials after clicking the link in the email. But when you do that you're not logging into your bank account, you're handing over your credentials to the attacker.[11]

A "denial of service" or DOS attack is used to tie up a website's resources so that users who need to access the site cannot do so. Many major companies have been the focus of DOS attacks in recent years. Because a DOS attack can be easily engineered from nearly any location, finding those responsible can be next to impossible.

Unlike a virus or malware, a DOS attack doesn't depend on a special program in order to run. Instead, it takes advantage of a natural vulnerability in the way computer networks communicate.

Here's an example: suppose that you wish to visit an e-commerce site in order to shop for a gift. Your computer sends a small packet of information to the website. This packet works as a "hello" – basically, your computer says, "Hi, I'd like to visit you, please let me in."

When the server receives your computer's message, it sends a short one back, saying, in a sense, "Okay, are you real?" Your computer responds – "Yes!" – and communication is established. The website's homepage then pops up on your screen, and you can explore the site. Your computer and the server continue communicating as you click links, place orders, and carry out other business.

In a DOS attack, a computer is rigged to send not just one “introduction” to a server, but hundreds or sometimes thousands. The server—which cannot tell that the “introductions” are fake sends back its usual response, waiting up to a minute in each case in order to hear a reply. When it gets no reply, the server shuts down the connection, and the computer executing the attack repeats, sending a new batch of fake requests.[12]

A distributed denial-of-service (DDoS) attack is an attack in which multiple compromised computer systems attack a target, such as a server, website or other network resources, and cause a denial of service for users of the targeted resource. [13]

The session hijacking is a type of web attack. It works based on the principle of computer sessions. The attack takes advantage of the active sessions.

The session refers to a certain time period that communication of two computer systems or two parts of a single system takes place. The session will be valid up to the end of the communication. In some cases the session is user-initiated. However, many of the active sessions will be hidden from the users. They will not know when a session starts and ends. The session is an important factor in Internet communications.

Coming to the session hijacking, as we’ve seen earlier, the attacker uses the active session for implementing the attack. For most Internet communications, authentication will be needed. Authentication can be done in different methods. The most used method is the user be asked to enter a predefined username and password by the website. When the user enters this credentials, the system will check the same with the stored details. If the entered details match with the stored details, the system grants access to the particular user to the particular database or part of the website.[14]

4.2 Arp Spoofing - Defenses

Most importantly, always make sure you're browsing securely. By encrypting the traffic between the network and your device using browsing encryption software, you can help fend off potential man in the middle attacks.

Always make sure the sites you're visiting are secure. Most browsers show a lock symbol next to the URL when a website is secure. If you don't see this symbol, check to see if the web address is preceded by "https." The "S" stands for secure, and this ensures your data won't be open to hacker interception.

Using a firewall is also a reliable way to help defend your browsing data. Although it's not foolproof, a firewall provides an extra layer of security when you're using public Wi-Fi. If you browse public Wi-Fi often, it's prudent to set up a virtual protected network (VPN). This type of network secures your traffic and makes it much more difficult for hackers to intercept it.

Keep your security solution software up to date. Cybercriminals won't stop adapting and honing their craft—and neither should the good guys. By ensuring your security solution is up to date, you always have access to the latest cutting-edge tools to keep a watchful eye on your online activity for safe, fun, secure browsing. [15]

The most effective way to protect against the impact of DoS attacks is to stop them before they even reach a company’s network. That means partnering with the contracted ISP to block the

attack at the gateway. This blunts their impact by protecting even network border devices from being overwhelmed by the flood of malicious traffic. Many ISPs offer a “clean pipes” service-level agreement that commits to a guaranteed bandwidth of legitimate traffic rather than just total bandwidth of all traffic.

In addition to this possibility it exist protection devices to further guard their networks against attack. These devices sit at the network perimeter and process traffic before it. They may be used in conjunction with a clean-pipes ISP service or a stand-alone solution when ISP protection is not available. Solutions in this category include the CheckPoint DDoS Protector and Radware DefensePro. [16]

The following are some of the ways to safeguard against session hijacking[17]:

- Use secure shell (SSL) to create a secure communication channel
- Use encrypted protocols that are offered at OpenSSH suite
- Pass authentication cookies over the HTTPS secure connection
- Implement the log-out functionality for each user to invalidate the session
- Generate different session ID after each successful login and logout
- Always pass the encrypted information between the users and the web servers
- Use string or long random variables as a session key
- Use different username and password for each account
- Configure the suitable internal and external spoof rules on gateways
- Do not transport session ID within the query string
- Limit incoming connections and Minimize remote access

5 Demonstration

In this section, I will demonstrate at the practical level how to make an email spoofing attack for this, we will use a PHP script. This script is very easy to use as I'll demonstrate below. This tool can be found in GitHub with the name email-spoofer being that the author of this tool has the name of Shumbham Badal in GitHub being, for this reason, the creator of this script.

Fig.1. Email Spoofer tool.

For the test, I will put a fake email so that it is presented with the same "spoof@gmail.com" and I will send a message "You were hacked!" and send it to one of my emails thus ensuring that it works.

Fig.2. Email Spoofer pre-filled

Fig.3. Email received.

With this test we can note that, in fact, this same script works well as it becomes easy to use, being only necessary to know the email of the person we want to attack. This script consists of 3 "pages" of PHP, the first one called index is the main page, in the other words, how the tool appears to us when opening. The second page consists of the direct-mailer which is the page where we will show the result of the sending, that is, if the email was sent or if there was an error. The third page is where the sending of this same email, is an intermediate page and is dubbed form-mailer. On the other hand, Gmail, for example, can send these emails directly to the spam box, ignoring them.

6 Conclusion

With this document, it is possible to realize that there are several strategies of attack to a certain user, in contrast, it is also demonstrated all the types of defenses for each one of these possible attacks to do. Besides these aspects, that give more emphasis to the defense as well as the attack by the hacker, one also begins to perceive better all types of ARP as well as the differences between each of the same.

Finally, with this final demonstration, we can see that this type of attack is easy to make the user aware so he can fight to always have the protection of the same in mind.

References

1. Whalen, Sean, “An Introduction to ARP Spoofing”, (2001).
2. Techopedia.com Homepage “Address Resolution Protocol (ARP)”, accessed (2018).
3. Cisco.com Homepage “IP Addressing: ARP Configuration Guide, Cisco IOS Release 15M&T”), accessed (2018).
4. study-ccna.com Homepage “ARP (Address Resolution Protocol) explained”, accessed (2018).
5. tcpipguide.com Homepage “ARP Caching”, accessed (2018).
6. GeeksforGeeks.org Homepage ” Computer Network | ARP, Reverse ARP(RARP), Inverse ARP(InARP), Proxy ARP and Gratuitous ARP”, accessed (2018).
7. Cisco.com Homepage “Proxy ARP”, accessed (2018).
8. Link4securenetwork.blogspot.com Homepage “SLARP (Serial Line Address Resolution Protocol)”, accessed (2018).
9. Omnisecu.com Homepage “ARP Spoofing Attack”, accessed(2018).
10. Techopedia.com “Spoofing”, accessed(2018).
11. Us.Norton.com Homepage “What is a man-in-the-middle attack?”, accessed (2018).
12. Us.Norton.com Homepage “DOS Attacks explained”, accessed (2018).

13. SearchSecurity.techtarget.com “Distributed denial of services (DDoS) attack”, accessed(2018).
14. Internetsaver.net Homepage “What is session Hijacking and how to prevent it?”, accessed(2018).
15. Kaspersky.com Homepage “Defend yourself from Man in the middle attack”, accessed(2018).
16. Biztechmagazine.com Homepage “The three elements of defenses against Denial-of-Service attacks”, accessed(2018).
17. Securitycommunity.tcs.com Homepage “Session Hijacking: Introduction and measures to safeguard”, accessed(2018).

Creating GDPR compliant interpretable models

Pedro Strecht[0000−0002−1077−0346]

INESC TEC/Faculdade de Engenharia, Universidade do Porto
Rua Dr. Roberto Frias, 4200-465 Porto, Portugal
pstrecht@fe.up.pt

Abstract. The enforcement of the General Data Protection Regulation in the European Union as of May 25, 2018 pressured organizations to take action regarding the way they process personal data, under threat of the application of very heavy penalties. The GDPR includes recommendations to be adopted in order for organizations to comply. Scientific research often makes use of interpretable models to describe or predict some phenomenon of interest. The datasets used to create them may contain information in the scope of personal data. This paper frames the creation of interpretable models with the privacy principles described in the GDPR and points out the specific safeguards to be deployed in the operations of data extraction and further processing to foster the conformity with it.

Keywords: GDPR compliance · Interpretable models · Safeguards

1 Introduction

The General Data Protection Regulation (hereinafter referred to as GDPR) [1] is a large and dense document, consisting of 173 recitals (guidances) and 99 articles (requirements), was approved by the European Commission (EC) on 27 April 2016 and is law from 25 May 2018 (replacing the previous EC Data Protection Directive of 1995). The GDPR deals with the protection of personal data of European Union (EU) citizens and also applies to organisations doing business with the EU.

A major change to previous legislation is the requirement to notify a supervisory authority of a personal data breach whenever there is a risk to the rights and freedoms of people. Organizations have to this in a period no longer than 72 hours. Fines for non-compliance are increased to a maximum of 4% of global turnover. Actual penalties will depend on a number of factors including the cause and size of the breach, the controls in place and the degree of co-operation with the supervisory authority.

The goal of this paper is to frame the creation of interpretable models with the privacy principles described in the GDPR and identifies the measures to promote the compliance with it. The remainder of this paper is structured as follows. Section 2 presents the fundamental concepts described by the GDPR. Section 3 introduces interpretable models and discusses the problem of exposing personal data in them. Section 4 recommends measures to adopt in order to promote the compliance with the GDPR. Section 5 concludes with a few remarks.

2 Fundamentals concepts of the GDPR

2.1 Key terms

The pursuance of the GDPR reinforced the need to clarify a number of key concepts, listed in Article 4, of which the following stand out:

- Personal data is any information relating to an identified or identifiable person, the data subject. One that can be identified, directly or indirectly, in particular by reference to an identifier. Examples are name, an ID number, location data, physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

- Processing is any operation which is performed on personal data, whether or not by automated means, such as collection, organization, storage, adaptation, retrieval, use, disclosure by transmission, or destruction.

- Controller is the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data. – Processor is a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller.

2.2 Principles relating to processing of personal data

In Article 5, the GDPR establishes seven privacy principles that underpin the legislation, laid out in Fig. 1. The principle of lawfulness, fairness and transparency states that the controller must have a lawful reason for collecting the personal data and must do it in a fair and transparent way (stating in clear terms to the data subject). The principle of purpose limitation declares that personal data can only by used for the reason it was collected. The principle of data minimisation enacts that the amount of collected personal data should be kept to the minimum necessary to perform the processing activities. The principle of accuracy determines that personal data must be kept up to date and any inaccuracies should be dealt with as soon as possible. The principle of storage limitation establishes that personal data shall be kept for no longer than necessary for the purposes for which it is being processed. The principle of integrity and confidentiality states that personal data must be protected from loss or tampering (safety) or unauthorized access (security). A further principle of accountability implies that the controller should be able to prove that is complying with the six previous principles.

2.3 Rights of the data subject

The GDPR foresees a number of rights of the data subject over his/her personal data. An organisation has to ensure that provides the mechanisms to allow him/her to exercise these rights, described in Articles 12 to 22 and summarized in Fig. 2.

Fig.1. Privacy principles

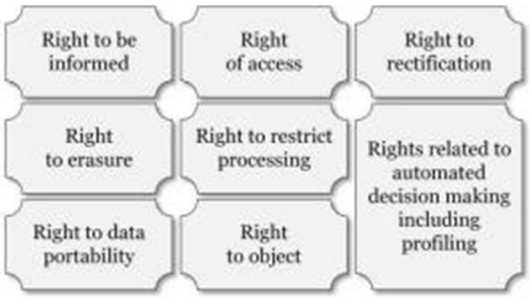

Fig.2. Rights of the data subject

The right to be informed means that the data subject has to be clearly told of what is the data to be collected and what we will be done with it while the right of access means that the data subject may inquire about details about the data that has been collected about him/her.

There are a couple of rights about data quality and maintenance, namely, the right to rectification meaning that the data must corrected by the controller as soon as any inaccuracy is discovered and even the right to erasure (or right to be forgotten) implying that the data subject may request for the data to be erased (if there are no longer lawful rights to hold it).

The data subject also has rights to control or even stop his/her personal from being processed. The right to restriction of processing means that the data subject may restrain the scope of the processing of his/her personal data by a number of reasons (although the controller may still hold it) while the right to object means that the data subject may invoke objections to stop the controller from processing his/her personal data altogether.

The right to data portability means that the data subject may receive the personal data concerning him/her in a structured, commonly used and machinereadable format. Is also implies that the data subject has the right to transmit it to another controller without hindrance from the previous controller. The rights related to automated decision making including profiling means that the data subject must be able to choose if decisions concerning his/her matters are made by human intervention instead of an algorithm.

2.4 Lawful basis of processing personal data (lawfulness)

A pivotal topic to comply with the GDPR is for the controller to ensure that it is processing data rightfully. According to Article 6, for the processing of personal data to be lawful, it must meet at least one of the criteria in Fig. 3. Therefore, it is the responsibility of the controller to establish which of the criteria applies in any given situation.

Consent of the data subject is a lawful basis if he/she has given permission to the processing of personal data for one or more specific purposes. According to Article 4, a consent has to be a freely given, specific, informed and unambiguous indication of the data subject's wishes by a statement or by a clear affirmative action, meaning agreement to the processing of personal data relating to him/her. Vital interests is a lawful basis clarified by Recital 46 as processing necessary to protect an interest essential for the life of the data subject or that of another natural person. Contract is a lawful basis when processing is necessary for the performance of a contract or pre-contractual arrangements between the data subject and the controller.

Compliance with a legal obligation is a lawful basis if processing is done to fulfill an obligation under the law of the country in which the controller carries on business. Public interest is a lawful basis if processing is necessary for the performance of public functions and powers that are set out in law or to perform a specific task in the public interest that is set out in law. Legitimate interests of the controller is a lawful basis if processing is necessary for the purposes of justifiable interest of the controller (as long as it does not affect the data subjects rights and freedoms).

Fig.3. Lawful basis of processing personal data

Although controllers often request consent from data subjects, most of the time it is unnecessary as a significant proportion of the personal data in organisations processes does not require it. Contractual (such as providing services to customers), legal (such as paying employees or dealing with the tax authority) and legitimate interests (such as collection of fingerprints for access to facilities) are more appropriate lawful basis. In fact, consent should be the last lawful basis to be invoked, because the process of obtaining and maintaining it involves changes to business processes and systems.

2.5 Personal data breaches

The GDPR defines a personal data breach in Article 4 as a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed.

According to the Guidelines on Personal data breach notification under regulation 2016/679 [2] by the Article 29 Working Party (WP29), it should be clear that a breach is a type of security incident. The GDPR only applies where there is a breach of personal data, consequently, while all personal data breaches are security incidents, not all security incidents are necessarily personal data breaches. In its Opinion 03/2014 report [3] on breach notification, the WP29 explained that breaches can be categorised as breach of confidentiality when there is unauthorized or accidental disclosure or access to personal data; breach of availability when there is unauthorized or accidental disclosure loss of access or destruction of personal data; breach of integrity when there is unauthorized or accidental disclosure alteration of personal data. Furthermore, Article 33 states that a personal data breach has to notified to a supervisory authority unless it is unlikely to result in a risk to the rights and freedoms of natural persons. At the other end, when it is likely to result in a high risk to the rights and freedoms of natural persons, the controller must also notify the data subject without undue delay, as pointed out in Article 34.

3 Personal data in interpretable models

Creating models is a scientific activity, that can be carried out to fulfill either research purposes or statistical purposes. The models are created through data analysis which implies the collection of data, assembled in data sets. If these contain personal data, then creating models can be considered "processing" as defined by the GDPR.

Usually personal data is laid out in physical documents or stored in operational systems. One example are Enterprise Resource Planning (ERP) software systems which offer integrated management of core business processes, often in real-time. Another are University Information Systems (UIS) in Higher Education Institutions. In both there are databases providing convenient sources of personal data, paving the way to the creation of prediction models, one application of automatic knowledge discovery from databases (or data mining) [4].

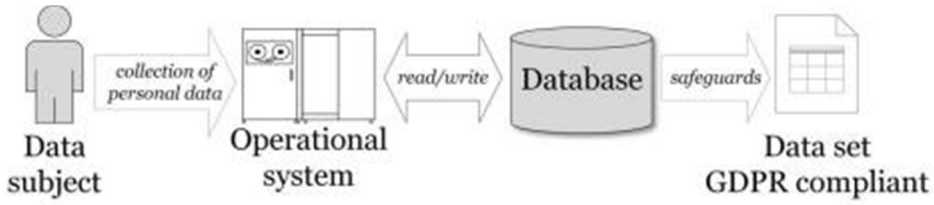

Fig.4. Collecting personal data to data sets

As depicted in Fig. 4, the process of collecting personal to data sets being by a data subject providing personal data to a controller which, in turn, processes into an operational system and stores it in a database. Managers, however, are able to recognize that a great wealth of information is being accumulated over the years and can be explored for more possibilities than the sheer storage or process support. Therefore, it is possible to create data sets from databases using tailored queries to extract relevant information for specific data analysis. Inevitably, this informations also contains personal data intertwined with data related to the process under analysis.

A challenge is to ensure that the assembled data sets are "GDPR compliant". In others words, the personal data included in them does not reveal who are the data subjects. The GDPR pinpoints that safeguards should be used when collecting data (discussed in subsection 4.1). Firstly, however, it is essential to introduce a few fundamental concepts about interpretable models and how they can expose personal data.

3.1 Interpretable models and related concepts

A variable x (also referred to as attribute or feature), takes values from a range of values (either discrete, as a set of values, or continuous, as a set of numbers limited by lower and upper bounds). An example (also referred to as an instance or observation), is a fixed ordered set of values of different variables. Examples may refer to entities or occurrences of an event. A dataset is an unordered set of examples. The dimension of the dataset is the number of variables in contains.

A training set, denoted as Tr, is a dataset used to create a prediction model. For this dataset it is useful to distinguish between a set of independent variables, denoted as X (eq. 1), and the target variable (y) to be predicted. The full set of variables (V) includes both (eq. 2).

$$X = \{x1,x2,...\} \quad (1) \qquad V = X \cup \{y\} \quad (2)$$

A prediction model, denoted as M, is a function that maps a set of independent variables to a target variable (eq. 3). The prediction made by the model for the target variable when, given an example, is denoted as ˆ y. A learning algorithm, is used to create a prediction model. The operation is referred to as TrainModel because it learns a model from a training set. It is necessary to specify the learning algorithmL, the training set Tr, the set of independent variables X and the target variable y (eq. 4).

$$\hat{y} = M(X) \quad (3) \qquad M = \text{TrainModel}(L,Tr,X,y) \quad (4)$$

In knowledge discovery from databases, models are created to predict future events or to understand the relationship between variables involved in a phenomena being studied. In the latter case, the characteristic of interpretability is essential, i.e., the models must be understood by humans. These are called white-box models which means that it is possible to navigate through a model to follow the reasoning behind a prediction, as illustrated in Fig. 5. The opposite are black-box models where the details of how the model makes predictions are either hidden or not readable by a human. Although standing very well in prediction quality, they do not have the characteristic of interpretability, i.e., it is not possible to "look inside the box" and understand how a prediction is made, as Fig. 6 suggests.

Decision trees [5] is a common example of an algorithm that induces whitebox models by creating tree-like structures where the independent variables are tested in nodes and the leaves hold the values of the target variable. Although there are several algorithms to create decision trees, the most popular are Classification and Regression Trees (CART) [6] and C5.0 [7].

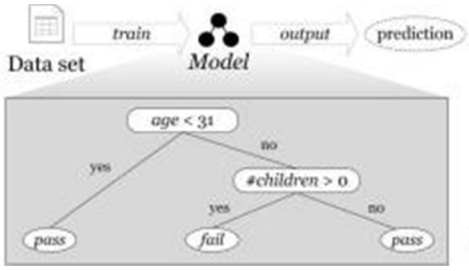

Fig.5. White-box model

Fig.6. Black-box model

Algorithms for constructing decision trees usually work top-down, by choosing a variable at each step that best splits the data set [9]. Different algorithms use different metrics for measuring what is considered "best". These generally measure the homogeneity of the target variable within the subsets. Some examples are Gini impurity, Information gain and Variance reduction. One of these metrics is applied to each candidate subset, and the resulting values are

combined to provide a measure of the quality of the split. In the example of Fig. 5, of all the independent variables in the data set, age and #children were selected to be splitting variables according to some metric. Correspondingly, 31 and 0 are split points, i.e., the values where a split in the tree navigation path takes place. Each path from the root node to the leaves defines a decision rule.

3.2 Exposing personal data in interpretable models

The decision tree in Fig. 5 relates to a course in an university where the goal is to predict if a student is going to pass or fail in it [8]. It is straightforward to grasp that the model first evaluates the age of the student. If it is under 31 then the model predicts that he/she passes, otherwise, it evaluates the number of children. If the student has children (regardless of how many), the model predicts that the student fails, alternatively he/she passes. Nonetheless, a word of caution about models is necessary. The predictions are based on the data used to create it, i.e., they describe the general behavior in a group of students taking a specific course in an academic year. They do not describe what happens with any student taking a course in the university. Decision trees expose personal data in the split points. By selecting a particular value for a variable as a split point, the model reveals that there are examples in the data set with that value. In the example above, the fail rule (age $\geq$ 31$\wedge$#children > 0) reveals that there is at least one student in these circumstances. This is a merely illustrative example of the problem as it is not foreseeable that disclosing the fact that a person's has children could put him or her in any way at risk. However, there may be other variables with the potential to do so.

There is no straightforward solution to the dilemma. On the one hand, including variables describing personal data are necessary because of their promising capability as explanatory variables relating a phenomenon of interest. On the other, the selection of some split points can unintentionally expose private information about the data subjects. Obviously, in the previous example, one can remove the variable #children. Still, that may not be the case in other circumstances where the inclusion of such variable may be instrumental to the quality of the model.

4 Towards creating models GDPR compliant

With the entry into force of the GDPR, it became mandatory to find measures that can reduce the risk of exposing information connected to personal data. In the context of creating interpretable models, these translate into taking particular attention in extracting variables from databases to data sets and also on the data pre-processing tasks.

Article 32, about security of processing, declares that taking into account the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons, the controller and the processor shall implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk. These measures are repeatedly referred to in the GDPR as safeguards with pseudonymisation and encryption of personal data as common examples.

There are, however, further safeguards that can be addressed in data sets towards GDPR compliance. A few more are presented next and are then discussed in the context of the privacy principles more closely related to them. Ultimately, even with a complete set of measures, a visual inspection by a human being will always be necessary to make sure that, due to unforeseen set of circumstances, there is no risk of a data subject being identified in a model.

4.1 Safeguards in the data sets

The following safeguards should be considered when preparing a data set for the creation of an interpretable model. Certainly not all of them will always be necessary and it depends on each problem at hand to decide which ones make sense to be applied.

Identifiers removal. Although it may look rather clear, it is important to stress out that any variables that uniquely identify a person have to be removed from the data sets. In fact, their inclusion does not even make sense as they have no role in a prediction model. Since the goal is to find patterns in the data and extract rules, a variable in which every value only occur in a single example can never qualify as explanatory. Examples are a citizen's card number, tax ID number, or social security number.

Replacement of variables. There are variables related to personal data that can be replaced by others (derived variables) to make it harder to trace to a data subject. The first kind is by calculating a new variable from existing data, e.g., replacing a date of birth by a person's age. The second kind is creating a categorical variable from a continuous variable. In this case, instead of the person's age, the value would be to an age range.

Decrease of granularity. One way to conceal personal data is to diminish the information's level of detail. An example is to replace an address by a parish, a county or a district. Another is to replace a job title by a career. In a university context, replacing the course by the scientific area it relates to.

Dimensionality reduction. Encompasses all processes for reducing the number of independent variables of a data set to obtain a set of principal variables, i.e., those that can actually be used by models. In the context of interpretable models a recommendation is to use feature selection techniques which, besides removing redundant variables, have the potential of simplifying the models and improve generalization by reducing overfitting (an analysis that corresponds too closely or exactly to a particular set of data, and may therefore fail to fit additional data or predict future observations reliably).

Data volume reduction. Consists on efforts for restricting the volume of a dataset into less training examples. One method is data compression in which an aggregation function or another more sophisticated method to downsize the dataset and train a model, described by Yael [10]. Another is instance selection in which examples are selected from a dataset if these are considered sufficient to train a model representative of what would be a model trained with all available data. These have been proposed by Liu [11] and Blum [12] as a strategy to deal with computationally intensive algorithms. The datasets are downsized so that learning is focused on a chosen set of informative examples.

Data anonymisation. Defined in Recital 26 as the process of removing personal identifiers, both direct and indirect, that may lead to an individual being identified. An individual may be

directly identified from their name, address, postcode, telephone number, photograph or image, or some other unique personal characteristic. An individual may be indirectly identifiable when certain information is linked together with other sources of information, including, their place of work, job title, salary, their postcode or even the fact that they have a particular diagnosis or condition. Once data is truly anonymised and individuals are no longer identifiable, the data will not fall within the scope of the GDPR and it becomes easier to use.

Data pseudonymisation. Defined in Article 4 as the processing of personal data in such a way that the data can no longer be attributed to a specific data subject without the use of additional information, as long as such additional information is kept separately and subject to technical and organizational measures to ensure non-attribution to an identified or identifiable individual. consists on a de-identification procedure by which personally identifiable information fields within a data record are replaced by one or more artificial identifiers, or pseudonyms. A single pseudonym for each replaced field or collection of replaced fields makes the data record less identifiable while remaining suitable for data analysis and data processing.

Data encryption. It is a process of cryptography which consists in encoding a message or information in such a way that only authorized parties can access it and those who are not authorized cannot. Encryption does not itself prevent interference, but denies the intelligible content to a would-be interceptor, therefore, is a security measure in interpretable models.

4.2 Compliance with the privacy principles

One way to prove compliance with the GDPR and its underlying demands is to align the processing operations with the privacy principles described in Article 5 (introduced in subsection 2.2). This section revisits a few of those principles in the light of the creation of interpretable models, in particular by discussing what needs to be done to foster conformity.

Principle of lawfulness, fairness and transparency. Fairness and transparency are directly linked with the right to be informed and the right of access. Therefore, it has to be traced to the time when the data was first collected for a specific purpose.

Lawfulness has to be guaranteed under a valid lawful basis. The controller may resort to legitimate interests as long as it ensures that the model does not allow the identification of any data subject. Consent, although remaining an alternative, may be infeasible due to frequently large number of examples that a data set needs to include in order to create reliable models. This is reinforced by Article 6, stating that the processing for a purpose other than that for which the personal data have been collected is not based on the data subject's consent. The controller shall, in order to ascertain whether processing for another purpose is compatible with the purpose for which the personal data are initially collected take into account, among others the existence of appropriate safeguards, which may include encryption or pseudonymisation.

Additionally, according to Recital 50, the processing of personal data for purposes other than those for which the personal data were initially collected should be allowed only where the processing is compatible with the purposes for which the personal data were initially collected. In such a case, no legal basis separate from that which allowed the collection of the personal data is required.

Principle of purpose of limitation. As creating models can be framed in the context of research or statistical purposes, it fits into the directives of Recital 156 which states that the processing of personal data for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes should be subject to appropriate safeguards for the rights and freedoms of the data subject. Processing may be carried out when the controller has assessed the feasibility to fulfill those purposes by processing data which do not permit or no longer permit the identification of data subjects, provided that appropriate safeguards exist.

Principle of data minimization. Keeping collected data to the strict minimum necessary must be looked upon both in the perspective of the number of variables used (dimension of the data set) and the number of examples (volume of the data set). Applying dimensionality reduction may cause the variables involving personal data to be rejected. However, there must be avoided any statistical procedures that change the variables themselves (also known as feature projection). An example is Principal Component Analysis (PCA) that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables (the principal components). If this is done, the pivotal characteristic of interpretability of the models is lost, because the variables become devoided of their business meaning. Data reduction techniques, although motivated by the problem of training models from very large datasets, hold the potential that a model can be obtained using less data, which is aligned with the principle of data minimization. Therefore, their deployment may help to decrease the volume of the data set necessary to train a model.

Principle of storage limitation. It is important to acknowledge that the GDPR does not set specific time limits for different types of data, but it makes organisations responsible for determining on how long they need to hold the data for their specified purposes. Recital 39 decodes, in a clear and concise manner, that the limits imposed by Article 5, by stating that the period for which the personal data is stored should be limited to a strict minimum and that time limits should be established by the data controller for deletion of the records (referred to as "erasure" in the GDPR) or for a periodic review. Article 5 identifies the circumstances when personal data may be kept for longer than necessary for the purposes for which it is being processed as "personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89 subject to implementation of the appropriate technical and organisational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject ('storage limitation')." . These measures could include the deletion of the data sets after creating the models, while keeping the latter.

4.3 Ensuring the rights of the data subjects

A model can be used in a decision support tool for automated decision making. The right to object and the rights related to automated decision making including profiling demand special attention concerning this situation. It is necessary to ensure that the data subjects who have invoked these rights are no longer included in the data sets used to create models. Thus, firstly it is necessary to scan the existing data sets in order to remove data subjects and then re-create the models. Next, in creating new data sets, it is mandatory to include a new condition in the extraction queries specifically to reject the data subjects who have given this indication.

5 Conclusions

Organizations have been preparing for the requirements of the GRPR, which has led to changes in processes and the introduction of measures as safeguards. The main consequence of non-compliance is the possibility of data breaches being punished with heavy fines. Creating interpretable models is a processing operation which, under certain circumstances, may inadvertently disclose personal data leading to the identification of a data subject. It is necessary for organizations pursuing such activities (e.g., including models in decision support systems) to take steps to prevent this from happening.

It is imperative to act both on the way personal information is extracted from databases and also the one already existing in data sets and physical supports. Also the rights of data subjects have to be fulfilled, by removing them permanently from those very data sets. Certainly, not a simple task, but at this point, organizations have no choice but to revisit all their processing operations involving personal data and making sure lawfulness is guaranteed.

References

1. European Union. Regulation 2016/679. Official Journal of the European Communities, 2014(March 2014):1–88, 2016.
2. Article 29 Data Protection Working Party. Guidelines on Personal data breach notification under regulation 2016/679. Technical Report October, 2017.
3. Article 29 Data Protection Working Party. Opinion 03/2014 on Personal Data Breach Notification. Technical Report March, 2014.
4. J. Han, M. Kamber, and J. Pei. Data Mining: Concepts and Techniques. Morgan Kaufmann, San Francisco, 2011.
5. J. Quinlan. Induction of Decision Trees. Machine Learning, 1(1):81–106, 1986.
6. L. Breiman, J. Friedman, R. Olshen, and C. Stone. Classification and Regression Trees. Chapman and Hall/CRC, 1984.
7. M. Kuhn, S. Weston, N. Coulter, and J. Quinlan. C50: C5.0 Decision Trees and Rule-Based Models. R package version 0.1.0-16, 2014.
8. P. Strecht, J. Mendes-Moreira, and C. Soares. Merging Decision Trees: A Case Study in Predicting Student Performance. In X. Luo, J. Yu, and Z. Li, editors, Advanced Data Mining and Applications, Lecture Notes in Computer Science, pages 535–548. Springer International Publishing, 2014.
9. Lior Rokach and Oded Maimon. Top-Down Induction of Decision Trees: A Survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 35(4):476–487, 2005.
10. B. Yael and T. Elad. A Streaming Parallel Decision Tree Algorithm. Journal of Machine Learning Research, 11:849–872, 2010.
11. H. Liu and H. Motoda. On Issues of Instance Selection. Data Mining and Knowledge Discovery, 6(2):115–130, 2002.
12. A. Blum and P. Langley. Selection of Relevant Features and Examples in Machine Learning. Artificial Intelligence, 97(1-2):245–271, 1997.

A data encryption application: Development Proposal

Diogo Vilas Boas

Lusofona University of Porto, Portugal
diogovilasboas98@gmail.com

Abstract. This paper follows the development of a particular software which uses a cryptographic algorithm to encrypt and decrypt data. Here, we will us e RSA (Rivest Shamir Adleman). The first part of the paper will discuss the mathematics involved in the creation of the keys and in the encryption and decryption of the data. While in the second part will be looking into the integration of mathematics involved when developing the software. This part will highlight some code of the project and provide an explanation of the methodology used. Concluding the paper with the results of performance tests done to the software and with a commentary on those results.

Keywords: RSA, Encryption, Decryption, Cryptography, Software, Public Key.

1 Introduction

Cryptography has as aim enabling users to maintain communication in a secure way over a channel which is not reliable to guarantee privacy and/or authenticity. In the ideal world, we wouldn't be able to decipher or make modifications without the permission of the person who encrypted the message [1].

It is not known for sure when the cryptography appeared, but there is some evidence of it in the earliest types of writing. In some way, every one, early on, wanted to have a secret language where they could communicate without the information being comprised. This was particularly useful in times of diplomacy and in times of war [2].

Normally, throughout history, you would have a password or a method, if it was used backward, you could decrypt the message that was encrypted. This wasn't the best way to secure a message because you would have to share that password or method so that other people could decrypt it [2]. The problem comes with this exchange. If it doesn't assure the secrecy of the password or method, encrypting it be comes useless.

In spite of Diffie Hellman method being secure to use it to communicate, it didn't implement digital signatures which means the receiver didn't know who or what the source of the encrypted messages was. So, when Ronald Rivest, Adi Shamir, and Richard Adleman read the paper of Diffie Hellman, they realized this flaw and started searching for a mathematical solution to solve this problem. And this was how RSA was born [3].

So, since RSA is such a good algorithm, it was proposed to me to develop one implementation of a software that uses RSA to encrypt and decrypt text, in order to learn a little bit more about the RSA and it’s challenges when anyone tries to implement it.

2 Mathematics of the encryption and decryption

RSA uses mathematics to encrypt and decrypt messages. Therefore, in this section, it will be focused on how those mathematic principles work and explain them.

2.1 Basic Concepts

Prime Numbers are numbers that are only divisible for themselves and by 1.They are like building blocks for numbers,almost all numbers can be created by multiplying prime numbers [4].

Euclidean Division is the process that divides two integers, which results in a quotient and remainder. It normally follows this structure [4]:

$$\frac{dividend}{divisor} = quotient + remainder$$

Modulo Operation also known by modulus gives the remainder of Euclidean division. After giving two positive numbers, it returns a number that ranges between 0 and a unit below of the divisor [4].

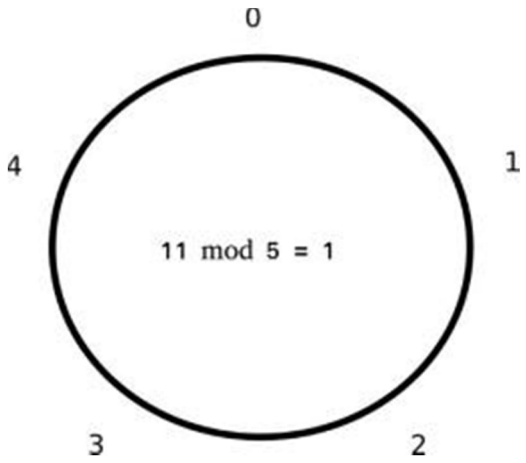

Fig.1. Modulo operation between 11 and 5.

$$dividend \bmod divisor = dividend - (quotient * divisor)$$

Looking at figure 1 and equation 2, we can notice that as the dividend increases the module numbers will repeat. In the case of figure 1 as we increase the the dividend of the module of 5, the module will repeat 0, 1, 2, 3, 4, in this exact order[4].

Greatest common divisor(gcd) says that for certain integers a and b are bigger than zero, than gcd(a,b) returns the largest integer that a and b are divisible by[4].

Euler's theorem says that if n and a are coprime then

$$a\varphi(n) \equiv 1 mod n$$

where φ(n) is Eule's totient function. The notation is explained in the article modular arithmetic. Which returns the number of cop rimes of n. Two numbers are coprime if the greatest common divisor of both of them is 1 [4].

Chinese Remainder Theorem(CRT) is a theorem that solves simultaneous linear congruence's, where the moduli are coprime [5].

The first step is to multiply the modili resulting a number N [5]:

$$N = x1 * x2 * x3...xk$$

Secondly, for each i= 1, 2 ... k, we calculate [5]:

$$yi = \frac{N}{xi}$$

Then, for the same i 's, we calculate [5]:

$$zi \equiv yi^{-1} mod xi$$

And finally:

$$x = \sum ai * yi * zi$$

Where x is solution of the system of congruence's, and x mod N is the only solution modulo N [5].

2.2 Generating Keys

RSA usually needs 5 numbers, p, q, N, e and d [4]. The p and q are really huge prime numbers where the ideal size would be 1024 bits. So the easiest way to generate this numbers is by generating a random number and then verifying if that number is prime, using primality test. If it is not a prime, repeats the process until it is prime [4].

The N is the result of the multiplication of p and q [4].

The e is a number between 1 and φ(N) such that [4]:

$$gcd(e,\varphi(N)) = 1$$

The d is the solution of [4]

$$d \equiv e^{-1} mod N$$

Now that we have the 5 numbers, the public key is the N and e and the private key is the d [4].

2.3 Encrypting and Decrypting

To encrypt message we need to transform the text or file to an array ofnumbers so that we can use RSA [4]. After doing this, we have our message(m) in a number form and then the encrypted cipher (c) is the result of [4]

$$c \equiv m^e mod N$$

3 Development Proposal

My development proposal uses python as main and only programming language and besides that uses PyQt5 which is a graphics interface framework. Now in the next sections, it will be explained the code from the project and what was the though process behind it.

3.1 Generating Keys

```
def generate_keypair(self):
    #a
    p = random.randint(2, 9007199254740991)
    q = random.randint(2, 9007199254740991)
    while not self.prime(p):
        p += 1
    while not self.prime(q) and p!=q:
        q += 1
    #b
    n = p * q
    #c
    phi = (p - 1) * (q - 1)
    #d
    e = random.randint(1, phi)
    g = self.gcd(e, phi)

    while g != 1:
        e = random.randint(2, phi)
        g,y,x = self.gcd(e,phi)
    #e
    d = self.multiplicative_inverse(e, phi)
    #f
    self.p= p
    self.q= q
    self.e= e
    self.d= d
    self.n= n
    #g
    self.publicKey=self.arrayToBase64([e,n])
    #h
    self.dP = d % (p-1)
    self.dQ = d % (q-1)
```

Fig.2. Function generate_keypair

The image above(Fig. 2.) represents the part of the code that generates the public and private keys. But there is many steps to being able to generate them, so it is divided by letters. From the letter a to letter b, two huge prime numbers(p and q) are created by randomly picking up a number between 2 and 9007199254740991 which is , using the random.randint, and then it verifies if it is a prime by using the primality algorithm that is represented by the function prime. Then, if it is not a prime number, it increments one until it is prime.

```
def prime(self, num):
    if num == 2:
        return True
    if num < 2 or num % 2 == 0:
        return False
    for n in range(3, int(num**0.5)+2, 2):
        if num % n == 0:
            return False
    return True
```

Fig.3. Function prime

From the letter b to c, the N is calculated by multiplying the previous prime numbers. From the letter c to d, we setup the phi variable by using the formula.

$$\varphi = (p - 1)(q - 1)$$

From the letter d to e, the e is calculated by choosing a number between 1 and phi(N), and then checking if the gcd(e,phi(N))=1. If it is not, it repeats the process until it is true.

```
def gcd(self,a, b):
    if a == 0:
        return (b, 0, 1)
    else:
        g, y, x = self.gcd(b % a, a)
        return (g, x - (b // a) * y, y)
```

Fig. 4. Function gcd

From the letter e to f, the d, the private key, is calculated by choosing a number that is a multiplicative inverse of e which is calculated with the function multiplicative_inverse.

```
def multiplicative_inverse(self,a, m):
    g, x, y = self.gcd(a, m)
    if g != 1:
        raise Exception('modular inverse does not exist')
    else:
        return x % m
```

Fig 5. Function multiplicative_inverse

From the letter f to g, we attach the numbers to the class, so that we can access them anywhere inside the class RSA(the class that is responsible to apply the RSA algorithm).

From the letter g to h, we create the public key that is the N and the e in base 64, using the function arrayToBase64.

```
def arrayToBase64(self,array):
    length= len(array)
    string=""+array[0].__str__()
    if(length>1):
        for i in range(1,length):
            string+=" "+array[i].__str__()
    return base64.b64encode(string.encode('utf-8'))
```

Fig.6. Function arrayToBase64

Lastly, from the h to the end of the function, the Chinese Remainder Theorem is setup. This will be explained in greater detail in the Decoding section.

3.2 Encrypting

```
def encrypt(self, publickey, plaintext):
    #a
    key, n = self.base64ToNumber(publickey)
    #b
    cipher = [pow(ord(char), key, n) for char in plaintext]
    #c
    cipher=self.arrayToBase64(cipher)
    return cipher
```

Fig.7. Function encrypt

The function above, encrypt (Fig. 7.) , it encrypts a string given a public key. From a to b, it takes the N and the key which is the e,that we mention above, using the function base64ToNumber because the public key is in base 64.

```
def base64ToNumber(self,msg):
    numbers = base64.b64decode(msg).decode('utf-8')
    numbers = ciphertext.split(' ')
    length=len(numbers)
    for i in range(length):
        numbers[i]=int(numbers[i])
    return numbers
```

Fig.8. Function base64ToNumber

From b to c, we encrypt each letter by turning the letter into a number, using the ASCII converter included in the python, and then raising it to the key. After this, it is done the modulus of N, and the result is a cipher encrypted using the RSA algorithm.

Lastly, from c to the end of the function, it converts the array of ciphers to base64 and returns it.

3.3 Decrypting

```
def decrypt(self,ciphertext):
    #a
    ciphertext=self.base64ToNumber(ciphertext)
    #b
    key=self.d
    n=self.n
    #c
    plain=''
    for i in range(len(ciphertext)):
        m1= pow(ciphertext[i], self.dP, self.p)
        m2= pow(ciphertext[i], self.dQ, self.q)
        qInv = (1/self.q) % self.p
        h = qInv * (m1 - m2)
        m = m2 + (h * self.q)

        plain += chr(int(m))
    return ''.join(plain)
```

Fig.9. Function decrypt

The image above show us the function decrypt, that decrypts a ciphertext given.

From a to b, it transforms the string ciphertext to numbers using the function base64ToNumber.

From b to c, it gets the private key and part of the public key and puts it in local variables.

From c to end of the function, it decrypts each numner by implementing the Chinese Remainder Theorem which divides the d into two smaller numbers(dQ and dP from generate_keypair) and then uses Garner's formula. So that, the operations be faster than powering it to d and moduling it to N [6].

4 Implementation and Results

In this section, the tests that were made to the software are going to be shared as their results. And then, I will give my interpretation about them.

4.1 First test

The first test was to encrypt the message "Hello World!!" using the public key of another window of the software. And decrypt it using the other window.

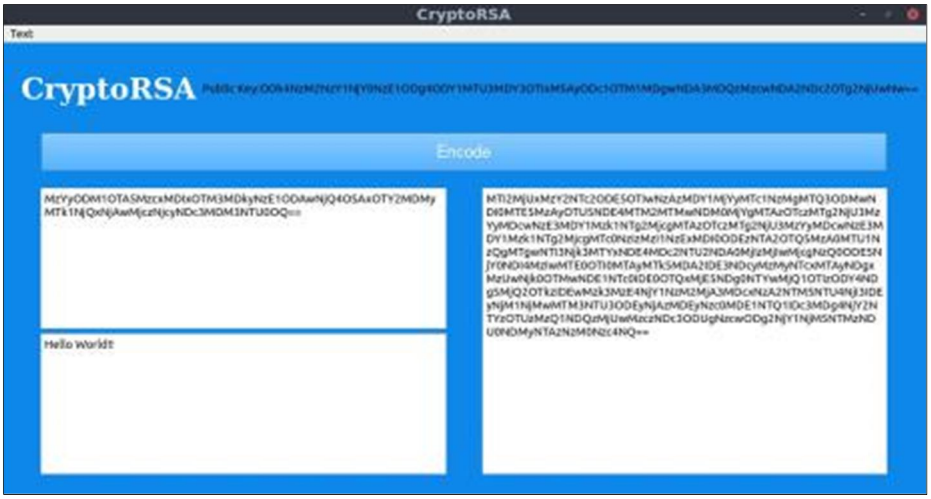

Fig.10. First part of the test

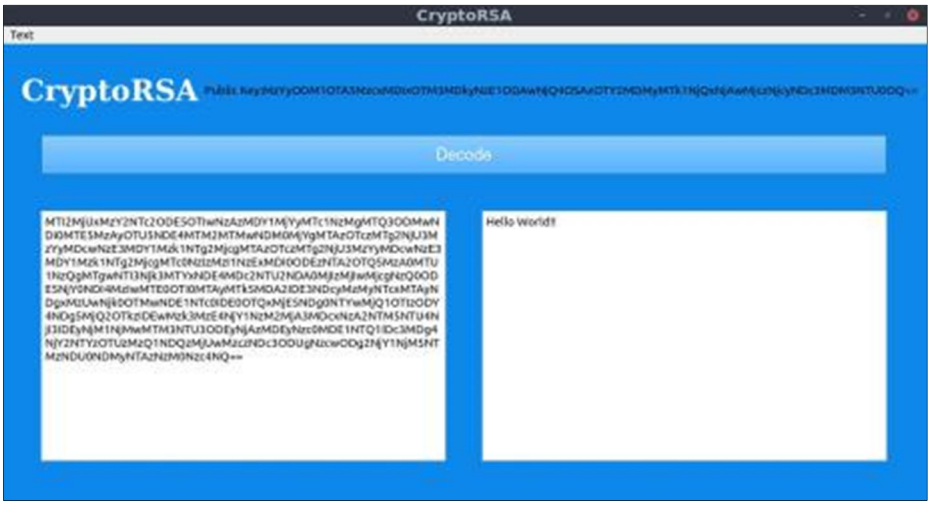

Fig.11. Second part of the test

The images above show the result of the test which successfully passed.

4.2 Second test

The second test consisted in picking the RSA class and adding a main, where it created to RSA objects(a and b) and then randomly generating a random message. And then encrypt the message using b public key using a. After this decrypts using b. Then checks if it the message decrypted is equal to the message originated. If it is, prints “yes”, it is not prints “not”. It repeats this process 100 times.

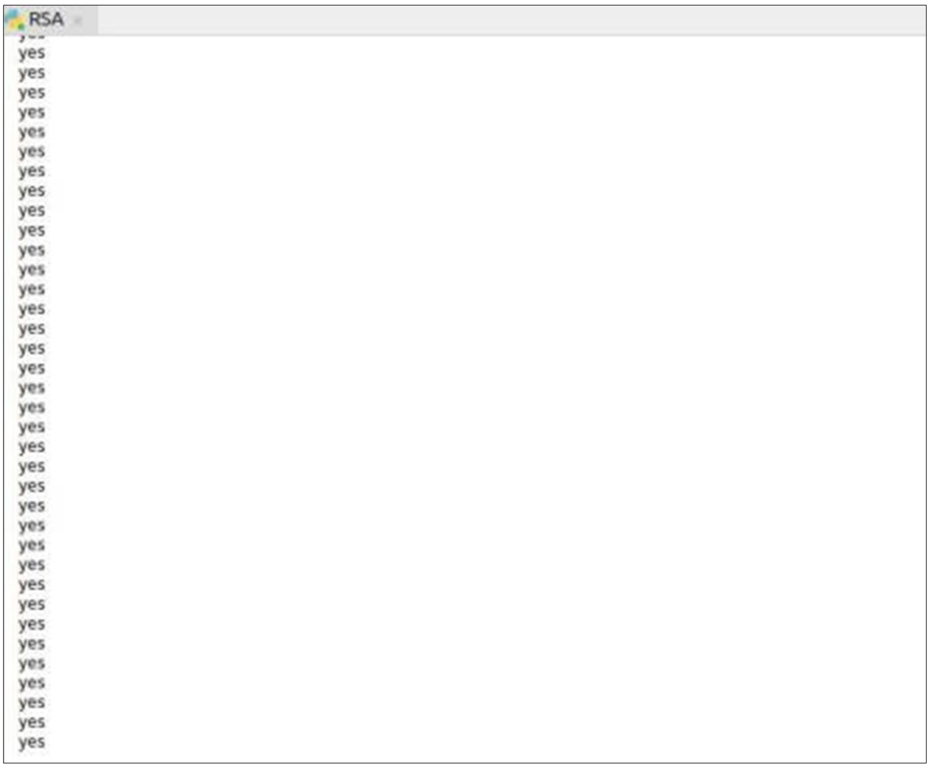

Fig.12. Part of the results of the second test

The image above shows part of the output of the test. The output was all "yes" which means the software passed the test successfully and it shows to be reliable.

4.3 Final Notes

The software seems to be pretty reliable when it comes to encrypting and decrypting. Although, it takes some time launch (5 seconds) because it has to generate the keys, it works in real time after launching. So we can conclude that this software was a success.

5 Conclusion

While in development, many adversities were faced. One of them was choosing the right programming language, where big numbers needed to be supported. If not, it wouldn't be possible making any operations with the numbers generated.

Another adversity was the mathematics of RSA. Although they are simple, they are not very intuitive. Which you need to spend some time understanding the mathematics involved. That is one of the reasons that I put an entire chapter on the mathematical concepts before explaining the method.

The last adversity, it was bugs in code. Although it is predictable that you will have bugs when developing a software, it does not mean they do not give some headaches. So, those where

treated with patience. Which means, many features I would have loved to put in code needed to be cut off, so that it was possible to finish the project in time.

References

1. Coron, J.: "What is cryptography?". IEEE Security & Privacy Magazine Volume 4 issue 1, (2016), doi:10.1109/MSP.2006.29
2. Davies, D.: "A brief history of cryptography". Information Security Technical Report Volume 2 issue 2 ,(1997) , doi: 10.1016/s1363-4127(97)81323-4
3. "Public Key Cryptography (PKC) History", https://www.livinginternet.com/i/is_crypt_pkc_inv.htm. Last acessed 21 Nov 2018
4. Blakley, G.; Borosh, I.: ”Rivest-Shamir-Adleman public key cryptosystems do not always conceal messages” Computers & Mathematics with Applic ations Volume 5 issue 3, (1979) doi:10.1016/0898-1221(79)90039-7
5. Schmid, H.; Mahler, K.: "On the Chinese Remainder Theorem", Mathematische Nachrichten, (1958), doi:10.1002/mana.19580180112
6. "Using the CRT with RSA", https://www.di-mgt.com.au/crt_rsa.html, Last acessed 21 Dec 2018

Papers in alphabetical order

Authors in alphabetical order

Conference EOI :10.11228/dpsc
DPSC2019 Proceedings EOI: http://eoi.citefactor.org/10.11228/dpsc

PRIVACY AND SECURITY CONFERENCE 2018

PRIVACYANDSECURITYCONFERENCE.PT