UNIVERSIDAD NACIONAL DE COLOMBIA

SOCIAL ENGINEERING: PSYCHOLOGY APPLIED TO INFORMATION SECURITY

Iván Del Pozo Falconí

Universidad Nacional de Colombia

Departamento de Ingeniería de Sistemas e Industrial

Bogotá, Colombia

2018

SOCIAL ENGINEERING: PSYCHOLOGY APPLIED TO INFORMATION SECURITY

Iván Del Pozo Falconí

Trabajo de profundización presentado como requisito parcial para optar al título de:

Magíster en Ingeniería de Sistemas y Computación

Director: Mauricio Iturralde, Ph,D.

Co-Director: Felipe Restrepo Calle Ph,D.

Línea de Investigación:

SEGURIDAD INFORMÁTICA

Miembro del grupo de Investigación:

PLaS – Programming Languages and Systems

Universidad Nacional de Colombia

Departamento de Ingeniería de Sistemas e Industrial

Bogotá, Colombia

2018

Agradecimientos

Agradezco al gobierno Colombiano y Ecuatoriano por ser patrocinadores de mi beca y mis estudios.

A mi familia por siempre estar pendiente de mí y por el apoyo incondicional que me brindaron a lo largo de toda mi formación personal y universitaria.

A mis directores y de más profesores que son parte del Departamento de Ingeniería que alguna vez tuve la oportunidad de ser su alumno, los cuales aportaron en mi formación.

A mis amigos cercanos y compañeros universitarios quienes compartieron conmigo días y noches difíciles de estudio y sacrificio, quienes también me alentaron a nunca desertar en ningún momento pese a todas las adversidades que se me presentaron para la realización de mi maestría.

Abstract— Psychology and computer science are two scientific disciplines that focus on identifying the particular characteristics of information processing. The first in the human being and the second in the construction of a technical tool that seeks to emulate the brain: the computer. That is why psychology is strongly tied to the moment for people to choose their passwords. Deceptive advertising often compensates (through money, products and free services or other self-esteem tests) to influence a product or service to appear on your social network. In order to increase its consumption among its followers and also to take personal information without your consent. Due to the increase of the use of social networks, our social engineering strategy can efficiently and effectively show that security is subjective and that a significant percentage of users are vulnerable to deceptive advertisement through the internet. This project is based on the need to prevent attacks of information subtraction by obtaining/decrypting the keys of access or in the worst case obtain directly their passwords to the different web services, bank accounts, credit cards of individuals, based on the information that people exposed or share on their social networks.

This paper also examines how attackers could obtain/decipher their passwords based on personal information obtained from deceptive advertisements implemented through a social network. The advantage of this approach also shows the user password composition providing a better vision of how hackers use the psychology applied to information security.

Key Words— Psychology, Social Engineering, Information security, Deceptive advertising, Passwords, Social networks.

Content

List of Figures

List of Tables

Chapter 1. INTRODUCTION

We presented the motivation to do this investigative work, the identification of the problem, the general and specific objectives. Describing the strategy and the methodology which will be used to develop an ethical attack on one popular social network.

MOTIVATION

Psychology is the scientific study of conduct and experience, of how humans and animals feel, think and learn in order to know how they adapt to the environment that surrounds them (Planck, Development, Wu, Meder, & Nelson, 2017).

Social engineering is the art of manipulating people so they give up confidential information (Ghafir, Prenosil, Alhejailan, & Hammoudeh, 2016); and Social engineering is based on psychology. One of the current techniques of social engineering is deceptive advertising, also known as false advertising, which refers to the use of confusing, misleading, or blatantly untrue statements when promoting a product with a fraudulent intention (Kotenko, Stepashkin, & Doynikova, 2011). That is the reason why we must understand the importance of the relation between social engineering and psychology, and in that way to be able to provide a better vision of how hackers use the psychology applied to information security based on scams or deceptive advertisings in order to be able to carry out illegal acts against us.

Security is subjective and each person will have a different perspective of it, because each one determines their level of risk and evaluates their compensatory strategies against the controls in order to be able to mitigate the risk. Since each person has different weaknesses, this one must be the one who meditates to manage his own measures of protection against the Social Engineering. For this reason it is important to make the reader aware of how someone who is malicious can take advantage of it and the value of information security.

The aim of this work is to increase overall information security readiness by addressing one of the most efficient attacks that exist, attacks against the human element. This is achieved by gaining a working knowledge of the threat named social engineering, evaluating all the most important and relevant attacks and evaluating the human behavior of a specific population where an ethical attack will take place, (Atkins & Huang, 2013) trying to identify strong and weak points about the implementation of this attack.

This project is based also on the need to prevent attacks of information subtraction by obtaining/decrypting the keys of access or in the worst case obtain directly their

passwords to the different services, bank accounts, credit cards or applications of individuals based on the information that a people exposed or share on their social networks. Also raising awareness and warning of the risks associated with misuse of the Internet and in social networks.

PROBLEM IDENTIFICATION

Psychology and computer science are two scientific disciplines that focus on identifying the particular characteristics of information processing, the first in the human being and the second in the construction of a technical tool that seeks to emulate the brain: the computer; that is why psychology is strongly tied to the moment for people to choose their passwords. Choosing a password has largely become a psychology test, with most office workers choosing a word that they believe to sum up their personality and that it will be easy to remember (Atkins & Huang, 2013).

As we may know Internet in general and Social Networks are one of the primary sources of Big Data, which refers to voluminous amounts of data which the majority of the successful organizations can potentially mine and analyze for business gains (Terzi, Terzi, & Sagiroglu, 2017). Rise of the social networking platforms are creating enormous amount of data where all people, children and adults emotions are constantly expressed in real-time" (Aziz, 2016). In times of Big Data and constant wireless connection, everybody must be forced to protect all their digital information and be aware of the importance of cyber security, which refers to mitigate the risk against unauthorized access for people or any attack (Stjohn-green, Piggin, Mcdermid, & Oates, 2015)

There have been several unknowns of Psychological phenomena, with respect of human behavior and psychologists have used various research strategies (Atkins & Huang, 2013) to approximate as close as possible to where it arises and how our thoughts are processed and that the relationship has with our behavior. Some studies refer to a comprehensive understanding of threats to human security (Heartfield, Loukas, & Gan, 2016) with an appropriate balance between structured improvements

to defend human weaknesses and security training and security awareness focused efficiently (Ferreira & Lenzini, 2015).

But more research is needed to better understand and explore the direct application of social engineering in the weakest link in the security chain; to psychologically manipulate the person by using social networks combined with deceptive advertisements and, in this way, obtain confidential data from it. So that is why we must have to know how vulnerable are people to deceptive advertising on social networks? And how attackers could obtain/decipher their passwords based on personal information obtained from such advertisements?

GENERAL OBJECTIVE

To conduct a social engineering study in order to identify how vulnerable people are to deceptive advertising on social networks; and how attackers could decipher their passwords based on personal information obtained from such advertisements.

SPECIFIC OBJECTIVES

a) To perform a bibliography study of the recent literature about social engineering, information security, and in general, psychology applied to information security and cognitive psychology.

b) To design a strategy to develop an ethical attack on one popular social network in order to collect specific users' information.

c) To perform statistical analysis of the results based on the collected information, in order to know how many people are vulnerable to deceptive advertising and how computer attackers can take advantage of the information achieved from the ethical attack to breakages users' passwords.

d) To give useful recommendations on the care of deceptive advertising, granting safety references between functionality and security when creating a password.

METHODOLOGY

This is an exploratory research project, which is quantitative, and it will be executed in four phases described below.

Phase 1. Phase of review of the literature and full understanding of all relevant necessary topics for the development of the research project.

Activities:

A1. Search, investigate and study the recent literature about social engineering, information security, and in general, psychology applied to information security in the best digital repositories and scientific journals of engineering.

Products:

P1. The development of a state of the art will be obtained with a vast knowledge of ethical hacking attacks and the respective understanding of each of the studied topics such as security information among others.

Phase 2. Design and implementation of the ethical attack on one social network. Scam baiters often use self-made documents to gather additional information about the scammers.

Activities:

A1. Design a strategy of an ethical attack that will be performed from the perspective of an attacker whose purpose is to determine the viability of an attack and the amount of impact, based on existing methodologies, with

defined standards so that the process of ethical hacking will be done in a logical and orderly way.

A2. Establish the dynamic for the ethical attack exposed in the social network.

A3. Implement the mentioned attack on an interface or announcement in responsive design in order to adapt to all mobile devices.

A4. Establish the best social network to put online this announcement.

Products:

P1. Strategy of an ethical attack that we will perform and the social network where this will take place.

P2. The ethical attack will be performed and implemented.

***Phase 3*.** Perform a statistical analysis in order to know how hackers use the psychology applied to information security.

Activities:

A1. Collect specific information on a database of random users with the consent of the users using a deceptive advertising on a popular social network.

A2. Define the criteria and the metrics for statistics that we will consider to evaluate this type of ethical attack based on the study of the recent literature.

A3. Perform an analysis of how vulnerable are people to deceptive advertising and how hackers can get advantage of that in order to decrypt your password.

A4. Make an analysis on how trust is gained, about the vulnerability in the information security of a person based on their personal information exposed on the social network.

A5._ Establish some patterns about how people select their passwords to the different services or applications of individuals based on the information that they expose on their social networks.

Products:

P1. Statistical analysis about how many people trust on deceptive announcements on social network using our database.

P2. Results about the vulnerability of the information security using our database

P3. Patterns between security and functionality of how people select their passwords to the different services or applications

Phase 4. Suggestions and recommendations. Providing a better vision of how hackers use the psychology applied to information, security for not only right people in the subject, but also raising awareness and warning of the risks associated with misuse of the Internet.

Activities:

A1. Based on the best practices we could recommend and give some tips of how to create a secure password, which information we can share on the social network in order not to be vulnerable to any social engineering attack.

A2. Based on the previous analysis, establish some patterns and give some suggestions and recommendations in order not to be the weakest link.

Products:

P1. Guideline of conclusions and recommendations for deceptive announcements on Internet.

P2. Guideline of conclusions and recommendations of how to create a secure password.

DOCUMENT STRUCTURE

The rest of this document is organized as follows:

Chapter II presents a Literature Review about all the relevant information about psychology, social engineering, information security, misleading and deceptive advertising, passwords and social networks.

Chapter III describes the proposed strategy which will be used to design and to develop an ethical attack on one popular social network.

Chapter IV presents the statistical analysis and numerical results about how many people are vulnerable to deceptive advertising; and how computer attackers can get advantage with the information achieved from the ethical attack to breakages user's passwords.

Finally in ***Chapter V***, given the statistical analysis, conclusions and recommendations are presented on the care of deceptive advertising, granting safety recommendations between functionality and security when creating a password.

Chapter 2. LITERATURE REVIEW

Review of the literature of relevant topics of psychology and social engineering for the development of the research project.

SOCIAL ENGINEERING AND PSYCHOLOGY

Psychological phenomena are unknowns related with human behavior, because each person is a completely different world from each other. However, all people have more affinity to certain common of influence patterns as is the case of social engineering.

Social engineering is based on psychology. There are different incentives and motivators in people who allow social engineers to take their victim to action. This is because there are hundreds of different tactics that compliance practitioners employ to produce yes as a response, the majority fall within six basic categories which will be explained below (Cialdini, 1984). Each of these six categories is governed by a primary and fundamental psychological principle which directs human behavior and, in so doing, gives the tactics their power to manipulate and know in better way the victim (Cialdini, 1984). The principles are:

- ✓ Consistency. We develop patterns of behavior that become habits.
- ✓ Reciprocation. To feel that we owe something or a favor to someone.
- ✓ Social Proof. When people are uncertain about to do a specific action, they tend to look to those around them to guide their decisions and actions.
- ✓ Authority. We are receptive to the orders and requests of figures with authority.
- ✓ Liking. People prefer to say ‘yes’ to those they know and like.
- ✓ Scarcity. The more rare and uncommon a thing, the more people want it. The less there is of something, the more valuable it is.

In the effort to be able to observe it in a clearer way, Figure 1 presents the six universal truths of influence below:

***Figure 1**: The six universal truths of influence*

Today, cybercriminals based on social engineering, seek to deceive their victims so they voluntarily give out their personal information. The vulnerability of legitimate users is stipulated by their needs (money, self-affirmation) and weak points of each (self-esteem, social approval) (Kotenko et al., 2011). Social engineering has been established based on psychological and security terms by various organizations and experts as taking advantage of vulnerable people's naivety via influence, in different aspects such as persuasion and manipulation to obtain vital information (Ghafir et al., 2016). It is an access attack that attempts to manipulate individuals into divulging confidential personal information or performing certain actions. Social engineers often rely on people's willingness to be helpful and trustable. Moreover also prey on people's weaknesses. By knowing the personal information or the likes of a person, they become defenseless so the attackers will know the structure of their passwords. Mostly they get their victim information by the social networks and

sometimes by deceptive advertising.(Huang, Yu, & Kao, 2017) This has been due deceptive advertisings and spam emails. The total amount of spam emails account for more than 75% of the total emails exchanged worldwide; (Zisiadis, Kopsidas, Varalis, & Tassiulas, 2011) current reports raise this number up to more than 90%. That is why a lot of time and dedication is devoted to research for anti-spam solution. Mostly to be followed by announcements of sophisticated methods to overcome them through the use of advanced software to reach the spammers (Zisiadis et al., 2011). Also we have advance fee fraud which gets its name because these schemes require the victim to pay the scammer in advance with the promise of receiving rewards later. This scam is neither the most costly nor frequent Internet crime; however, it remains to be the most ubiquitous and well-known of all cyber-crimes (Atkins & Huang, 2013).

Social network users such as Facebook, Instagram, Youtube, blogs, twitter and other users of the current social media use the social networks that give them name as a world window in which they expose multiple aspects of their day to day, for various purposes; some of them, clearly commercial. Nowadays, cybercrime is introduced by hackers through social networks, because they have the ability to identify various ideas to damage everything regarding computer software and hardware and also identified methods to forward spam messages for advertisement purpose in an illegal way (Bhise, 2016). Misleading advertising often compensates (through money, products and free services or other advantages) to influence a product or service to appear on your social network, in order to increase its consumption among its followers. And cyber fraud is a problem that leads to weight problems, such as economics. It affects particular people and business through identity theft, the creation of botnets and the spread of viruses all of which are interconnected manifestations of Internet threats in general not only for social networks (Bhise, 2016).

SOCIAL ENGINEERING AND DECEPTIVE ADVERTISING

Responsibility for misleading advertising operates with the single demonstration that advertising does not correspond to reality or because it is insufficient has the ability to mislead or confuse the consumer but who is responsible

for misleading advertising on social networks? Or how to recognize if an ad is true or misleading? That is why many tests have been carried out and many methods have been designed to distinguish between true and fraudulent advertising. In order to carry out these studies, the literature have made ethical attacks with random users, creating advertisings, apps or in a social network post. For example:

- A complete analysis to distinguish legitimate and malicious posts using Netbeans IDE is presented in (Ekta Science, 2016). With 91.01% of accuracy identifying legitimate posts. This technique can find legitimate posts effectively.

- In (Huang et al., 2017), The authors demonstrate an intimate knowledge about deceptive advertisement. They implement a system which can efficiently detect deceptive ads and phone scams by taking advantage of our unified framework on deep neural network with convolutional neural network. The work challenges is in the appliance of the same system throughout the different internet services.

- In (Wang, Han, & Chen, 2016), the authors carried out a series of preliminarily studies on users' preference distribution. Collecting 479,048 user's information. It present 6,276,422 particular preference items in total. Facebook classifies users' preferences into 11 types as Music, TV, Movie, Activity, Book, Interest, Athlete, Game, Team, Sport and the people who Inspire you. But it not present nothing regarding vulnerability on social networks.

- In (Aziz, 2016), the authors wrote a C# console application based on Graph API to download all the posts that is available on FoodBank since the beginning. This data downloaded is not completely analyzed.

- "*The current study focuses on the analysis of social network activities of the course implemented in the 2015 spring semester (the course instructor has adopted Facebook as a platform for course-related discussion since 2012). A total of 505 discussion messages posted during*

the course (from January to May 2015) were collected from 109 participants in the Facebook discussion forum." (Chan et al., 2016) .

- "*Gather Facebook user data from participants via their access token, i.e., 'like' pages and 'share' page posts. For each user, compute the frequency of each behaviors on the pages with respect to each category. It is noted that there are 204 page categories provided by Facebook. In addition, obtains also their user preferences to be used as class attributes. Then, use One-vs-Rest (OvR) approach to convert a multiclass dataset into binary class dataset. In this context, there are 16 classes (desirable user preferences). Therefore, 16 class attributes are created.*" (Rachsuda, 2017)

- "*SOCIAL ENGINEERING - SKILLFUL MANIPULATION OF USERS*" (Zingerle, 2014)

  - Method 1: Fake Form Elicitation
  - Method 2: Spear-Phishing money transfer
  - Method 3: Phishing web service attack

  The methodologies mentioned have a very strong documentary structure, and are specially designed for in-depth Information Security Audits

  - OSSTMM (Open Source Security Testing Methodology Manual).
  - ISSAF (Information Systems Security Assessment Framework).
  - OWASP (Open Web Application Security Project).

- "*Scamming or 419 fraud is usually in the form of "Advance Fee Fraud" (named after the relevant section of the Criminal Code of Nigeria that deals with such crimes). It begins when the target receives an unsolicited fax, e-mail, or letter often concerning Nigeria or another African nation containing either a money laundering or other illegal proposal. With the increase in use of internet facilities for electronic commerce, online banking, social networking and other financial transactions phishing attacks*

are also on the increase. Spamming was said to be one of the most prevalent activities on the Nigerian Internet landscape accounting for 18% of all online activities amongst others". (Longe, Mbarika, Kourouma, Wada, & Isabalija, 2010)

- Phishing is a form of social engineering in which the attacker (or phisher) fraudulently retrieves confidential or sensitive information by imitating a trustworthy or public organization. Two basic methods are commonly employed by phishers to steal valuable personal identification (APWG, n.d.). (Atkins & Huang, 2013)
    - *"The first method is the technical artifice method, which involves infecting personal computers with malicious software. This software is capable of recording keystrokes entered by the user, and sending that information to the phisher. This software can also redirect Internet users from legitimate websites to false ones via a remote connection"*. (Atkins & Huang, 2013)
    - *"The next method that phishers employ is social engineering, which, is defined by Yoo (2006) as "gaining intelligence through deception or also as using human relationships to attain a goal" (p. 8). Phishers using social engineering techniques employ deceptive devices to trick Internet users into a situation where they are willing to disclose sensitive information. Usually, the social engineering methods launch a false e-mail urging the receiver to click on a linked website appearing to come from a genuine business. After clicking the link, the user is actually brought to a fraudulent site asking for personal financial information such as credit card or bank account numbers. Phishers then use the records they obtained to swindle money from the credit card or bank account, or even apply for a new credit card with a false identity."* (Atkins & Huang, 2013)
        - *To examine the deceptive operations and techniques used in phishing and advance-fee e-mails, the study has collected a*

sample of 200 fraudulent e-mails related to the two types of scam. (Atkins & Huang, 2013)

SOCIAL ENGINEERING AND PASSWORDS

When you want to elaborate a bit more complex passwords, the user relies, even unconsciously, on symbolic references such as his birthday, his children's or the date of his wedding. In this way, one makes it easy for hackers to access sites such as Facebook, see some of these data and, from there, search for the combination of entry to personal services. After entering the password it is important to check that it does not contain personal tracks. Regarding the username, the "professionals in breaking keys" know that almost everyone uses the same one that has in their email address. So it is therefore appropriate to be much smarter, be one step forward and shield what is now almost like an open book (Heartfield et al., 2016).

Most people's passwords have a great relationship with their tastes, personal information, ideologies. For those with weak awareness of information protection and vulnerability through internet their security issues will be worse. For example, According to the Survey of Security Awareness of Netizens which took place in China on 2015, the misbehaviors of netizens include the followings (Zou, 2016):

- The account password is rarely changed. 81.64% netizens in China didn’t change password regularly.
- 75.93% netizens used the same password for different accounts and services for convenience.
- 44.42% netizens chose their birthday, id number, phone number or name spelling as passwords. Looking for a balance between functionality and security

These problems also bother netizens in many other countries. So it can be deduced that this pattern does not only happen in China. If not also on a global scale. Anyone with an understanding of social engineering can deduce that all the passwords of the people are easily vulnerable if they expose too much information about their tendencies and activities on the social networks, like Facebook which

provides valuable information of page category over two hundred relating to user preferences (Rachsuda, 2017); and also it depends on your geographic location. The geographical location is an extremely significant factor in people's preference distribution (Wang, Han, & Chen, 2016). Hackers know a lot of social engineering, and manipulate computer users, through psychological elements that influence an individual to do a particular task of which he is not aware. Despite users having different levels of computer experience, familiarity with technology, backgrounds, ideologies, religions, and gender, it is apparent that social networking sites are not restricted to any particular type of user. Nowadays, it should not be unexpected that social networking sites present a huge market for information security threat agents such as phishers and hackers. These are some types of social engineering attacks (Cisco Networking Academy, 2017):

- ✓ Pretexting - This is when an attacker calls an individual and lies to him in an attempt to gain access to privileged data. An example involves an attacker who pretends to need personal or financial data in order to confirm the identity of the recipient.
- ✓ Tailgating - This is when an attacker quickly follows an authorized person into a secure location.
- ✓ Something for Something (Quid pro quo) - This is when an attacker requests personal information from a party in exchange for something, like a gift.

The importance of this attack factor are two points in particular, the first one is the large amount of personal information stored by social networks. Every social networking site like Facebook, Instagram, Twitter are mostly used in expressing the opinions about a particular entity of life or just post photos of our daily activities (Islam & Dey, 2016). Currently the internet is almost within reach of all the people of the world. And approximately two billion Internet users worldwide using social networking sites (SNSs), it is common to find individuals active on at least one social network (Frauenstein & Flowerday, 2016). Moreover a hacked email account represents a huge risk because it is used to reset passwords to other web services or applications

often resulting in identity theft (Zingerle, 2014). And the second one, is that our mobile devices or computers, store a lot of personal information such as contacts, photos, videos, conversations but mainly they store social network passwords, emails, bank accounts, credit cards numbers and even geo location, which makes us extremely vulnerable to any kind of fraudulent action (Zingerle, 2014).

SOCIAL ENGINEERING IN RELATION TO PEOPLE'S CONFIDENCE

In effort to identify what social engineering consists of, we created a mind-map, which includes different, and possible influencing factors. The main aim of the mind-map, Figure 2, is to illustrate an understanding of the complex issues that in some way can be argued to be connected to or influence social engineering and the human element of security. There are up to six different disciplines: sociology, psychology, economics and management, security, law and education. All are related to each other and each point addresses a sensitive psycho-social and computer security information in order to gain the trust of the people.

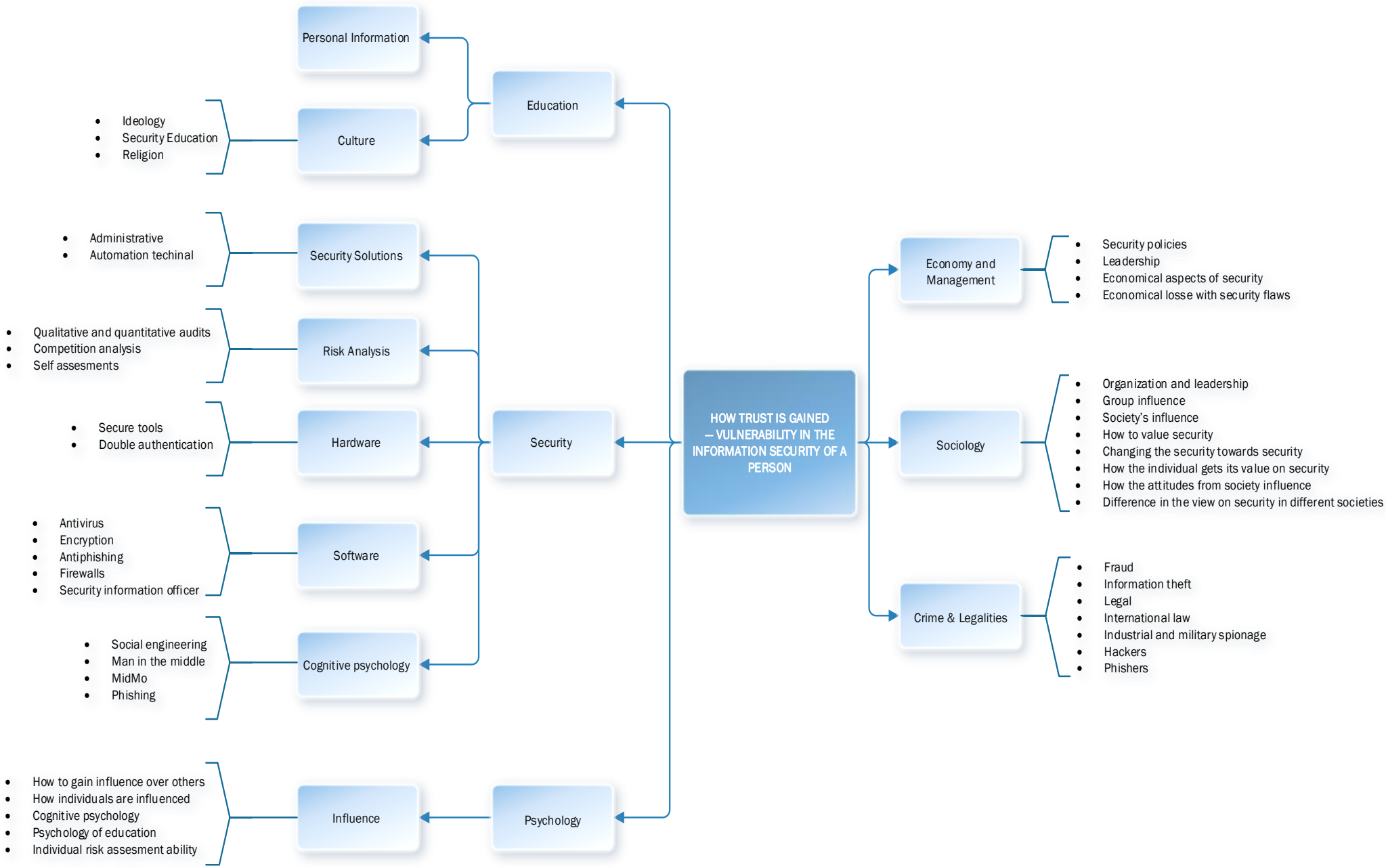

***Figure 2**:How trust is gained – Vulnerability in the information security of a person*

Within the branch of social engineering, there are different types methods reported but what is lacking is a unifying effort to understand these methods in the aggregate. So that it is why we bring up the valuable work of (Frauenstein & Flowerday, 2016), presented in Table 1, in their effort to present the Taxonomy of social engineering persuasion techniques with a comparison of persuasion principles as is shown below.

Principles of Influence	Psychological triggers	Principles of Scams
Authority	Authority	Social Compliance
Social Proof	Diffusion Responsibility	Herd
Linking and Similarity	Deceptive Relationship	Deception
Commitment and Consistency	Integrity and Consistency	Dishonesty
Scarcity	Overloading	Time
Reciprocation	Reciprocation	Need & Greed
	Strong Affect	Distraction

***Table 1**: Taxonomy of social engineering persuasion techniques.*

Taken from (Frauenstein & Flowerday, 2016)

As we can appreciate it, taxonomy is always required when it is necessary to provide a clear and consistent overview of a phenomenon is evident that there are common techniques overlapping in each of the principles used by social engineers without any obvious overlap. The social methods used to get around technological countermeasures is a dynamic process, which uses different types of technologies or social medium, or finally with dominant attitudes, playing with the cognitive psychology of the victim and having a better approach for direct contact. The motivated attacker will always try to face the victim and steal personal information.

Unlike traditional hacking methods, where the attacker needs to be technically equipped to carry out an sophisticated attack, a social engineer needs to focus on his social skills, in their behavior, in their tastes, in their ideology (Grande & Guadrón, 2015) in order to carry out a successful attack. Recently techniques of persuasion have been considered in information security. These Principles of Persuasion in Social Engineering (PPSE) are summarized in Table 2 based on the analysis made by (Ferreira & Lenzini, 2015), where identifies, the elements which reflect the effectiveness of social engineering using phishing, and manually quantifies them

within most relevant attitudes or qualities, where the victim may be much more vulnerable.

Most elements recognized as effective in phishing commonly use persuasion principles such as authority and distraction which represent dominant attitudes, people of strong character, which will be able to manipulate the victims at their whim.

PPSE	ACRONYM	DESCRIPTION
AUTHORITY	AUTH	Society trains people not to question authority so they are conditioned to respond to it. People usually follow an expert or pretence of authority and do a great deal for someone they think is an authority.
SOCIAL PROOF	SP	People tend to mimic what the majority of people do or seem to be doing. People let their guard and suspicion down when everyone else appears to share the same behaviours and risks. In this way, they will not be held solely responsible for their actions.
LIKING, SIMILARITY & DECEPTION	LSD	People prefer to abide to whom (they think) they know or like, or to whom they pretend to be similar to or familiar with, as well as attracted to.
COMMITMENT, RECIPROCATION & CONSISTENCY	CRC	People feel more confident in their decision once they commit (in public or in writing) to a specific action and need to follow it until the end. This is true whether in the workplace, or in a situation when their action is illegal. People have tendency to believe what others say and need, and they want to appear consistent in what they do. When they owe a favour, there is an automatic response for repaying it.
DISTRACTION	DIS	People focus on one thing and ignore other things that may happen without them noticing; they focus attention on what they can gain, what they need, what they can lose or miss out, or if that thing will soon be unavailable, has been censored, restricted or will be more expensive later. These distractions can heighten people's emotional state and make them forget other logical facts when making decisions.

***Table 2**: Principles of Persuasion in Social Engineering.*

Taken from (Ferreira & Lenzini, 2015)

IMPORTANCE OF SECURITY INFORMATION

It is extremely important to save the information and data that are handled both personally and business. Nowadays security informatics and everything that concerns security of information, becomes a key knowledge domain in computer science and engineering education (Chan et al., 2016). Therefore the most valuable treasure that exists today is the data, and the knowledge that can be obtained from the other person because of the rapid change and development of web and social network sites, which support social interactions between course instructor and students can be a good venue supporting the teaching and learning of information security. If a person can guess our password, spy on our instant messaging applications and common mail or at least induce us to tell it, we are exposed not only to heavy jokes within social networks, but mainly to computer fraud linked to the financial, labor and personal information sectors that can be used for criminal acts which may be the cause of our being deprived of freedom (Del Pozo & Iturralde, 2015).

Formally the desired properties for creating a secure communication and password are: confidentiality, authentication, data integrity, non-repudiation and availability. All these must be guaranteed during the exchange of personal information through the network (Cisco Networking Academy, 2017):

- **Confidentiality**. Confidentiality is the property that prevents the disclosure of information to unauthorized persons or systems. Broadly speaking, it ensures that the access to information is exclusively to those who have authorization.

- **Authentication**. It is the property that identifies the generator of the information. For example if someone send you a message, it must be sure that who receive the message is the correct one and not a third person in the middle of the conversation.

- **Non-repudiation**. Provides protection against the interruption, by any of the entities involved in the communication, of having participated in all part of the communication.

• **Integrity**.- Data integrity ensure that both parts involved in the connection have the true information, and the content of their communication is not altered either maliciously or by accident during transmission. In this case, it is necessary to verify that there have been no modifications, insertions, omissions, repetitions or reordering of the information.

• **Availability**: It guarantee that the user who is attempting to access the requested service can afford unique access and the credentials of the passwords are the same and that they have the appropriate access rights and carry out their income from one properly.

Inappropriate access to information may occur and security is not monitored closely. Therefore, with the design and proper supervision, electronic records can offer controls that protect more protection than the protected information of the records. Because the security of all our personal and business information is based on the use of secure systems, reliable networks, algorithms and databases, it is important to complete the entire security chain and help mitigate the risk of the weakest link, the people (Ghafir, Prenosil, Alhejailan, & Hammoudeh, 2016) Nowadays it is common for the security officer of each company, or designer of each service, net-works, applications and data that are monitored to be able to detect it on real time and in the best stage mitigate threats (Heartfield et al., 2016).

However, it is important to emphasize that the cheapest security is just common sense. With the application of common sense, the risk is not completely mitigated; but it decreases significantly and facilitates the work of the investigation in the event of a crime that could be committed.

IMPORTANCE OF UNDERSTANDING THE RELATIONSHIP OF PSYCHOLOGY WITH SOCIAL ENGINEERING

While technological know-how certainly plays a large role in enabling attackers to hack any given code system, public or private network or individual, what is often overlooked is that some tricks of the trade, like social engineering, are also psychological games (Longe et al., 2010). Cybersecurity attacks are increasingly based primarily on social engineering techniques, the use of psychological

manipulation to trick people into disclosing sensitive information or inappropriately granting access to a secure system. And that is why it makes it one of the most complex techniques to avoid and is undetectable or questionable given that it handles aspects of psychology that could not be factually proven (Grande & Guadrón, 2015). That means that protecting and defending against these kinds of attacks is, in turn, part mental as well and it depends of ourselves.

A word, a state or registration in a social network, a fact that for those who mention it may not be of any importance, for an expert in social engineering it may be the key that opens the Pandora's Box of personal security and the protection of the sensitive information (Frauenstein & Flowerday, 2016). A self-inflicted abduction to our privacy based basically on our ignorance and lack of skill for social relations and on the capacity of the social engineer to take advantage of them based on using human psychological characteristics, such as: curiosity (Cialdini, 1984), which moves us to look, to respond and touch where we should not; fear, we seek help in any way or fall easier in the traps because we cannot reason with peace; trust, we feel safe at the slightest sign of authority.

Using traditional non-aggressive presence techniques, people would be monitored, home surveillance, immersion in buildings, access to agendas and Dumpster Diving (search for information such as tickets, receipts, account summaries, etc. in the trash of the researched) would be the first records to be reviewed.

But in more up-to-date and technological aggressive methods, the work of the experts becomes more intense, and that is where identity theft comes up (Ghafir et al., 2016) (posing as IT, technical services, security personnel, presentation of deceptive advertising, etc.). The combination of this last group of techniques together with the exploitation of 3 psychological factors mentioned above about the affected person can be highly effective in face-to-face work between victim and victimizer.

As mentioned before, there are studies which refer to a comprehensive understanding of threats to human security, the social engineering as a silent attack, common phishing attacks, security training and security awareness focused efficiently. But more research is needed to deal specifically with the direct application

of social engineering; to psychologically manipulate the person by using social networks combined with deceptive advertisements and in this way obtain confidential data from it.

Chapter 3. PROPOSED STRATEGY

A description of the proposed strategy and ethical social engineering attack that will be developed on one popular social network is presented, such as the implementation and the results.

To examine the deceptive operations and techniques used in social engineering and phishing, the study has designed a strategy of an ethical attack that will be performed from the perspective of an attacker whose purpose is to collect personal information through deceptive advertising in order to use for different purposes.

Below is presented the strategy that would be implemented in order to perform our study. The entire flow will be divided into 6 stages which will explain each of them in detail.

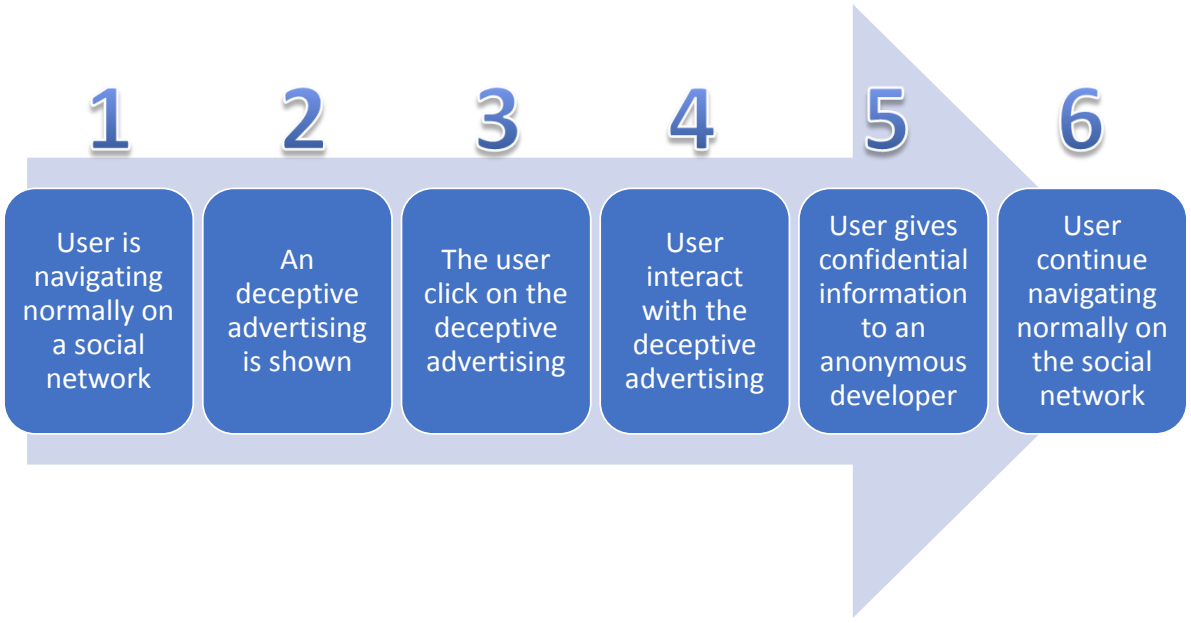

***Figure 3**: Proposed strategy implemented for an ethical attack*

STAGE 1 - USER NORMAL NAVIGATION

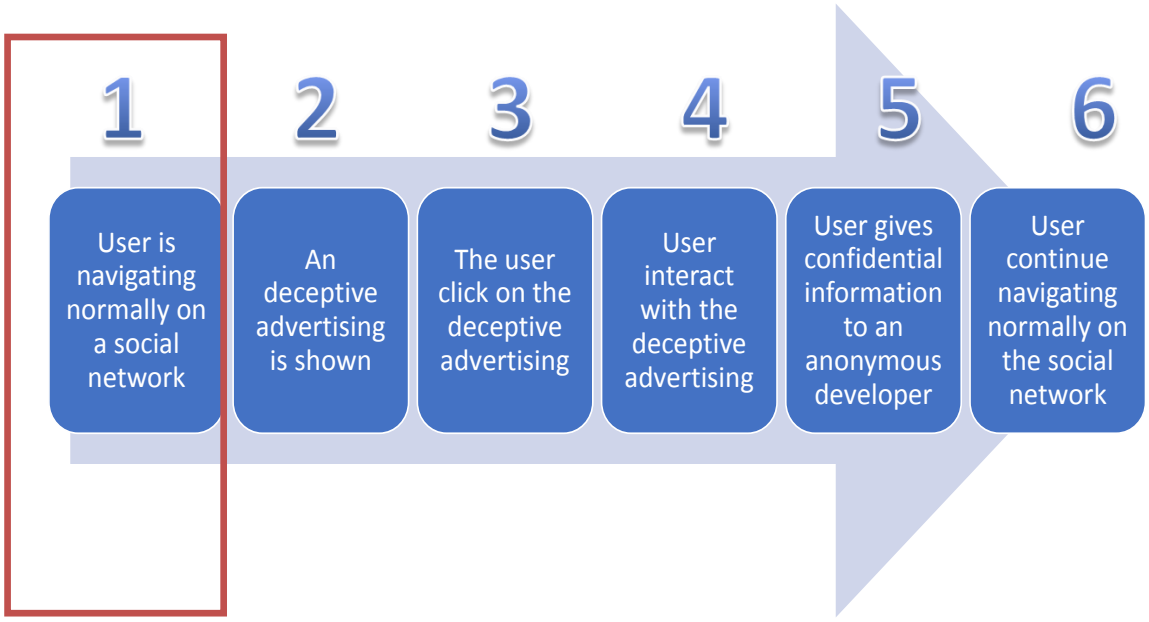

***Figure 4**: Proposed strategy implemented, Stage 1 - User normal navigation*

The Internet has undoubtedly revolutionized the world as it was known some decades ago. It is a global phenomenon, closely linked to communication, which greatly influences almost all the areas of society. Their creators never imagine that in just 20 years it would be an invention as essential as the telephone or television. In principle, the main purpose of the Internet was communication, creating an efficient channel of communication allowing the human communicates with others without barriers, which does not take into account the space, or borders, distances, societies (Smith, 2016). The Internet is full of communication in all areas, from advertising, interviews, articles, videos, chats, emails, but also data, music, documents, books, images, among others because almost any type of visual and aural communication that we can imagine we can find on the world wide web.

With all the facility that the Internet gives us, social networks emerged. Facebook was born in 2004 only for Harvard Students. On 2006, Facebook was opened to everyone at least 13 years old with a valid email address, and it converted on one of the most successful social networks on the Internet, which would let us generate user profiles in order to establish links with old friends and colleagues, as

well as offering a multitude of games to entertain the user. Little by little, social networks began to be known and nowadays many activities are managed through social networks. To be more specific and to enter into context, some relevant data of the internet is presented (Smith, 2016):

- The world population in March 2016 was 7.4 billion.
- Internet has 3.17 billion users.
- There are 2.3 billion active users on social networks.
- 91% of retail brands use two or more social media channels.
- Internet users have 5.54 social media accounts on average.
- Social network users grew 176 million last year.
- There are 1 million active users of social networks on new mobile phones every day. That is, 12 every second.
- Facebook Messenger and WhatsApp handle 60 billion messages daily.

Within the most popular social networks, we find the following, each with its respective number of users (Smith, 2016):

- Facebook: 1.71 billion users
- Wechat: 1.12 billion users
- WhatsApp: 900 million users
- Weibo: 600 million users
- Instagram: 400 million users
- Twitter: 320 million users
- Google+: 300 million users
- LinkedIn: 300 million users
- Flickr: 112 million users
- Pinterest: 100 million users

- Snapchat: 100 million users
- MySpace: 50.6 million users

Based on these statistics and according to the study conducted, it can be concluded that, having so many people navigating through the Internet and specifically a large percentage of users interacting with social networks, it is very possible that many people perform or at list try some type of illicit activity and a huge percentage of users that could be vulnerable of this types of illicit activities. In particular for this study, we select the social network with more users: **Facebook.**

STAGE 2 - ATTRACTING USER ATTENTION

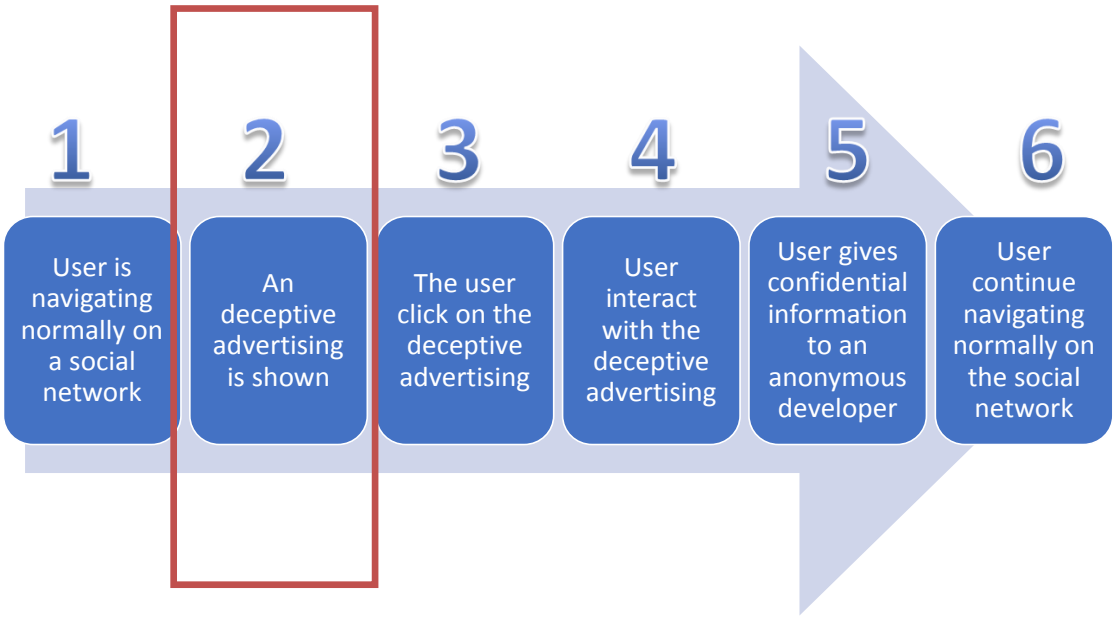

***Figure 5**: Proposed strategy implemented, Stage 2 - The deceptive advertising*

On the Internet it is usually shown with all kinds of deceptive Banners whose purpose is for the user to click on them. Advertising on Facebook is already starting to be used by some companies that are timidly trying this tool to attract customers (Smith, 2016). The advertising on Facebook consists of the purchase of small advertising spaces that appear on the right side of the Facebook page or while one is reviewing the news on the wall the advertising is shown and that are accompanied by a small image and a short descriptive text. By clicking on this ad you can take us to a

fan page of a company, to your wall or to an external page. Facebook allows companies to hire these advertising spaces and offers the possibility of segmenting the view of these ads only to people who meet the profile of the potential client (Smith, 2016). With Facebook, a specific audience can be assigned to a campaign in order to have a higher return on investment. Between the options it is possible to direct the announcements that are for the men, for the women or for a determined age segment or based on your cookies or what you search on the internet, Facebook select the best articles for you (Smith, 2016).

So based on how Facebook present advertisement and according to the study conducted, it can be concluded that, present and deceptive advertising on the main wall is viable hogging a large percentage of users.

STAGE 3 - THE DECEPTIVE ADVERTISING

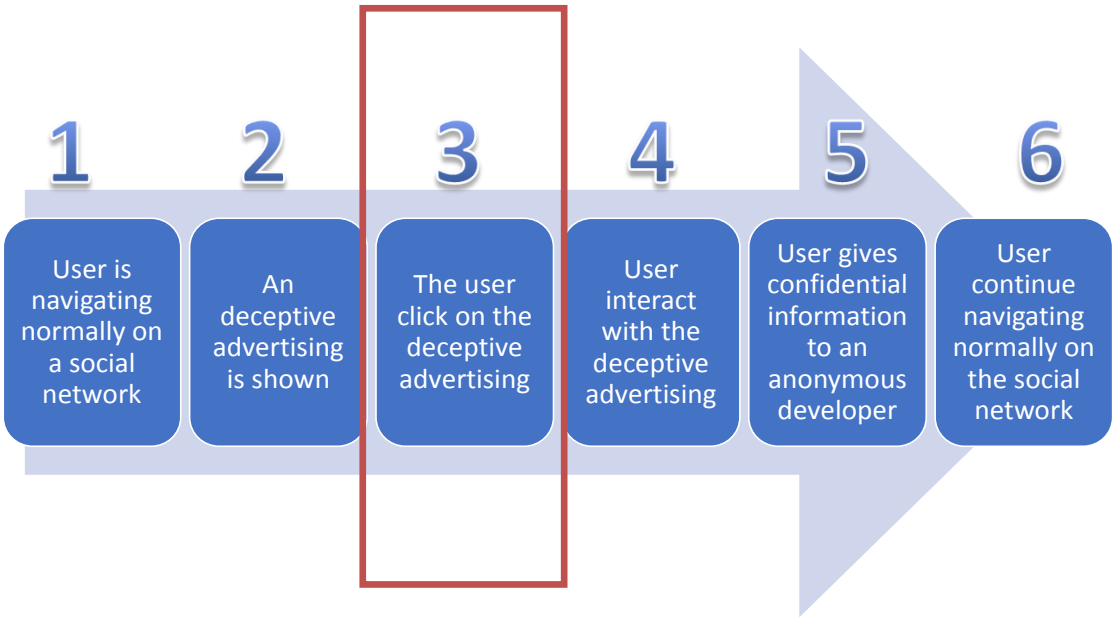

***Figure 6**: Proposed strategy implemented, Stage 3 – Click on the deceptive advertising*

As we mentioned before, social engineering use different methodologies in order to attract our attention so we act in the way they want (Zingerle, 2014); for example, convincing us of the need to forward an email to our entire contacts list, or to download and open a file that is enclosed under some excuse or inciting us to provide sensitive information such as access codes to a certain service or account.

In order to capture the user attention, this take advantage of the curiosity and curiosity that comes from knowing the deeds of famous people, humanitarian catastrophes, current affairs or simply use the emotions of people, their self-esteem and what they want to sell to the world of themselves. (Ferreira & Lenzini, 2015)

Many people publish on their social accounts in order to show their friends something they are not, something they lack or just they post because of social pressure. Our moods, photographs, videos, everything you upload directly to your accounts in different social networks, says a lot about who you are, your personality, your intelligence, the content of your publications reflects a projection of yourself and

your hobbies, (Rachsuda, 2017) so if you know a person both in real life and in social networks you can make a critical judgment and determine the truth of a study like this.

Everything you share can determine your fears, your complexes, traumas, paranoia, your ego and even more importantly, your values and education (Chan et al., 2016) , so it is advisable to analyze what you want to upload to any social network so as not to expose yourself, one thing is to publish some achievement that cost you a lot of work and share that triumph of which you fill with pride with all your friends or something you want to enjoy in a group, and another thing is to be bragging about your day to day, no matter how good or bad be.

Among the most recurrent activities that people post on social networks in order to draw attention to the rest and raise their self-esteem are (Smith, 2016):

- Upload all the purchases you've made
- Photos in the "Gym"
- Share of what you are going to eat and / or drink
- Share your location or your recent trips
- Selfies
- Stream live
- Post amorous disappointments

Based on these statistics presented and in particular for this case of study, the social engineering attack will be performed based on the most recurrent activities that people share and post on Facebook, focused on "*Share of what you are going to eat and / or drink*". This because as we see with our thoughts. If we attach great importance to the opinion and criteria of others it is easier to influence us, and this is how social engineering works.

Society has a strong influence on our behavior and attitude, especially when it comes to collective actions. "*The decisions we make as our own, or the way we direct our lives, are conditioned by society*" (Cialdini, 1984). All this is the result of what is right to do, what we think we should do according to the opinion of the majority or the

procedures we are accustomed to perceive. Since childhood we felt an imperative need to belong to a group with the family and the class group. In addition, we have dependence or interdependence on these, as they provide us with security and social inclusion, even leaving aside our values or ideologies, making us vulnerable to manipulation by third people.

So as we mentioned before, we will perform and advertising focused on "*Share of what you are going to eat and / or drink*", this advertising will be named as "¿Qué tan buen borracho eres?", which means "What a good drunk are you?" This is a test which makes reference first if you like to drink yes or no, and if it is positive is to say you like to drink, to determine if you have a good behavior while you are drunk. With this, we seek to access as many users as possible.

***Figure 7**: Deceptive Advertisement posted on Facebook*

STAGE 4 - INTERACTION WITH THE DECEPTIVE ADVERTISING

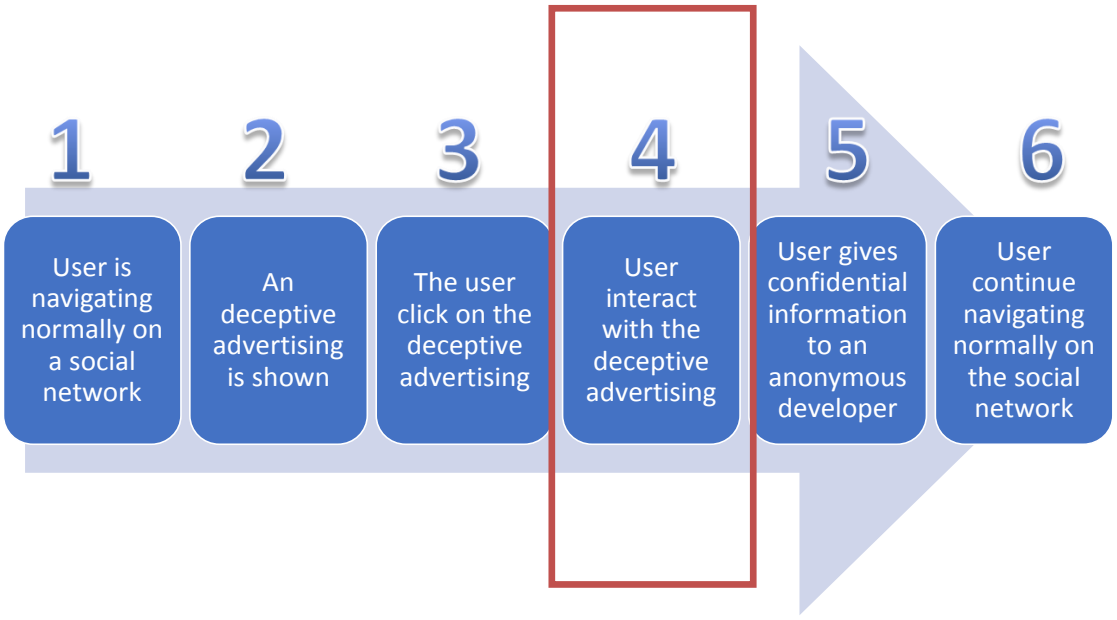

***Figure 8**: Proposed strategy implemented, Stage 4 - Interaction with the deceptive advertising*

As is shown in the previous image, the advertisement presents an initial message, which it is intended to persuade the user, capture his attention and encourage him to do this little test. This advertisement, which we show a screenshot in Figure 9 says:

> *"When we have drinks, we usually act in particular ways that reflect our personality. Knowing our way of thinking, feeling, acting, the consequences of our actions, even when we do not have our full awareness is very important.*
>
> *By evaluating these 10 questions, we will help you deduce and measure What a good drunk are you, and we will give you practical tips to control your actions."*

***Figure 9**: Index page of the deceptive Advertisement posted on Facebook*

Once the user clicked on the advertisement, the user will be interacting with four different screens. Once the user clicked on the *"CLICK AQUI"* button, an alert window is presented in order to notify the user that just for this case, this test has also academic purposes.

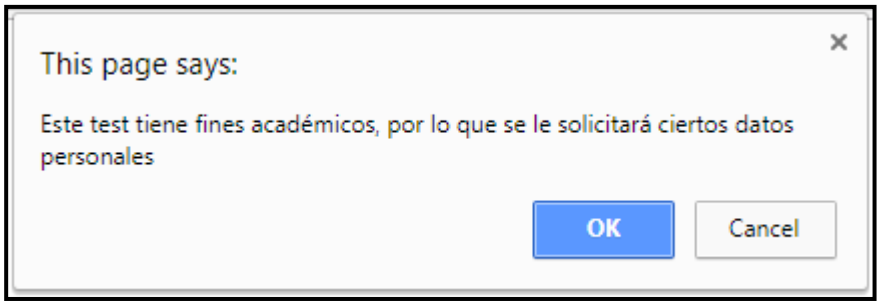

***Figure 10**: Alert window shown when the user clicked on the "CLICK AQUI" button*

As is shown in the previous image, the advertisement presents an initial message, which it is intended to persuade the user, capture his attention and encourage him to do this little test; but in turn it mentions: The test has academic purposes, so certain personal data will be requested.

When writing a survey our objective is to obtain certain information. To obtain the information we seek, it is important to use the most appropriate type of question in each case. *"The ability to ask good questions is essential for cognition and decision-making, because the choice of query determines what information is acquired, influencing all subsequent inferences and decisions."* (Planck et al., 2017). Although there are numerous reasons for asking questions the information we receive back (the answer) will depend very much on the type of question we ask (Planck et al., 2017)

Questions, in their simplest form, can either be:

- Open
- Closed
- Mixed

And each of them can be formed as:

- *Unique choice*
- *Multiple choice*
- *Ranking*
- *Scale*
- *Funnelling*

On the Web there are endless tools to conduct surveys online. These allow creating a questionnaire with a high level of design and with the possibility of putting several types of questions. Then at the end of the questionnaire a link is created to share the survey (via email, Facebook, or through online advertising in search engines and related pages). The type of predetermined questions when creating an online survey is *"closed questions formed by multiple-choice option"* (Smith, 2016). Here the user is entertained, it's easy and it does not require a lot of as other types of questions.

Based on the literature review and according to the study conducted, it can be concluded that this test will consists in fill a form of ten basic mixed questions formed by multiple-choice option and filling information in a text box in order to persuade the

use, and through this form collect the information need for the study. This test will be performed in Spanish because the majority of the persons which will take the test speak Spanish.

***Figure 11**: Form with 10 questions, second page of the deceptive Advertisement posted on Facebook*

The questions are:

1. ¿Que tan fiestero eres?
2. ¿Si sales 3 veces en el mes a tomar, cuántas de ellas te emborrachaste?
3. En el caso de que llegases a emborracharte, ¿Qué haces?
4. ¿Tu grupo y círculo más cercano de amigos toma?
5. ¿Cuál es tu fecha de nacimiento?
6. ¿Cuál es tu bebida favorita?
7. ¿Cuál es tu hobby o pasatiempo favorito?
8. ¿Cuál es tu color favorito?
9. ¿Cómo defines tu personalidad?
10. ¿Cuál es tu nivel de educación?

Translated to English, they are:

1. How much do you like party?
2. If you go out 3 times in the month to drink, how many of them did you get drunk?
3. In case you got to get drunk, what are you doing?
4. Does your group and closest circle of friends drink?
5. What is your date of birth?
6. What is your favorite drink?
7. What is your favorite hobby?
8. What is your favorite color?
9. How do you define your personality?
10. What is your level of education?

Question 1, 2, 3, 4, 9 and 10 are distractors to divert the user's attention and give a meaning that is a formal survey; and question 5, 6, 7 and 8 will be used to obtain personal information, which for this case of study will be used to know the type of people who are vulnerable to deceptive advertising. All this information will be saved into a text file.

***Figure 12**: Results of your test, third page of the deceptive Advertisement posted on Facebook*

Once the user finished the test, the user will have a result that he can share in Facebook if he wants.

The four different possible results that will be shown, depend on what the user selected, will be:

- Una persona tranquila (A calm person)
- Un buen bebedor social (A social drinker)
- Un borracho con episodios vergonzosos (A drunk person with shameful episodes)
- Un borracho agresivo (An aggressive drunk person)

STAGE 5 - RETRIEVE CONFIDENTIAL INFORMATION

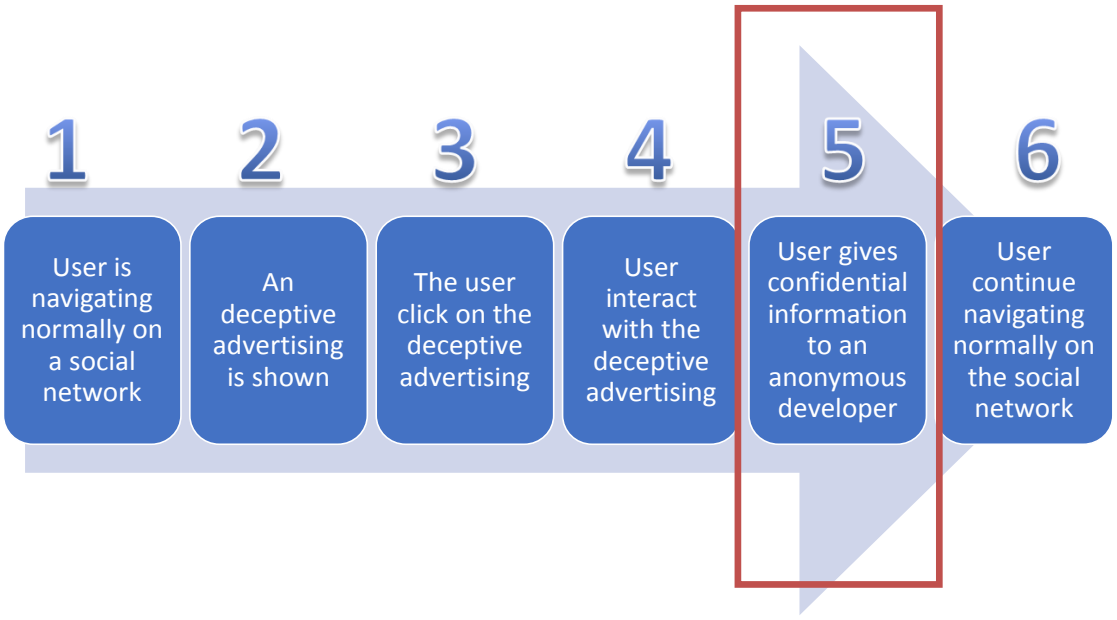

***Figure 13**: Proposed strategy implemented, Stage 5 - Retrieve confidential information*

Once user finishes interacting with all the deceptive advertisement he will find a share button “COMPARTIR”. Normally, we tend to promote the social plugin "like" but since social networks goal is to be everywhere and interact with more and more users we must use the "share" button, In addition, sharing a content also shares its source, something basic to increase our presence within the network and increases confidence of the source of the test, with to the user.

If the user clicked on the “COMPARTIR” button a pop up window will appear so the user is asked to write his email and his password, so it means that he will gives all the personal and confidential information, which will be saved on a second text file on the server.

***Figure 14**: Facebook Scam window, last page of the deceptive Advertisement posted on Facebook*

At this part of the of the proposed strategy, the user does not realize that everything is false despite that the screen is exactly the same as the original Facebook trusted page. And this is how deceptive advertising works. Because of false or inaccurate data, or because of its ambiguity, omission, similarity with reality or other circumstances, induces or may mislead its addressees about essential elements of the confidential information of the people.

STAGE 6 - USER CONTINUE HIS NORMAL NAVIGATION

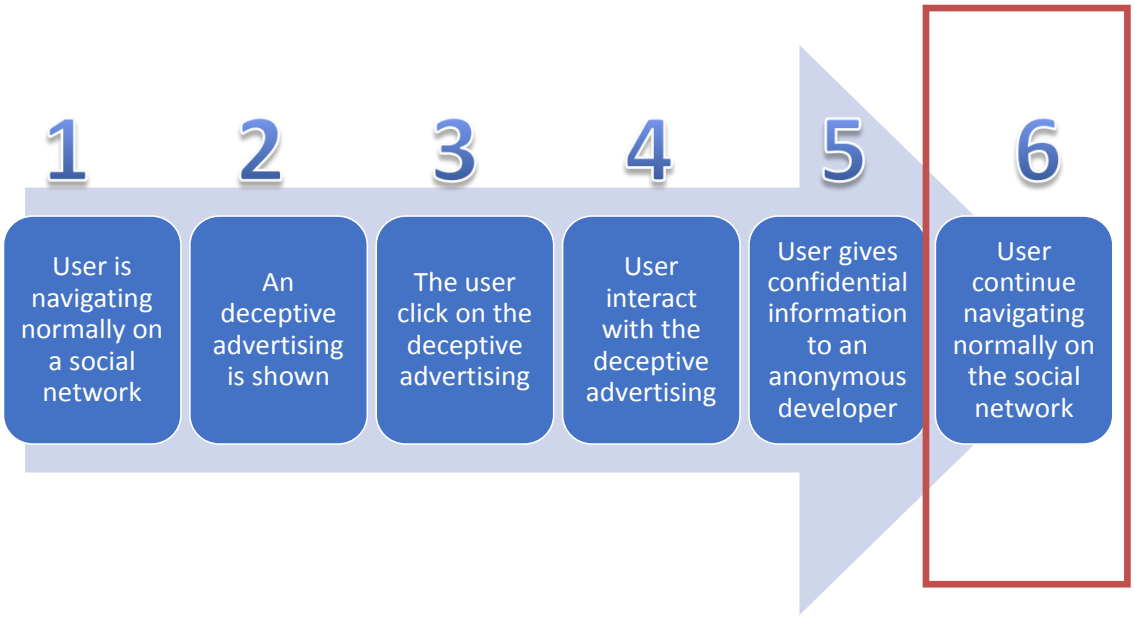

***Figure 15**: Proposed strategy implemented, Stage 6 - User continues his normal navigation*

The enormous importance of the Internet in all areas of human activity at this time is indisputable, every day there are millions of commercial transactions that take place in the network, the exchange of information between different companies and the millions of social contacts all kinds that are offered on the network.

For this case on Stage 6 the ethical attack will finish and the user will continue navigating normally on the social network without realizing that he has just been a victim of theft of extremely valuable confidential information.

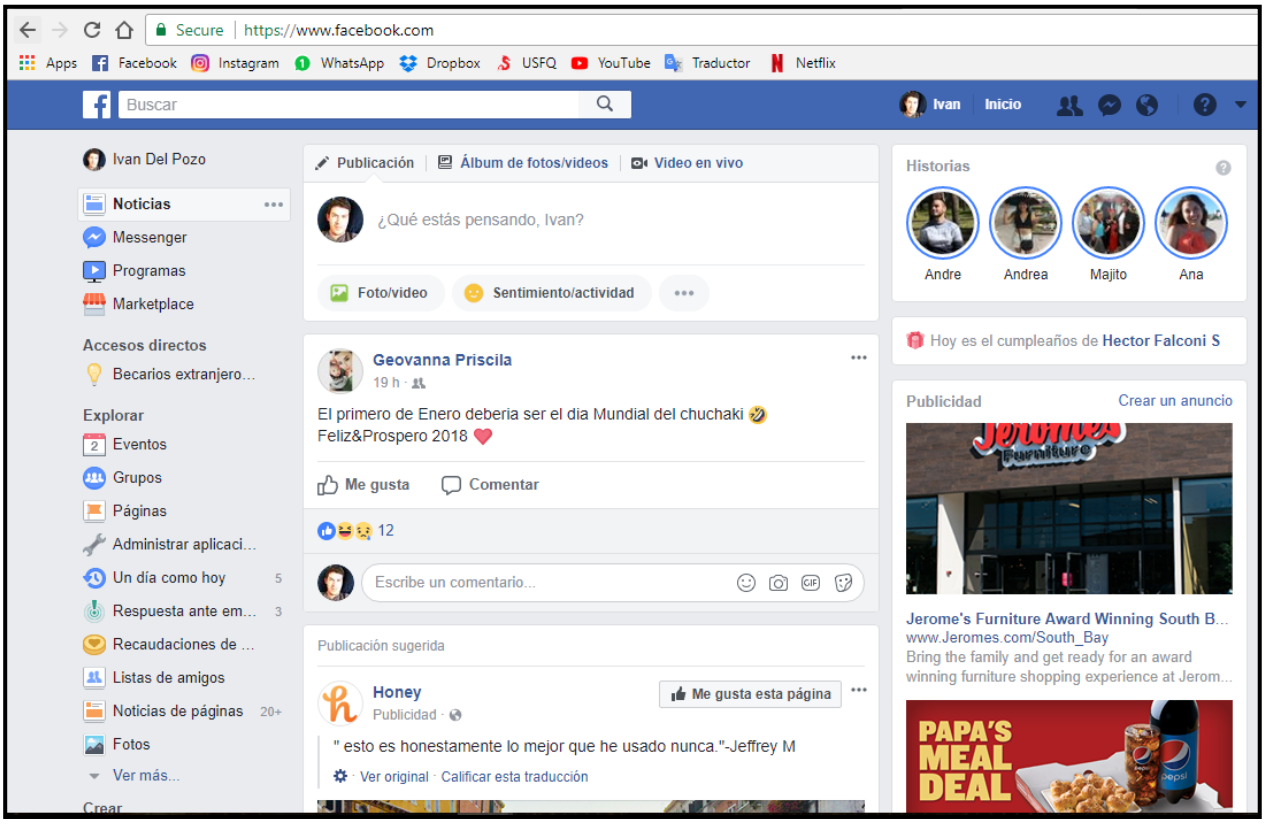

***Figure 16**: Facebook window, user comeback to the Facebook Page*

Social networks have become a gateway to the exchange and an immediate possibility to meet new people, with the main ingredient that we can build true virtual community where you have a lot of options to share and know new friends, games, confidential information, exchange experiences, which later derive in meetings that exceed the virtuality of the relationship and that's where we find that just as there are very special youth relationships, other people are also using these networks to contact others with false information and impersonation to achieve their purposes. Therefore, the importance of learning to identify safe and reliable navigation sites.

IMPLEMENTATION

Unlike the publication of a book, or the production of a material to be broadcast on television, the implementation of a page on the Web is relatively simple if you have the adequate tools. On the other hand, the cost demanded by this action is currently relatively low (or null in some cases) and it can be accessed by more users in any part of the world.

This strategy implemented understands about usability and how to create a site that customers want to navigate around because it is so easy to do.

In the other hand it takes into account the best tools to generate and provide an experience of an easy navigability, usability, information architecture, design, interactivity and the interaction of media such as text, image, audio, web links and video (among others). In order to capture the user's attention and in that way get their information, which is the main purpose of the deceptive advertisement. So the website was made to work in:

- With different types of technologies
- In different monitors and platforms
- Accessible for any type of user

All the design of the complete web page which contains the deceptive advertisement was made in HTML5 and CSS3, which is compatible with the majority of all modern browsers; you can make it as responsive design and enables connectivity and access to several new interactive tools (Deng, Wu, Yan, & Zhang, 2017).

Unlike where is the logic for the analysis of the information obtained, which is developed in the C programming language. The general architecture of proposed strategy implemented is shown in the Figure 17.

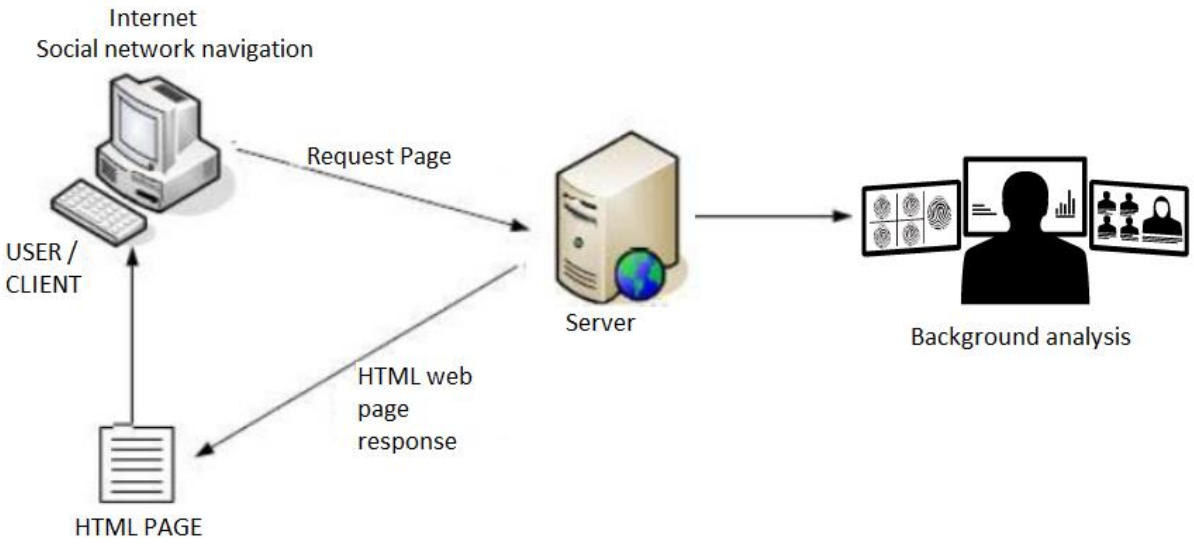

***Figure 17**: General architecture of proposed strategy implemented*

As we can appreciate is a normal Client – Server architecture with the additional component that on the server. All the personal information collected from the test of the deceptive advertisement is subjected to an analysis that the user never finds out and in this way to be able to fulfill the purposes of the attacker. The background analysis will be done manually in the computer´s attacker.

As we mentioned before we will obtain two test files. One corresponds to the entire form of the test and the second file corresponds to the personal information such as the email and the password of the user. This second file is the one which will be evaluated and where the complete analysis will be done, which will be compared with a previously preset dictionary (the greater the dictionary words, the more accurate the statistics will be) and it will show us different individual and global statistics, in order to know all the information and training the passwords of the users.

The parameters and statistics that will be evaluated will be:

1. Length of the password
2. Number of lowercase characters
3. Number of uppercase characters
4. Number of digits

5. Number of specials characters
6. Composed by two or more words
7. Root of the password
8. Match with information of the form (hobby/ favorite drink/ birthday / color)
9. Composed by the email user

Chapter 4. STATISTICAL ANALYSIS

Statistical analysis of the results obtained to determinate about vulnerability of deceptive advertising and about password elaboration.

Once the strategy to develop an ethical attack on one popular social network was designed, the next step is that based on the information collected, perform statistical analysis of the results obtained in order to know how many people are vulnerable to deceptive advertising and how computer attackers can get advantage with the information achieved from the attack to breakages or guess users passwords.

DECEPTIVE ADVERTISING

The deceptive ad was published in a particular Ecuadorian Fan Page. This fan page has as particularity that it publishes false news, jokes and ridicule of characters of show business or politics. As a result, this Fan Page that has a large number of followers, which is perfect to carry out our implementation of malicious advertising.

ANALYSIS ABOUT VULNERABILITY OF DECEPTIVE ADVERTISING

Cybercrimes share some similarities with crimes that have existed for centuries before the advent of the cyber space. So for the purpose of our research, the first analysis takes place in stage 3 (Figure 6), during just one week in order to know the population that is vulnerable on deceptive advertisement, and at list clicked on a not trusted website.

$$\forall\, n \in \mathbb{N}\ \exists\{n | n\ = total\ population\}$$

Where:

$$n = 228250\ people$$

228250 Represents the number of people how follows the page which was used for sharing the ad. From of all this population we have:

$$\forall\, m \in \mathbb{N}\ \exists\{m | m\ = users\ which\ have\ accessed\ to\ the\ not - trust\ website\}$$

Where:

$$m = 11645\ people$$

So we can say that:

$$\Rightarrow m * \frac{100}{n}$$

$$\Rightarrow 11645 * \frac{100}{228250}$$

$$5{,}101\ \%$$

Complementing the previous analysis, we will do the confidence interval analysis. The confidence interval is a range around the estimate obtained where, with the level of significance set, we have the confidence to find the true value of the estimated parameter. And this with all the intervals that we can build from all the samples of the same size (Xiao & Wang, 2016). Confidence is measured in terms of probability and represents how often the true percentage of the entire population who would select an answer lies within the confidence interval. (Xiao & Wang, 2016).

For this case we will use 95% confidence level. A confidence interval with a confidence level of 95% does not mean that the probability of finding the population parameter between these margins is 0.95. What it really means is that if we extract a certain number of samples of the same size from a population with a constant value parameter. 95% of the confidence intervals constructed from those samples will contain the value of the parameter we are looking for and the 5 remaining percentage will not contain it (Xiao & Wang, 2016).

With 95% confidence level and using the Binomial proportion confidence interval with the Wilson score interval, which is an improvement over the normal approximation interval in that the actual coverage probability is nearest to the nominal value (Qian, Shen, Mo, & Chen, 2016). We have that:

Where:

$$\hat{p} = 0{,}05101\ ;\ z = 1{,}961$$

$$\frac{\hat{p} + \frac{z^2}{2n} \mp \frac{z}{1 + \frac{z^2}{n}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z^2}{4n^2}}}{1 + \frac{z^2}{n}}$$

$$= 0{,}051007 \mp 9{,}021 * 10^{-4}$$

And represented on a confidence interval graphic as shown in Figure 18 below:

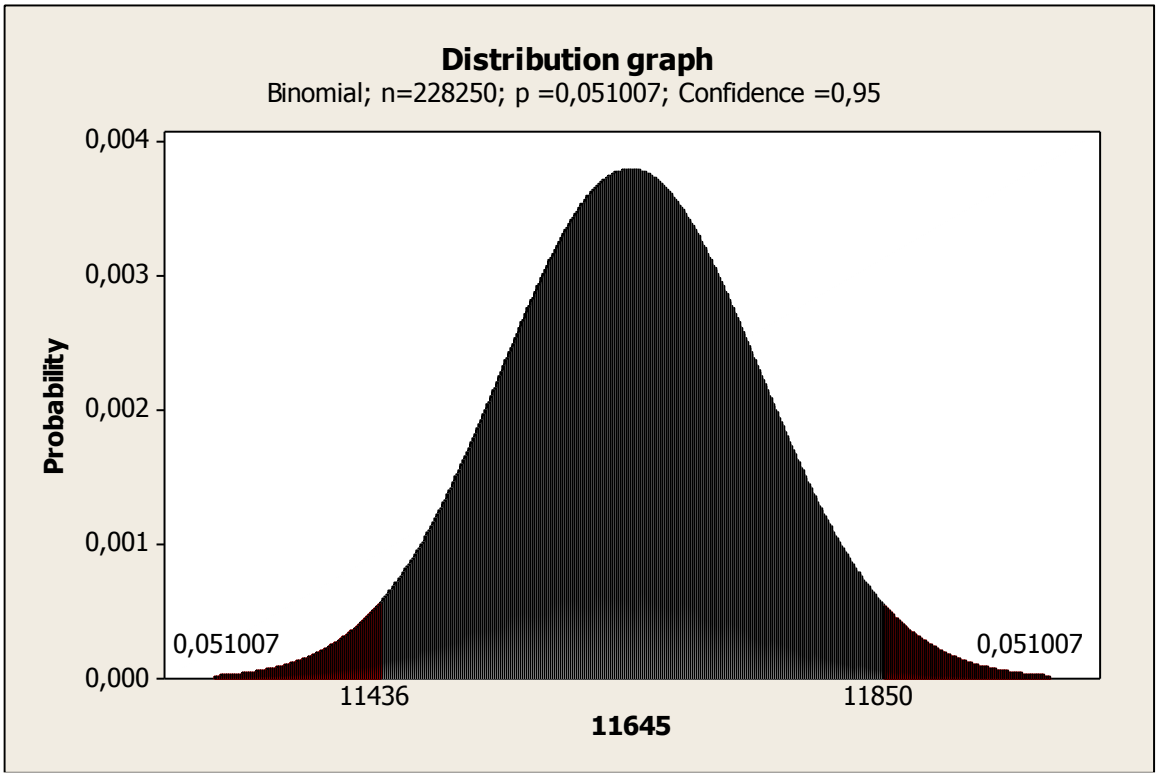

***Figure 18**: Confidence interval range*

There are many and a varied victims of cybercrime and this number represents the percentage of people that are potentially vulnerable to be good targets of social engineering attacks. In just one week; almost the twentieth part of the population fell into the lie of deceptive advertising which is the first step for computer frauds. These people are more vulnerable to getting carried away by the visual things, fooling their mind playing with their psychology to get information easily. To believe in any advertisement in any of the advertising distribution methods depends on whether we

can distinguish between realistic advertising, well-applied advertising, and advertising that is simply misleading. Moreover, understanding who the criminal is likely to target can assist in taking preemptive actions to forewarn and prepare for all forms of attack.

Based on the previous analysis and taking into account that Facebook has *1.71 billion users* (Smith, 2016):

$$if\ n\ \rightarrow 1.71\ billon\ Facebook\ users \Rightarrow k\ \rightarrow\ 87\ \ million\ users$$

This is a greater number of users to access to this fake web site. However, it should be considered that the study was carried out in Latin America. Where people have a different education level from other continents and countries of the first world and where religion prevails against science. So the impact of deceptive advertisement in social networks is not easy to determine and generalize. Moreover it should be considered that if this deceptive test is left for longer on the web it can be distributed more quickly to a larger number of people. But what can be concluded is:

$$\forall\ n\ \wedge m \in \mathbb{N}\ \exists\{n|n\ = total\ population\} \wedge \exists\{m|m\ = users\ which\ have\ accessed\ to\ the\ not-trust\ website\}$$

For the second purpose of our research, this analysis takes place in stage 4 and 5 (Figure 7 and Figure 8) in order to determinate how many people give confidential information to the computer attackers.

So

$$\forall\ j \in \mathbb{N}\ \exists\{\ j|\ j\ = users\ who\ did\ the\ questionnaire\}$$

$$j = 9823$$

$$\forall\ k \in \mathbb{N}\ \exists\{k|\ k\ = users\ who\ tried\ to\ "share"\ the\ results\ and\ re\ login\ on\ Facebook\}$$

$$k = 127$$

Taking in account the previous conclusion, it should be considered that if this deceptive test is left for longer on the web, there is a greater probability that more users will deliver the confidential information of the email and the password.

ANALYSIS ABOUT PASSWORD ELABORATION

The passwords are one of the pillars of cybersecurity, since they are the way we have for the device or service with which we interact to identify us, but the user is still not aware of the need to have a robust password in his digital life to protect his personal and confidential information.

In order to analyze the password in detail, we used the information retrieval technique “Single pass in memory indexing” using inverted index with counts. It uses term instead of termID’s, writes each blocks dictionary to disk and then starts a new dictionary for next block (Manning, C. D., Raghavan, P., & Schütze, H. 2008). Additionally this implementation supports better ranking algorithms and works as follows:

```
1 output_ file = NEWFILE ()
2 dictionary = NEWHASH ()
3 while (free memory available)
4     do token ← next (token_stream)
5           if term (token) /∈ dictionary
6             then postings_list = ADDTODICTIONARY(dictionary, term(token))
7           else postings_list = GETPOSTINGSLIST(dictionary, term(token),
            counter)
8           if full(postings_list)
9             then postings_list=DOUBLEPOSTINGSLIST(dictionary, term(token),
            counter)
10    ADDTOPOSTINGSLIST (postings_list, docID(token))
11 sorted_terms ← SORTTERMS (dictionary)
12 WRITEBLOCKTODISK (sorted_terms, dictionary, output_ file)
13 MERGEBLOCKS (f1, . . . , fn; fmerged)
14 return output_ file
```

Single pass in memory indexing using inverted index with counts is called repeatedly on each that we will consider as token, until the entire collection has been processed.

All tokens are processed one at a time as we can appreciate it at line 4. During each successive call of Single pass in memory indexing using inverted index with counts. When a term, or a word or character, occurs for the first time, it is added to the dictionary, and a new postings list is stored as we can appreciate it at line 6. The call in line 7 returns this postings list for subsequent occurrences of the term. Each postings list is dynamic and it is immediately available to collect postings. This has two main advantages: It is faster and saves memory because there is no sorting required, and we keep track of the term a postings list belongs to, which makes it easier for us to get the number of repetitions of each term.

When the memory had run out, the index of the block, which consists of the dictionary plus the postings lists to disk at we can appreciate at line 12. We sort the terms at line 11 in order to facilitate the final merging step. So the last step of Single pass in memory indexing using inverted index with counts is then to merge the blocks into the final inverted index at line 13. So then our output file is ready to be used for the attacker.

Once all the output files are saved, and taking in account the previous analysis $k = 127$; we have this passwords from the users:

```
Mom123
karinamorenof
1949
and271091
davidobregon123
carladp2295
ligacampeon
0987598064m
4,17132E+12
Lucas2017
Reyesfernan
Mundial2018
eve_90
Exodia
Hellokitty2
mcest1192
AnaGV
Miclaveentuculo@miverga.com!nohagascaeralagenteeningenieriasocial
clave1992
Delangel21
DianaEnviado
Interdin.1
Casalindaenquito
Noentiendo
Vane92
Qwerty
Caballoelpoder
Caballos
matiysofi
Santiagoh93
```

ggaattooddoo88
1714302344
EC91S07H17
0984251047mcfv
Tu mama
Mafesafe28
1234n
LulÃ°
1809cefb
vivichester
caballos
Derbi
Cantinero
Cantinero
Jrjjjjj
A09091996
898978
Lulita2020
13razonesporque
delfin
Camille23
Bucky10
Jessicarico
hola
Lileyasiqueiros
56a04k78b77f
Alejandrafiva
Latreguaggm
Fotos y videos
Codeblue
Sisisisisisisi7
Economia2018
984255459
Fack u
EstefiChiri25
Dolores
DLFigueroa@81
bucaramanga
El nombre de mi mejor amiga
mamita
Medschool
Toyotacorazon
joel.ch.1088
BSCtupapa
Alejorr906
diosesamor
Junior
charito2000
celica
celica94
deamoemosho
maisadp
Divino
90020158
Für immer
2523386
Argentina18
correo17
Mila
0874329133##)
Medschool
Fbdani77
Isabella2017#
mges@ecuador.ec
emilio
nogracias

```
alesandro
valentina
37619394428
papaDios
Nuevanueva12345
Fico
Bello
Wingo13
Salesianos.com
15ars
RiobambaPoloClub
Olmedo
Imoreno
Alejandromofi
Izamar
23031999
Miriampvj95
Fjaoo
ccbrera
monita123
GabyMoralesC
rodriline
RaptorF150
superorejon
holanegrita
Tigresa
Rafael3012.
DesarrolloAplicaciones
pilotoPorsiempre
A01202340_06s
iKSS36sTyyy
```

As general data, the single pass in memory indexing using inverted index with counts program used shows that:

```
Mom123:
NOT found in the dictionary
Contains the words / terms:

karinamorenof:
NOT found in the dictionary
Contains the words / terms: en, ka, no, amo, mor, ore, ore, rin, amor, more,
more, oren, reno, moren, moreno,

1949:
NOT found in the dictionary
Contains the words / terms:

and271091:
NOT found in the dictionary
Contains the words / terms: and

davidobregon123:
NOT found in the dictionary
Contains the words / terms: da, id, vi, ido, vid, obre, obre, rego, avido,

carladp2295:
NOT found in the dictionary
Contains the words / terms: ad, la,

ligacampeon:
NOT found in the dictionary
```

Contains the words / terms: pe, aca, liga, peon, acampe, acampe, campeo, campeo, campeon,

0987598064m:
NOT found in the dictionary
Contains the words / terms:

4,17132E+12:
NOT found in the dictionary
Contains the words / terms:

Lucas2017:
NOT found in the dictionary
Contains the words / terms: as, lucas,

Reyesfernan:
NOT found in the dictionary
Contains the words / terms: es, fe, ye, rey, yes, reyes,

Mundial2018:
NOT found in the dictionary
Contains the words / terms: al, di, mu, un, dia, mundial,

eve_90:
NOT found in the dictionary
Contains the words / terms: ve,

Exodia:
NOT found in the dictionary
Contains the words / terms: di, ex, dia, odia,

Hellokitty2:
NOT found in the dictionary
Contains the words / terms: el, he, ll, lo, el, ello,

mcest1192:
NOT found in the dictionary
Contains the words / terms: ce, es, ces,

AnaGV:
NOT found in the dictionary
Contains the words / terms: ana,

Miclaveentuculo@miverga.com!nohagascaeralagenteeningenieriasocial:
NOT found in the dictionary
Contains the words / terms: al, as, cu, en, ge, ha, la, lo, mi, mi, ni, no, oh, so, te, tu, te, tu, ve, ala, ala, aso, aso, ave, cae, cia, era, gas, gen, ria, ver, caer, ente, eral, haga, lave, lave, rala, rias, caera, clave, clave, erala, gente, hagas, socia, agente, asocia, social, ingenie, ingenie, ingenieria, ingenierias,

clave1992:
NOT found in the dictionary
Contains the words / terms: la, ve, ave, lave, lave, clave, clave,

Delangel21:
NOT found in the dictionary
Contains the words / terms: de, de, el, ge, la, el, del, dela, angel,

DianaEnviado:
NOT found in the dictionary
Contains the words / terms: ad, di, en, vi, ana, dia, via, envia, enviad, enviado,

Interdin.1:
NOT found in the dictionary

Contains the words / terms: di, te, te,

Casalindaenquito:
NOT found in the dictionary
Contains the words / terms: al, as, da, en, asa, sal, casa, sali, linda, quito, quito, alinda,

Noentiendo:
NOT found in the dictionary
Contains the words / terms: en, no, ti, noe, tiendo, entiendo,

Vane92:
NOT found in the dictionary
Contains the words / terms: va, van,

Qwerty:
NOT found in the dictionary
Contains the words / terms:

Caballoelpoder:
NOT found in the dictionary
Contains the words / terms: al, de, de, el, ll, lo, el, pode, pode, cabal, poder, caballo,

Caballos:
Se encuentra en el diccionario
No es un Verbo
Contains the words / terms: al, ll, lo, os, los, cabal, caballo,

matiysofi:
NOT found in the dictionary
Contains the words / terms: fi, so, ti,

Santiagoh93:
NOT found in the dictionary
Contains the words / terms: oh, ti, san, tia, santiago,

ggaattooddoo88:
NOT found in the dictionary
Contains the words / terms:

1714302344:
NOT found in the dictionary
Contains the words / terms:

EC91S07H17:
NOT found in the dictionary
Contains the words / terms:

0984251047mcfv:
NOT found in the dictionary
Contains the words / terms:

Tu mama:
NOT found in the dictionary
Contains the words / terms: tu, tu, ama, mama, mama,

Mafesafe28:
NOT found in the dictionary
Contains the words / terms: es, fe, esa, esa,

1234n:
NOT found in the dictionary
Contains the words / terms:

LulÃ° :

NOT found in the dictionary
Contains the words / terms:

1809cefb:
NOT found in the dictionary
Contains the words / terms: ce,

vivichester:
NOT found in the dictionary
Contains the words / terms: ch, es, he, te, te, vi, che, este, este, vivi, este, ester,

caballos:
Se encuentra en el diccionario
No es un Verbo
Contains the words / terms: al, ll, lo, os, los, cabal, caballo,

Derbi :
NOT found in the dictionary
Contains the words / terms: de, de,

Cantinero:
Se encuentra en el diccionario
No es un Verbo
Contains the words / terms: ro, ti, can,

Cantinero:
Se encuentra en el diccionario
No es un Verbo
Contains the words / terms: ro, ti, can,

Jrjjjjj:
NOT found in the dictionary
Contains the words / terms:

A09091996:
NOT found in the dictionary
Contains the words / terms:

898978:
NOT found in the dictionary
Contains the words / terms:

Lulita2020:
NOT found in the dictionary
Contains the words / terms:

13razonesporque:
NOT found in the dictionary
Contains the words / terms: es, por, que, que, razon, porque, porque, razone, razone, razones,

delfin:
Se encuentra en el diccionario
No es un Verbo
Contains the words / terms: de, de, el, fi, el, del, fin,

Camille23:
NOT found in the dictionary
Contains the words / terms: le, ll, mi, mi, mil,

Bucky10:
NOT found in the dictionary
Contains the words / terms:

Jessicarico:

NOT found in the dictionary
Contains the words / terms: es, si, si, rico,

hola:
Se encuentra en el diccionario
No es un Verbo
Contains the words / terms: la, ola,

Lileyasiqueiros:
NOT found in the dictionary
Contains the words / terms: as, ir, le, os, ro, si, si, ya, asi, ley, que, que,

56a04k78b77f:
NOT found in the dictionary
Contains the words / terms:

Alejandrafiva:
NOT found in the dictionary
Contains the words / terms: al, fi, le, va, jan, aleja, alejan, alejandra,

Latreguaggm:
NOT found in the dictionary
Contains the words / terms: la, tregua,

Fotos y videos:
NOT found in the dictionary
Contains the words / terms: de, de, id, os, vi, tos, vid, foto, ideo, ideo, fotos, video,

Codeblue:
NOT found in the dictionary
Contains the words / terms: de, de,

Sisisisisisisi7:
NOT found in the dictionary
Contains the words / terms: si, si,

Economia2018:
NOT found in the dictionary
Contains the words / terms: mi, mi, no, con, eco, mia, cono, economia,

984255459:
NOT found in the dictionary
Contains the words / terms:

Fack u:
NOT found in the dictionary
Contains the words / terms:

EstefiChiri25:
NOT found in the dictionary
Contains the words / terms: ch, es, fi, ir, te, te, este, este, este,

Dolores:
Se encuentra en el diccionario
No es un Verbo
Contains the words / terms: es, lo, ore, ore, res, dolo, olor, ores, dolor, lores, olores,

DLFigueroa@81:
NOT found in the dictionary
Contains the words / terms: fi, ro, roa, figueroa,

bucaramanga:
NOT found in the dictionary

Contains the words / terms: ama, ara, aman, cara, rama, manga,

El nombre de mi mejor amiga :
NOT found in the dictionary
Contains the words / terms: de, de, el, me, mi, mi, no, el, miga, amiga, mejor, nombre, nombre,

mamita:
NOT found in the dictionary
Contains the words / terms: mi, mi, mita,

Medschool:
NOT found in the dictionary
Contains the words / terms: ch, me,

Toyotacorazon:
NOT found in the dictionary
Contains the words / terms: yo, ora, oyo, tac, taco, razon, corazon,

joel.ch.1088:
NOT found in the dictionary
Contains the words / terms: ch, el, el,

BSCtupapa:
NOT found in the dictionary
Contains the words / terms: tu, tu, papa, papa, tupa,

Alejorr906:
NOT found in the dictionary
Contains the words / terms: al, le, alejo, alejo,

diosesamor:
NOT found in the dictionary
Contains the words / terms: di, es, os, se, se, amo, dio, esa, mor, ose, ose, esa, amor, dios, oses, dioses, sesamo,

Junior:
NOT found in the dictionary
Contains the words / terms: ni, un, uni, unio, junio,

charito2000:
NOT found in the dictionary
Contains the words / terms: ch, ha, rito,

celica:
NOT found in the dictionary
Contains the words / terms: ce, el, el,

celica94:
NOT found in the dictionary
Contains the words / terms: ce, el, el,

deamoemosho:
NOT found in the dictionary
Contains the words / terms: de, de, os, amo,

maisadp:
NOT found in the dictionary
Contains the words / terms: ad,

Divino:
Se encuentra en el diccionario
No es un Verbo
Contains the words / terms: di, no, vi, vino,

90020158:

NOT found in the dictionary
Contains the words / terms:

Für immer:
NOT found in the dictionary
Contains the words / terms: me,

2523386:
NOT found in the dictionary
Contains the words / terms:

Argentina18:
NOT found in the dictionary
Contains the words / terms: en, ge, ti, gen, tina, argentina,

correo17:
NOT found in the dictionary
Contains the words / terms: reo, corre, correo,

Mila:
NOT found in the dictionary
Contains the words / terms: la, mi, mi, mil,

0874329133##):
NOT found in the dictionary
Contains the words / terms:

Medschool:
NOT found in the dictionary
Contains the words / terms: ch, me,

Fbdani77:
NOT found in the dictionary
Contains the words / terms: da, ni, dan,

Isabella2017#:
NOT found in the dictionary
Contains the words / terms: be, el, la, ll, el, bel, abel, ella, sabe, bella, isabel,

mges@ecuador.ec:
NOT found in the dictionary
Contains the words / terms: ad, cu, es, ge, ges, ecuador,

emilio:
Se encuentra en el diccionario
No es un Verbo
Contains the words / terms: mi, mi, lio, lio, mil, mili,

nogracias:
NOT found in the dictionary
Contains the words / terms: as, no, cia, gracia, gracias,

alesandro:
NOT found in the dictionary
Contains the words / terms: al, es, le, ro, esa, les, san, esa, lesa,

valentina:
Se encuentra en el diccionario
No es un Verbo
Contains the words / terms: al, en, le, ti, va, len, tina, vale, valen, valentin,

37619394428:
NOT found in the dictionary
Contains the words / terms:

papaDios:
NOT found in the dictionary
Contains the words / terms: ad, di, os, dio, dios, papa, papa, adios,

Nuevanueva12345:
NOT found in the dictionary
Contains the words / terms: nu, va, eva, van, nueva,

Fico:
NOT found in the dictionary
Contains the words / terms: fi,

Bello:
Se encuentra en el diccionario
No es un Verbo
Contains the words / terms: be, el, ll, lo, el, bel, ello,

Wingo13 :
NOT found in the dictionary
Contains the words / terms:

Salesianos.com:
NOT found in the dictionary
Contains the words / terms: al, es, le, no, os, si, si, ano, les, nos, sal, anos, sale, sale, sales,

15ars:
NOT found in the dictionary
Contains the words / terms:

RiobambaPoloClub:
NOT found in the dictionary
Contains the words / terms: lo, rio, rio, club, polo, apolo, bamba,

Olmedo:
Se encuentra en el diccionario
No es un Verbo
Contains the words / terms: me,

Imoreno:
NOT found in the dictionary
Contains the words / terms: en, no, mor, ore, ore, more, more, oren, reno, moren, moreno,

Alejandromofi:
NOT found in the dictionary
Contains the words / terms: al, fi, le, ro, jan, romo, aleja, alejan, alejandro,

Izamar :
NOT found in the dictionary
Contains the words / terms: ama, iza, mar, amar,

23031999:
NOT found in the dictionary
Contains the words / terms:

Miriampvj95:
NOT found in the dictionary
Contains the words / terms: ir, mi, mi, ria, iria,

Fjaoo:
NOT found in the dictionary
Contains the words / terms:

ccbrera:
NOT found in the dictionary
Contains the words / terms: era,

monita123:
NOT found in the dictionary
Contains the words / terms: ni, monita,

GabyMoralesC:
NOT found in the dictionary
Contains the words / terms: al, es, le, les, mor, ora, mora, oral, moral, orales, morales,

rodriline:
NOT found in the dictionary
Contains the words / terms: ro, dril,

RaptorF150:
NOT found in the dictionary
Contains the words / terms: apto, rapto, raptor,

superorejon:
NOT found in the dictionary
Contains the words / terms: pe, ro, su, ore, ore, pero, supe, orejon, supero, supero,

holanegrita:
NOT found in the dictionary
Contains the words / terms: la, ola, hola, grita,

Tigresa:
Se encuentra en el diccionario
No es un Verbo
Contains the words / terms: es, ti, esa, res, esa, gres, tigre, tigres,

Rafael3012.:
NOT found in the dictionary
Contains the words / terms: el, el, rafael,

DesarrolloAplicaciones:
NOT found in the dictionary
Contains the words / terms: de, de, es, ll, lo, ro, des, esa, ion, loa, rol, esa, iones, plica, rollo, rollo, sarro, aplica, aplicacion, desarrollo, desarrollo, aplicaciones,

pilotoPorsiempre:
NOT found in the dictionary
Contains the words / terms: lo, pi, si, si, por, loto, topo, piloto, piloto, siempre,

A01202340_06s:
NOT found in the dictionary
Contains the words / terms:

iKSS36sTyyy:
NOT found in the dictionary
Contains the words / terms:

This shows that none of the password is a unique dictionary word (compared with our dictionary). All of them are composed by two terms or more. As global statistics we have:

```
Total number of entries: 127

From the list given 3.15% of the entries had 4 characters.
From the list given 3.937% of the entries had 5 characters.
From the list given 11.02% of the entries had 6 characters.
From the list given 7.874% of the entries had 7 characters.
From the list given 9.449% of the entries had 8 characters.
From the list given 14.96% of the entries had 9 characters.
From the list given 8.661% of the entries had 10 characters.
From the list given 16.54% of the entries had 11 characters.
From the list given 3.937% of the entries had 12 characters.
From the list given 7.087% of the entries had 13 characters.
From the list given 3.937% of the entries had 14 characters.
From the list given 4.724% of the entries had 15 characters.
From the list given 2.362% of the entries had 16 characters.
From the list given 0.7874% of the entries had 22 characters.
From the list given 0.7874% of the entries had 28 characters.
From the list given 0.7874% of the entries had 30 or more characters.
From the list given 51.97% of the entries had 1 capital letters.
From the list given 3.15% of the entries had 2 capital letters.
From the list given 3.937% of the entries had 3 capital letters.
From the list given 1.575% of the entries had 4 capital letters.
From the list given 2.362% of the entries had 1 lowercase letters.
From the list given 2.362% of the entries had 2 lowercase letters.
From the list given 4.724% of the entries had 3 lowercase letters.
From the list given 8.661% of the entries had 4 lowercase letters.
From the list given 11.02% of the entries had 5 lowercase letters.
From the list given 13.39% of the entries had 6 lowercase letters.
From the list given 10.24% of the entries had 7 lowercase letters.
From the list given 6.299% of the entries had 8 lowercase letters.
From the list given 7.874% of the entries had 9 lowercase letters.
From the list given 3.937% of the entries had 10 lowercase letters.
From the list given 5.512% of the entries had 11 lowercase letters.
From the list given 4.724% of the entries had 12 lowercase letters.
From the list given 4.724% of the entries had 13 lowercase letters.
From the list given 0.7874% of the entries had 14 lowercase letters.
From the list given 1.575% of the entries had 15 lowercase letters.
From the list given 0.7874% of the entries had 20 lowercase letters.
From the list given 0.7874% of the entries had 21 lowercase letters.
From the list given 0.7874% of the entries had 30 or more lowercase letters.
From the list given 2.362% of the entries had 1 numbers.
From the list given 14.96% of the entries had 2 numbers.
From the list given 3.937% of the entries had 3 numbers.
From the list given 11.02% of the entries had 4 numbers.
From the list given 0.7874% of the entries had 5 numbers.
From the list given 2.362% of the entries had 6 numbers.
From the list given 0.7874% of the entries had 7 numbers.
From the list given 3.937% of the entries had 8 numbers.
From the list given 0.7874% of the entries had 9 numbers.
From the list given 3.937% of the entries had 10 numbers.
From the list given 0.7874% of the entries had 11 numbers.
From the list given 9.449% of the entries had 1 special characters.
From the list given 3.15% of the entries had 2 special characters.
From the list given 3.15% of the entries had 3 special characters.
From the list given 0.7874% of the entries had 5 special characters.
From the list given 0.7874% of the entries had 6 special characters.
From the list given 16.54% of the entries were only lowercase letters.
```

```
From the list given 6.299% of the entries were only numbers.
```

And represented on column chart, with all password characteristic percentage, as shown in Figure 19:

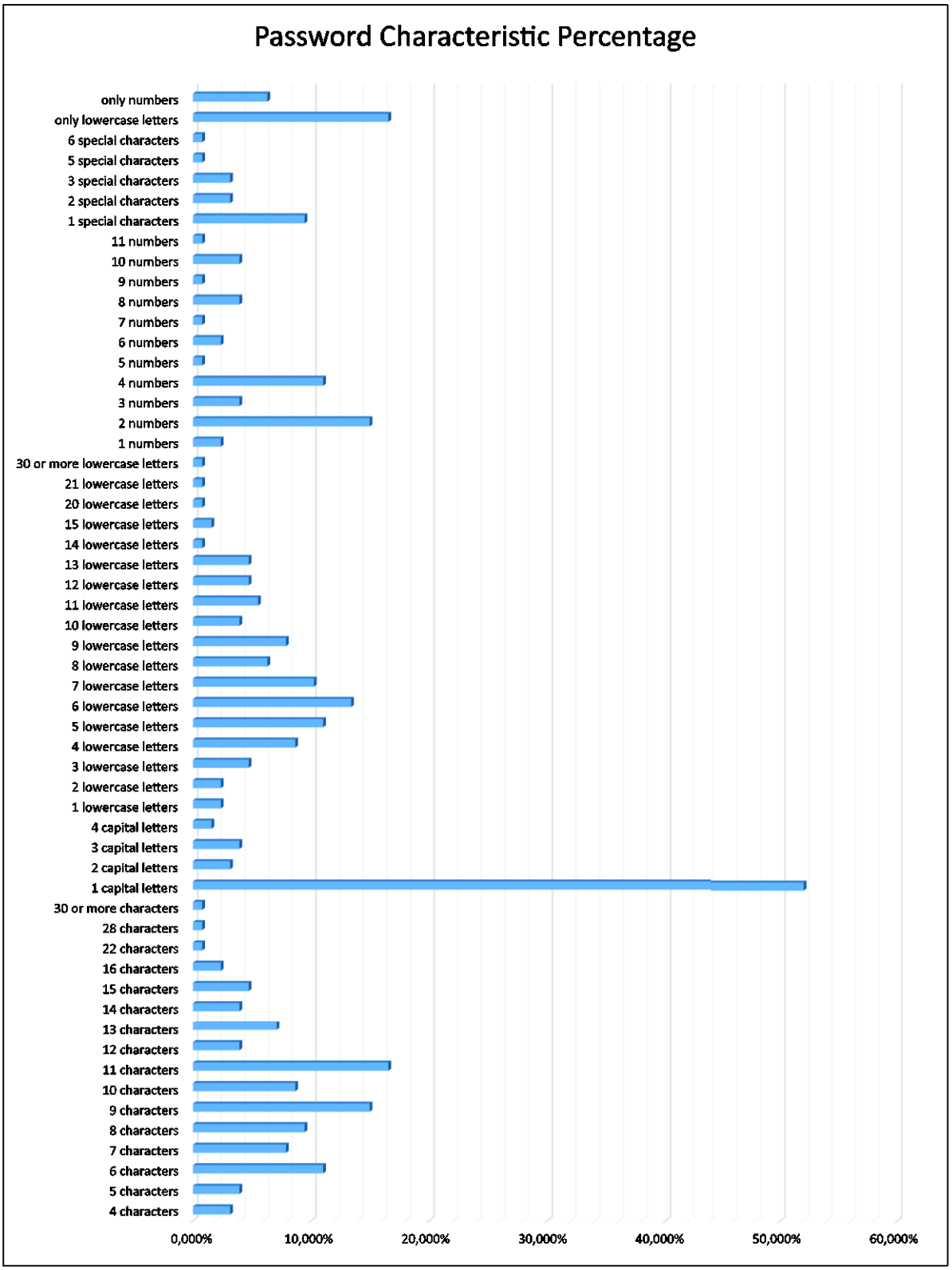

***Figure 19**: Password characteristic percentage*

This represents that the majority of the population $k = 127$ uses lowercase letters for their passwords; and specifically $16{,}54\%$ uses just only lowercase letters and $51{,}97\%$ use 1 capital letter in their password. A standard password of eight characters can be discovered in less than one minute by a current computer and if we add that it only uses letters, and only small letters, it makes it easier for the attacker to generate a much more efficient comparison dictionary, without much work or effort required. Analyzing each password individually we have to:

PASSWORD	Length of the password (# characters)	Number of uppercase characters	Number of lowercase characters	Number of digits	Number of specials characters	Composed by two or more words	Root of the password	Match with information of the form	Composed by the email user
Mom123	9	1	2	3	3	F	Mom	F	T
karinamorenof	13	0	13	0	0	T	Karina	T	T
1949	4	0	0	4	0	F	1949	T	F
and271091	9	0	3	6	0	T	and	T	F
davidobregon123	15	0	12	3	0	T	david	T	F
carladp2295	11	0	7	4	0	T	carla	T	F
ligacampeon	11	0	11	0	0	T	liga	T	F
0987598064m	11	0	1	10	0	F	987	F	F
4,17132E+12	11	1	0	8	2	F	4171	F	F
Lucas2017	9	1	4	4	0	F	Lucas	F	F
Reyesfernan	11	1	10	0	0	T	Reyes	T	F
Mundial2018	11	1	6	4	0	T	Mundial	T	F
eve_90	6	0	3	2	1	T	eve	T	F
Exodia	6	1	5	0	0	F	Exodia	T	F
Hellokitty2	11	1	9	1	0	T	Hello	F	F
mcest1192	9	0	5	4	0	T	m	T	F
AnaGV	5	3	2	0	0	T	Ana	F	F
Miclaveentuculo@miverga.com!nohagascaeralagenteeningenieriasocial	65	1	61	0	3	T	Mi	F	F
clave1992	9	0	5	4	0	T	clave	T	F
Delangel21	10	1	7	2	0	T	Del	T	T
DianaEnviado	12	2	10	0	0	T	Diana	F	F
Interdin.1	10	1	7	1	1	T	Inter	F	F
Casalindaenquito	16	1	15	0	0	T	Casa	F	F

Noentiendo	10	1	9	0	0	T	Noe	F	F
Vane92	6	1	3	2	0	T	Vane	T	F
Qwerty	6	1	5	0	0	F	Qw	T	F
Caballoelpoder	14	1	13	0	0	T	Caballo	T	F
Caballos	8	1	7	0	0	F	Caballos	T	F
matiysofi	9	0	9	0	0	T	mati	T	F
Santiagoh93	11	1	8	2	0	T	Santi	T	F
ggaattooddoo88	14	0	12	2	0	F	g	T	F
1714302344	10	0	0	10	0	F	1714	T	F
EC91S07H17	10	4	0	6	0	F	E	F	F
0984251047mcfv	14	0	4	10	0	F	984	F	F
Tu mama	7	1	5	0	1	T	Tu	F	F
Mafesafe28	10	1	7	2	0	T	Mafe	T	T
1234n	5	0	1	4	0	F	1235	F	F
LulÃº	8	1	2	0	5	F	Lul	F	F
1809cefb	8	0	4	4	0	T	1809	F	F
vivichester	11	0	11	0	0	T	vivi	F	F
caballos	8	0	8	0	0	F	caballos	T	F
Derbi	6	1	4	0	1	F	Derb	T	F
Cantinero	9	1	8	0	0	F	Cantin	T	F
Cantinero	9	1	8	0	0	F	Cantin	T	F
Jrjjjjj	7	1	6	0	0	F	Jr	F	F
A09091996	9	1	0	8	0	F	A	F	F
898978	6	0	0	6	0	F	8989	F	F
Lulita2020	10	1	5	4	0	T	Lul	F	F
13razonesporque	15	0	13	2	0	T	13	F	F
delfin	6	0	6	0	0	F	delfin	F	F
Camille23	9	1	6	2	0	T	Camil	T	F
Bucky10	7	1	4	2	0	T	Buc	T	F

Jessicarico	11	1	10	0	0	T	Jessica	T	T
hola	4	0	4	0	0	F	hola	F	F
Lileyasiqueiros	15	1	14	0	0	T	Lil	F	F
56a04k78b77f	12	0	4	8	0	T	56	F	F
Alejandrafiva	13	1	12	0	0	T	Alejandra	T	T
Latreguaggm	11	1	10	0	0	T	Lat	F	F
Fotos y videos	14	1	11	0	2	T	Fotos	T	F
Codeblue	8	1	7	0	0	T	Cod	F	F
Sisisisisisisi7	15	1	13	1	0	T	Si	F	F
Economia2018	12	1	7	4	0	T	Economia	T	F
984255459	9	0	0	9	0	F	9842	F	F
Fack u	6	1	4	0	1	T	Fac	F	F
EstefiChiri25	13	2	9	2	0	T	Fac	T	F
Dolores	7	1	6	0	0	F	Dolor	F	F
DLFigueroa@81	13	3	7	2	1	F	D	T	F
bucaramanga	11	0	11	0	0	F	buca	T	F
El nombre de mi mejor amiga	28	1	21	0	6	T	El	F	F
mamita	6	0	6	0	0	F	mam	F	F
Medschool	9	1	8	0	0	T	Med	F	F
Toyotacorazon	13	1	12	0	0	T	Toy	T	F
joel.ch.1088	12	0	6	4	2	T	joel	F	F
BSCtupapa	9	3	6	0	0	T	B	T	F
Alejorr906	10	1	6	3	0	T	Alej	F	F
diosesamor	10	0	10	0	0	T	dios	T	F
Junior	6	1	5	0	0	F	Junior	F	F
charito2000	11	0	7	4	0	T	char	F	F
celica	6	0	6	0	0	T	cel	F	F
celica94	8	0	6	2	0	T	cel	F	F
deamoemosho	11	0	11	0	0	F	dea	T	F

maisadp	7	0	7	0	0	T	mai	F	F
Divino	6	1	5	0	0	F	Divino	F	F
90020158	8	0	0	8	0	F	9002	F	F
Für immer	10	1	6	0	3	T	F	F	F
2523386	7	0	0	7	0	F	2523	F	F
Argentina18	11	1	8	2	0	T	Argentina	T	F
correo17	8	0	6	2	0	T	correo	F	F
Mila	4	1	3	0	0	F	Mill	F	F
0874329133##)	13	0	0	10	3	T	874	F	F
Medschool	9	1	8	0	0	T	Med	F	F
Fbdani77	8	1	5	2	0	T	F	F	F
Isabella2017#	13	1	7	4	1	T	Isabel	T	F
mges@ecuador.ec	15	0	13	0	2	T	m	T	F
emilio	6	0	6	0	0	F	emilio	T	F
nogracias	9	0	9	0	0	T	no	F	F
alesandro	9	0	9	0	0	T	ale	F	F
valentina	9	0	9	0	0	F	valentina	T	F
37619394428	11	0	0	11	0	F	3761	F	F
papaDios	8	1	7	0	0	T	papa	T	F
Nuevanueva12345	15	1	9	5	0	T	Nueva	F	F
Fico	4	1	3	0	0	F	Fic	F	F
Bello	5	1	4	0	0	F	Bello	F	F
Wingo13	8	1	4	2	1	T	Win	F	F
Salesianos.com	14	1	12	0	1	T	Sale	F	F
15ars	5	0	3	2	0	T	15	F	F
RiobambaPoloClub	16	3	13	0	0	T	Rio	T	F
Olmedo	6	1	5	0	0	F	Olmedo	T	F
Imoreno	7	1	6	0	0	T	I	F	F
Alejandromofi	13	1	12	0	0	T	Alejandro	F	F

Izamar	7	1	5	0	1	F	I	F	F
23031999	8	0	0	8	0	F	2303	F	F
Miriampvj95	11	1	8	2	0	T	Miriam	T	T
Fjaoo	5	1	4	0	0	F	F	T	T
ccbrera	7	0	7	0	0	T	c	T	F
monita123	9	0	6	3	0	T	mon	F	F
GabyMoralesC	12	3	9	0	0	T	Gab	F	F
rodriline	9	0	9	0	0	T	rodri	F	F
RaptorF150	10	2	5	3	0	T	Rap	T	F
superorejon	11	0	11	0	0	T	super	F	F
holanegrita	11	0	11	0	0	T	hola	F	F
Tigresa	7	1	6	0	0	T	Tigre	F	F
Rafael3012.	11	1	5	4	1	T	Rafael	T	F
DesarrolloAplicaciones	22	2	20	0	0	T	Desarrollo	F	F
pilotoPorsiempre	16	1	15	0	0	T	piloto	T	F
A01202340_06s	13	1	1	10	1	F	A	F	F
iKSS36sTyyy	11	4	5	2	0	F	i	F	F

Table 3: *Analysis about password elaboration*

Where we can conclude that:

$$p = 53\ people\ \ (Match\ with\ information\ of\ the\ form)$$

And

$$q = 8\ people\ \ (Composed\ by\ the\ email\ user)$$

So for the first parameter we have that:

$$\Rightarrow p * \frac{100}{k}$$

$$\Rightarrow 53 * \frac{100}{127}$$

$$= 41{,}73\ \%$$

And for the second parameter we have that:

$$\Rightarrow 8 * \frac{100}{k}$$

$$\Rightarrow 8 * \frac{100}{127}$$

$$= 6{,}29\ \%$$

As we can appreciate passwords that contain personal information (the birthday, human name, or pet name, or a personal identification number) may or may not be found through an attack based on dictionary passwords. However, if the attacker knows him personally (or is motivated enough to investigate his personal life), he may be able to guess his password with little or no difficulty.

In addition to dictionaries, many password crackers also include common names, dates and other information in your password search. Therefore, even if the attacker does not know that his dog's name is such, they may still discover that his password would be based on his pet, with a good password decoder. Because you upload many photos of your pet, at home the puppy has a preferential place and better treatment than even some people, and so many more things.

The advantage that hacker would have of the personal information obtained is huge, for this case it represents that he easily will guest the password of the *42.10%* from the 19 people who gave the confidential information. If people continue using passwords based on their hobbies, tastes, ideologies, birthdays, id numbers, among others, there will be no need to extract the password directly. Because the attacker know the behavior of the victim, and only to analyze them, the attacker can deduct passwords for the rest of services.

On the other hand, many times people for ease, comfort to remember passwords use the same users of your email in order not to get confused. Which for this particular case represents the *15.78%* from the 19 people who gave the confidential information. At first glance it seems that it is a low range, nevertheless we are generating a huge security gap for any person, if in fact we get to know that we are bad for technological tools, that we are easily psychologically manipulable and that we also have bad memory and we get confused with all our passwords; we are easy targets to be victims of an attack of social engineering.

Chapter 5. CONCLUSIONS

Conclusions and recommendations are presented based on the research. On the care of deceptive advertising, providing a better vision of how hackers use the psychology applied to information raising awareness and warning of the risks associated with misuse of the Internet through social networks. Finally safety recommendations between functionality and security when creating a password are mentioned.

GENERAL CONCLUSIONS

The human mind becomes a resource for storing sensitive information and can be attacked by cybercriminals through social engineering at an electronic and personal level.

In this exploratory research project through the study, tests and implementation made, it appears that people is vulnerable to deceptive advertisements, and presenting even greater risk, they give their personal information to unknown sources or people. As seen in the evidence previously discussed, psychology plays an important role in both conceptions of social engineering, since it is from the use of psychological techniques that the implementation of them is possible, and how easy is to manipulate a person in order to do what we want to do.

Furthermore, Social Engineering is an art that few develop because not all people have social skills. Human nature plays a role in shaping social life while the social structure in turn, with its habits, norms and customs also exerts an influence on people, which handled in dangerous hands triggers an attack of social engineering.

Moreover this project also shows us that our passwords are still very functional rather than secure. Passwords offer the first line of defense against unauthorized access to third parties to any application, or web service, but people create their passwords based on their personal information or easy remembering. However, there is no single method that is the best to define an adequate balance between security and convenience of direct access first and foremost, the level of security is inherent and depends of our self, which are influenced by several factors depending on the medium we are in.

RECOMMENDATIONS

To evaluate if the advertising is misleading we must refer to the review of key points that are easy to identify and that we must take into account:

- ✓ The domain of the website is external; because it does not have the corresponding security credentials.
- ✓ If the website asks you for personal data in confusing tests of dubious origin and to share them you are even asked to rewrite your password even though you were already browsing normally in the social network.
- ✓ Products, services or incentives are offered free of charge when the delivery of the same is subject to the fulfillment of any condition by the consumer that is not indicated in the commercial advertisement.
- ✓ Necessary information is omitted for the adequate understanding of commercial advertising.
- ✓ When the indispensable information for the adequate handling, maintenance, form of use, use of the good and / or service as well as precautions on possible risks, is not in Spanish language.

According to good practice, using the same password for all your online accounts is like using the same key for all your locked doors, if an attacker was to discover your password he would have the ability to access everything you own. We use so many online accounts that need passwords; it is too much to remember. As the secure keys are very difficult, not to say impossible, to remember, one option to avoid reusing passwords or using weak passwords is to use a password manager.

Save the keys in a text document. As the secure keys are very difficult, not to say impossible, to remember, the logical thing is to have them written in a text document, which will be used to store the passwords of all personal services. Each time you enter a service, you will have to resort to this document. It may be heavy, but it is safer. A password manager stores and encrypts all of your different and complex passwords. The manager can then help you to log into your online accounts

automatically. You only need to remember your master password to access the password manager and manage all of your accounts and passwords because if it is lost, the rest of the keys will be lost. The problem is that users tend to trouble remembering a password that are considered to be secure such passwords usually have a long string and appear random. So, most users tend to create simple and short passwords, representing that insecure sea. As a result, most of the time, the level of password usability has not yet reached a suitable optimization for a secure password. Through our study we are able to draw a number of recommendations and some tips for choosing a good password:

- ✓ Do not use dictionary words or names in any languages.
    - The letters can be combined with robots to find the key. Regarding words, they always have a symbolic connection with the subconscious, so that someone who knows a little the user can guess the clues if you think of the name of your partner, their children or their pets.
- ✓ Do not use common misspellings of dictionary words related with your personal affinities, hobbies, or information.
    - Mixing letters and numbers is the safest solution because two classification systems come together, which greatly expands the combinations. Anyway, a "hacker" who has some personal data about the user and a lot of psychology can guess the keys if there has been no care in making them. You have to be aware that, in an automatic way, you always look for easy-to-remember combinations related to people and important dates.
- ✓ Never use only numbers.
    - Even if keys with eight or more digits are used, if only numbers are used, it is a matter of time before a robot finds the password and enters the person's pages.
- ✓ Do not use just one capital letter

- It is reasonable to conclude that with the half of the population using one capital letter in their password, attackers will induce always to put the first letter of the password in capital letter. So it will be better to put more than one capital letter throughout the entire password.

✓ Do not use computer names or account names

- Therefore, after writing the password it is important to check that it does not contain personal clues.

✓ If possible use special characters, such as ! @ # $ % ^ & * ( )

- A trick that will allow to use letters and numbers related to the life of the user without danger is to intersperse symbols like "#", "$", "&" or "%" between the characters of the password. Its presence is much harder to discover for hackers and robots.

✓ Use a ten or more character length password

- The fewer characters according to a key, the easier it is to break it for a hacker, since the number of possible combinations is less. "Weak" combinations are considered smaller than eight digits, which can be identified with programs that generate random combinations (called robots), which is known as "brute force".
- The longer length makes passphrases less vulnerable to dictionary or brute force attacks.

✓ Changes periodically your relevant passwords

- It is advisable to change the most important passwords every so often.

✓ Beware of open sessions.

- On many occasions, users keep sessions of different online services open in the browser. In this way, in case of losing the computer or leaving the session open in a public terminal or a third party, it can endanger your privacy and security, by facilitating access to your account. To minimize this risk, a recommendable option is to leave

all the services of habitual use, be it the electronic mail, the different social networks where one participates or the platforms where documents are kept to synchronize them, etc.

FURTHER WORK

This proposed strategy has been implemented for obtaining user email and user password for one social network. Further work must focus on banking transactions or online payments, in order to verify if, when there is money in between, people are psychologically not so easy to manipulate, reject deceptive advertising and, as a result social engineering loses its influence.

REFERENCES

Atkins, B., & Huang, W. (2013). A Study of Social Engineering in Online Frauds. *Open Journal of Social Sciences*, *1*(3), 23–32. https://doi.org/10.4236/jss.2013.13004

Aziz, M. T. (2016). Sentiment Analysis On Facebook Group Using Lexicon Based Approach, 8–11.

Bhise, K. (2016). A Method For Recognize Malignant Facebook Application, 41–44.

Chan, R. Y., Ho, K. M., Jia, S., Wang, Y., Yan, X., & Yu, X. (2016). Facebook and Information Security Education What Can We Know from Social Network Analyses on Hong Kong Engineering Students ?, (December), 303–307.

Del Pozo, I., & Iturralde, M. (2015). CI: A new encryption mechanism for instant messaging in mobile devices. In *Procedia Computer Science*. https://doi.org/10.1016/j.procs.2015.08.381

Deng, X., Wu, T., Yan, J., & Zhang, J. (2017). Combinatorial Testing on Implementations of HTML5 Support. *Proceedings - 10th IEEE International Conference on Software Testing, Verification and Validation Workshops, ICSTW 2017*, 262–271. https://doi.org/10.1109/ICSTW.2017.47

Ekta Science, C. (2016). Reputation Based Technique to Distinguish Posts in Facebook Social Network.

Ferreira, A., & Lenzini, G. (2015). An analysis of social engineering principles in effective phishing. *2015 Workshop on Socio-Technical Aspects in Security and Trust*, 9–16. https://doi.org/10.1109/STAST.2015.10

Frauenstein, E. D., & Flowerday, S. V. (2016). Social Network Phishing : Becoming Habituated to Clicks and Ignorant to Threats ?, 98–105.

Ghafir, I., Prenosil, V., Alhejailan, A., & Hammoudeh, M. (2016). Social Engineering Attack Strategies and Defence Approaches. *2016 IEEE 4th International Conference on Future Internet of Things and Cloud (FiCloud)*, 145–149. https://doi.org/10.1109/FiCloud.2016.28

Grande, C. E. L., & Guadrón, R. S. (2015). Social Engineering : The Silent Attack, (Concapan Xxxv).

Heartfield, R., Loukas, G., & Gan, D. (2016). You are probably not the weakest link: Towards practical prediction of susceptibility to semantic social engineering attacks. *IEEE Access*, *4*, 1–1. https://doi.org/10.1109/ACCESS.2016.2616285

Huang, T. H.-D., Yu, C.-M., & Kao, H.-Y. (2017). Data-Driven and Deep Learning Methodology for Deceptive Advertising and Phone Scams Detection. Retrieved from http://arxiv.org/abs/1710.05305

Islam, S., & Dey, J. J. (2016). Supervised Approach of Sentimentality Extraction, 383–387.

Kotenko, I., Stepashkin, M., & Doynikova, E. (2011). Security Analysis of Information Systems taking into account Social Engineering Attacks. https://doi.org/10.1109/PDP.2011.62

Longe, O. B., Mbarika, V., Kourouma, M., Wada, F., & Isabalija, R. (2010). Seeing Beyond the Surface, Understanding and Tracking Fraudulent Cyber Activities. *International Journal of Computer Science and Information Security*, *6*(3), 12. Retrieved from http://arxiv.org/abs/1001.1993

Planck, M., Development, H., Wu, C. M., Meder, B., & Nelson, J. D. (2017). Asking Better Questions : How Presentation Formats Influence Information Search, *43*(8), 1274–1297.

Qian, Z., Shen, B., Mo, W., & Chen, Y. (2016). SatiIndicator: Leveraging User Reviews to Evaluate User Satisfaction of SourceForge Projects. *Proceedings - International Computer Software and Applications Conference*, *1*, 93–102. https://doi.org/10.1109/COMPSAC.2016.183

Rachsuda, J. (2017). User preferences profiling based on user behaviors on facebook page categories, 248–253.

Stjohn-green, M., Piggin, R., Mcdermid, J. A., & Oates, R. (2015). Combined Security and Safety Risk Assessment – What Needs To Be Done for Ics and the Iot . *The 10th International Conference on System Safety and Cyber Security Conference*. https://doi.org/10.1049/cp.2015.0284

Terzi, D. S., Terzi, R., & Sagiroglu, S. (2017). Big data analytics for network anomaly detection from netflow data. *2017 International Conference on Computer Science*

and Engineering (UBMK), 592–597. https://doi.org/10.1109/UBMK.2017.8093473

Wang, L., Han, X., & Chen, L. (2016). An Empirical Study on Preference Distribution of Facebook Users. *2016 Intl IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld)*, 1069–1073. https://doi.org/10.1109/UIC-ATC-ScalCom-CBDCom-IoP-SmartWorld.2016.0167

Xiao, J., & Wang, H. (2016). An Approch to Predict the Confidence Interval of Web Services QoS Based on Bootstrap. *2016 9th International Conference on Service Science (ICSS)*, 100–107. https://doi.org/10.1109/ICSS.2016.20

Zingerle, A. (2014). How to obtain passwords of online scammers by using social engineering methods. https://doi.org/10.1109/CW.2014.54

Zisiadis, D., Kopsidas, S., Varalis, A., & Tassiulas, L. (2011). Mailbook. A social network against spamming, (December), 11–14.

Zou, H. (2016). Protection of personal information security in the age of big data, 5–8. https://doi.org/10.1109/CIS.2016.141