PASSWORD STRENGTH AND MEMORABILITY

Bachelor`s thesis in Informatics (15 credits)

Hanna Julkunen
Josefin Ceder Molander

Fall 2016: 2016KANI16

UNIVERSITY
OF BORÅS

Title: Password strength and memorability

Year: 2016

Author/s: Hanna Julkunen & Josefin Ceder Molander

Supervisor: Peter Rittgen

Abstract

The society today is dependent on information technology and with the help of the technology makes it easier to access information. Due to the constantly growing network environment, various techniques of accessing and handling information have developed. One of the most used solutions to access and protect information is by using a password. The purpose of a password is to protect sensitive and important data from unauthorized users who intentionally or accidentally access the system. This can lead to unsolicited modifications of the original data as well as unauthorized access of confidential information. Humans are those who design the information security, but at the same time the ones who are the weakest link in the security chain. To prevent unauthorized access it is important to have a strong and tamper proof password. A good password should be easy to remember, hard to guess by others and be difficult to predict by a person or software. The goal in this study is to find a good balance between a memorable and a secured password. The study will compare three types of constructions for password, own set, modified dictionary and association against each other to find the one which is the strongest and the most memorable.

Keywords: (Password strength, Memorability, Password meters, Levenshtein distance, Experiment)

Acknowledgements

First of all we want to thank our supervisor Peter Rittgen, who has guided and supported us through the whole writing process from the start to the end. He has given us advice and kept us on the right path. We would also like to thank Stravroula Wallström who has helped us with the analytic part of this thesis when using SPSS. She helped us to find the relations and reading the statistics correctly in the quantitative analysis. Lastly we want to thank all of our participants, who gave us their time, not once but twice. They gave us the opportunity and the ability to complete this experimental study.

Table of Contents

1 INTRODUCTION ..... - 1 -

1.1 RELATED RESEARCH ..... - 2 -
1.2 PROBLEM DISCUSSION ..... - 2 -
1.3 PURPOSE ..... - 3 -
1.4 RESEARCH QUESTION ..... - 3 -
1.5 TARGET GROUP ..... - 4 -
1.6 DELIMITATIONS ..... - 4 -

2 THEORETICAL FRAMEWORK ..... - 5 -

2.1 INFORMATION SECURITY ..... - 5 -
2.2 THREATS AGAINST INFORMATION SYSTEMS ..... - 6 -
2.3 DICTIONARY HACKING ..... - 7 -
2.4 PHISHING ..... - 7 -
2.5 PASSWORD MANAGEMENT ..... - 8 -
2.6 ASSOCIATION PASSWORDS ..... - 8 -
2.7 PASSWORD METERS ..... - 8 -
2.8 JOHN THE RIPPER ..... - 9 -
2.9 ENCRYPTION ..... - 9 -
2.10 PHYSICAL HARDWARE CONTROLS ..... - 9 -
2.11 PASSWORD STORAGE ..... - 10 -
2.12 STRING-EDIT DISTANCE ..... - 10 -
2.13 HUMAN BEHAVIOR ..... - 11 -
2.14 COGNITIVE PSYCHOLOGY ..... - 11 -
2.15 MEMORABILITY ..... - 13 -
2.16 THE APPROACH TO ADDRESSING SOME OF THE LIMITATIONS IN THE THESIS ..... - 13 -

3 RESEARCH METHOD ..... - 14 -

3.1 RESEARCH APPROACH ..... - 14 -
3.2 RESEARCH DESIGN ..... - 14 -
3.3 DATA COLLECTION ..... - 15 -
3.3.1 Literature review ..... - 15 -
3.3.2 Collection of empirical data ..... - 16 -
3.4 SAMPLING METHOD ..... - 17 -
3.5 DATA ANALYSIS ..... - 19 -

4 RESULT AND ANALYSIS ..... - 21 -

4.1 THE EXPERIMENT ..... - 22 -
4.1.1 Password construction 1, Own set password ..... - 22 -
4.1.2 Password construction 2, Modified password ..... - 23 -
4.1.3 Password construction 3, Association password ..... - 24 -
4.1.4 Password construction comparison ..... - 25 -
4.2 THE SURVEY ..... - 27 -
4.2.1 Survey 1 ..... - 28 -
4.2.2 Survey 2 ..... - 31 -

5 DISCUSSION ..... - 34 -

5.1 METHOD DISCUSSION ..... - 34 -
5.2 RESULT AND ANALYTICAL DISCUSSION ..... - 34 -
5.3 CONCLUSION ..... - 36 -
5.4 CONTRIBUTION ..... - 36 -
5.5 FUTURE RESEARCH ..... - 37 -

REFERENCES I

APPENDIX ..... **IV**
APPENDIX 1 IV
APPENDIX 2 VII
APPENDIX 3 IX
APPENDIX 4 X

Figure of contents

FIGURE 1 - THE CIA TRIAD ..... - 5 -
FIGURE 2 - LEVENSHTEIN DISTANCE MATRIX ..... - 11 -
FIGURE 3 - STAGES OF MEMORY ..... - 12 -
FIGURE 4 - MULTI-STORE MODEL OF MEMORY ..... - 12 -
FIGURE 5 - SWEDEN'S POPULATION PYRAMID ..... - 18 -
FIGURE 6 - PASSWORD STRENGTH AND MEMORABILITY ..... - 26 -
FIGURE 7 - SCATTER PLOT ON PASSWORD STRENGTH AND MEMORABILITY MEAN ..... - 27 -
FIGURE 8 - HOW MANY PASSWORDS DO YOU USE IN YOUR EVERYDAY LIFE? ..... - 28 -
FIGURE 9 - DO YOU USE THE SAME PASSWORD ON SEVERAL ACCOUNTS? ..... - 28 -
FIGURE 10 - HOW DO YOU STORE YOUR PASSWORD? ..... - 29 -
FIGURE 11 - DOES ANYONE KNOW ONE OR MORE OF YOUR PASSWORDS? ..... - 30 -
FIGURE 12 - HOW OFTEN DO YOU CHANGE YOUR PASSWORD? ..... - 30 -
FIGURE 13 - WHAT DOES YOUR PASSWORD CONSISTS OF? ..... - 31 -
FIGURE 14 - HOW EASY WAS THE PASSWORD TO REMEMBER? ..... - 32 -
FIGURE 15 - HOW EASY DO YOU THINK THE PASSWORD WAS TO TYPE IN ON A COMPUTER? ..... - 32 -
FIGURE 16 - DO YOU THINK YOU WILL USE PASSWORD CONSTRUCTION 2 AND/OR 3 IN THE FUTURE? ..... - 33 -

Table of contents

TABLE 1 - CROSS TABULATION PASSWORD CONSTRUCTION 1 ..... - 23 -
TABLE 2 - CHI-SQUARE TEST PASSWORD CONSTRUCTION 1 ..... - 23 -
TABLE 3 - CROSS TABULATION PASSWORD CONSTRUCTION 2 ..... - 24 -
TABLE 4 - CHI-SQUARE TEST PASSWORD CONSTRUCTION 2 ..... - 24 -
TABLE 5- CROSS TABULATION PASSWORD CONSTRUCTION 3 ..... - 25 -
TABLE 6 - CHI-SQUARE TEST PASSWORD CONSTRUCTION 3 ..... - 25 -
TABLE 7- PASSWORD STRENGTH AND MEMORABILITY MEAN ..... - 26 -

1 Introduction

It is important to keep personal and organizational information safe and protected. Passwords have become the most used authentication to protect resources from unauthorized access (Merkow & Breithaupt 2014; Vu et.al 2007).

There are three different types of threats (Kritzinger & Solms 2010; Mitrovic´ 2005), physical, logical, human and organizational threats. The physical threats include theft, vandalism and natural disasters which affect the physical equipment. The logical threats are driven by unauthorized users, manipulation or obstruction of service which focus is on the software. The human and organizational threats could be: poorly structured work routine, or no routine at all, unstructured division of labor and lack of insufficient knowledge, which can damage the data.

Even though there are many threats against information systems, there are controls which protect the confidentiality, integrity, availability, authenticity and nonrepudiation (Merkow & Breithaupt 2014; Dhillon 2007). Three types of defense controls are encryption, software controls and physical hardware controls. The encryption, scramble the data into unreadable text, which needs an encryption key to unlock and read the data. Software controls are crucial for companies, where they observe and protect their software. These controls are divided into three groups: The first one is a software development control, which is a process of computer testing, programming and maintenance. The second are the software, which operates an inbuilt limitation in the system, for example a protection for every user. The third, defense control is physical hardware controls and these controls include locked doors or guards at the front door (Dhillon 2007).

A personalized password is a simple way for the user to create, implement and remember. A disadvantage with these kinds of passwords: they make an easy target for crackers (Mitovic´ 2005). If an unauthorized person has cracked a password and successfully entered a system, login, account etc. it is not unusual that the cracker can try to use the same password elsewhere. When reusing the same password, for instance on a more secure system, file or account can make it vulnerable for the user's sensitive data (Ives, Walsh & Schneider 2004).

There are measurements for helping and encourage users to create and strengthen a better password, by using a password meter. Password meters are indicators that measures the password the user types in, and gives suggestions and guidance in the creation of the password. It shows the given password how weak or strong it is, also the mixture of characters, numbers, signs, symbols etc. (Ur et. al. 2012).

Furthermore, there are various types of techniques and studies of how to make a better and stronger password. It has been positively successful in theory and in practice, for example exchange letters to numbers: Freedom becomes Fr33dom (Yan, Blackwell, Anderson & Grant 2004).

The human memory is quite limited and do often remember sequences that are familiar, or if it has been repeatedly memorized over and over again. When people do not remember a certain password, it is common that they write it down physically on a piece of paper or have it stored in their phones (Sasse, Brostoff & Weirich 2001). These options are not a secured viable solution. There are some mnemonic methods that the password user can adapt, to help them

memorize random sequences or in this case, a strong memorable password (Yan, Blackwell, Anderson & Grant 2004).

1.1 Related research

Users often create passwords that are easy to remember which also makes them easy to crack. Lyastani, Acar, Fahl and Backes (2016) describes that the use of mangling rules are a great technique to create stronger passwords based on a simple dictionary word using John the Ripper which is a software for password cracking.

The memorability is usually tested over time for testing the long term memory. In Lyastani, Acar, Fahl and Backes (2016) article, they tested their participant's memorability after one week. But in their study they used repetition to enhance the memorability, where the participants had to write down the passwords a few times before handling it in. Roediger and Karpicke (2006) did also use the time limit of one week in their research when they tested the participants' memory. Their opinion was that one week was enough time to test the long term memory. In our study we have tested the long term memory as well after one week. But we have used the measurement of the Levenshtein distance (which measures the distance between two strings) when testing our participants' memory.

There are several ways to test the password's strength. In the study of Vu et.al (2007), they tried to use the password strength to crack them. They examined how big of difference strength of the passwords was. To verify the password's strength, they counted the timespan it took to crack the different passwords that were used in their study.

Another way to measure the strength of the passwords is by using a password meter. A password meter is coded to calculate different types of criteria a strong password should contain. Ur et. al. (2012) made a study analyzing different password meters' strengths and weaknesses. The password meters are designed to tell the strength of the password by the measurement of very weak to very strong, bad to excellent or by crack-time. In this report the measurement for testing the password strength will be by using a password meter, with the inspiration from the study of Ur et. al. (2012).

1.2 Problem discussion

It is known that the human is the weakest link in information security and the problem of weak passwords lies in the basis of the user's knowledge, motivation and effort (Sasse, Brostoff & Weirich 2001). To strengthen and change the information security behavior, we need to find new mechanisms which can help the users to select a better password. The user must choose mechanisms which are both hard to guess by others, difficult to crack by cracking tools and still be memorable (Forget, Chiasson, Oorschot & Biddle 2008).

Arvidsson, Eriksson and Nerén (2013) discuss that many users have a tendency to care about security regarding passwords but do not understand if they are weak or strong. If the users are handed a more complex and secure password, the consequence will be that it will make it harder for them to remember, rather than if they have created it by themselves. Sasse, Brostoff and Weirich (2001) do also bring up the problem with password memorability and that a complex password is easily forgotten.

Therefore more research is needed, regarding the user's password memorability, their habits and behavior regarding what different password creation should be conducted. This is to get a broader and better understanding about the subject and to find a reachable balance between a strong yet memorable password for the user.

1.3 Purpose

The purpose of this thesis is based on the problem discussion above. The purpose is to find a balance between password strength and memorability. Hopefully to counterbalance with a complex secure password and be relatively easy for the user to remember for a long period of time. When creating a strong password it will not only keep the intruders away, it will also increase the safety of people's data integrity (people's integrity alone implies personal integrity) and their personal information.

1.4 Research question

Based on the problem discussion and purpose described in this research, there is a strong interest to investigate and examine both the memorability and strength of the user's chosen password. A good password should be easy to memorize, be hard to crack and be suitable for the password user (Forget, Chiasson, Oorschot & Biddle 2008). While searching through facts and theories there were a lot of diverse procedures of password creation that were found interesting for this experimental study. After choosing from several different password technics for a stronger password it came to a conclusion to narrowing them down to three passwords construction types: modified password, associations password and own set password. Therefore, the research question in this thesis is:

How can the best trade-off be created between password strength and memorability when a group of individuals construct a password?
The definition "trade-off" in this research question, is to measure the combination of password strength and memorability through different password constructions.

In the research question, the independent variable is the individual. The password constructions criterion depends on what type of password that has been created by the individual. This makes the *password* dependent to the *individual*. The *password strength* and the *memorability* are dependent on the *password* which is created and also dependent on each other.

The appearance of the password and its content is vital for the strength of the password. The content of the password do also affect the memorability. If the password structure is too complicated and illogical, the harder it gets for the memory to remember. The password strength and the memorability are on the same level, and they are both dependent on how the individuals construct a password.

1.5 Target group

The aim of this study is to be beneficial for the everyday user, both in private and professional. It is mainly for those who lack knowledge in information security to help them create a stronger password that they can remember. In this report it is important to help the user create a password that they can remember for their own safety (Mitrovic´ 2005).

1.6 Delimitations

The research was carried out at the University of Borås and nearby areas in the region of Västra Götaland of Sweden. The delimitation in this report focused on the participants where they have been a mix of adults from various generations and from different places. The researchers chose participants who had some kind of experience in using different systems, which requires a login using passwords.

Due to different kind of password types, it was necessary to set boundaries. This report was only focused on passwords for accessing email, social media, website accounts etc. Passwords containing only numbers, such as Bank ID and pin codes will not be addressed because we are not presenting passwords, only constructions.

2 Theoretical framework

2.1 Information security

The definition of information in the context of information security is based on knowledge, which can make it an asset for both individuals and organizations. The information is varied and can be used in different ways, such as for communication; it can also be stored and be processed.

The majority of our information can be extremely important for the general population and in some occasions vital, according to Informationssakerhet (2015). This type of information needs to be secured, which is why information security was formed (Merkow & Breithaupt 2014).

Information security means the protection information and keeping it safe from unauthorized people. There are three types of goals of protecting an information system, and the information security is trying to achieve at least one of those goals. Confidentiality, integrity and availability are the three goals. These three creates the CIA triad, seen in Figure 1.
The CIA triad is often used for illustrating the standard baseline of an information security plan (Merkow & Breithaupt 2014).

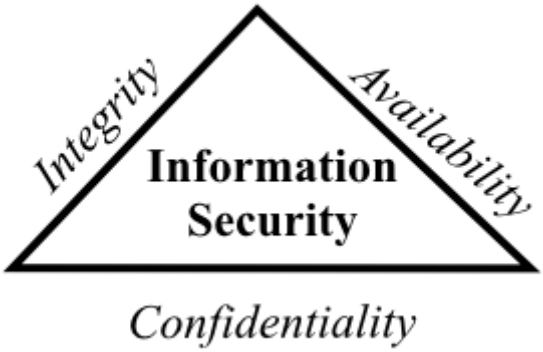

Figure 1 - The CIA triad

- *Confidentiality:* The information should be inaccessible to unauthorized people, a prevention of information leaking out.

- *Integrity:* The information should not be modified in an unapproved way. It is about preventing unauthorized alteration or modification of the information.

- *Availability:* The accessibility to the information, correct information should only be available to the right person or persons.

These three principles are linked together and are dependent on each other, for data not to be intentional or accidental misused (Merkow & Breithaupt 2014).

Authentication is so only the legitimate user can access the hidden data (Merkow & Breithaupt 2014). The most commonly used authentication tool is password for entering an information system. Passwords are commonly insecure in the way that the constructions of them are too weak, and they are easily broken with the right tools. Passwords are an example of a single authentication but there are also multifactor authentication as two factor authentication and three factor authentications. Multifactor authentication gives the information owner an ease if there are several steps for entering the information in process to keep unauthorized users off. Examples on a two factor authentications are when the user has a physical device, a key-card for example, connected to a pin-code. So if an unauthorized person would get access to the physical device they would not be able to access the information without the pin-code. Three factor authentications are when another factor is added to increase security. This could, for example be adding fingerprint or a card access on the door to the room where the information is stored electronically (ibid).

2.2 Threats against information systems

Today people have become very dependent of the internet, both in organizations and in their private life. Not only has the internet given us freedom of widening our knowledge, it has also created a territory of risks where unauthorized people have accessed information for their own gain. The intrusions and threats against information systems, personal information, company information, have increased due to the growth of the internet since it has become more common and public. But this has also led to gaining more knowledge and awareness of preventing these kinds of situations (Kritzinger & Solms 2010).

Threats and intrusions can occur differently, according to Kritzinger and Solms (2010) there are two types of threats: physical and nonphysical threats. Both could be caused accidentally, intentionally or by a natural disaster. The physical threats can be someone or something damaging the computer system, the information system or/and the hardware physically, for example through vandalism, theft etc. The nonphysical threats, also called logical threats include threats and intrusions targeting the software, for instance a virus infecting the system, or an unauthorized person breaking in and stealing important and sensitive information in a database.

If an intrusion would succeed into a system, the purpose of the intrusion determines the consequences. According to Jouini, Rabai and Aissa (2014) some consequences could be;

- Destruction or corruption of information. The intruder steels information or manipulate system operations.
- Disclosure of information. The cracker publishes information which may be sensitive.
- Password change and denial of network. The password and log-ins might be changed or the network is blocked so the user cannot access the system.

The goal of the information security is to reduce the threats as much as possible. Since there are so many different kinds of threats against the information system it is important to analyze where in the system it is vulnerable, but also the risks and how prepared an organization or a person is willing to lose (Nationalencyklopedin).

2.3 Dictionary hacking

A dictionary attack utilizes words in a pre-arranged list, similar to a dictionary. It can contain words, names or phrases and often succeed to crack passwords because users tend to choose passwords that are phrases or words. Dictionary attacks can be prevented by using passwords that for example are not a word found in a dictionary, a name or in password from a list of common passwords (Wikipedia).

Narayanan and Shmatikov (2005) conclude in their article that as long as a password can be stored in the human memory it makes it vulnerable of smart-dictionary attacks. Defense against online and offline dictionary attacks with human-memorable passwords are difficult if the attacker has an advanced dictionary-list. For a password to be safe from a smart-dictionary attack Narayanan and Shmatikov (2005) suggest that face or geometric images can be used as an alternative instead of passwords.

2.4 Phishing

Phishing is when the hacker tries to collect personal or financial information by sending out an email from what looks like a reliable source. The email looks authentic as it if is sent from the user's bank or a credit card provider, where they want the user to update their personal information, such as their login account and password information, according to Britannica Academic (2016).

The amount of fraud using phishing has expanded over the last years combined with the increased usage of internet. More people are using the internet to bank errands, for example to pay their bills. The customer's independent has contributed the banks to offer free online-service. Today internet frauds are massively expanding and instead of robbing a bank criminals target their customers. Even though it is the private customers which get affected, it indirectly affects the banks that cannot protect their customer's assets (Lance 2005).

According to Lance (2005) there are three methods in phishing which are more common. The first and most common one is *impression*, which is when fake websites are created. These fake websites often has the same pictures and layout as a real website to attract customers and to look reliable.

The second phishing method is *forwarding*. This is when the phisher send out an email where account information to a specific account is required to be filled in by a link. The email graphics is supposed to be similar to the real website, commonly Amazon, e-Bay and PayPal is used. If the user fills in their information they are then transferred to the real website but the information they filled in about their account. The victims rarely notice that they have been exposed for a phishing attack before it is too late. Nonetheless there is one weakness with this phishing method which is the HTML which might block these emails as spam (Lance 2005).

The third phishing method according to Lance (2005) is *popup attack*. The popup attack appears when clicking a website link in an email which result in a popup window occurring. Behind the actual popup there is a hostile target that is collecting data. Today these attacks are less common because many browsers have popup blocks installed by default.

2.5 Password management

Burnett and Kleiman (2006) have three theories on how to strengthen a password; rule of complexity, rule of uniqueness and rule of security. These three elements combined create a stronger and safer password.

A theory for keeping data safe can be explained by Burnett and Kleiman (2006) where they mention the rule of complexity. A complex password according to the writers should contain three elements, for example characters, letters, numbers, words or phrases. To increase the complexity numbers should not be in the beginning or the end of a word and the password should contain less than 50% of numbers.

Burnett and Kleiman (2006) second theory is the rule of uniqueness where a password should be unique for every system where it is used and distinct among passwords. To create a unique password one should avoid common words, dictionary words, names or numbers related to you. To continuing keeping a password unique it should be changed between every three to six months and if there are any suspicious of a password has been cracked the password should be changed at once.

The third theory Burnett and Kleiman (2006) wrote about is the rule of security which is about how to keep a password safe. To keep the password and maintain data confidentiality the password should not be shared with anyone and the password should not be saved on web browsers. For further safety users should also always change password that is automatically assigned (ibid).

2.6 Association passwords

One type of password creation can be an associative password where the user chooses a single-word and then use it to associate it with when creating a password. An alternative technique can be a passphrase, to form a longer password by building a set of words from an associated sentence to create one whole password. For instance, “I love food” → “Ilovefood”. Another method can be a mnemonic-based password where the user takes inspiration from the passphrase but to only use the first letter of each word to create a random sequence of characters to design a password, for example: I love to ski at Seven Springs!” → “Ilts@7S!” according to Taneski, Brumen and Heričko (2014).

2.7 Password meters

Password meters are used to help the user develop safer passwords by giving different criteria for stronger passwords. The structures of password meters are different and the criteria depend on the programmer. The most common criterion used in a password meter is built on the word’s length, if it contains uppercase letters, lowercase letters, special characters and sometimes spaces. Some of them also contain dictionaries with common passwords which are considered less safe because they are easier to crack (Carnavalet & Mannan 2015).

There are many different kinds of password meters and 14 different meters have been investigated (Ur et. al 2012). They have concluded that different meters have different criteria for creating a stronger password, and it is not always reinforced if the password meter is bad. A bad password meter can sometimes only count the length of the password, for instance typing in the password "passwordpassword". This seems to be an "excellent" password by the meter because it has many characters, when in real life it is not.

Most of these meters can improve a stronger password when giving example suggestions to the user, helping them to mix capital letter or numbers in the password. Some of the meters also go through a database to see if the password is a common hacked word and if it is, the meter informs the user (Roediger, H.L. & Karpicke, J.D. 2006.) The password meter is a good aid or the users when creating a better password since they can clearly see by the scale if the password is good or bad, weak or strong.

2.8 John the Ripper

John the Ripper is an open source system which is designed to crack password encryption. It is commonly used by system administration to recover passwords in a computer system and to find weak passwords, according to Sykes, Lin and Skoczen (2010).

John the Ripper works almost the same as a dictionary attack. Companies have for example certain password criteria which must be followed. If a password is forgotten and the employee remembers part of it, can John the Ripper be of help. The function then takes into account the criteria to be included and the twists and turns the word until it is correct. Cracking a trivial password with the John the ripper method is possible, but to crack a more advanced and stronger password needs more resources (Sykes, Lin & Skoczen 2010).

2.9 Encryption

Encryption is when the data is transformed from clear text to encrypted text, this is to protect and prevent intrusions of sensitive data (Mitrovic´ 2005). When encrypting data, it is usually made by a program or one or several persons. They are also the ones who know how to decrypt the information. When creating encrypted information, there are keys made for decrypting the code or information. This key transforms the information back to clear text again. There are, as mentioned before, programs that can decrypt the data or information without these keys. The security of the data is only as good as the protection itself (Dhillon 2006).

2.10 Physical hardware controls

To prevent physical damage and stop unauthorized people from accessing hardware, physical locks can be used on doors and or guards can be placed in front of the building. This can minimize the risk of intrusions (Dhillon 2006). Physical computer inventory such as hard drives, fans and other components can also be damaged or exposed for theft if wrong people can access it. If someone manages to get physical access to the IT-resources the integrity, confidential information can be harmed according to Mitrovic´ (2005).

Physical damage can also be exposed by the climate and environment with for example fires, flooding or earthquakes that may harm the computers (Mitrovic´ 2005).

2.11 Password storage

The best and the most secure way to store your password is in your own memory. But, the stronger and more complicated the passwords are, the more you forget. Therefore, creating an extremely strong password cannot be successful if the memory fails. Consequently, will the password users choose an easy password instead, and if the password does not have any personal association it is more likely that they will write them down physically. This will instantly lower the security, when it becomes visible to others (Vu et. al 2007).

There are other technical solutions for helping the users to storage and remember their passwords, as an alternative for remembering several password. As mentioned before, it is not recommended to reuse your passwords, but to remember a lot of passwords are not manageable by everyone. An article made by Gehringer (2002) describes an alternative solution, software called “Wallet”. This system stores all of your passwords in one place in a secure way, and you only need to remember one password for accessing the software system. In today's technology other options have developed, where you can save your passwords and personal information. There are several companies that offer these kinds of services online, through their websites, for example LastPass.

Nevertheless, the memory is a big issue in human behavior and a study that Sasse, Brostoff and Weirich (2001) made showed, as previously mentioned, that people remember passwords easier when they can associate them with something. A suggestion from the authors when creating a secure password, which is both strong and memorable, is to exchange some letters with numbers and symbols in the password-word, for example: BF$gat0 for Black Forest Gateau. This is a good and simple technique for making it easier for the users to remember, yet it is still secure. According to Burnett (2006) it is important for the users to be able to memorize their password, not learn to memorize it.

2.12 String-edit distance

String-edit distance measures the distance between two text strings. It calculates the minimum amount of changes between two words to transform one string to the other by using insertions, deletions and substitutions (Ristad & Yianilos 1998). The Levenshtein distance is one type of string- edit distance and has been used for example with checking spelling, language and comparing DNA (Likhitha, Ninitha & Kanchana 2016).

If two words are identical to each other the return of Levenshtein distance will be 0 and if there are a difference between the strings for every insertion, deletion and substitutions the distance increases with +1 (Schepens, Dijkstra & Grootjen 2012).

An example from Navarro (1999) in Figure 2 shows how the Levenshtein distance works as a dynamic programming algorithm. Two words are tested against each other, surgery and survey, and the matrix helps to see how big the distance are between them. In this case there are distances of 2 between the two words.

		s	u	r	g	e	r	y
	0	1	2	3	4	5	6	7
s	1	**0**	1	2	3	4	5	6
u	2	1	**0**	1	2	3	4	5
r	3	2	1	**0**	1	2	3	4
v	4	3	2	1	**1**	2	3	4
e	5	4	3	2	2	**1**	**2**	3
y	6	5	4	3	3	2	2	**2**

Figure 2 - Levenshtein distance matrix

If for example the words Books and Backs are to be calculated against each other this formula will appear:

1-2/5= 0, 6

The number 2 symbols that there are a difference of two between the words, which means two letters are not equal. Number 4 symbols the length of the password and number 1 is to get the similarity in percent. So in this case there are 60 % similarities between the words Books and Backs.

2.13 Human behavior

There are many different factors and theories that can explain the human or in this case the user's behavior when it comes to creating and managing their own passwords, such as the human memory, lack of motivation, knowledge, training etc. Humans play the biggest part in many security failures. The users are the weakest link in the "security chain". The statement of blaming the users' behavior will not lead to a more secure system, therefore it is necessary in the human-computer interaction to isolate and identify the causes of the users' behavior before designing a security system (Sasse, Brostoff & Weirich 2001).

2.14 Cognitive psychology

A short definition of cognitive psychology is that it is focusing on our mental processes; the way humans process information. The mental process includes memory, perception and emotions, and the cognitive perspective gives also an understanding about the human behavior (Study).

Simply psychology describes; being able to remember and maintaining information on past experiences over time are based on that we have a functioning memory. The memory can be divided in three different parts:

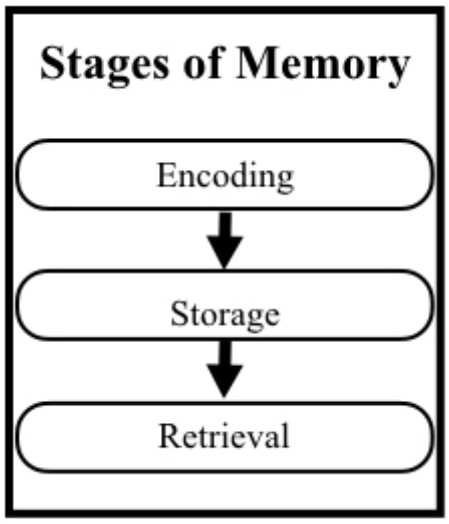

Figure 3 - Stages of memory

- *Encoding* (sensory input), what we see, hear and understand.
- *Storage*, how long and how much we can remember.
- *Retrieval* (output), what we actually can remember

For a better understating of the human memory can be seen in Figure 4. The term of information processing through a system, known as the "Multi-store model of memory" by Shiffrin and Atkinson (1968).

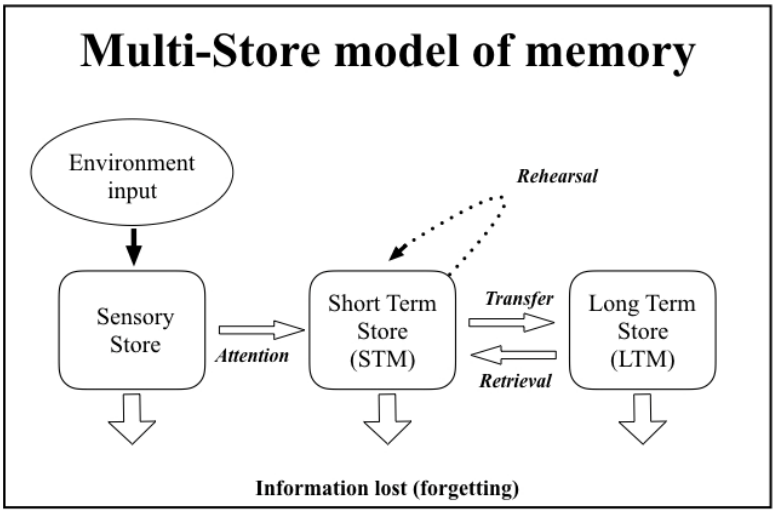

Figure 4 - Multi-store model of memory

The information enters the sensory memory store from the environment input, and then to the short term memory. It transfers from the short term memory to the long term memory, only if that information is rehearsed otherwise it gets lost.

The information operates through the "memory stores"; the sensory store, STM and LTM. These three stores have different capacity and attributes:

- *Sensory store:* The durations are about ¼ to ½ second and the capacity is all of the sensory experience.
- *Short Term Memory (STM):* The duration is about 0-18 seconds and the capacity is about 7 +/- 2 items.
- *Long Term Memory (LTM):* The duration and the capacity are unlimited.

2.15 Memorability

When it comes to the research of the password users, it is important to find a good balance between a strong and memorable password in the long run. Therefore, the focus will be about the long term memory. Because the human memory is limited, and it is difficult to remember random sequences of symbols, characters, words etcetera, the techniques of how to memorize needs to be better (Simply psychology 2009). The human's memorability has a big impact on the creation and remembering a password. When defining the term memorability it is described as the quality of how easy it is to remember, or even the importance of remembering. Regarding passwords and the human memory it is known and proven that people have a hard time remembering strong and complex passwords. Therefore many users create simple and memorable passwords, which makes is also easy for others to guess and hack (Yan, Blackwell, Anderson & Grant 2004).

If a person's memory was endless or if we as humans did not have to remember passwords at all, there would be extremely secure and unhackable passwords. These passwords could contain a string with random selected characters and would be as long as the system would let it - the most theoretical and hardest password ever. But in the reality, this is not going to work (Yan, Blackwell, Anderson & Grant 2004). Blaming human behavior will not lead to a more secure system, therefore it is necessary in the human-computer interaction to identify the causes of why, before designing a security system (Sasse, Brostoff & Weirich 2001).

2.16 The approach to addressing some of the limitations in the thesis

The theoretical framework reflects various factors and situations when using a weak password, such as, what the consequences and the threats are. We have also included research of how to prevent weak passwords, for example Burnett and Kleiman (2006) three theories, about how to create strong passwords, with the help of complexity, uniqueness and security.
The methodological in this thesis is based on the theoretical framework, to potentially achieve a good reliable research study. When measuring our variables, the password strengths and the memorability, against each other, we used a password meter and the Levenshtein distance. Research regarding the human memory and the psychological aspect has been taken in consideration as well, where the focus has been on the long term memory, when designing the experiment.

3 Research method

3.1 Research approach

When approaching a quantitative research method, the research was founded on previous work from other scientists (Oates 2006). In this study, after gaining more knowledge and a well based theoretical background from other studies and experiments, a research approach was developed. A deductive work became an explicably approach for this thesis. We initially started to collect information about the subject of the report and then worked from there. A deductive research approach is commonly used in a quantitative research, where the technique is for testing an existing theory, according to Bryman and Bell (2015).

The quantitative method is to process, describe and analyze data, according to Bryman and Bell (2015). The quantitative data can transform into numbers or statistics, and by collecting numbers through the quantitative data as evidence. Furthermore, searching for patterns from the collected data can generate the result of this study and the opportunity to draw conclusions (Oates 2006).

3.2 Research design

To find the answer for this thesis, a quasi-experiment was used. When an experiment is using this kind of method, the chances are higher to achieve a strong and valid result (Bryman & Bell 2015). A quasi-experiment is according to Bryman and Bell (2015) similar to a classic experiment, but with a freer approach.

A Classic experiment design includes three parts:

1. **Observation** to the dependent variable which may be two or more times. First there are a pre -observation before the experiment and then, there are the observations.
2. **Experimental treatment** which is manipulation of the independent variable.
3. **Timing** is when to do the experiment, in relation to the dependent variable.

The difference between a classic experiment and a quasi-experiment is that the quasi-experiment does not need any control groups or manipulation (Bryman & Bell 2015).

The quasi-experiment design in this research had a time series design, where it involved one experimental group, consisting participants. The group was observed and tested, both before and after the experiment. A time series design is the most common and simplest design used in a quasi-experiment (Robson 2011), and was the experimental design chosen for this research. This study did only need one experimental group, where the participants created all three password constructions at once, and not in separate groups for every password. This gave the opportunity to work more efficient, and gather as many participants as possible.

Within this experiment, there was some kind of manipulation, where the participants had to follow the chosen password constructions and the criteria when performing the experiment. Why this experiment did not use control groups, was because we wanted all of the participants to be part of the same experiment.

What control groups does, are that they have two equally big groups. One of these groups use manipulation, where the participants has the same attributes, in form of gender, age, previous experiences and so on. The other group has no manipulation connected to the independent variable. In an experimental study, the independent variable is the one usually tested, to see if it's affecting the dependent variable (Oates 2006). The outcome of the control groups are measured afterwards, where the results between the two groups analyses, to see if the right participants and the right criteria has been correctly chosen (Oates 2006). Using manipulation on the independent variable is for detecting any differences between the variables, and also see if the independent variable has some influence on the dependent one (Bryman & Bell 2015).

The independent variable in the research question is the *password construction*. Followed by the dependent variable, the *password* created from it. The password has in its turn two dependent variables, which we are focusing on in this study, the *password strength* and the *memorability*. To clarify the variables relations to each other; the password construction is independent against the created password; the created password is dependent on the password construction. Furthermore, the password strength and memorability are dependent on the created password.

Another method, such as observational notes, is common during an experiment (Robson 2011), was not used in this experimental study. Instead of taking notes, the problems that occurred during the experiment were solved directly or right after the experiment.

3.3 Data collection

We initially conducted a literature review to get a deeper understanding of the subject. The early study has also been the basis to form our experiment and survey. The data collection has also been an important part during the research process.

3.3.1 Literature review

Before beginning the research, a study among scientific articles and theses was made to gain inspiration to a research subject. We mainly used electronic information and especially Summon, which is a web based search engine that contains scientific literature at the University of Borås. Physical literature has also been applied in the research process throughout the whole report, especially in the methodology and the theoretical chapter.

A critical point of view was approached when finding sources, the same strategy was applied when searching for literature and information. With a critical point of view, we as researchers had to critically analyze and evaluate other studies in order to achieve a quality research.

In the third section in this report, theoretical framework, we summarized the important theories from different sources that could strengthen this report. Wikipedia is used in the theory, not as a scientific source, but as a definition of terms.

3.3.2 Collection of empirical data

This thesis conducts two empirical studies, experimental and self-completed questionnaire. The surveys were a requirement for the experiment to be successful, because it gave more useful information to our study.

Before performing the experiment we invited six test participants to do a test experiment. By reason of to see what outcome we could get, and to see what the possibilities and disadvantages the test experiment could give. After the test we recognized some weaknesses;

- Parts of the experiments description were not clear enough, with lacking of information, and had to be improved.
- New questions were added when we discovered that the information we got out of the participants was not sufficient.
- The first survey added a possible response; "When the system asks me to change" to the question "How often do you change your password?"

The real experiment compared three different password techniques to see which one of our three password constructions would be the strongest and the most memorable.
The three password constructions:

1. Own set password, chosen by the participant
2. Modified password
3. Association password

To begin with, the participants created a password on their own requirement. This was, for the researchers to get an understanding of the participants' everyday password creation. The other two password constructions were a guidance to help the participants to create passwords on our criteria, with the aim of creating a stronger and a more secure password.

The second password construction, modified password, was inspired by Sasse, Brostoff and Weirich (2001). The authors described and suggested a method when replacing letters into numbers and/or special characters to build a stronger password in a simple way, a technique which makes it harder to be hacked by a dictionary attack.

The third password construction was based on associations, where the user creates a meaningful sentence with a date (numbers) included. From that specific sentence a password would be designed. The users had to use the first letter in every word in that sentence, also the numbers or dates to create a password. This results a password containing random letters with numbers (Yan, Blackwell, Anderson & Grant 2004).

After the experiment a comparison was made, to compare the three passwords that the participants created. With the help of a password meter we could measure and compare the strength of the passwords. A week later we had a follow-up meeting to test the participant's memorability, which was measured with Levenshtein distance. The participants were not allowed to store their passwords in any other form than in their memory, which was clarified in the criteria in the experiment.

According to Bryman and Bell (2015) a self-completed questionnaire is most used through email or a postal questionnaire to receive responses. Although, an experiment was performed in this thesis, the decision was to only collect answers and information from the surveys face-to-face, combined with the experiment.

For further data collection two surveys were made. The first survey was performed right after the first part of the experiment. This survey assembled information about the participant's password habits and attitudes. The last survey performed after experiment two, the follow-up. The motive of the last survey was to gain more knowledge about the participant's experience of the whole experiment, and discovering if they were using one or all of the password constructions. The questions for our surveys were inspired by the questionnaires that Karlsson and Stålnacke (2002) and Ahlberg, Blomberg, Davidsson (2008) used in their studies.

The questions in the survey only conducted closed questions, except name, email, age and comments in the experiment. The first part of the questionnaire contained check- or multiple answer questions, which was created in a vertical format. The second survey had a horizontal format and used the Likert scale with seven levels. The Likert scale is preferred when the questions are a statement and when questioning about the level of agreement or disagreement (Bryman & Bell 2015).

3.4 Sampling method

There are two types of samples according to Bryman and Bell (2015), which are probability sample or non-probability sample. Probability sampling is when the sample of the population has been chosen randomly, and every unit of the population has a known chance to be part of the sample. The probability sampling is most likely to be used when the sampling involves employees in a company. It is also possible to make conclusions from the random sample to the population that was selected.

The sampling method chosen in this report is a non-probability sample. A non-probability sample is when the sample has not been assigned randomly. Instead, some people of the population are more likely to be specific selected, than others (Bryman & Bell 2015).

According to Oates (2006), the non-probability sample is used when the researchers believes that it is not necessary to have a representative answer, whether the time and cost to get a valid sample is too excessive. To get a valid sample, Oates (2006) describes that you need at least 30 respondents to achieve a statistical analysis; otherwise it is not a reliable result. Non-probability sample was chosen for this study in ability to limit the research. The resources to have a completely random sample could not be fulfilled, due to the scope out of this research.

Bryman and Bell (2015) also covers the two main types of non-probability sample, convenience sample and quota sample. The convenience sample is when the researcher uses simply available sample people via accessibility. A Quota sample's aim is to produce a sample that reflects a population in terms of categories as gender, age, education etc. We used both of these methods. The convenience samplings were used when we chose who to do the experiment on. We chose students and other companies where we had connections, or were geographically located close to us. We used the quota sample to try to cover variety categories of people, as equally of men and women, different age groups and different education degree of the participants. The

variation of the participants is important for getting the sampling to be both reliable and valid, which are an important part of research according to Bryman and Bell (2015).

In the experiment, 41 people between the ages 18-66 participated. The study took place both at the upper secondary and postsecondary schools, to reach out to younger adults. For variety the ages, the study performed on various workplaces, for example at Drivhuset and Hällingsjöhus to reach the older generation. The experiment has been carried out in different segments to get a wider variation of the population.

Statistics Sweden (2015) has made a population pyramid, see Figure 5, of the Swedish population, where all of the population is divided into age groups. Several years ago, the population was more shaped as a pyramid, but today with increased healthcare and less children born per woman, it more shaped as a tower. There are some age groups that are larger. The reason why, was because there were more children born in Sweden in that period of time. The ages in 2015 were between, 20-24, 45-49 and 65-69.

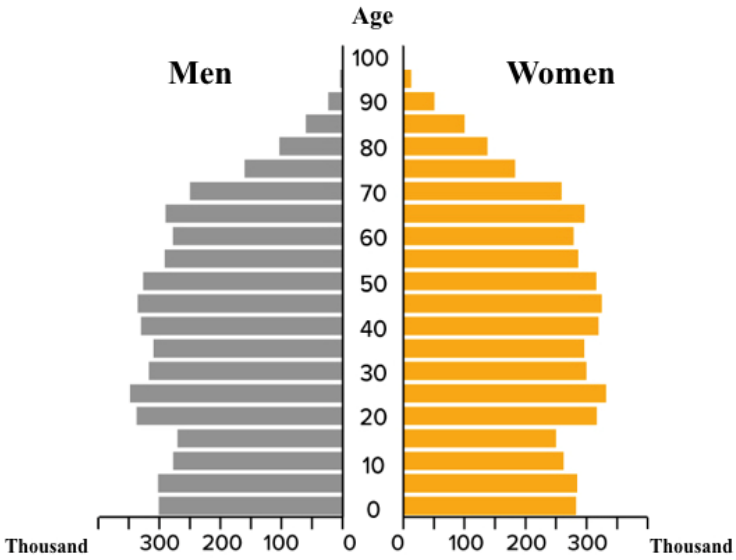

Figure 5 - Sweden's population pyramid

We have tried to get participants from various ages to participate in the experiment, and to get close to an equal number of men and women to participate.

The participants in the experiment were anonymous. The reason why we collected their name and email were to connect the first part of the experiment together with the second one. Their personal information, will not exploit in the report in any way, in protection of their integrity. It the experiment, it was very clearly described and well informed that the participants should not write down passwords they use in their life.

3.5 Data analysis

The experiment analyses were conducted in three steps. First we tested the passwords' strength, afterwards the memorability and finally we tested the passwords' strengths trade-off with the memorability.

When the passwords were tested, we used a password meter on My1login to measure the strength. Before choosing this specific password meter, we tested five different types, in order to narrow them down to only one. We tested them by using the same passwords to see what kind of criteria that were fulfilled. My1login was the meter with the most mixed criterions. It was the one meter which met with our theories that we have found, when measuring the strength of a password.

Many of the tested meters were just counting characters, and some of them did not increase the strength on the password when special characters were entered. My1Login criteria used numbers, special characters, lowercase letters and uppercase letters. The website, My1Login, counted the characters that were entered to evaluate the safety of the password. It could also read if it was a commonly used dictionary word, which could lower the password's safety. One of Burnett and Kleiman (2006) criteria to create a safe password is that it should contain 14 characters, be absent of personal information and have a good mixture of letters and numbers. Based on these criteria, this password meter was the closest one.

The password meter, My1Login also showed how safe a password was. It shows and calculates the estimated cracking time on the entered password. The cracking time went from milliseconds to billions of years. We began to compile all the different time units to one unit - hours, from the created passwords. The reason of converting to hours was for easily handle the data in SPSS, and furthermore cluster the data - *hours* into groups.

When adding all the data together to calculate the mean value of password strength, we noticed that the numbers were too big to handle. We had to divide the extremely large data - *hours,* for every created password construction, with 10 000 000, this to achieve a more manageable number for analyzing the mean value between password strength and memorability. With lower numbers for password strength, a comparison against the memory was easier because the variable numbers were closer.

To test the memorability of the passwords, we used Levenshtein distance, a type of string-edit distance. The Levenshtein distance made it possible for us to compare two strings- *the passwords,* and find the difference between them (Ristad & Yianilos 1998). This tool helped us to compare the passwords, of how well the participants remembered the created password one week later during the second experiment. If the participants did not remember 100% of the passwords, the Levenshtein distance helped us to calculate the similarity of the passwords, in percentage.

After both parts of the experiment were performed, the collected information was inserted into SPSS to find a result. The first step for analyzing the information and result was to manage and handle the raw data, and reducing and summarizing the information to make it a useful data (Bryman & Bell 2015).

With the help of SPSS, the three different password constructions were tested against the memorability, to see what the significance could be, and to see if there were any relations with another (Bryman & Bell 2015). The reason why 30 participants are a minimum, according to Bryman and Bell (2015), is because less is considered not statistically measurable. A cross tabulation needs an expected count more than 5 in every square. In this evaluation in SPSS, there were four squares in every cross tabulation. The 41 participants in this research satisfy the count more than 5.
Pearson´s r is a method used to measure the strength of the relationship between the variables; the strength of the password and the memorability (Bryman & Bell 2015).

The experiment had some invalid information, for instance when the criteria for creating a password by our participants were not fulfilled. SPSS contributed to distinguish inaccurate and absent information with features in the program.

Furthermore, in a quantitative analysis are diagrams often used. The diagrams show and describe a clearer overview of a result. Therefore, we found bar charts and pie charts the most suitable in this research, though the majority of our data has ordinal variables, for example our age groups (Bryman & Bell 2015).

4 Result and Analysis

For achieving the best potential experiment as possible in a limited time frame, a test experiment was performed. The test experiment was exactly designed as the real experiment, but with a smaller group consisting of six participants which was held in the University of Borås.

The test experiment was a useful preparation for the actual experiment and received positive feedback. However, the experiment needed to be improved and there were some weak points and confusion in the description. Some of the altered changes were for instance a better explanation in an easier terminology when describing the test. This became helpful in the real experiment though some of the participants in the older age range had difficulties to understand when it was not in their native language. Other adjustments was to add a few more questions, such as adding one more alternative to "How often do you change your passwords?" with the answer alternative "when the system asks me to change my password". This was one of the major improvements in the first survey. Another notification was to reconstruct the organization of managing and timing the experiment to make it more efficient for both the participant and the researchers. For the test experiment part two, the follow up, went as planned and got the collected data that was needed for the experiment.

After we gained a slight experience from the test experiment the execution of the real experiment began. As mentioned before the experiment was divided into two sections in which 41 participants with different education background and between the ages 18 to 66 years old were conducted. The experiment operated on various places, such as different workplaces and schools, to get a broader variation on ages and people.

The experiment began with meeting the participants. The meeting occurred in groups of two to eight people at the time, and the time frame between the first and second experiment was approximately one week. The purpose of meeting the participants physically was not only to observe but to be available if there were any questions or unclearness regarding the experiment. There were a few minor difficulties when some of the participants could not meet physically and had to do the experiment over the internet, through an email connection. This worked surprisingly well, and made it also possible to reach out to other potential participants.

Both parts of the experiment initiated with an introduction (see appendix 1 and appendix 3), the experimental assignment and ended with a survey (see appendix 2 and appendix 4). The first part of the experiment took about 10-15 minutes per person and the second part took 5-10 minutes.

Each experiment was performed individually and throughout the whole experiment the participants wrote down their password creations for hand and afterwards they did the surveys on the computer via google forms. Furthermore, we informed the participants to not write down passwords they use in their real life, but create passwords that resemble the kind of password they could use.

4.1 The experiment

To be able to get a valid result we needed to cluster the data and structure our variables into groups in SPSS, and there from measure the significance. The variables, password strength and the memorability needed to be divided and structured.

First we clustered the password strength to a 1 - 5 scale; 1=very weak, 2= weak, 3 = medium, 4 = strong and 5 = very strong. These were selected by how long the cracking time for a password was.

In contempt to, was this not enough divided and to get a valid significance, a furthermore division was needed. The password strength needed to be compressed into two groups, with the clusters included. Which transformed into; 1 = very weak, weak, medium *(cluster 1-3 in group one)* and 2 = strong, very strong *(cluster 4 -5 in group two).*

Medium was divided as a weaker password to get a better balance between the numbers. The cluster name seen as a headline in the different cross tabulation is PW_VSS_MVWW in SPSS, which stands for = Password very strong, strong versus medium, weak and very weak.

The memorability cluster was divided in two parts; where 100% = 1 and 99, 99% - 0% = 2. The name of the cluster is m_100%remember_and_others. For further explanation, the 100% defines the fully recall, and anything less is partial recall (99%).

The reason of the extreme division was to not get a count less than 5 in every cell. If the password strength would have an all the five groups the result in SPSS would not be valid on any of the password constructions because of lacking of participants in every cell.

4.1.1 Password construction 1, Own set password

In the first password construction the significance shows 83, 7 % which means that there are 83, 7 % chance of difference. The significance should be maximum 5 % in ability to not be a coincidence according to Norusis (2008). In the case of Table 1, the significance is too high, which means that there are no connection between the password strength and the memorability from password construction 1, Own set password. But even if the significance is not there, it shows a tendency to a pattern which can be seen in the cross tabulation. The weaker passwords and the stronger passwords have equal amounts of respondents remembering them to 100%, which means that the self-constructed passwords do not affect the memory associated with what strength the password has.

Table 1 - Cross tabulation password construction 1

PW1_WWWM_VSS * m1_100%remember_and_others Crosstabulation

			m1_100% remember_and_others		
			100%	All other	Total
PW1_WWWM_VSS	Very weak, weak, medium	Count	13	7	20
		% within m1_100% remember_and_others	50,0%	46,7%	48,8%
	strong, very strong	Count	13	8	21
		% within m1_100% remember_and_others	50,0%	53,3%	51,2%
Total		Count	26	15	41
		% within m1_100% remember_and_others	100,0%	100,0%	100,0%

Table 2 - Chi-square test password construction 1

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,042[a]	1	,837		
Continuity Correction[b]	,000	1	1,000		
Likelihood Ratio	,042	1	,837		
Fisher's Exact Test				1,000	,547
Linear-by-Linear Association	,041	1	,839		
N of Valid Cases	41				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 7,32.

b. Computed only for a 2x2 table

4.1.2 Password construction 2, Modified password

In the second password construction, the significance is 82, 3 % which means that password construction 2, the modified password, has a very high coincidence. The chance has a too great of impact in the modified password construction for it to have a connection in SPSS. When analyzing the cross tabulation in Table 3 the majority of the participants remembered their passwords, both those who created a weaker password, as those who created a stronger password.

Table 3 - Cross tabulation password construction 2

PW2vwwm_svs * M2_100%_other Crosstabulation

			M2_100%_other		
			100%	All other	Total
PW2vwwm_svs	Very weak, weak, medium	Count	9	6	15
		% within M2_100%_other	39,1%	42,9%	40,5%
	strong, very strong	Count	14	8	22
		% within M2_100%_other	60,9%	57,1%	59,5%
Total		Count	23	14	37
		% within M2_100%_other	100,0%	100,0%	100,0%

Table 4 - Chi-square test password construction 2

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,050[a]	1	,823		
Continuity Correction[b]	,000	1	1,000		
Likelihood Ratio	,050	1	,823		
Fisher's Exact Test				1,000	,546
Linear-by-Linear Association	,049	1	,825		
N of Valid Cases	37				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 5,68.

b. Computed only for a 2x2 table

4.1.3 Password construction 3, Association password

Password construction 3 differs from the other two password constructions, the own set and the modified password. It shows a better number in the significance. The significance in the association password construction was 12, 6% which makes the possibility for it to be a chance only 12, 6%. According to Norusis (2008) the significance should be 5%, but in this relatively small study, is 12, 6% significance a good number, even though it is not optimal.

In the cross tabulation in Table 5, the amount of the participants who remembered their password to 100% did not differ when creating a weak and strong password. Of the participants who did not remember their passwords fully, which was 73, 9 % did actually create a strong password.

Table 5- Cross tabulation password construction 3

PW3vwwm_svs * M3_100%_other Crosstabulation

			M3_100%_other		
			100%	All other	Total
PW3vwwm_svs	very weak, weak, medium	Count	8	6	14
		% within M3_100%_other	50,0%	26,1%	35,9%
	strong, very strong	Count	8	17	25
		% within M3_100%_other	50,0%	73,9%	64,1%
Total		Count	16	23	39
		% within M3_100%_other	100,0%	100,0%	100,0%

Table 6 - Chi-square test password construction 3

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2,345[a]	1	,126		
Continuity Correction[b]	1,421	1	,233		
Likelihood Ratio	2,337	1	,126		
Fisher's Exact Test				,179	,117
Linear-by-Linear Association	2,285	1	,131		
N of Valid Cases	39				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 5,74.

b. Computed only for a 2x2 table

4.1.4 Password construction comparison

Between the three different password constructions, the third, association, was the only one with a statistically significant, and was the one that showed a potential tendency for being significant. The other two constructions were not shown to be statistically significant to get a statistic support. Although, when analyzing all the three different password constructions in a cross tabulation. It showed that it is possible to see a pattern between the passwords strength and memorability. Both password construction 1, own set, and password construction 3, association password, for example, had equally many participants, where they remembered a weak or strong password perfectly. So in these cases the strength of the password did not affect the memory.

For further comparison, to find out which password construction is the best, a calculation on both variables was needed. Password strength was measured in hours and then divided in 10 000 000 because of the high numbers. The memorability was measured in percentage and calculated with the help of string-edit distance. In Table 7, the calculated result is shown from the three different password constructions.

Table 7- Password strength and memorability mean

	Password strength	Memorability
Pw1	139,1897641	76,195122
Pw2	2,1367472	70,8292683
Pw3	938,5173512	71,7073171

The result from the password strength among the password constructions was various, and the third password construction had by far the highest password strength. The memorability result was close in mean to each other, and all three constructions were around 70 %, and the construction with the highest memorability was the Own set password.

Figure 6 shows the relationship between the passwords strength and the memorability of the different password constructions in a bar chart. The result from this bar chart shows that the password construction 3, association password is the one with the best trade-off between a password's strength and memorability.

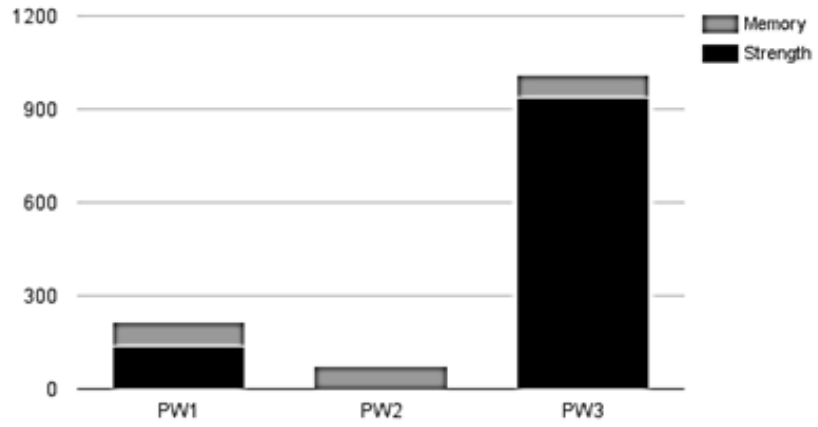

Figure 6 - Password strength and memorability

Figure 7 shows the password construction's strength and memorability put in a scatter plot. The figure describes how the password constructions relate to each other. Password construction 2, modified dictionary password, was the weakest one in both the variables. The most memorable one was password construction 1, the own set password, and lastly the third password construction, association password, got the highest password strength.

In this chart password construction 1, own set, and password construction 3, association are both strong but in different areas, which make it difficult to see which one of them that has the most potential and strongest trade-off.

Figure 7 - scatter plot on password strength and memorability mean

4.2 The survey

After every experiment a self-completed questionnaire was filled in by the participants to get more information about their password habits and their experience from the different password constructions. Survey 1 contributed with information about the participant's habits regarding passwords and the security pattern they have with them. This contributed with information and gave a better understanding of what type of passwords our participants have, and how they store them. This knowledge gave us the opportunity and the ability to compare their results with the existing theory. Survey 2 had the purpose of finding out what the participants thought of the different password constructions and how well they did remember.

4.2.1 Survey 1

Figure 8 shows how many passwords the participants use in their everyday life. According to Burnett and Kleiman (2006) a password should be unique for every account and not be reused. Figure 8 and Figure 9 are an insight of how the various participants' password management is. The few amount of passwords the participants have, might be an effect that they reuses them on several sites and accounts.

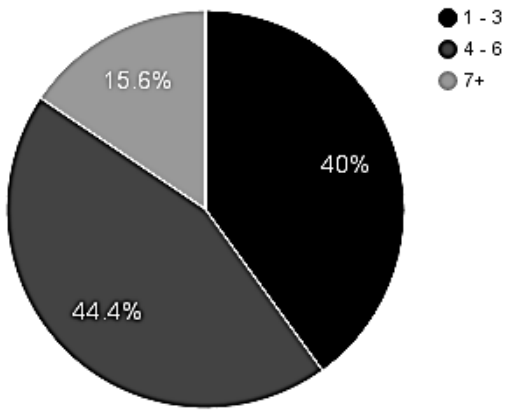

Figure 8 - How many passwords do you use in your everyday life?

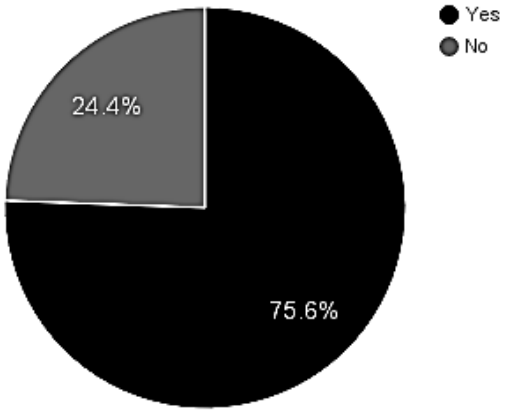

Figure 9 - Do you use the same password on several accounts?

To understand how the participants in the experiment memorize their passwords, the question presented in Figure 10 was needed for a broader understanding. According to Mitrovic´ (2005) storing a password physically is a risk because of the exposure to unauthorized people. Electronic storage is a safer way to store passwords because they are often already guarded by a password. The best and safest way is, if it can be remembered, to store a password by the user's memory, because it cannot be accessed by any other person.

The question shown in Figure 10, resulted in that there were 3 participants who did not store any passwords in their memory. Since the question could be answered with several options, shows that many participants store their password in several ways. It might be that password which is advanced or seldom used need to be written down. 9 of the participants stored their passwords physically, which are 22 % in total.

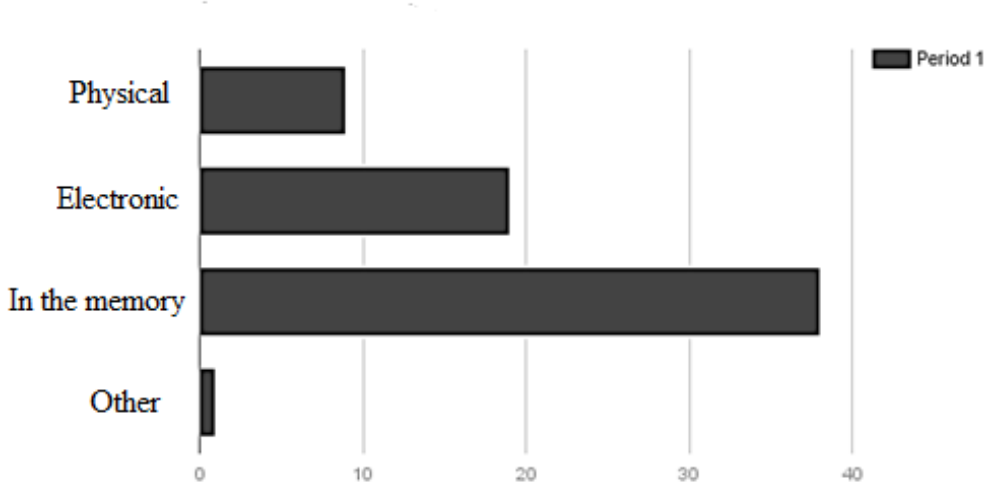

Figure 10 - How do you store your password?

Burnett and Kleiman (2006) argue that a password should remain a secret and it is only the user who should know it. A password should never be shared to prevent it from leaking. In

Figure 11, 57, 8% of the participants have shared their password to someone in their surroundings.

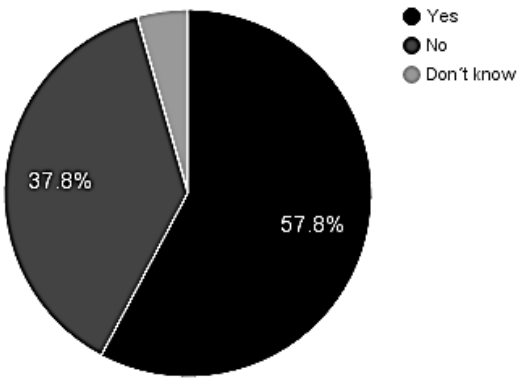

Figure 11 - Does anyone know one or more of your passwords?

According to Burnett and Kleiman (2006) a password should be changed every six month, or at least once a year. This increases the security, and the prevention from being hacked. There a few good features in some systems and accounts, where the system regularly reminds their users to change their passwords from time to time. In Figure 12, the majority of the participants changed their password only when the system asked them to, and the majorities never do change. There are exceptionally few who actually change their passwords frequently. Nevertheless, there are some participants who actually change their passwords once a year.

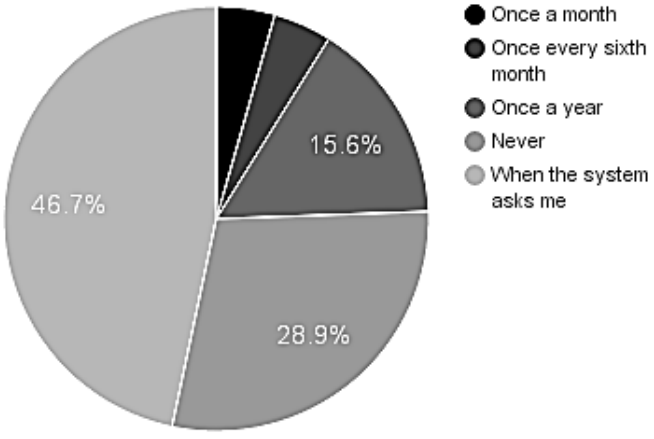

Figure 12 - How often do you change your password?

Burnett and Kleiman (2006) describes that a password should apply different constructions when creating a password, in possibility to make it stronger. The more different parts a password consists of, the stronger it gets and the harder it is to crack. The question in Figure 13

gives an overview of what the participant's passwords consist of and also an insight on the password strength. Few of the participants do use special characters, which contributes to a stronger password A simple dictionary attack does seldom use special characters in their prearranged word lists. They use more simple and common words (Narayanan & Shmatikov 2005).

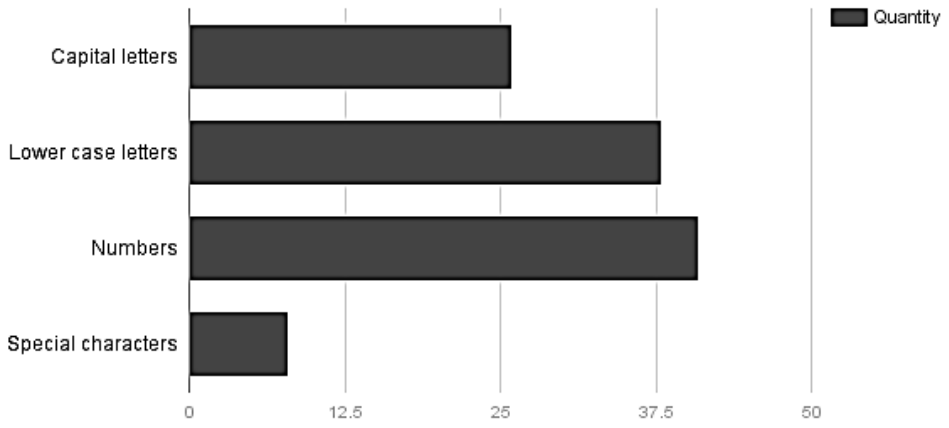

Figure 13 - What does your password consists of?

4.2.2 Survey 2

The first survey considers password security and strength. The second survey reflected the whole experiment and the memory. In Figure 14, the participants responded to how easy they thought the passwords were to remember. The memory can be affected of how well they remembered the password, at once or if they needed some extra time to rethink.

All of the three password constructions have been placed next to each other, which are shown in the Figure 14. They are also on the same scale to easily compare them against each other.
Drawn from Figure 14, the participants thought the own set password was the easiest to remember. The modified dictionary password and the association password got a more diverse result and the participants' assumptions were quite similar when it came to how well they did remember. The overall thought was that the modified password was generally difficult to remember, and the association password was hard to remember, which shows on the level 5 and 6 in the scale.

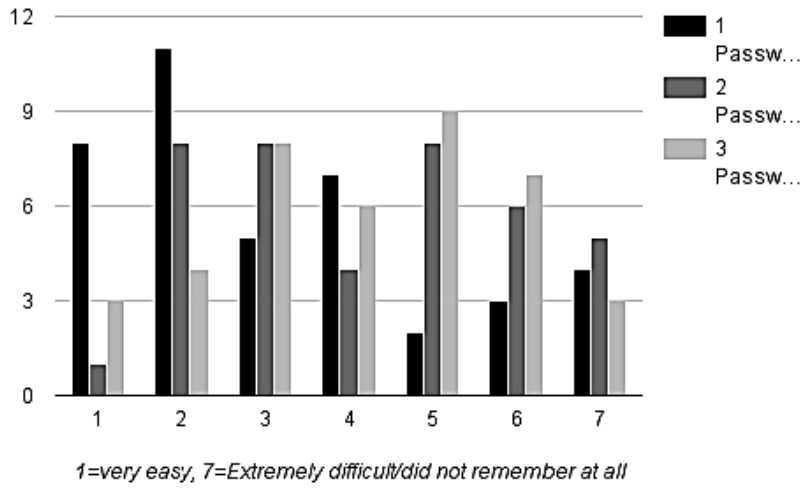

Figure 14 - How easy was the password to remember?

According to Burnett and Kleiman (2006) an effortless password is important for the user when typing it in. It must go fast, and not be complicated, otherwise will the user not be as motivated, and the risk is that they will forget it.

In Figure 15, the three different password constructions are compared to each other. To see which one of the constructions the participants prefers, and which one who is the most suitable one.

The password they constructed by them self, Own set, was the one they assumed to be the easiest one to type in on a computer. The password which would be the most difficult one to manage while typing in was the modified dictionary password according to our participants.

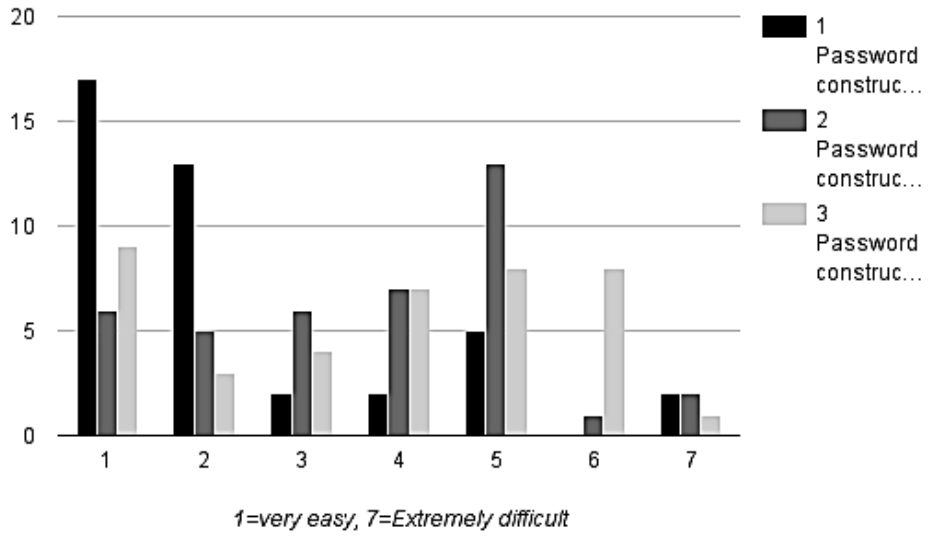

Figure 15 - How easy do you think the password was to type in on a computer?

If the participants would consider using one of the two password constructions created by the researchers, was an important question. This could lead to an increasing safety of the participant's future passwords, viewed in Figure 16.

Based on Figure 16, two of the participants do already use one or both of the different techniques that were performed in the experiment, and totally 40 % of the participants could think of using these creations when creating a password in the future.

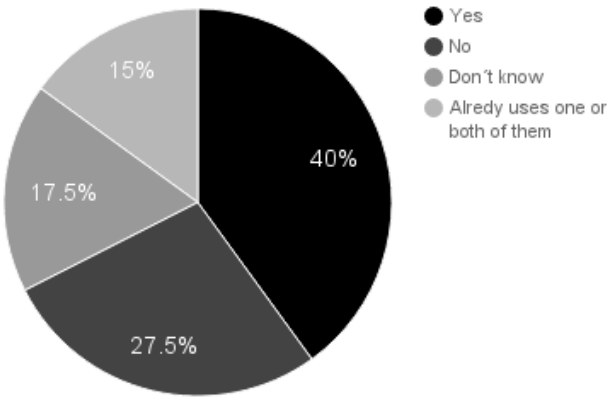

Figure 16 - Do you think you will use password construction 2 and/or 3 in the future?

5 Discussion

5.1 Method discussion

In the experimental approach, there were 41 participants included and according to Bryman and Bell (2015) a quantitative analysis need a minimum of 30 participants to be able to be measured in SPSS. When the experiment's main variables, password strength and memorability, were measured towards each other with a chi-square test, there was no statistical significance there. Looking at the cross tabulation of the different password constructions a tendency to a pattern could be seen but it could not be proved with statistics. The reason why this might be the case is that the sampling was too small. With a higher sample in the experiment could result in a significant number.

The approach of the method could be conducted in a qualitative way as well. Instead of a survey after the experiments, the participants could be interviewed about their password habits, individually or in groups. According to Recker (2013), the questions can be more or less formal. If less formal questions are applied, they are usually semi-structured, which makes the interview more flexible and the interviewee can answer more freely. A disadvantage with interviews could be that the respondents will not be fully truthful which could affect the result negatively (Bryman and Bell 2015). Therefore, in this case by using surveys can be an advantage, where the respondents can be anonymous and may answer the question more truthfully.

Choosing a different method is an option, which could be done. The reason why, was because we, as researchers, found that this report did not require any in-depth study of the participants. The main focus was the result from the experiment. The research focused on the understanding the habits and how to improve passwords by using simple constructions. A qualitative approach is possible for this study, but in the research executed in this thesis was better suited with a quantitative method because of the possibility to measure the numbers.

5.2 Result and analytical discussion

The results and the analysis rely on the experimental study. The findings from Survey 1 turned out to be that some of the participants do use strong passwords in their everyday life. The majority of the participants do not have and use a lot of passwords, and 75% of them reuse their passwords. An interesting part was that 84, 4% of the participants do use their memories to store their passwords, which was a lower number than expected. Many of them did share their passwords with others, and a quarter of the participants saved their passwords physically on a paper. They rarely changed their passwords, but when they did, was when the system asked them to. Almost all of the participants did not use special characters in their passwords, which can affect them negatively according to Burnett and Kleiman (2006). Using a special character can minimize the risk of being hacked and be a target for a dictionary attack. It is an easy way of strengthening a password (ibid).

Some factors based on the result from Survey 1 regarding the participants' passwords in their daily life, can be that they choose few passwords that are convenient and easy enough for them to remember. The reason of the reusing their passwords can be that they already find them safe and strong, and get an impression that they do not need more and better passwords. This can also be an explanation of the reason why they do not change their passwords so often. The

majority of the participants only change their passwords when they get reminded by the system. According to Burnett and Kleiman (2006) a password should be changed every six months and be unique for every account that the user has in ability to keep it safe. Even though their passwords seemed weak and vulnerable when analyzing Survey 1, it was surprising to notice that the password construction 1 in the experiment was more secure than expected, and the remembrance was high as well. A reason of their ability to recall could be the low criteria of the password construction - where it should at least contain 8 characters with both letters and numbers. This gave the participants the option to create their password freely through their own capacity and set their own security on the password.

In survey 2 the participants answered a question on if they thought the password was easy to remember, showed in Figure 14. This question can be linked to the result from the Levenshtein distance where their memories were measured. We compared the password similarities to each other to measure the memorability. The measurement of the memory is close related to the result of how the participants thought they remembered. The first password construction might be the one with the highest memory rate, because the participants could set their own password related to their capacity and could choose what to relate to. The other two constructions were likely to be unfamiliar for them, which can be a contributing factor for the participant to forget (Sasse, Brostoff & Weirich 2001). However, this research shows a tendency that the memorability is quite high in these two constructions as well. One reason for the high remembrance of the different password constructions could be the amount of guidance they got, which helped them along the way. For instance, they got the same password examples in experiment 1 as they had at experiment 2.

The participants got a question about how easy they assumed the three different password constructions would be to type in on a computer, seen in Figure 15. Password construction 2, modified dictionary password, was the password the participants thought would be the hardest to type in on a computer. One reason for this could be the extra effort to type in a special character, because it requires a button combination. Figure 13 represents what the participants passwords consists of. Few of the participants use special character in their everyday password, which may be the reason why they think they have difficulties remembering it. A good password should be simple to type in and easy to remember and still be strong. Using a special character in a password increases the strength (Burnett and Kleiman 2006).

Other aspects to take in consideration which affected the result of this experiment can be the participant's education and their IT knowledge (Bryman and Bell 2015). Each participant has different kind of knowledge and experience regarding password security and creation of a secure password. When performing the experiment the size of the groups affected the result as well. In every experimental occasion there were between two and eight participants. In groups above four participants misunderstandings of the task often occurred. The reason for the misunderstandings in the larger groups might be because of less private instructions (besides the tasks described in the experiment) could be given. With larger groups the researchers perceived the environment more stressful and disordered, which more often resulted in invalid answers. After this problem was discovered, and to prevent these kind of situations, only smaller groups of two to four persons participated each time .

Reviewing the result, it showed that the memorability was quite similar between all of the constructions and password construction 1, own set, was the one with the highest memory rate. The password strength between the constructions showed a larger difference. Association

password was the one with the highest password strength by far. The other two constructions had lower numbers in password strength and had a more similar result (see Figure 6).

5.3 Conclusion

The conclusion is based on the research that has been presented in the result and analysis, and it will be presented in this chapter along with the answer of this research question in this thesis. The research questions purpose was to find out which one of association-, modified dictionary- and own set password that got the best trade-off between password strength and memorability.

The conclusion of the three password construction (own set, modified and association), was that the own set password construction was the one with the highest memory rate. The modified password construction was the weakest one of all, on both password strength and memorability. The association password construction created the strongest passwords, and was still quite memorable.

The password construction with the best trade-off between password strength and memorability, according to the researcher, was password construction 3, the association password. Because it had by far the highest password strength and it was relatively close in the memorability rate compared with the other constructions which can be seen in Table 7.

We also concluded, drawn from the result in survey 1, that the majority of the participants do obtain weak passwords. Almost all literature that has been applied in this report, for gaining knowledge about information security has repeatedly stated that the humans are the weakest link in an information system. This conclusion was based on theories about password security and the result from survey 1. The majority of the participants do have few passwords, which they share with people in their surrounding and reuses. When creating a new password the passwords strength and memorability should be considerate, a safe password should be complex, yet easy to remember. The trade-off between the passwords strength and memorability is an important key for a successful and safe password.

5.4 Contribution

This research has contributed to science with new information and knowledge, both from the theoretical framework and from the analysis of the executed experiment. Instead of changing the whole password authentication systems, has this study given other authentication solutions were the password users can adapt from. Password strength and memorability could be seen as guidance, and using a construction of a combination between a strong and a memorable password are beneficial for everyone who are interested, both organizations and home users.

A similar research has been performed, where Acar, Fahl and Backes (2016) tested the memorability of passwords. The difference was that they did only research the participants who remembered the passwords to 100%, and those who did not. Where in this thesis we reflected and analyzed the remembrance a bit further. We calculated the similarities between the created password strings. The reason for that was the experience we had with our participants, where we noticed that they did often forgot just one character in their passwords, which makes the distance between the passwords very small, and the similarity to each other was around 99%. The remembrance was quite high, and in fact that we, as researchers took account in what

proportions the participants actually remembered. Where we reflected around the subject, and we consider that this aspect has contributed to science as well.

5.5 Future research

In future research a larger sample of participants might be able to give the research a significant number and statistical confidence. In this research it was not possible, probably because of the sizes of the sample. For future study, the researchers recommend to perform an experiment in a larger scale, where more research time is required and with at least 100 participants.

If a larger sample size is used, we also recommend adding more password constructions to expand the variety of finding and for a wider knowledge of the connection between password strength and memorability. It would be interesting to see how a password created by a John the Ripper application affects the participant's memory as well.

We think that the variables, password strength and memorability, measured in this research also will suit a research of a larger scale. The Levenshtein distance was a good choice to measure the memorability; it gave an exact number based on the similarities between the passwords that were made in the experiment. The password meters are on the other hand under constant development, and will expectantly be better evolved and more reliable over the time. We recommend that future studies to examine different alternatives before choosing one or several password meters, this to get one suitable for their password construction or constructions when testing the password strength.

A suggestion for future research is to choose various samples, for example stratified sampling, of the participants when performing a study. While conducting this research, we came to realize that different age groups and education levels obtained diverse results in the experiment. A broader sample span is a strong recommendation.

References

Ahlberg, M. Blomberg, M & Davidsson, P. (2008). *Passwords – The Achille's heel of information systems?* Borås: University of Borås.
URI: http://hdl.handle.net/2320/4079

Bryman, A. & Bell, E. (2015). *Business research methods.* 4.th edn., Oxford: Oxford University.

Burnett, M. & Kleiman, D. (2006). *Perfect passwords: selection, protection, authentication, Syngress*. Rockland: Mass.

Carnavalet, X. & Mannan, M. (2015). "A Large-Scale Evaluation of High-Impact Password Strength Meters", *ACM Transactions on Information and System Security (TISSEC).*
DOI: 10.1145/2739044

Dhillon, G. (2006). *Principles of information systems security: text and cases.* Danvers: Mass.

Encyclopædia Britannica Inc (2016). Phishing.
http://academic.eb.com/levels/collegiate/article/471407
[2016-12-12]

Forget, A., Chiasson, S., Van Oorschot, P.C. & Biddle, R. (2008). Persuasion for Stronger Passwords: *Motivation and Pilot Study" in Springer Berlin Heidelberg*. Berlin: Heidelberg.

Fåk, V. (n.d.) Nationalencyklopedin - *IT-säkerhet.*
http://www.ne.se/uppslagsverk/encyklopedi/lång/it-säkerhet
[2017-01-06]

Informationssäkerhet. (2015). Varför är informationssäkerhet viktigt?
www.informationssakerhet.se/Om-informationssakerhet-kon/vad_ar_informationssakerhet/
[2016-12-15]

Jouini, M., Rabai, L. & Aissa, A. (2014). *Classification of security threats in information systems*, pp. 489-496.

Karlsson, K. & Stålnacke, E. (2002). Data och systemvetenskap: *Säker konstruktion och hantering av lösenord*. Luleå: Luleå tekniska universitet.

Kritzinger, E. & von Solms, S-H. (2010). Cyber security for home users*: A new way of protection through awareness enforcement, Computers & Security,* pp. 840-847.

Likhitha, C., Ninitha, P. & Kanchana, V. (2016). DNA Bar-coding: *A Novel Approach for Identifying an Individual Using Extended Levenshtein Distance Algorithm and STR analysis, International Journal of Electrical and Computer Engineering*, pp. 1133-1139.

McLeod, S. A. (2009). Simply psychology - Short Term Memory.
www.simplypsychology.org/short-term-memory.html
[2017-01-06]

Merkow, S.M. & Breithaupt, J. (2014). *Information security- principles and practices.* 2nd edn., Indianapolis: Pearson.

Mitrović, P. (2005). *Handbok i IT-säkerhet.* 4th edn., Sundbyberg: Pagina.

Narayanan, A. & Shmatikov, V. (2005). *Fast dictionary attacks on passwords using time-space tradeoff.* p. 364.

Norusis, M. (2008). *SPSS Statistics 17.0 Guide to Data Analysis*. Chicago: Prentice Hall Inc.

Oates, B.J. (2006). *Researching information systems and computing.* SAGE: London.

Recker, J. (2013). *Scientific research in information systems: a beginner's guide.* Berlin: Springer, Heidelberg.

Ristad, E.S. & Yianilos, P.N. (1998). Learning string-edit distance: *IEEE Transactions on Pattern Analysis and Machine Intelligence.* pp. 522-532. DOI**:** 10.1109/34.682181

Robson, C. (2011). Real world research: *a resource for users of social research methods in applied settings.* 3.rd edn., Chichester: Wiley.

Roediger, H.L. & Karpicke, J.D. (2006). The Power of Testing Memory: *Basic Research and Implications for Educational Practice.* pp. 181-210.
DOI: https://doi.org/10.1111/j.1745-6916.2006.00012.x

Roundy, L. (n.d.). What is Cognitive Psychology? - *Definition & Theories*.
http://study.com/academy/lesson/what-is-cognitive-psychology-definition-theories-quiz.html
[2017-01-08]

Sasse, M A., Brostoff, S. & Weirich, D. (2001*).* Transforming the 'Weakest Link': *a Human/Computer Interaction Approach to Usable and Effective Security*. BT Technology Journal, pp. 122-131. DOI: 10.1023/A:1011902718709

Schepens, J., Dijkstra, T. & Grootjen, F. (2012). Distributions of cognates in Europe as based on Levenshtein distance, *Bilingualism: Language and Cognition*, pp. 157-166.
DOI: http://dx.doi.org/10.1017/S1366728910000623

Shiffrin, R.M. & Atkinson, R.C. (1969). Storage and retrieval processes in long-term memory: *Psychological Review*, pp. 179-193.

Statistics Swedish. (2015).
http://www.sverigeisiffror.scb.se/hitta-statistik/sverige-i-siffror/manniskorna-i-sverige/befolkningspyramid/
[2016-12-07]

Sykes, E.R., Lin, M. & Skoczen, W. (2010). Journal of Physics: Conference Series*: MPI Enhancements in John the Ripper*. DOI: 10.1088/1742-6596/256/1/012024

Taneski, V., Brumen, B. & Heričko, M. (2014). Password security — No change in 35 years? *Information and Communication Technology, Electronics and Microelectronics*, pp.1360-1365. DOI: 10.1109/MIPRO.2014.6859779

Ur, B., Kelley, P.G., Komanduri, S., Lee, J., Maass, M., Mazurek, M.L., Passaro, T., Shay, R., Vidas, T., Bauer, L., Christin, N. & Cranor, L.F. (2012). How does your password measure up? *The effect on strength meters on password creation.* Carnegie: Mellon University.

Vu, K.L., Proctor, R.W., Bhargav-Spantzel, A., Tai, B., Cook, J. & Schultz, E. (2007). International Journal of Human - Computer Studies: *Improving password security and memorability to protect personal and organizational information*, pp. 744-757. DOI: http://dx.doi.org/10.1016/j.ijhcs.2007.03.007

Yan, J., Blackwell, A., Anderson, R. & Grant, A. (2004). Usability and Security: *Password Memorability and Security: Empirical Results,* pp.25-31. DOI: 10.1109/MSP.2004.81

Appendix

Appendix 1

Introduction

We are two students, Hanna and Josefin, from the University of Borås and are currently writing our thesis in informatics. You are going to take part of an experiment connected to our thesis which is about passwords and memorability. The purpose of this experiment is to see what kind of passwords that is both safe and easy to remember.

In this experiment you are going to create **three <u>different</u> passwords** for a more secure login, example a password to your e-mail. We will not put any interest in bank- account passwords because we find the criteria too complicated for this study.

This experiment will be performed in two occasions. This first time you will create three different passwords, and the second time, the follow up, is to see if you can **remember the passwords** you have created.

For the experiment to work you are challenged to remember all of the three passwords in your head, and you are not allowed to write the passwords down.

**** Do not write a password you use and have in your real life****

Part 1- Own password

Password example: HannaJosefin16

Create a password that you find safe and easy to remember.

Criterias:

- It should be **8-14 characters long**
- No characters of "Å, Ä, Ö"

Password: _____

REMEMBER THIS PASSWORD UNTIL NEXT TIME
Part 2 - Modified dictionary password

Password example: R35turang#

For creating a stronger password, should numbers and character be included. A trick is to choose a word that you can remember and modify it to make it safer by changing some letters to number and characters. This makes it harder for example a dictionary attack to crack your password.

Criterias:

- The password should be **8-14 characters long**
- You should have at least one capital letter
- Change at least two letters to numbers *(Ex: E=3, 4=A, S=5, I=1)*
- You should at least have one sign character *(#¤%&$)*
- No characters of “Å, Ä, Ö”

Password: _____

REMEMBER THIS PASSWORD UNTIL NEXT TIME

Part 3- Assocation password

Password example:I got married in Las Vegas December 2015 and got divorce 2016 -----> IgmiLVd15agd16

Association password can for example be a memory or a moment that you compress to a password. It can also be an association to a certain website, which can make it easier to remember.

Criterias:

- Create a sentence that is meaningful for you
- The password should be between **8-14 characters long**
- Should contain numbers
- Use at least one capital letter and one lowercase letter
- Special characters are optional *(#¤%&&$)*
- No characters of "Å, Ä, Ö"

Write down the memory sentence:

_____

- Take the first letter in the in every word and the two last numbers in the year or dates to create the password:

Password:_____

REMEMBER THIS PASSWORD UNTIL NEXT TIME
You will now do a survey on a computer

Appendix 2

Survey after experiment 1

1. **First and last name:**_____
   (Name is needed to connect your answers for the next part of the experiment)

2. **Age**: _____

3. **Gender:** [ ] Male [ ] Female [ ] Other

4. **Your highest completed or ongoing education?**
   [ ] Primary education (högstadiet)
   [ ] Secondary education (gymnasiet)
   [ ] Post-secondary education (högskolan/universitet)

5. **How many passwords do you use in your everyday life?**
   [ ] 1-3 [ ] 4-6 [ ] 7+

6. **Do you use the same passwords on several accounts?**
   [ ] Yes [ ] No [ ] Don't know

7. **How do you store your password (multiples answer is okey):**
   [ ] In memory
   [ ] Physical(ex. on paper)
   [ ] Electronic(ex. in your phone/computer)
   [ ] Other:_____

8. **How often do you change your password?**
   [ ] Once a week
   [ ] Every month
   [ ] Every six months
   [ ] Once a year
   [ ] Never
   [ ] When the system asks me to change

9. **Does anyone know one or more of your passwords?**
   [ ] Yes [ ] No [ ] Don't know

10. **Does your password contain any personal information?**
    [ ] Yes [ ] No [ ] Don't know

11. **How many characters does your password contain?**

[ ] 0-5
[ ] 6-8
[ ] 9-12
[ ] 12+

12. **What does your password consist of? (multiples answer is okey).**

[ ] Capital letters
[ ] Lowercase letters
[ ] Numbers
[ ] Special characters (eg. # * /)

Please write down you email for further connection if we can not meet for the follow up.

Thank you for your cooperation, See you in 1 week!

Appendix 3

Experiment part 2
- The follow up

Last time we meet you created three different passwords. We would like you to write the passwords down from last time.

Part 1. Own password:_____
Password example: HannaJosefin16

Part 2. Modified dictionary password:_____
Password example: R35turang#

Part 3. Association password:_____
Password example: I got married in Las Vegas December 2015 and got divorced 2016 ---> IgmiLVd15agd16

Appendix 4

Survey 2

Circle the answer you think is right

Password from part 1, Own password

1. How easy was the password to remember?
1 being very easy and 10 extremely difficult/didn't remember at all

1 2 3 4 5 6 7

2. How easy do you think this kind of password will be to type in on a computer?
1 being very easy and 10 being very hard

1 2 3 4 5 6 7

Password from part 2, Modified dictionary password

3. How easy was the password to remember?
1 being very easy and 10 extremely difficult/didn't remember at all

1 2 3 4 5 6 7

4. How easy do you think this kind of password will be to type in on a computer?
1 being very easy and 10 being very hard

1 2 3 4 5 6 7

Password from part 3, Association password

5. How easy was the password to remember?
1 being very easy and 10 extremely difficult/didn't remember at all.

1 2 3 4 5 6 7

6. How easy do you think this kind of password will be to type in on a computer?
1 being very easy and 10 being very hard

1 2 3 4 5 6 7

Summary question:

7. Do you think you will use the part 2 and/or part 3 password constructions in the future when creating a password?

[ ] Yes [ ] No [ ] Don't know [ ] Already using one of them

8. If you have any further comments or you want to add something:

_____

_____

_____

_____

THANK YOU FOR YOUR TIME!

// Hanna Julkunen & Josefin Ceder Molander

University of Borås is a modern university in the city center. We give education programs and courses in business administration and informatics, library and information science, fashion and textiles, behavioral sciences and teacher education, engineering and health sciences.

At the **Department of Information Technology,** we have focused on the students' future needs. Therefore, we have created programs in which employability is a key word. Subject integration, wholeness and contextualization are other important concepts. The department has a closeness, both between students and teachers as well as between industry and education.

Our **courses and programs** with a major in informatics are centered around basic concepts as system development and business development. In our wide range of specializations there is everything from programming advanced systems, analyze the needs and requirements of businesses, to conduct integrated IT and business development, with the common purpose of promoting good use of IT in enterprises and organizations.

The department is carrying out IT-related **research** within the university's research area called Business and IT. In terms of field, the research activities are mainly within **computer and systems science**. Particular areas of focus are **data science** and **information systems science**. Both scientifically and professionally-oriented research are performed, which among other things is manifested through that research is often conducted based on domain specific needs of business and government organizations at local, national and international arena. The professionally-oriented research is also often manifested through our participation in the Swedish Institute for Innovative Retailing (SIIR), which is a research center at the University with the aim of contributing to commerce and society with the development of innovative and sustainable trade.

UNIVERSITY
OF BORÅS

VISITING ADDRESS: JÄRNVÄGSGATAN 5 · POSTAL ADDRESS: ALLÉGATAN 1, SE-501 90 BORÅS
PHONE: + 46 33 435 40 00 · E-MAIL: INST.HIT@HB.SE · WEB: WWW.HB.SE/HIT