Bachelor Degree Project

UNIVERSITY OF SKÖVDE

PASSWORD PRACTICE

The effect of training on password practice

Bachelor Degree Project in Computer Science
G2e 15 ECTS
HT 2015

Niklas Ekström

Supervisor: Manfred Jeusfeld
Examinator: Jianguo Ding

Abstract

There are several concerning issues with passwords today; one of them being weak passwords, but password management also plays a big role e.g. when the users reuses passwords over several services or don't change their passwords on a regular basis. With the usage of passwords for several aspects of our daily lives comes the responsibility of trying to mitigate these issues, a role that often falls on to the users themselves. The usage of guidelines has proved helpful in this regard but still lacks important aspects. This paper suggests the usage of education in the form of a lecture to help with the problem. In this paper we conducted a study of password leaks, a literature analysis of the area around passwords and perform some qualitative interviews with different kinds of people with varying education and usage of passwords. The results from these studies will then lay the foundation for the lecture in the experiment part of the paper, two experiment groups will be used, one given a lecture as education on the matter and one control group not given any education. The study has showed that the usage of a lecture can help increase the entropy, average length of user's passwords. These results can be interpreted together with another study that did a similar experiment to that a lecture can be a more efficient way to teach users about passwords.

Keywords: Password, Lecture, Interviews, Experiment, Entropy

Table of content

1. Introduction

Passwords are a part of people's life; billions of people come in contact and use passwords each day both for work and personal life. But with the high usage of passwords and the data that they help protect, also comes the necessity to pick a good password. Criminals have in recent years started to focus on online crimes because of the growth of e-business (Power, 2000) and the fact you can attack someone from anywhere across the globe. According to Brottsförebyggande rådet (BRA) the amount of reported data breaches increased by 99% during the first half of 2012. Anton Färnstrom, a statistician working at BRA, claims that there is an increase of reported computer breaches; statistics shows that over the last five years the increase is an average of 35% per year (BRA 2014).

2. Background

There are several other problems with passwords other than a passwords strength. It is also important how users store their passwords, how often they change them and that they do not reuse the passwords for several different services.

According to Scarfone & Souppaya (2009) recommendations of the National Institute of Standards and Technology, a password can be defined as a secret string of characters, numbers and special characters applied by a user to authenticate its own identity. The authentication process can either be this string or any other means of authentication, like fingerprints or vocal patterns. There are several different factors of authentication, by using just a single authentication factor only one of these will be used. With two-step authentication an extra layer of authentication can be used such as Radio-frequency identification (RFID) chip or an authentication code generated by a mobile application or being sent by an automated email to the user. By using more authentication factors it will be harder for malicious users to gain access to the system that it helps protect. More authentication methods also make it more difficult for the user itself to authenticate to the system, which could be why some users do not take advantage of two-step authentication when it exists. A password can consist of both a passphrase, which is a longer combination of several words and numbers or a mix of characters, numbers and special characters. A personal identification number (PIN) can also be used as a password, which consists of a series of digits and does not include characters. This is most commonly used together with banking cards or access cards. This is a way of two factor authentication where just having the PIN number won't be enough to gain access to the data it helps to protect.

The background section will also include some more information about strong passwords and what different organizations consider being a strong password. It will also examine what additional guidelines they provide for users who are registering an account such as password management, and general information about password management and two studies of leaked passwords. The last part of the background will study two similar papers on the area of password management.

2.1 Strong password

A strong password according to Fordham (2008) is a password that is extremely difficult to guess, no matter how much information the guesser knows about the user's personal life. The password should appear to be a random mixture of letters, numerals and special characters that appear as gibberish for someone else than the owner. The password should also be unique for the service it is used for, reusing the same password for several services should be considered a weakness, Furnell (2007). The password should also be changed on a regular basis; it should be changed before a malicious user can get access by using a brute force attack.

A study made by Lucas (2009), who runs the website called Lockdown which focuses on computer security, tested how different passwords stand up against a brute force attack. It was found that when using an 8 character long password combined of both upper- and lowercase passwords, which results in 200 billion possibilities, it would take a computer with a Pentium 100 processor which is capable of 10,000 guesses per second 242 days to crack the password by testing all possible combinations. A dual processor personal computer (PC) which is capable of testing 10,000,000 passwords per second could do it in 348 minutes. This shows that the capability to crack a password is both dependent on the strength of the password and the capability of the attacker.

Adding special characters or numbers to a password will add even more cardinality to the password and make it even harder to brute force.

2.1.1 Password guidelines

There are several guidelines available online on what the user should think about when picking their password; this section will go through three different ones. One is supplied by the University of Skövde for new coming students. The second one is from National Aeronautics and Space Administration (NASA) which is the guidelines supplied for employees at the organization. The last one is from Microsoft and are the guidelines provided when registering a new e-mail account on their service. The reason these three guidelines were picked is because they all protect varying types of data.

The University in Skövde (2005) lists their guidelines as:

- The password should be 8 characters long.
- Do not use words that exist in a dictionary.
- Do not use any information that can be linked back to you.
- The password should to be easy to remember.
- The password should contain at least one number and at least two characters; one upper- and one lowercase.

- No special Swedish characters (ÅÄÖ).

No information regarding how often the password should be changed or about reusing passwords was included in the guidelines. These rules are also not enforced by the system in any way, and the user is able to pick a password not fitting to the rules.

NASA guidelines defined by Moyer (2014a) seem to have a more strict policy on passwords than the University in Skövde and it includes:

- A minimum of 12 characters.
- Include at least three of the following types of characters:
    - Uppercase letters
    - Lowercase Letters
    - Numbers
    - Special characters (e.g., !@#)
- Do not use a password that can be easily guessed (Music bands, user id, 1234, abc)
- Do not reuse any of your previous 24 passwords.
- You must change your password every 60 days.

For NASA employees the password has to be 4 characters longer than for the students at the university, it also has to include a special character. Also, password management is discussed in the guidelines like changing the password on a regular basis and not reusing a recent password (Moyer 2014a). The password guidelines are provided for users registering accounts for the *High-end computing capability* service. The users also have to use two-step authentication where they use a RSA SecurID Fobs which is a device that generate a number every 30 seconds which also has to be entered at logon (Moyer 2014b).

The data that the passwords protect are very different and that is probably why the NASA guidelines are stricter than the guidelines supplied by the University of Skövde. A compromised account at NASA could be more costly for the company than if the same thing happens for a student at the University of Skövde.

According to Craddock (2013) there are 400 million Outlook email accounts registered, one reason could be that Windows 8 requires a Microsoft account which also could be used for the Outlook account. According to the Microsoft (n.d) guidelines the keys to a strong password are:

- Whenever possible, use eight characters or more.
- Don’t use the same password for everything.
- Change your password often. (3 months)

- Use a great variety of characters in your password.

Microsoft also provides some tips on creating a long, strong password that is easy to remember:

1. Start with a sentence or two (Passwords are safe).
2. Remove the spaces between the words in the sentence (Passwordsaresafe).
3. Turn the words shorter or intentionally misspell a word (Passwarsaafe).
4. Add length by using numbers that are meaningful to you (Passwarsaafe2015).

Microsoft also lists some common pitfalls to avoid for example using words that exist in dictionaries, sequences of repeated characters or personal information.

Following the guidelines from all three sources, a strong password can be summarised into:

- Use more than 8 characters, the longer the better but not so long that you will have problems remembering it. Put several words together to easier remember the password.
- Don't use common words or words that can be linked back to you or personal interests.
- Use a mixture of special characters, lower- and uppercase letters and numbers.
- Do not reuse the same password on several services, and do not reuse recently used passwords.
- Change your password on a regular basis.
- Use a passphrase if able to.

In 2007 a study was conducted by Furnell (2007) of the guidelines given by the top 10 visited websites listed on the *Alexa Global top 500 Websites*. The study's intent was to see what kind of guidelines they were giving the users when registering new accounts on the websites. Furnell (2007) investigated if the websites gave any guidance to selecting passwords, if they had any restrictions on the passwords to stop the user from making poor choices and if any form of assistance was available if the user forgot the password. He found that most websites provided little to non-guidance on how to pick a good password but restrictions on what could not be used. He raises the question on how users will be able to pick good passwords if websites don't emphasise to use them. He argues that with the restrictions given regarding what to not include in the password but no information on why it is restricted, the websites fail to provide the user with information on why these constrains are required.

2.2 Password management

Password management includes how the users store their password, and manage the aspects surrounding it. How often do they change it? Do they reuse the password on several other services? The management of passwords is the first line of defence in picking a strong password. Even if the user has what is considered a strong password but stores it on a note on the screen a malicious user only need to gain physical access to this note to be able to learn the users password and then it does not matter how strong the actual password is. Same with changing the password and reusing the same password over several services, if one service gets compromised services with the same password are also compromised.

A survey conducted by Florêncio & Herley (2007) on half a million users over a 3 month period where they recorded user's password management. The authors found that an average user has around 25 accounts that require passwords and have an average of 6.5 different passwords. The survey also discovered that a user uses an average of 8 passwords per day.

According to Florêncio & Herley (2007) a user with 30 different accounts does not have a problem remembering their passwords; the biggest problem is remembering which of his 5-6 passwords were used on what service. Users seem to tackle this problem by writing their password down on a piece of paper, trial and error tries or password resets.

In an article written by Harrison (2006) he mentions that what security professionals see as responsible behaviour with password management users only see as an obstacle in the way of the task they are trying to perform. They might not know why they have to change the password on a regular basis or why they need to use different password on all their services and only see this as an annoyance. Harrison argues that the users will find ways to circumvent password changes, so they don't have to remember an additional password e.g. to make additional password changes so they can reuse old passwords.

Cheswick (2013) believes that a way to address this problem is by making the authentication procedure less tedious and more fun. When entering a password and making a typographical error the users should not be punished for this, entering the same incorrect password twice on a service should only count as one try. Also the password management for personal accounts might differ for users than the management for passwords for their workplace.

2.3 Studies of leaked passwords

This section will include a study made by *The Imperva Application Defense Center* (2014) (here by referred as ADC) of 32 million passwords. It will also include a section where a study was made off 10 million passwords leaked early February of 2015.

2.3.1 RockYou leak

In 2009, 32 million *RockYou* accounts where leaked onto the Internet, they were according to *ADC* stored in clear text and extracted by using an Structured Query Language (SQL)

injection vulnerability. *ADC* then analysed the strength of these passwords and found that over 49% of the passwords used less than the recommended 8 characters as a password. 41.69% of the passwords only consisted of lower case characters. The *ADC* found that only 0.2% used what they consider a strong password which they defined by the NASA guidelines that a strong password consists of eight characters or longer, a mixture numbers, special characters and upper and lowercase letters. These NASA guidelines are from the publication "*NPG 2810.x Guidelines for Passwords*" which differs from the guidelines from NASA used in the section 2.1 *Strong passswords*, and was used for a timekeeping system used by NASA.

They also found out that the 5000 most popular passwords on the website were used by 20% of the users with the most popular password being "*123456*" which was used by 290 731 users. There were 61 958 that used the word "*password*" as their password, 5 out of 20 of the most common passwords on in the leak were first names.

2.3.2 Study of the xato leak

In February 2015 Mark Burnett shared his own collection of ten million passwords to be used for academic purpose. This paper will also conduct its own analyse of these ten million passwords to look for patterns of the passwords.

To remove any forms of illegal use of the passwords the domain portion of the email addresses have been removed, the password samples are also collected over incidents occurring over a 10 year period so they cannot be tied back to a specific company. Mark Burnett has also manually reviewed much of the data to remove information that can be linked back to a specific individual and also any forms of information connected to credit cards or financial account numbers. Mark also consider these passwords to be dead passwords which cannot be defined as any form of authentication because they will not allow a user to authenticate with them which will make them useless for illegal purposes.

After analysing the file by sorting it the file the following data was collected which is shown in table 1.

Password	**Number of occurrences**	**Password**	**Number of occurrences**
1. 123456	16 147	**6. qwerty**	2 789
2. password	6 370	**7. 1234**	2 295
3. 12345678	4 340	**8. 1111111**	1 818
4. 123456789	3 534	**9. klaster**	1 697
5. 12345	3 458	**10. 1234567**	1 510

Table 1. Top 10 passwords from Xato leak

The most common password seemed to be "123456" which does not follow the guidelines for a strong password. Not a single one of the top 10 passwords can be considered as a strong

password from the guidelines described earlier. It is not until the 1439st most common password that a password fulfils the criteria previously defined as strong with having both upper and lowercase, numbers and not being a common word. This password is "Soso123aljg" with 86 occurrences. The results are also similar to the study discussed earlier that was done by *ADC*.

2.4 Password memorability

According to Yan et al. (2000) many of the problems with password are linked to the human memory. If this limitation did not exist the maximally secure password would consist of the highest entropic value possibly and the maximum of characters the system allows i.e. a totally randomly generated password. The entropy value is described by Shannon (1950) as a statistical parameter of how much information is produced by calculating the cardinality and the length of a word; it is described in depth in section *6. Password entropy*. The problem with a password of the highest entropic value possible is that it would be close to impossible for most people to remember, this would lead to that a lot of people would write the password down which would just open up to the possibility of an authorized person to get access to the password. According to Adams & Angela (1999) users lack the security knowledge of what defines a strong password and most organisations use user-generated passwords instead of system-generated passwords which puts the responsibility of creating the password on the user which can lead to weaker passwords because of users lack of security knowledge combined with the problem of users remembering strong passwords could lead to weaker passwords. This leads to a trade-off on passwords, either the passwords are system generated and the users might have a problem remembering them which could lead to them writing it down or the users get to select passwords which could lead to passwords that are considered weak because the users might choose passwords that are weak or directly connected to them.

3. Problem definition

The purpose of this report is to find out if bad password habits can be changed by receiving training on what defines a strong password. The report will focus on students at the University of Skövde but the general principles should be applicable in other contexts as well.

3.1 Question formulation

The questions this study aspires to answer are:

What problems exist in password management today? And are these only limited to certain user groups?

What defines a strong password and is bad password management related to gender, education level and the usage of passwords in daily life?

Can you change bad password behaviour by giving a lecture on the matter?

The first question will be answered by doing the literature study and by conducting interviews that are more deeply described in section 4.1 *Interviews*. The second question will also be answered during these interviews. The third will be answered by conducting the experiment described in section 4.2 *Experiment* and with the interviews from 4.1 *Interviews*. The main focus of the report will be on the experiment and not the interviews.

3.2 Motivation

With the study of leaked passwords in section 2.*3 Studies of leaked passwords* it's proven that a lot of people neglect the guidelines provided by most websites on how to make a strong password. Are guidelines the wrong way to tackle the problem? Is it better to give a small lecture on the matter? Looking at the survey performed by Zviran & Haga (1999) bad passwords existed even at the Department of Defence; one can hope that they have a stricter policy now. With the study performed at the University of Maribor by Taneski et al. (2014a) a lot of the answers they got were "No answer" which could be because of the method they used to collect the data. The usage of two separate groups one with the lecture and one without the lecture and the collection of passwords through an experiment that will be used for this report will then be compared to the results and method from Taneski et al. (2014a) results.

3.3 Objectives

There are some objectives for the study that need to be fulfilled:

1. The first step will be to do a literature study to get more background on the subject before designing the questions and conducting the interviews.

2. The second step will be to write the background and method from the data collected at the literature study.

3. The third step will be to design the questions for the interviews conducted.

4. The fourth step will be to conduct the interview to get more information if bad password habits exist for several different types of groups of people.

5. The fifth step will be to analyse the data from the interview and design the lecture for the experiment from data collected from the interview and the literature study. The websites for the experiments will also be designed from the data collected at the interviews.

6. The sixth step will be to conduct the experiment using the two different experiment groups.

7. The seventh step will be to analyse the data collected at the experiment.

The first and second step will try to give the report a better foundation and also help to interpret the data from the other steps. Steps 3-5 are there to help with the qualitative

interviews that will be the foundation for the experiment, the interviews will be used to help understand how the users think when they pick the password for the design of the lecture. Steps 6 and 7 are used for the experiment part of the report. *Figure 1* is a diagram of how the questions will be answered with the steps.

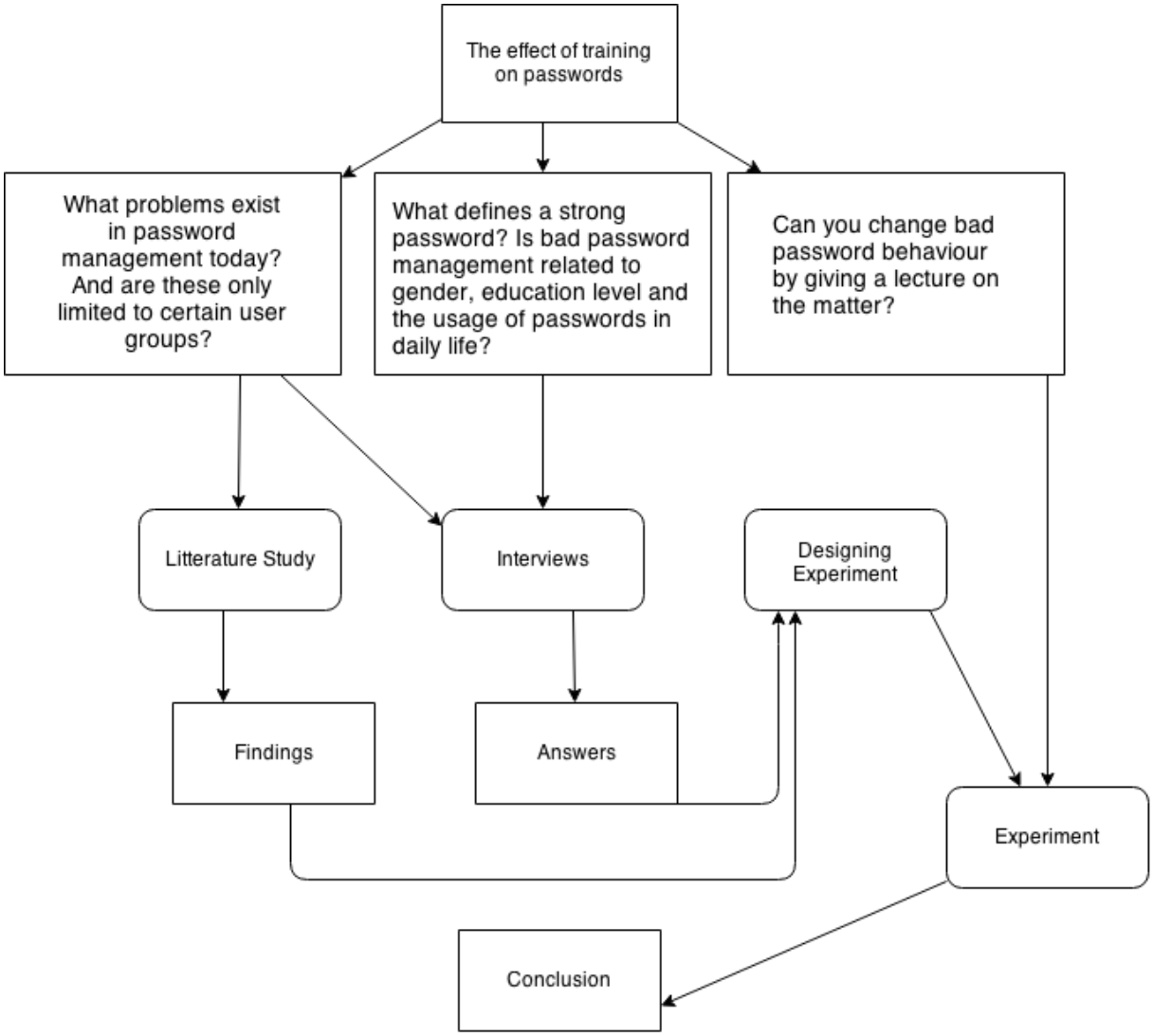


Figure 1 Map over objectives

3.4 Boundaries

The boundaries for this report are that in the experiment part it will only focus on students at the University of Skövde. The only difference with the students that will be considered is education type, no focus on the age of the participants. The report will only focus on what kind of passwords users pick on websites, application passwords will not be considered. The websites designed for the experiment will not take design into too big of a consideration other than they will try to resemble actual websites of the purpose they are designed for. The interviews will be performed on people not studying at the university.

3.5 Study of similar work

This section will look at two similar studies performed at the University of Maribor in 2014 and the Department of Defence in 1999. Both studies looked at password strength, frequency of changing passwords and password composition. They both gathered the data using a questionnaire.

3.5.1 University of Maribor Study

In 2014 a study was performed by Taneski et al. (2014a) at the University of Maribor where they used an online questionnaire to determine the characteristics of textual passwords. They had a group of 33 undergraduate students at the Faculty of electrical engineering and computer science at the university conduct the survey. The survey consisted of two phases, in phase one the students performed the questionnaire without any education in password security. After the first phase the students attended a lecture designed by the authors which consisted of topics on how to create a strong password and how to manage them. After the lecture they had a two week period before they contacted the students again to ask them to perform the second part of the survey. They then compared the data from the two different sections the questions on the survey included:

- Average password length.
- Password change frequency
- Password memorability and write-down.

The author found that students had an improvement in the characteristics of the passwords in the second phase. And that the overall average password length used by the students had increased. They also found that from the first phase that a lot of the users never changed their passwords since their first and even despite the lecture on the importance of a frequent password change they did not change it on phase two.

3.5.2 Department of defence

In a study performed by Zviran & Haga (1999) which did a survey of password security the subjects for this survey were computer users at the Department of defence in California. The questionnaire was distributed to two thousand users, 49.9% (979) answered the survey. The authors identified that according to several sources that an acceptable password should have between 6 and 8 characters. They found that 47% of the respondents to the survey had a password shorter than this. Only 14,1% of the users had a password that consisted of 8 characters or more that is todays suggested standard by many guidelines. They found that 79,6% of the users also never changed their password and that 14.9% changed it on an annual basis, only 5,5% changed it several times a year. 80% of the users conducting the survey also had a password that only consisted of alphabetic characters. 78% of the users based their password on a combination of meaningful details, like the data they protected or personal information.

The author believes that the findings can be explained by the fact that a user picks their password before knowing the type of data it's going to protect and that new users often lack information security consciousness. They found a correlation that the frequency of changing passwords is affected by the level of data it protects.

It's important to take into consideration that this survey was conducted almost 16 years ago and the usage of services that use a password has gone up, which could result in better passwords since many websites give tips about strong passwords and force users to pick a password with certain amount of characters and mixtures of upper- and lowercase letters, numerals and special characters. The problem is that with more passwords the user might reuse the same password as discussed by Furnell (2007) which can bring a greater security threat if one password is leaked.

The result from these two experiments will be compared to the findings in this report in the conclusion.

4. Method

There will be two methods used in this paper; one is a qualitative interview with different people with varying education, work, age and gender. The other part will be an experiment where users will pick passwords for three different services; the participants will be divided into two different groups. One group will receive a lecture on what defines a strong password, information about password management and some information on how to create a passphrase. The other group will just pick their passwords without any form of extra education given on the matter other than the education they might already have received from their personal life. The passwords will then be analysed to see if they are considered strong or weak passwords based on the parameters collected in the pre study and their entropy value will be calculated. The entropy value will be calculated using Shannon's formula (1948) and by using the *zxcvbn* algorithm by Wheeler (2012). The goal of the experiment is to see if the group with a password education picked stronger passwords to see if the pattern of bad passwords can be changed with a simple lecture on password information. According to statistics provided by Universitets- och högskolerådet (2015) there were 19,356 students at the University of Skövde fall 2014, due to time restrictions a larger sample group of the students will not be used. A sum of 40 students will be used for the experiment, 20 of them will get the education and the remaining will not receive any education on the matter. They will be chosen randomly on campus were the experiment will be performed. If there is trouble finding students on campus willing students studying at the university will be contacted over social media as well.

4.1 Interviews

The interviews will be conducted with people of different educations, ages and genders to see if the problem exists all throughout the population. The questions included will not be yes or no questions but so called *open questions*. The specific questions will be discussed in part

4.1.2 Questions. The interviews will be conducted over Skype and in person with a written transcript for each interview that will later be used to analyse the given answers.

4.1.1 People included in the interview

The interview will try to include a wide variety of people from different educations, ages and genders. To try to cover all types of people the participants will be:

Gender	**Age**	**Password habits**	**Why**
Female or Male	50-65	Works with passwords on a daily basis in line of work.	Would be interesting with the older generation that was not born when passwords where used each day and see how they think about passwords and password management.
Female & Male	20-30	Works in a line of work were passwords are not a part of daily work.	The newer generation that use passwords since they were young. But don't use it for their work but only for personal use.
Female & Male	20-30	Student with education in information technology (IT).	Students with an education in IT and information security. And presumably have a more educated stand on the matter.
Female or Male	15-18	Only uses passwords for personal media or services.	Younger generation that don't use passwords presumably for other than personal services.

Table 2. Interview selection

4.1.2 Questions

The questions are designed as described by the book Intervjuteknik Häger (2001. P 57) who suggest the use of "open questions" which is the opposite of "closed questions" which are normally answered with a *Yes* or *No* answer. Open questions should start with for example the word "*Why, how and what*" that will help with getting a longer answer from the person being interviewed.

Questions are also good if they give room for a follow up question this will help with getting a better quote from the interviewee (Häger, 2001. P 60). Directional questions should also be avoided to get a correct answer from the interviewee (Häger, 2001. P 63).

What do you consider a good password?

This question will help to find out what the user considers to be a strong password, what do they define as a strong password? Does it meet the criteria earlier discussed in the background part of the report?

How often do you change your password and why?

The information technology (IT) security company Symantec (2010) performed a study back in 2010 with a series of password related questions. There were 446 respondents to the survey; one question from the survey was "How often do you change your password?" 4%

(20) answered once per month, 17% (78) changed it quarterly. 63% of the respondents changed their password "Not very often" which could be an indication that passwords are only changed on a need to basis. This question will also be asked during the interviews, with a follow up question of "Do you change it more often on certain services?".

Do you reuse the same password for several websites?

As the study by Florêncio & Herley (2007) showed most users have a set of 25 passwords and reuse these on different services, is this because the user does not have the cognitive function to remember more passwords? If the user does not reuse the same password for several sites a follow up question could be "Is there a third party tool you use for this or what kind of method do you use to remember?"

Do you trade password strength for your own convenience for picking an easy to remember password?

A big problem with picking a strong password is that with more characters it can be harder for the user to memorize. That is why this question is proposed to see if the users overlook strength of the password for an easy to remember one. If this is a problem it would be good to include the section from Microsoft's guidelines about how to create a strong password.

Do you think the design and purpose of a website motivates when you pick a password?

Users seem to pick stronger passwords for certain websites that store more sensitive data, is this in any way connected to the design of the website. Or is it connected to the data they store on the website itself and the purpose the website is made for? A follow up question could be "What parts of a website makes you pick a stronger password for this type of website?". The answer from this question could help with designing the websites to look more secure even though this is not the goal of the report as mentioned in the boundaries section.

Are you aware of online crimes? What type of crimes do you know of and do you consider them when using the internet?

Are the users aware of that online crimes even exist? And do they ever consider them when using the internet.

Have you ever been the victim of any online crimes? (Hacking, identity theft)

With the high rise in crimes committed online according to BRA (2014) it is important to know if the interviewee has been the victim of any crime and maybe have changed password behaviour because of this.

What do you know about password guidelines and do you take them into consideration when registering an account?

As previously mentioned websites often give guidelines on how to pick passwords, are people aware of these guidelines and what they include or do they just scroll through them?

Does your workplace have an IT policy and do you fully understand and follow all the parts of it?

Lots of workplaces force their users to sign an IT-policy which dictates how the IT infrastructure can and cannot be used. This question is asked for the purpose of getting a understanding if the interviewee understands the policy and follow it. This question will have another variation when asked to people not currently working, like the student under 18 years old, it will not be asked to the interviewees not working with a job were passwords are required.

4.2 Experiment

The experiment will have two different test groups, one that will get an education based on information collected from the literature study and interviews. The other group will perform the experiment without being given the education on security and management. The second group will function as the control group while the first group given education will be the experimental group that the lecture will be tested against. Wohlin et al. (2012) describes this setup as a "one factor with two treatments" where the two treatments are the new and the old method. It's important that both groups perform the same experiment and there is no difference between the two groups other than the treatment given by the lecture.

4.2.1 Experiment Design

Three websites will be created, one that looks like a bank, one that looks like a pizzeria and one that will look like an email website. The bank website should try to encourage the user to pick a stronger password then the other two websites. The usage of three websites will also allow the study if the users pick the same password for the three websites, as well as a deeper understanding of how they create their passwords. What type of structure do they use? Is there a pattern to the passwords? The websites will be created using pre-fabricated templates that get changed to fit the purpose. The websites are created with basic Hypertext markup language (HTML) code, cascading style sheets (CSS) and JavaScript, the username and password that the user register will be posted to a file on the webserver without any form of encryption of the password, the file will however not be accessible by other people. The users will be identified by being given a username that will be connected to which group of the two groups they belong to. The passwords will then be studied to see if the group with the education picked a stronger password then the group without the education.

The experiment will take place in the University of Skövde and the experiment groups will consist of students currently studying at the university. They will be in a room somewhere at the university and only be told that they are supposed to pick passwords for the services with no further information about that the password should have any special characteristics like being 8 characters long.

As previously mentioned the passwords will be graded by calculating the entropy using two different formulas that will be more deeply described in sections *6. Password entropy* and *6.1 Zxcvbn algorithm by Dropbox*.

4.2.2 Website design

The three websites that were designed using pre-existing templates that were subjected to some design changes to make them fit the experiment more. All websites had a homepage that was the first thing the user saw when entering the website. Each website also included a register page that could be easily accessible from the homepage. The user only had the ability to enter a password and a user name on each website. The first website has an option were the user could inform that they are currently studying a "computer education". The users were first shown a disclaimer page and then sent to the first website which was the banking website. After submitting the form with the username and password they were automatically sent to the next page until they complete the experiment. Print screens of the websites will be included in the *Appendix 2 - Websites*.

4.*3 Validity*

According to Berndtsson et al. (2008) it is important to consider the various threats to validity and reliability to the project. Not having an appropriate account for the threats may lead to lower quality of the project or that people question the overall quality. This section will contain information about the threats against the validity of both the experiment and interviews.

According to Wohlin (2012) there are four different types of validity, *Conclusion validity*, *Construct validity*, *internal validity* and *external validity* and Wohlin explains them like following:
Conclusion validity is connected with problems that affect the ability to get a correct conclusion from experiments. An example is 'fishing' for results by trying to lead the results to a specific outcome or having a bad experimental setting which could interrupt or affect the experiments result. Internal validity is aspects were the result can be affected by another factor. Construct validity is connected to the design of the experiment and faults against it. External validity threats limit the ability to generalize the results in an industrial practice according to Wohlin (2012).

4.3.1 Threats against experiment

It's important to think about the construct validity when constructing the experiment One big problem can be "Hypothesis guessing" which is according to Wohlin (2012) when the user tries to figure out what the purpose of the experiment is and try to perform it based on its hypothesis. Another threat against the experiment can be if the user feels threatened to showcase what kinds of passwords they would typically pick and instead pick a random password. These threats can be eliminated by telling the user to pick a password based on the characteristics they normally use for a password, but not a password they already use, also to pick a password they seem fit for the service they are registering on. All the students participating in the experiment must get the same information about how to perform the experiment and given no extra information before so the results cannot be bias in anyway. To also try to remove the threat of external validity the experiment will be conducted in a room with a closed door to remove sources that can disturb the experiment such as other people. To

not fish for any results which will affect the conclusion validity the websites for the experiment will all be the same for both groups conducting the experiment.

4.3.2 Threats against interview

For the interviews the two that are important to keep in mind are *Construct validity* which is that the interviewee interprets the question in another way than it's meant to be asked. The other one is *Internal validity* that one factor is affected by the third factor. This will be taken into consideration if any of the interviews pose this issue.

The subjects might not fully understand the question that is being asked which would count as a *Construct validity*. This can be avoided by designing the questions to be easy to understand even without any technical knowledge about passwords or general IT education. The other issue might be that the users can be affected by an internal validity because of something related to the IT-policy at the work place which would be linked to *Internal validity*. Also during the interviews it's important that no questions are asked in such a way that they will fish for an answer from the interviewee. The interviewees will also be conducted on a time of the day picked by the interviewee so they don't feel tired or not fit enough to answer the questions. The selection of the interviewees is not random but they are picked from people that fit into the area know or referred by other people (i.e the male under 18 subject for example is a friend of a family member). This should not affect the outcome of the interviewees more than that they might be more open and answer the questions more truthfully because they might feel some form of trust. The interviews will be used to design the experiment and lecture as previously stated so the selection should not affect this in anyway. The questions are also short so the interview will be fast so the interviewee does not get bored or exhausted during the interview.

4.3.3 Ethics

With the handling of people's password ethics aspects are an important part, users are informed during the experiment that their passwords will not be displayed in the report. Only characteristics of the passwords will be discussed and measured in the final report. All the participants of the experiment will read through a disclaimer page at the start of the experiment that tells them of this fact. The passwords themselves are only studied by the author and no one else will see the composition of the passwords in a textual form, they are also being stored on a computer and can only be accessed by using the said computer. The participants in the interviews will not be named by name as well and the participants of experiment will be given a username to use.

5. Good password versus strong password

What gets defined as a strong password does not automatically define a good password. Even if the password fulfils all the qualities identified earlier as a strong password, it is not defined as a good password if the user cannot remember it.

According to Taneski et al. (2014b) a password with higher entropy will make it more difficult for the user to memorize. After analysing several papers Taneski et al. (2014b)

identified several different methods for creating memorable passwords these are *Cognitive passwords, associative passwords, passphrases & mnemonic-based passwords*. A *Cognitive password* is a authentication mechanism that gives the user a series of question selected by the user to answer as authentication instead of a textual password. Which can be easier for the user to remember than strong textual password, a possible issue is that it will take the user longer to authenticate. An *associative password* is an alternative where the user will be given a single-word and then type out whatever the user associate with that word. A *passphrase* is a set of words that together form a long password, it could be easier to remember for the user and more difficult to guess for someone who does not know the passphrase. Passphrases have also shown to be more resistant to brute-force attacks as well. A *mnemonic-based password* is an alternative to a passphrase were only the first letters of each word form a sequence of what looks to be random characters, but for the user will be easy to remember (Taneski et al. 2014b).

This will be taken into consideration when designing the lecture for the experiment group, it is important that the lecture does not just focus on what a strong password is but also give the student information on how to create the strong password. Passphrases will be explained during the lecture.

6. Password entropy

According to Ma et al. (2010) entropy is a quality indicator for passwords and high entropy can give a better quality. The entropy only establishes the boundary for the amount of guesses needed to crack the password. There are several ways of cracking a password according to the article; one could use dictionary words, applying different variations to the dictionary words or by brute forcing. The article mentions that the quality of a password depends on the time it will take to find the right match by using these methods.

Entropy will show the passwords variation expressed as bits. It's calculated by a formula provided by Shannon (1948). Where C stands for the password cardinality which is the amount of different elements in a set, by using the values in *Table 3* for the cardinality and L stand for the length of the password and the formula is:

$$E = \log_2(C^L)$$

Symbols	Cardinality
a-z	26
A-Z	26
0-9	10
Special Characters e.g. +,\/`~!@#$%^&* ()-_=;:'",<.>?	30

Table 3: Cardinality

For example a password that consists of 8 characters and upper- and lowercase characters and numbers will give the equation:

$$E = \log_2(62^8) = 47.6$$

By using this formula it's evident that increasing the length is more important than increasing the cardinality of a password. If we use the formula to test this by using the two passwords: "A13F=;54d!" and "IhaveAOldHorse123". The first password if we use the table above will have a cardinality of 94 and a length of 10, while the other password has a cardinality of 62 and a length of 17.

By calculating this we will get entropy of 65.5 bits respectively 101,2 bits which shows that a longer password should take the computer longer time to try out and calculate. But the practice that is being taught today is to increase the cardinality and not the length of passwords. There is a Munroe (2011) quote about how passwords are designed today that fits well:

"Through 20 years of effort, we've successfully trained everyone to use passwords that are hard for humans to remember, but easy for computers to guess." - Munroe

By using the resources from the website *passwordstrengthcalculator.com* we can calculate how long it would take a computer to try all combinations to brute force the password. Take notice that this value will be the maximum amount of time a computer will need to test all possible combinations.

By using a supercomputer which has a lot of allocated power for calculation, the password "A13F=;54d!" can be cracked in a maximum of 9 minutes according to the website. By using a normal desktop computer it would take a maximum 125 days for the same password. The other password "IhaveAOldHorse123" which has more characters but a lower amount of cardinality would take the same supercomputer a maximum 937,243 years and a desktop computer an unimaginable amount of time. This is however by calculating the entropy value

using only Shannon's formula by for example a dictionary as explained in the following sections this value could be lowered a lot.

6.1 Zxcvbn algorithm by Dropbox

In 2012 the *Dropbox* team developed an algorithm they call *zxcvnbn* Wheeler (2012). The algorithm uses the Shannon formula as explained in the previous section to calculate entropy, combined with checking the words against different dictionaries such as common English words, movies and television shows, spatial patterns, most commonly used passwords, first names and surnames and sequences of letters. It takes the word and split it up in different patterns and estimates the entropy for each pattern. Then the sum of the entropy is calculated by adding the different patterns together. It always uses the lowest of the entropy summations as the estimate. By looking at sequences of words and patterns in the password structure it can often calculate a lower entropy value than what the normal entropy value would be by using the Shannon formula. If the algorithm does not find a pattern in the sequence of the password it will be marked as a random string of characters and have to be brute forced by guessing all the characters.

By taking for example the password "UmustP4ssG0ToCollect132GHJ" it starts of by having to brute force the letter *U* because this does not exist in a dictionary giving this an entropy value of 6.5. It then moves on to finding the word *must* in a dictionary giving it the entropy value of 7.9. The word *P4ss* is a commonly used word for passwords and is found in the algorithms dictionary of common passwords and is given the entropy value of 7.1. The word *G0* is detected as the common word *Go* but with a 0 instead of an *O* and is also matched and given the value of 7.8. The two following words *To* and *collect* are also dictionary words and are given the values 2.5 and 12.4 respectively. The sequence of numbers *132* is detected as a sequence and given the entropy value of 9.9 and the last sequence of letter *GHJ* is recognised as a spatial pattern and given the value of 11.7. The values summed up together give the entropy value of 66.3, when using the Shannon formula to calculate the same passwords entropy the calculation will be:

$$E = \log_2(62^{26}) = 154.8$$

The *zxcvbn* algorithm is evaluated in a paper written by de Carné de Carnavalet (2014) which tested several similar password checkers the author explains the strength of the algorithm by saying:

> *"Zxcvbn considers the composition of a password more thoroughly than all other checkers in our test, resulting into a more realistic evaluation of the complexity of a given password. In this regard, it is probably the best checker."* – Xavier de Carné de Carnavalet 2014.

He also points out some weaknesses with the using the algorithm such as reversing a word gives a higher entropy value than it should give because the algorithm can't detect reversed words. It also only uses the English dictionary as its only common word dictionary (though

there is a German version that is made by an independent developer); also the dictionary seems to be lacking a lot of words that should be regarded as common words.

The result gathered in the experiment will be ran through the Shannon formula and the *zxcvbn* algorithm to give an perspective of both the *raw entropy value* collected from the Shannon formula and the entropy value from the *zxcvbn* algorithm to give an value that matches how an attacker would detect the entropy.

The version of *zxcvbn* program used for evaluating the experiment had the commit number "0064153c1b" on Github for reference.

6.1 Understanding password cracking

Most websites will not store passwords as plain text but will use password hashing to turn them into a large set of letters and numbers. Akins (2012) explains how to test against a hash, by taking a dictionary of words you can hash all these words by using the same hash the website has and then try to find a matching hash to detect a password. He tried an attack against the 6.5 million leaked passwords from the LinkedIn breach and by using his list of the top 7184 passwords he found 3854 matching hashes in 14 seconds. Then he tried matching them against an English dictionary containing he was able to recover 22,572 passwords in 15 seconds. All using a *HP pavilion g6* computer with a Intel core I3 2 core processor running at 2,3 gigahertz and an solid-state drive to be able to load the passwords and hashes faster. By doing additional testing against a wordlist of 18 million words from several different languages he was able to crack 390,000 of the hashes in 2 minutes and 9 seconds. By trying to brute force a list you generate all possible password combinations by testing all characters against the hash one by one. By letting his password cracker try to brute force for 48 hours he collected 2 million passwords just by testing different combinations. He states that if an attacker has access to more computer resources they could distribute cracking between these computers and speed up the process. By using what is called a *rainbow table* the cracking could be done a lot quicker, by using a *rainbow table* hashes are partially generated ahead of time and this saves the time it would take to start over again. Take in mind that the hash the website is using often has to be known by the attacker to be able to crack the passwords quicker, if this is not known some cracking programs have a big list of known hashes which it tests against. To protect against testing against hashes and slow down the process an administrator could use a technique called *salting* were they add extra characters to the hash to make the testing take far longer time than it should do. The strength of how strong a password is often measured by its entropy value.

Most websites would not allow a user to do extended requests against their web service and this would only work against an offline version of their database. But if an attacker gets access to a user's email and password by example studying the LinkedIn breach, this password could be tried against other services. Because as previously mentioned a lot of users reuse the same password over several services and are often very bad at changing passwords.

7. Interview Responses

After conducting the interviews the conclusion was made that most users consider an easy to remember password more important than a password with a high cardinality and a long length. Also none of the users changed their password on a regular basis and all of them reuse passwords for several services. It also seems like they pick easy to remember passwords rather than strong passwords and only the two male's and the female with an IT-education use certain stronger passwords for certain services. All the users seem to be aware of online crimes and some have been affected by online crimes but it seems like it has not changed their behaviour of passwords. Most interviewee that has an IT policy or guidelines at their workplace seems to know about it and understand most parts of it as well. The answers are entered into a diagram below with an X to mark out that they mentioned this particular answer in their response.
The responses during the interview will be used for the design of the lecture that will be given during the experiment.

	Good password = high variance	Good password = Easy to remember	Good password = is long	Change password	Reuse passwords	Strength for convenience	Purpose changes passwords	Aware of E-crime	Knows of Guidelines
Female age 55		X		(X)	X	X		X	X
Man age 16	X	X			X	X		X	X
Female age 24 it edu	X		X		X	X	X	X	X
Male age 23 it edu	X		X		X	X	X	X	X
Female age 25		X			X	X		X	N/A
Male age 28	X	X	X		X	X	X	X	N/A

Table 4: Interview responses

The only person that changed their passwords on a regular basis was the female over 50+ that uses passwords in her daily line of work. She mentioned that her workplace forces her to change every 7 months, but that this does not apply to her personal passwords. After

conduction the interviews it's obvious that users need to know about the dangers of reusing passwords, what a weak password is, what defines a strong password and how to create an easy to remember strong password. This will be used as the main focus for the lecture created for the experiment part.

8. Experiment Results

The experiments were conducted during a three week period at the University of Skövde. Students who attended at the school were contacted using social media and were approached at the schools facility's and asked to conduct the experiment. Each student was assigned with a username that signified which of the two experiment groups they were a part of. This fact was not known by the students themselves but was used by the author to be able to tell the different groups apart when looking at the results of the experiment. Each experiment was performed individually and the students were told to use passwords they seemed fit for the service they were shown. The three websites were designed to give a serious look as previously mentioned the websites created were a bank, pizzeria and an email service. The students where showed a disclaimer page at the start of the experiment with information about that they should not use the same passwords for the experiment as they use in real life but a password that resembles the type of password they would deem fit the kind of service. The lecture group were shown a lecture that included information about:

- What is a weak password?
- What is a strong password?
- Information regarding password management
- Information about what types of passwords people use (gathered from previous password leaks) but also information about how weak passwords affect services (breaches).
- Information about passphrases and how to create a sufficient passphrase.

The lecture was shown as slides with a voice track included and spanned over 3 minutes and 24 seconds. After students were given the lecture they conducted the experiment. The experiment was the same for both students given lecture and the students not given the lecture. The lecture slides can be viewed in *Appendix 4 – Lecture* this is however not with the included voice track that was played together with the slides for the students given the lecture.

After the lecture was performed the results will be calculated, factors such as average value, entropy value by Shannon's formula and by using the *zxcvbn* algorithm mentioned earlier.

The average password length between the two groups was higher in the group given education. With an increase of up to ~4 more characters on the banking service.

The average passwords given by users differ a lot because some users used 25+ long passwords and some uses only used passwords with the length of 4 characters. As the result shows a lot of users in the group without lecture also choose passwords that are longer than

the average password length previously mentioned in papers which could be contributed to that some users tried to "figure out" the experiment.

By using the formula by Shannon (1948) for calculating the entropy of the values the Swedish characters "ÅÄÖ" which are used in some of the passwords are reformed into A and O instead to help with the calculation because otherwise when using the *zxcvbn* algorithm they will be used as special characters because they don't fit into the alphabetical pattern system used by the algorithm which only uses the English alphabet of 26 characters and not the Swedish alphabet with 29 characters and instead assign them as special characters. Most websites would not allow the Swedish characters as well; this should have been considered when designing the experiment and not be allowed but during the creation of the experiment the author failed to notice this problem until after the whole experiment was conducted.

As mentioned earlier the formula used for the calculation of the password entropy is:

$$E = \log_2(C^L)$$

Where C stands for the cardinality of the password and L stands for the length of the password.

The results from the entropy value calculated using the formula above on users given the lecture is presented in Table 5:

	Bank	Pizzeria	Email
Subject 1	48,3	74,1	53,6
Subject 2	123,9	123,9	123,9
Subject 3	101,2	117,4	130,5
Subject 4	150	101,7	131
Subject 5	123,9	78,3	113,1
Subject 6	97,9	97,9	91,3
Subject 7	89,3	117,4	104,4
Subject 8	83,4	77,4	95,3
Subject 9	83,4	95,3	71,5
Subject 10	107,2	113,1	95,3
Subject 11	108,1	108,1	108,1
Subject 12	75,1	59,5	53,6
Subject 13	195,7	114,4	71,8
Subject 14	120,8	123,9	137
Subject 15	119,1	59,5	119,1
Subject 16	58,7	58,7	47,6
Subject 17	47,6	47,6	47,6
Subject 18	136,9	125	101,2
Subject 19	68,4	53,6	46,5
Subject 20	83,4	119,1	95,3

Table 5: Entropy values for group given lecture

The highest recorded entropy value was 195,7 bits and that password was 30 characters long and included lower- and uppercase characters, numbers and special characters. The lowest recorded entropy was 46,5 bits was 9 characters long and used only uppercase letters and numbers. The website with the highest average entropy value was the banking website which was also the hypothesis of the author. The email had the lowest average entropy value of the group which was a bit odd because the hypothesis was that this would rank higher than Pizzeria. This could also be connected to that this was the last website of the experiment, and the websites should have been displayed in a random order to counteract this problem. The difference however was very slim and could also not mean anything.

The result from the entropy value calculated on the group not given the education is presented in table 6:

	Bank	Pizzeria	Email
Subject 1	113,1	107,2	83,4
Subject 2	52,2	89,3	78,3
Subject 3	41,7	47,6	47,6
Subject 4	42,3	52,2	41,7
Subject 5	41,7	77,4	47,6
Subject 6	41,7	47,6	47,6
Subject 7	47,6	41,7	59,5
Subject 8	53,6	53,6	53,6
Subject 9	85,5	79,8	85,5
Subject 10	119,1	19,9	125
Subject 11	104,4	104,4	104,4
Subject 12	71,8	41,7	35,7
Subject 13	156,6	78,3	59,5
Subject 14	59,5	71,5	47,6
Subject 15	47,6	53,6	53,6
Subject 16	65,5	47,6	59,5
Subject 17	22,8	22,8	22,8
Subject 18	83,4	65,5	91,3
Subject 19	65,5	65,5	23,3
Subject 20	57,2	78,3	82,6

Table 6: Entropy values for group without lecture

The highest recorded value for this group was 156,6 bits and this password consisted of 24 characters with upper- and lowercase letters, numbers and special characters. The lowest entropy value recorded was 19,9 bits of entropy and this was a 6 character long password only numbers.
The average entropy value for this group was also higher on the bank than the other two services, it however had a higher value for the email than for the pizzeria, but the difference was too low to signify anything.

The average entropy value for the group given the lecture is higher than for the group not given education on the matter. This is a result of more cardinality and a longer password length.

Looking at users that reused the same password over two or more services the group with the lecture had 2 out of 20 users that reused the exact same password and one of the two just added an extra character to the password which could also help an attacker get the password for other services. The group not given education had four users that reused the same password over two or more services with three of them having exacting matching password over at least two sites and two of them had the same password for all three websites which non in the group given education had. The group given education also had a higher cardinality with 43,3% of the 60 passwords supplied by 20 users in this group used special characters. In the other group this number was 25% of the 60 passwords used special characters.

By using the *zxcvbn* algorithm developed by the Dropbox team which calculates the entropy using the Shannon formula and then divides the passwords into sections and looks for patterns such as common used passwords, dictionary words, sequences of characters, spatial patterns and dictionaries of personal names and movies which could be tools used by an attacker when trying to brute force a password the entropy numbers are reduced in the group given lecture to the values presented in table 7:

	Bank	Pizzeria	Email
Subject 1	31	59,5	38,5
Subject 2	95,9	95,9	95,9
Subject 3	78,3	87,3	94,7
Subject 4	115,3	74,4	105,3
Subject 5	88,8	55,4	90
Subject 6	63,5	63,7	60,4
Subject 7	68,5	82,1	73,5
Subject 8	64,1	58,6	76,2
Subject 9	60,8	67,3	48
Subject 10	78,6	80,3	62,8
Subject 11	83,2	83,2	83,2
Subject 12	60	42,1	39,5
Subject 13	140	85,1	48,8
Subject 14	93,3	118,5	109,2
Subject 15	90,7	40,5	90,7
Subject 16	40,2	34,1	27,3
Subject 17	38,3	37,6	38,2
Subject 18	102,2	100,4	73,3
Subject 19	51,1	37,6	35,7
Subject 20	63,4	95,9	79,2

Table 7: Entropy values for lecture group using *zxcvbn* algorithm

This for example lowers the entropy value of the password that had 195,7 bits using the Shannon formula down to 140 bits. The password that had the lowest entropy before of 46,5 is now lowered to 35,7 bits. In some cases applying the formula took away up to 35 bits.

When running the same algorithm against the group not given a lecture the values are reduced to the values in table 8:

	Bank	Pizzeria	Email
Subject 1	87,3	84,2	64
Subject 2	37,9	67	60,6
Subject 3	26,5	32,6	30,2
Subject 4	52,4	35,7	32,8
Subject 5	26,2	59,4	35
Subject 6	27,5	31,7	33,3
Subject 7	33,8	26,5	40
Subject 8	35,1	35,1	35,1
Subject 9	66,7	65,5	66,7
Subject 10	92,3	9,7	98
Subject 11	77	82,7	75,5
Subject 12	46,1	26,4	21,4
Subject 13	112,4	53,3	42,4
Subject 14	40,8	54,2	34,4
Subject 15	33,5	38,3	40
Subject 16	49,7	32,5	41,7
Subject 17	13,9	13,9	13,9
Subject 18	62,7	49,1	66,7
Subject 19	47,9	47,9	11,6
Subject 20	43,4	49,8	61,6

Table 8: Entropy values for group not given education using *zxcvbn* algorithm

Here the only one of the entropy values stays over 100, this is the same password that has the highest entropy value when applying only the Shannon formula as well, then it had a entropy value of 156.6 after the *zxcvbn* algorithm the value is 112,4 bits. The lowest recorded entropy value after using this algorithm is 9,7 bits of entropy, this value was 19,9 bits before applying the algorithm.

By using the formula the password entropy value for users given education is lowered which is displayer by the diagram below. The values are the average values of all the values combined.

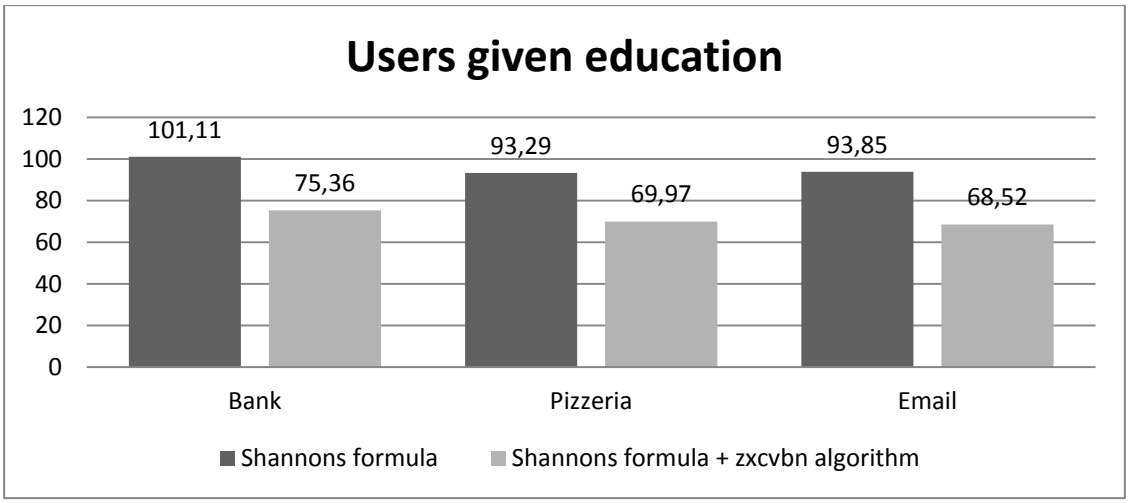


Diagram 1 Users given education

By using comparing the two entropy values for the group not given education the average values are:

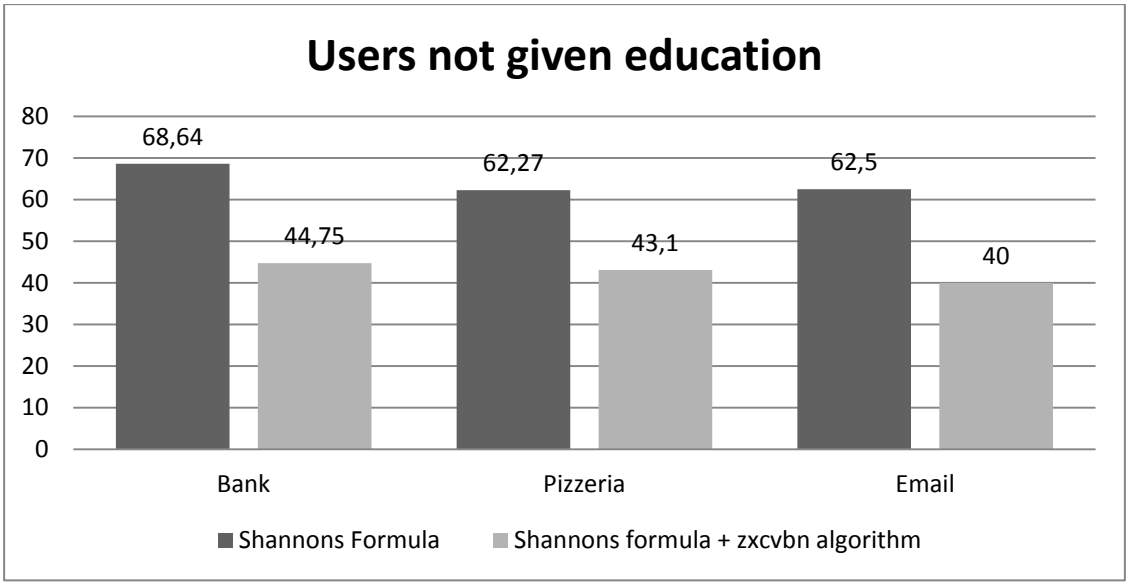


Diagram 2 Users not given education

The results show that the group given the education had a higher entropy value than the group not given the education. What has to be taken into consideration looking at the data from the *zxcvbn* algorithm is that the dictionary it uses is an English dictionary and will not include Swedish words which have been used by most of the users that included passphrases in their passwords.

The values for the group given education are higher in both aspects, with just the Shannon formula and by applying the *zxcvbn* algorithm. Diagram 3 shows a comparison between the two groups.

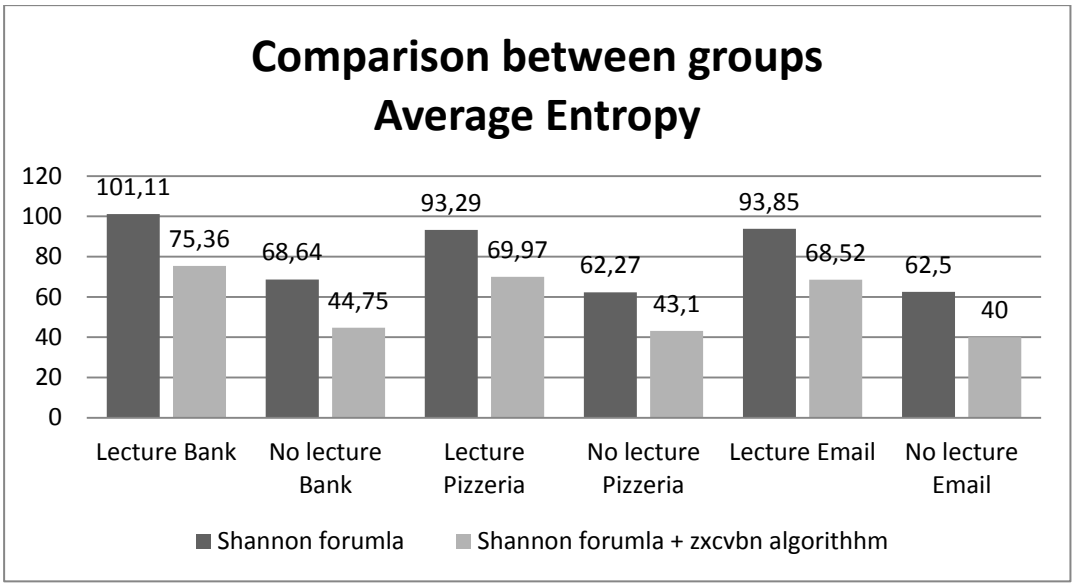


Diagram 3 Comparison between groups

Another variable recorded was if the users had an IT education, this was submitted when registering on the first website. It was not specified what an IT education implied for the users which should have been included in the experiment. It should have referred to users with an education in some sort of area that included computer security but students which study for example computer game development at the university may have not understood this. All users who selected IT education will however be treated the same in the following calculations.

A comparison was made between the groups on the bank website to see if the users with IT education had stronger entropy than the users without IT education; the results are presented in Diagram 4.

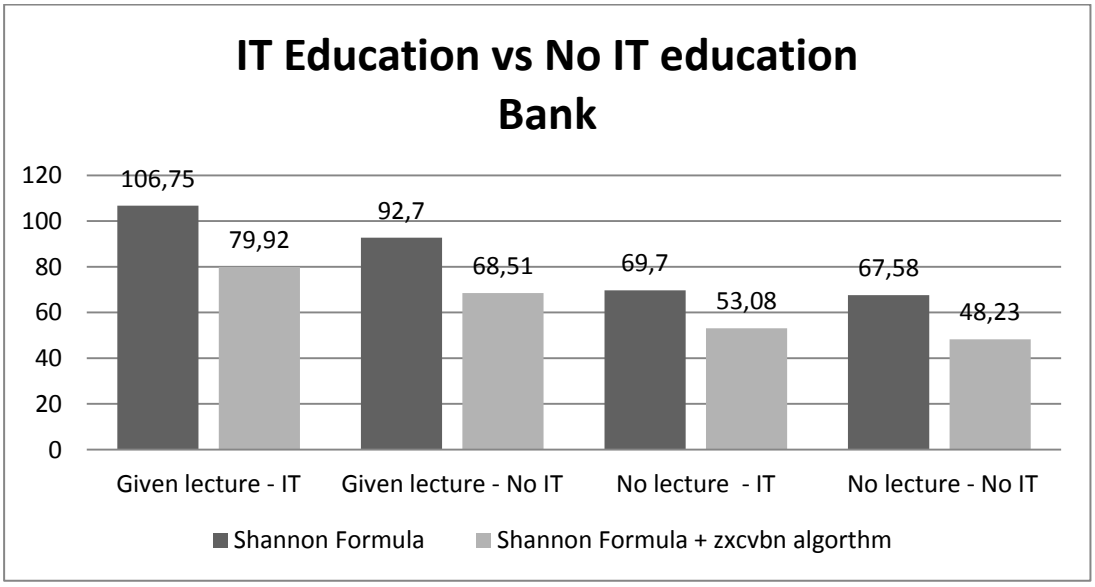


Diagram 4 Comparison on bank website, IT education vs No IT education.

The first column shows the users given lecture and that reported that they had an IT education. Both the entropy value with the Shannon formula and the entropy value with the

zxcvbn algorithm are showed next to each other just like previous Diagram. The results show that users with an IT education had a higher average entropy value than the users that did not have an IT education.

9. Comparison to similar work

The results of this study will be as previously mentioned compared to two other password related studies, one at the University of Maribor and the other one at the department of defence. Take note that both differ in execution and what they look for than study performed in this paper and this will be discussed and taken into consideration during the comparison.

9.1 Comparison to the Maribor study

In the Maribor study made by Taneski et al. (2014a) they studied the characteristics of user's password on services such as their university account and for social media. They used a sample group that was similar to the one of this study. 33 Students at the university, the big difference between the studies performed by Taneski et al (2014a) and the study performed in this paper is that they performed the questioner first and then gave a lecture and then went back two weeks later an performed the questioner again. This would have been an interesting method to trying as well in this study but was not done because of the lack of time and it would have been harder to get students to perform the experiment once more two weeks later because of exam weeks. Also a major difference is that they used a questioner and not an experiment phase which could be more of a real life example.

They found that the average password length had an increase between the two phases of the experiment similar to the one found in the experiment performed in this study at the university of Skövde. They found that the average password for example a Facebook account increased from 10.33 to 12.05 and the password for a notebook computer increased from 8.79 to 11. Looking at the increase of the notebook computer the average password length was 31,5%, and by looking at the increase found by this study at the University of Skövde the average password for the bank was 11.45 without education and 16.3 with education which is an 37,03% increase. They did however not look at the entropy of the passwords used but only presented the average length, frequency of password change, and how many users wrote their password down. They did however mention that they found a raise in the cardinality of the passwords between the first and second phase. This result is also similar to the one found in this study. They also found that many users stated that they would change their password after the lecture because they now understood their passwords are weak and could compromise their accounts. After performing the interviews some interviewees also stated that they would now rethink their password structure.

9.2 Comparison To department of Defence study

In 1999 Zviran & Haga (1999) performed a study of password at the Department of defence, they found that a very small number of users used what today is the suggested to be a strong password according to the guidelines they used, 47% used a password length shorter than 8 characters but in the study done in this experiment this number was found to be 35% of the users that did not get an education on passwords. This number was calculated to look at all the

passwords that users used, if one of the three passwords were under 8 characters they were added to this category even if the other 2 were over 8 characters in length. This could be because the other study was performed back in 1999 and many websites will not allow a password of less than 8 characters and the users could have been moulded over time to use stronger passwords. Maybe 16 years of password guidelines and enforcements of this rule by websites and applications could have had an effect on the users? The result from the study performed in this experiment is not big enough to make this a fact but the study by Zviran & Haga should be performed again on the same sample group to see if the fact has changed.

10. Discussion

By reviewing the guidelines set by many websites for passwords it seems like developers are on the right track of trying to make users pick stronger passwords. But a big problem is that the websites do not tell the users why they need to use stronger passwords just that they should. With users not being told why they need to use it, research has shown that users often just add some characters to the password and reuse the same password over several services.

The results by doing the experiment have shown that by giving a lecture on password related matters the entropy of the passwords can be increased. However the experiment was performed directly after the students got the lecture so there was no memory decay for the users. But this does not have to be an issue, if company's make their workers partake in education right before they start registering their company accounts it could affect their users to pick stronger passwords. If they then have a 6 month period between changing passwords the user will use the strong password they presumably over this 6 month period and be influenced about using this kind of password later on. This is only a hypothesis and has to be tested with more future work, by following a group of users over an extended period of time. But by looking at the experiment done in the Maribor study they also showed positive results after a 2 week memory degradation time.

Also the sample group used for the experiment was only students at the University of Skövde and the results can be affected by this, but during the interviews even students with an IT education claimed they sometime pick easy to remember and weak passwords for the convenience of having easy to remember passwords.

The experiment could have been broadened to include an application as well were the users register accounts. Because users might differentiate between application logins and web logins and it would be interesting to test this theory, also the users should have answered some sort of questionnaire after the experiment with questions similar to the ones asked in the interview.

The design of the websites were not the primary goal of this experiment but it could also be a good idea to look more into the matter of "how the design" and the purpose of a website can affect the behaviour of a user when picking a password. One could also test if a SSL certificate implemented on a website will affect the strength of a user's password.

Another problem that showed itself when interpreting the data was that the variable of if the users had an IT education was not defined enough and should have indicated that the users had some sort of computer security education.

11. Final conclusion

This paper had a few research questions it wanted to try to answer:

- What problems exist in password management today? And are these only limited to certain user groups?
- What defines a strong password and is bad password management related to gender, education level and the usage of passwords in daily life?

By looking at the literature studied for this report there are several problems with passwords today. One of the biggest issues are weak passwords by viewing the studies of the password leaks one can notice that a lot of users have very weak and short passwords, also the results in this study showed that some of the users had really weak passwords, Some just 4 characters long and several were just lowercase letters. The management of passwords is also an issue and after conducting the interviews this seems to not be limited to one user group. None of the people interviewed changed their passwords on a regular basis and all of them reuse the same password for several services. Overall it seems like the behaviour around passwords are very poor in all user groups. Also articles studied for this paper shed some light on more matters that effect passwords picked by the users. The Florêncio & Herley (2007) study showed that users reused the same password over roughly four services.

- Can you change bad password behaviour by giving a lecture on the matter?

The goal of this paper was never to try to change all the behaviours regarding passwords but to collect more information about them. Also to try to see if the overall password entropy could be increased by informing the users more about the problems that exist around passwords and try to help them understand how to make longer and stronger passwords. The experiment gave a higher entropy value and a longer average password length and fewer users also reused the same password over the three services than the group not given the education. So it seems like a lecture on the matter can affect the password behaviour of users, this combined with the results of the Maribor study shows that maybe it's better to teach our users by informing them about matters regarding passwords instead of using guidelines. The average password length was also increased after users being given a lecture on the matter. In school we get education about how to do dishes and cook food to prepare us for the life of an adult by partaking in cooking classes, with passwords being a big part in our life it would also be good to talk some about this during the school year. Not just passwords but computer security in general, how users data is stored online, what rights we have and do not have when using the internet. There is already computer education on the syllabus but this only focuses on how to use a computer and write on a keyboard but does not go into detail about some of the more important parts of computers.

The lecture used in this paper could also have been a longer version with more real life examples and more ways of how to create a password, but also with more information about two-step authentications for example.

12. Contribution

The results of this study can be used to argue for more research in the area instead of trying to affect user by using password guidelines. It can also be used by companies affected by weak user's passwords by giving their users a lecture on the matter. A similar study was performed by Taneski et al. (2014a) which evaluated the changes by using a questionnaire. This paper instead used an experiment which gives more of a real life scenario result than the Maribor study, the results from this study included with the results of the Maribor study shows that training does prove to an increase of password strength.

13. Future work

Future work would be a longer study with time included between the experiment and the lecture, maybe a bigger sample group and several different groups with different kinds of educations. It would also be preferable to look at a more mixed population of people also working with passwords instead of just students. Also results like frequency of password changing and reuse of passwords would be interesting to study in the future. The lecture itself could be designed in several different ways to see what kind of lecture material that could change the behaviour the most.

References

Berndtsson, M., Hansson, J., Olsson, B., Lundell, B. (2008) *Thesis Project.* Springer-Verlag London.

Brottsförebyggande rådet. (2012). *Något fler anmälda brott första halvåret - Brå*. [online] Bra.se. Available at: https://www.bra.se/bra/nytt-fran-bra/arkiv/press/2012-07-12-nagot-fler-anmalda-brott-forsta-halvaret.html [Accessed 3 Mar. 2015].

Cheswick, W. (2013). *Rethinking Passwords*. Association for Computing Machinery. Communications of the ACM. 56, 40.

Craddock, D. (2013). *One year since the preview of Outlook.com – thank you for helping us build the world's fastest growing email - Office Blogs*. [online] Office Blogs. Available at: http://blogs.office.com/2013/07/31/one-year-since-the-preview-of-outlook-com-thank-you-for-helping-us-build-the-worlds-fastest-growing-email/ [Accessed 3 Mar. 2015].

de Carné de Carnavalet, X. (2014). *A Large-Scale Evaluation of High-Impact Password Strength Meters*. Concordia University.

Florencio, D. & Herley, C. 2007, "A large-scale study of web password habits", *16th International World Wide Web Conference, WWW2007,* pp. 657.

Fordham, D. R. (2008). *How Strong Are Your Passwords?* Strategic Finance. 89, 42-47.

Furnell, S. (2007). An assessment of website password practices. *Computers & Security.* 26(7-8):445–451

Haley, K. (2010). *Living with Passwords*. [online] Symantec.com. Available at: http://www.symantec.com/connect/blogs/living-passwords [Accessed 3 Mar. 2015].

Harrison, W. (2006). *Passwords and Passion*. IEEE Software. 23, 5-7.

Häger, B. (2001). Intervjuteknik. Stockholm, Liber AB.

Lucas, I. (2009). *Password Recovery Speeds*. [online] Lockdown.co.uk. Available at: http://www.lockdown.co.uk/?pg=combi&s=articles [Accessed 5 Mar. 2015].

Ma, W., Campbell, J., Tran, D., & Kleeman, D. (2010). *Password entropy and password quality*. *In Network and System Security* (NSS), 2010 4th International Conference on (pp. 583-587). IEEE.

Microsoft. (n.d). *Create strong passwords*. [online] Available at: http://www.microsoft.com/en-gb/security/online-privacy/passwords-create.aspx [Accessed 3 Mar. 2015].

Moyer, M. (2014a). *Password Creation Rules - HECC Knowledge Base*. [online] Nas.nasa.gov. Available at: http://www.nas.nasa.gov/hecc/support/kb/Password-Creation-Rules_270.html [Accessed 3 Mar. 2015].

Moyer, M. (2014b). *RSA SecurID Fobs - HECC Knowledge Base*. [online] Nas.nasa.gov. Available at: http://www.nas.nasa.gov/hecc/support/kb/RSA-SecurID-Fobs_58.html [Accessed 3 Mar. 2015].

Munroe, R. (2011). Password Strength [Comic strip]. *Massachussetts: Xkcd.com*. Available at: https://xkcd.com/936/ [Accessed: 11 April 2015]

Power, R. (2000). *Tangeled Web: Tales of Digital Crime from the Shadows of Cyberspace*. Macmillan Press Ltd., Basingstoke, UK, UK.

Scarfone, K., & Souppaya, M. (2009). *Guide to enterprise password management (draft).* NIST Special Publication, 800, 118.

Sato, T., Kikuchi, H. (2012). *Synthesis of Secure Passwords*. Information Security (Asia JCIS), 2012 Seventh Asia Joint Conference on , vol., no., pp.35-37.

Shannon, C.E. (1948). *A mathematical theory of communication.* Bell System Technical Journal, vol.27, no.3, pp.379-423.

Shannon, C.E. (1951). *Prediction and entropy of printed English*. Bell System Technical Journal, vol.30, no.1, pp.50-64.

Taneski, V., Brumen, B. & Heričko, M. (2014a). *The effect of educating users on passwords: A preliminary study*. *In Third Workshop on Software Quality Analysis, Monitoring, Improvement and Applications* (p. 107).

Taneski, V., Brumen, B. & Heričko, M. (2014b). *Password security — No change in 35 years*?. *Information and Communication Technology, Electronics and Microelectronics, 37th International Convention on* , vol., no., pp.1360,1365, 26-30 May 2014

The Imperva Application Defense Center (ADC). [pdf] (2014). *Consumer Password Worst Practices*. Available at: http://www.imperva.com/docs/wp_consumer_password_worst_practices.pdf [Accessed 4 Mar. 2015].

University of Skövde. [pdf]. (2005). Regler för datornanvändning. Available at: http://www.his.se/PageFiles/1906/datorregler200508.pdf [Accessed 3 Mar. 2015].

Wheeler, D. zxcvbn: realistic password strength estimation. Dropbox blog article (Apr. 10, 2012). https://tech.dropbox.com/2012/04/zxcvbn-realisticpassword-strength-estimation/.

Wohlin, C., Runesson, P., Host, M. (2012). *Experimentation in Software Engineering*. Springer-verlag Berlin and Heidelberg.

Zviran, M., & Haga, W. J. (1999). *Password security: an empirical study*. Journal of Management Information Systems, p. 161-185.

Appendices

Here are the appendixes for the paper.

Appendix 1- Average password lengths

Here is the length of all the passwords for the users given the lecture with the average value of the password length:

	Bank	Pizzeria	Email
Subject 1	**8**	13	**9**
Subject 2	19	19	19
Subject 3	17	18	20
Subject 4	23	16	22
Subject 5	19	12	19
Subject 6	15	15	14
Subject 7	15	18	16
Subject 8	14	13	16
Subject 9	14	16	12
Subject 10	18	19	16
Subject 11	17	17	17
Subject 12	12	10	9
Subject 13	30	18	11
Subject 14	19	19	21
Subject 15	20	10	20
Subject 16	9	9	8
Subject 17	8	8	8
Subject 18	23	21	17
Subject 19	12	9	9
Subject 20	14	20	16
Average	**16,3**	**15**	**14,9**

Table 9: Average password length for group given education

Here is the length of all the passwords for the users not given the lecture with the average value of the password length:

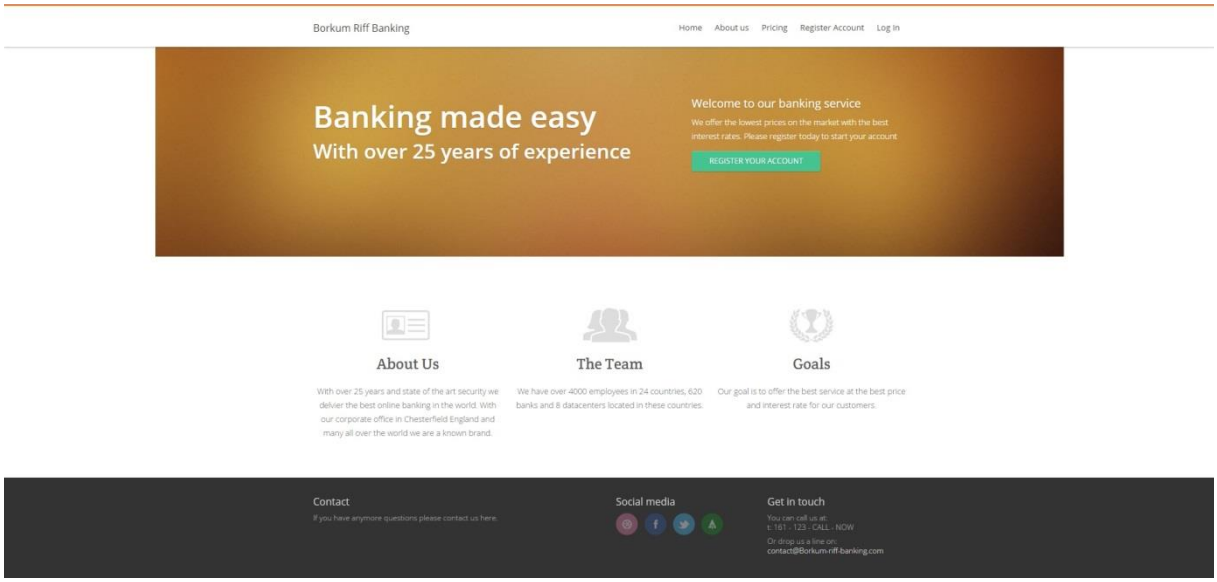
	Bank	Pizzeria	Email
Subject 1	19	18	14
Subject 2	8	15	12
Subject 3	7	8	8
Subject 4	7	8	7
Subject 5	11	13	8
Subject 6	7	8	8
Subject 7	8	7	10
Subject 8	9	9	9
Subject 9	15	14	15
Subject 10	20	6	21
Subject 11	16	16	16
Subject 12	11	7	6
Subject 13	24	12	10
Subject 14	10	12	8
Subject 15	8	9	9
Subject 16	11	8	10
Subject 17	4	4	4
Subject 18	14	11	14
Subject 19	11	11	7
Subject20	9	12	13
Average	11,45	10,4	10,45

Table 10: Average password length for group not given education

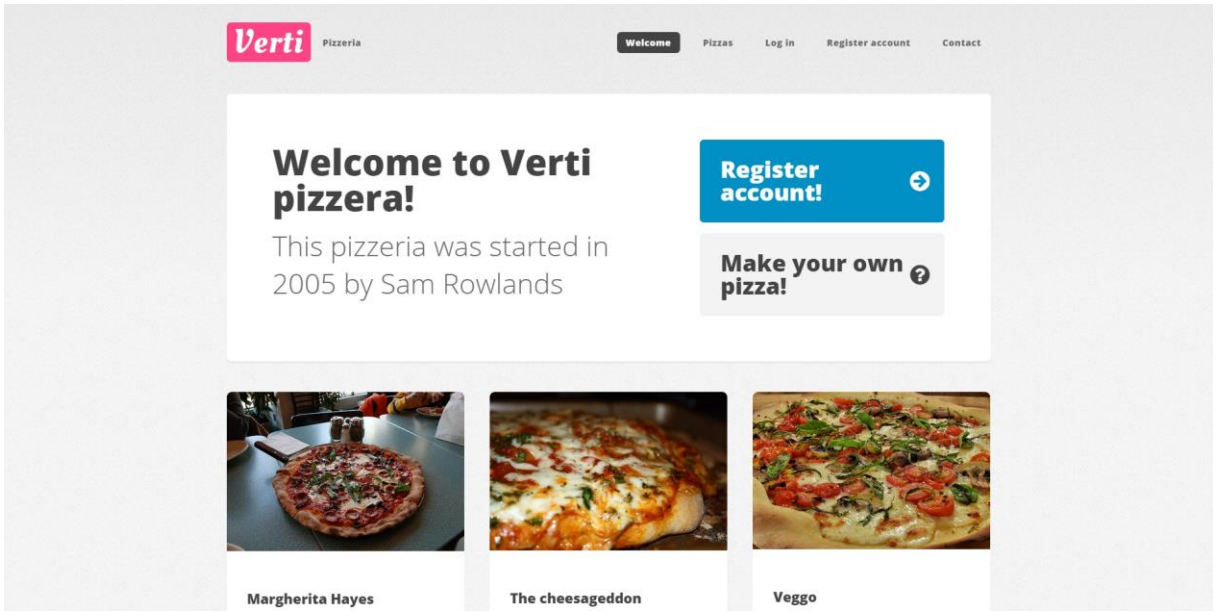
Appendix 2 - Websites

These are the websites used for the experiment

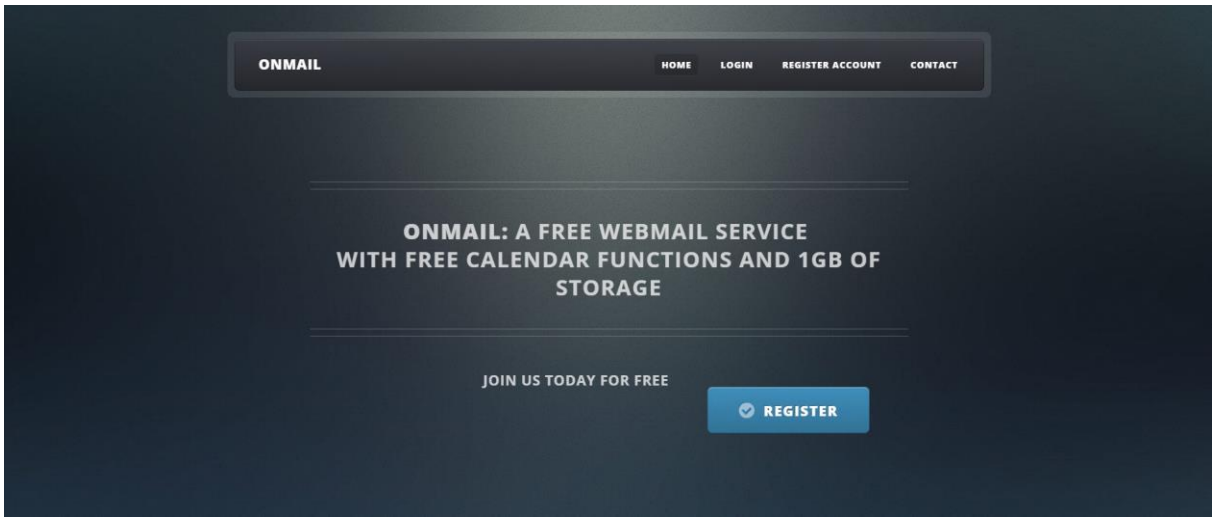
The online banking service:



The online pizzeria:



The email service:



Appendix 3 – Interview responses

Interview 1

Gender: Male Age: 20+ Password habits: Does not work with passwords

What do you consider a good password?

Ett bra lösenord är ett lösenord som

1. Jag kan komma ihåg
2. Inte är för enkelt för någon annan att knäcka liksom, det ska ju inte vara för få bokstäver men även siffror och helst unikt för mig

How often do you change your password and why?

Jag byter inte lösenord så ofta, jag kör oftast med ett jättelångt krypterat lösenord via en hemsida som heter lastpass, jag skriver sällan in mer än ett lösenord och jag byter aldrig inne på den sidan.

Do you reuse the same password for several websites?

Ja det händer, jag använder lastpass till allt. Mycket saker som inte heller är internetrelaterade också

Do you trade password strength for your own convenience for picking an easy to remember password?

Blandat, det är ju ett flummigt lösenord som man alltid kommit ihåg själv som är ett behagligt och enkelt lösenord för mig, men för någon annan så blir det ett fullständigt nonsens lösenord för någon annan men för mig är det något som jag lätt minns.

Do you think the design and purpose of a website motivates when you pick a password?

Jai bland kanske, det kan göra det. Men vanligtvis har jag en speciell typ av lösenord som jag använder och bygger utifrån. Jag brukar titta ifall dom har den gröna hänglåset på en sida så känner jag mig alltid tryggare med att använda den sidan.

Are you aware of online crimes? What type of crimes do you know of and do you consider them when using the internet?

Ja det gör jag, jag har det I åtanke när jag använder internet.

Have you ever been the victim of any online crimes? (Hacking, identity theft)

Nej jag har inte blivit utsatt för någon form av bedrägeri.

What do you know about password guidelines and do you take them into consideration when registering an account?

Jag känner till d om och vet vad som gäller när man ska välja lösenord men jag är lat och ibland så väljer jag något som är lätt att komma ihåg eller brukar använda.

Does your workplace have an IT policy and do you fully understand and follow all the parts of it?

Not applicable.

Interview -2

Gender: Male Age: 20+ Password habits: IT Education - University

What do you consider a good password?

En blandning av stora och små bokstäver med nummer inkluderat och minst 8 tecken långt.

How often do you change your password and why?

Jag byter faktiskt aldrig systematiskt, bara ifall jag har glömt mitt lösenord eller ifall en hemsida påminner mig om att byta.

Do you reuse the same password for several websites?

Ja det gör jag, det är lättare att komma ihåg för mig. Det sitter som ett muskel minne.

Do you trade password strength for your own convenience for picking an easy to remember password?

Jag försöker ha starka lösenord men samtidigt försöker jag välja ord som jag kan relatera till med nummer bakom, så det är en blandning av båda.

Do you think the design and purpose of a website motivates when you pick a password?

Ja det gör det, jag använder lättare lösenord på hemsidor som inte är så viktiga, som inte lagrar någon viktig data för mig.

Are you aware of online crimes? What type of crimes do you know of and do you consider them when using the internet?

Ja, känner nog till det mesta, försöker alltid att undvika självklara fishing scams när jag använder internet.

Have you ever been the victim of any online crimes? (Hacking, identity theft)

Ja, fick en gång mitt paypal konto stulet och dom spenderade pengar. Men jag fick tillbaka det efter 3 dagar igen.

What do you know about password guidelines and do you take them into consideration when registering an account?

Jag använder bara guidelines när en hemsida kräver det, ibland blir ju lösenords mätaren röd för att jag inte uppfyllt något krav, då går jag tillbaka och läser igenom dom mer.

Does your workplace have an IT policy and do you fully understand and follow all the parts of it?

Ja, men jag har inte läst igenom den.

Interview - 3

Gender: Male Age: under 18 Password habits: No education / work

What do you consider a good password?

Jag vet att lösenord som har många olika variationer, tecken, siffror och stora och små bokstäver är starka.

How often do you change your password and why?

Jag byter aldrig mina lösenord, jag har bytt några gånger till smågrejer. Bara nära jag blivit tvingad att byta lösenord.

Do you reuse the same password for several websites?

Ja det gör jag, har två tre olika lösenord som jag använder, jag har det så för säkerhetsskull ifall någon får reda på mitt lösenord så ska jag fort kunna ta mig in på alla mina tjänster och byta.

Do you trade password strength for your own convenience for picking an easy to remember password?

Jag vill ha ett långt lösenord med visa olika monster, då känns dom bra. Så jag tror inte att jag byter det.

Do you think the design and purpose of a website motivates when you pick a password?

Nej jag har alltid samma typ av lösenord för allt, har inte extra starka för olika tjänster.

Are you aware of online crimes? What type of crimes do you know of and do you consider them when using the internet?

Jag har aldrig blivit påverkad av det men jag känner till det, men jag känner inte att jag är rädd för det.

Have you ever been the victim of any online crimes? (Hacking, identity theft)

Nä det har jag inte blivit utsatt för, dock har folk försökt scamma mig.

What do you know about password guidelines and do you take them into consideration when registering an account?

Jag har läst igenom dom, men ifall det inte är tvingat så brukar jag inte följa dom.

Does your workplace have an IT policy and do you fully understand and follow all the parts of it?

Frågan omformulerades till ifall det fanns någon IT policy på skolan.

Jag har inte läst den men jag vet att det finns, så jag vet inte om jag förstår de olika delarna av den. Men vi var tvungna att acceptera regler när vi började.

Interview - 4

Gender: Female Age: 20+ Password habits: Does not work with passwords

What do you consider a good password?

Något som inte är för personligt, det ska vara svårt att gissa för personer som försöker få tillgång till mitt konto.

How often do you change your password and why?

Jag brukar byta mitt lösenord när tillexempel Microsoft säger åt mig att göra det, eller när jag tror att någon fått tillgång till mina konton. Men annars så brukar jag inte byta dom.

Do you reuse the same password for several websites?

Ja det gör jag, det brukar vara lättare att komma ihåg dom även ifall jag vet att det är en dålig idé.

Do you trade password strength for your own convenience for picking an easy to remember password?

Ja, jag har hellre ett svagt lösenord som är lätt att komma ihåg och påminner om ett lösenord jag tidigare använt.

Do you think the design and purpose of a website motivates when you pick a password?

Nej det tror jag inte, jag använder ofta samma lösenrod på olika hemsidor så jag tror inte att hemsidan påverkar lösenordet på något sätt.

Are you aware of online crimes? What type of crimes do you know of and do you consider them when using the internet?

Jag har en gång fått min email hackad, det var riktigt jobbigt. Men det hade inget med att göra att jag var oförsiktig, flera hundra emailkonton och lösenord blev läckta. Men jag har en bra lösning för paypal som jag tycker är smart, eftersom att jag vet att det är ett viktigt konto så

har jag en speciell email kopplad till det kontot. Detta kontot använder jag aldrig för andra tjänster utan bara för paypal. Detta kontot har även ett annat lösenord än mina andra konton. Sedan vet jag ju att det finns flera andra typer av scams där ute, jag får ofta email från tex "The International Cricket Foundation" som säger att jag har vunnit 100.000.000 dollar.

Have you ever been the victim of any online crimes? (Hacking, identity theft)

Ja om jag nämnde tidigare, så har min email blivit hackad.

What do you know about password guidelines and do you take them into consideration when registering an account?

N/A

Interview - 5

Gender: Female Age: 20+ Password habits: It education

What do you consider a good password?

- ett lösenord som har både stora och små bokstäver, siffror och specialtecken.

How often do you change your password and why?

- Privat ändrar jag aldrig mina lösenord, även om jag vet att man borde göra det för att minska risken att bli hackad.
På jobbet gör jag det när datorn säger till en gång i månaden att vi måste byta inlogg till datorn och lösenord till våran mejl. Men då byter jag oftast bara ut en siffra i lösenordet.

Do you reuse the same password for several websites?

- ja.. även om jag vet att det också är fel. brukar köra på två olika lösenord (fler än så kommer jag inte att komma ihåg).
sedan har jag ett helt annorlunda när jag behöver lämna lösenord på en hemsida som känns "oseriös" för att undvika att de kommer åt alla mina sidor med det lösenordet.

Do you trade password strength for your own convenience for picking an easy to remember password?

- Ja, det gör jag redan. Jag kommer inte ihåg för svåra eller krångliga lösenord. Då slutar det med att jag måste ha en lapp uppskriven med mitt lösenord vilket inte är en jättebra idé.

Do you think the design and purpose of a website motivates when you pick a password?

- Ja, som jag sa tidigare så använder jag två lösenord på nätet + ett helt annorlunda som jag har på sidor som ser eller inte känns seriösa.

Are you aware of online crimes? What type of crimes do you know of and do you consider them when using the internet?

- Det jag tänker på är att någon kan hacka sig in på min bank och ta pengar, det tänker jag på om jag behöver handla på nätet. men det hindrar mig inte från att handla, jag skulle dock inte skriva in mina bankkorts uppgifter om sidan känns oseriös eller jag inte hört talas om den innan.
jag känner även till identitetsstölder men det oroar jag mig inte för.

Have you ever been the victim of any online crimes? (Hacking, identity theft)

- När jag var yngre blev min mejl hackad och de bytte lösenord så jag inte kom in. Då hade jag inte så stor användning av den så det var ingen större fara, skapade bara en ny med ett mer seriöst namn.
När jag var tonåring var även en sida som hette "Efterfest", där la man upp bilder och chattade. Jag kommer ihåg att allas lösenord hade kommit ut och att de spreds på internet, så det slutade med att en bekant till mig var en av alla de som hade fått allas lösenord i ett excelark. Där såg jag att de faktiskt hade mitt lösenord men jag hade redan ändrat till något annat, men minns att jag tyckte det var obehagligt att så många olika kunde kolla vad man hade/haft för lösenord.

What do you know about password guidelines and do you take them into consideration when registering an account?

- Jag vet vad man ska ha för lösenordssäkerhet men ska jag vara ärlig så använder jag det inte, måste jag ha siffror så lägger jag till tex: 123 i slutet på mitt vanliga.

Kräver det en stor bokstav så tar jag det i början. så jag och lösenordssäkerhet är inte jättebra, men efter denna intervjun tänker jag nog gå och uppdatera mina lösenord.

Interview - 6

Gender: Female Age: 50+ Password habits: Use passwords in a daily line of work

What do you consider a good password?

Ett lösenord som är lätt att komma ihåg, som man inte själv glömmer av. Jag väljer oftast någonting med samma ord så byter jag ut siffrorna.

How often do you change your password and why?

Vi måste byta på jobbet var sjätte vecka, mina personliga lösenord byter jag aldrig och har aldrig behövt byta. Mina personliga skulle jag bara byta ifall jag behövde byta. Det kommer upp ett meddelande på jobbet att det är dags att byta lösenord och då gör jag det.

Vi får inte återanvända samma lösenord, men efter ett tag så kan man återanvända ett gammalt lösenord igen. Men vi ska inte göra det egentligen. Vi har även ett kort som man sätter i tangentbordet som också har ett separat lösenord.

Do you reuse the same password for several websites?

Jag återanvänder samma lösenord för allt, kan ha något som skiljer sig men då är det bara sifforna som jag byter, detta är för att jag ska komma ihåg mitt lösenord. Då jag känner att jag problem att komma ihåg lösenord och får ofta återställa lösenordet ifall jag inte kommer ihåg det.

Do you trade password strength for your own convenience for picking an easy to remember password?

Ja det gör jag verkligen.

Do you think the design and purpose of a website motivates when you pick a password?

Nej det gör det inte jag har samma lösenord på allt.

Are you aware of online crimes? What type of crimes do you know of and do you consider them when using the internet?

Jag känner till det men tanker mer på den fysiska säkerheten igenom att låsa datorn. Så jag tror jag vet mer om den fysiska brottsligheten än att bli utsatt för brott över internet. Men man har ju läst mycket om det på internet om bedrägeri och sådant.

Have you ever been the victim of any online crimes? (Hacking, identity theft)

Inte vad jag vet, det är ju det som är så svårt att man vet ju inte förens det är försent känns det som.

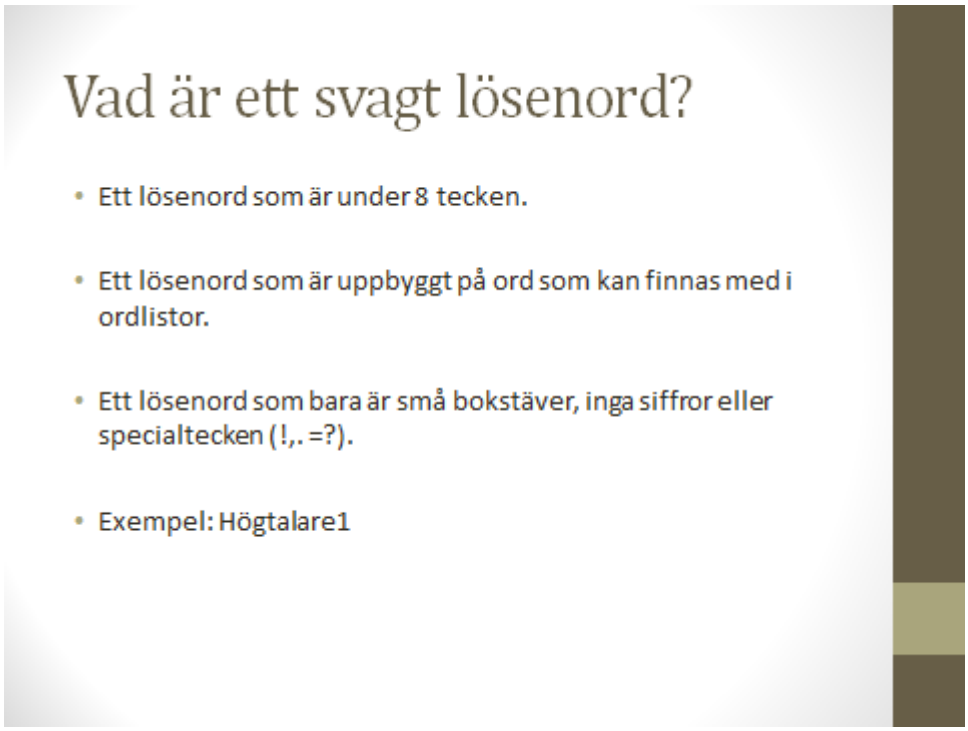
What do you know about password guidelines and do you take them into consideration when registering an account?

Våra riktlinjer är att det ska vara siffror och en storbokstav och minst 7 tecken. Det är väl det jag har koll på från jobbet, sedan vet jag att vi ska undvika sitt eget namn. Men i övrigt så har jag ingen koll på sådant på internet, när det är utanför jobbet så använder jag mitt vanliga lösenord.

Does your workplace have an IT policy and do you fully understand and follow all the parts of it?

Ja det har vi, jag känner väl att jag förstår den, att man inte ska visa när man loggar in och byta lösenord när folk ser dom.

Appendix 4 - Lecture



Vad är ett starkt lösenord

- Ett lösenord som har 8 eller mer tecken
- Det borde inkludera små och stora bokstäver men även siffror och specialtecken.
- Ett lösenord som inte är byggt på information som kan kopplas till dig exempel är:
  - Månaden du är född
  - Dina barns namn.
  - Favorit fotbollslag.

Lösenordshantering

- Ge det aldrig till en annan person.
- Skriv aldrig ner ditt lösenord på ett papper.
- Du ska försöka byta lösenord regelbundet och inte bara när du behöver byta lösenord.
- Återanvänd inte samma lösenord på alla hemsidor.

Vad för lösenord väljer folk?

- Vid en analys av 32 000 000 läckta lösenord så kom forskare fram till att:
  - 290 791 lösenord bestod av "123456".
  - 61 958 lösenord bestod av ordet "password".
  - 13 340 800 lösenord innehöll bara små bokstäver.
  - Bara 64 000 av lösenorden innehöll 8 tecken och en blandning av små och stora bokstäver, siffror och specialtecken.
- 76% av alla intrång på konton sker på grund av stulna lösenord eller för att användare har undermåliga lösenord.

PassPhrase

- En Passphrase är en serie av ord som tillsammans formar ett lösenord.
- Längre lösenord är svårare att knäcka med brute force.
- Ett långt lösenord med en blandning av flera olika tecken ger en hög lösenords entropi.
- Det är viktigt att de ord som används för lösenordsfrasen inte är direkt sammankopplade då det kan bli lättare att knäcka.

Passphrase Exempel

- Ett exempel för en Passphrase kan vara:
- BankMannenSpelarPåFläkten451_
- Lösenords fraser kan vara lättare att komma ihåg än svåra lösenord som (Ytm43_!.). Då man kan skapa sig en minnesbild av en bankman som spelar på en fläkt.
- För att göra sitt lösenord ännu säkrare kan man ta bort bokstäver ur ordet eller byta ut dom mot siffror. Detta kan dock göra ordet svårare att komma ihåg så det är viktigt att inte överdriva.
- Exempel: B4nkManneSpelarPåFläkte451_

Appendix 5 – Summary

There are several concerning issues with user's passwords today; it's a known fact that users pick weak passwords and very often manage their passwords in a bad way. This report will do a study of what users are being taught today by evaluating password guidelines provided by three different organizations for users when registering accounts on different services. Then a study of passwords leaks from two different leaks will be performed to see what kinds of passwords users pick. Two methods will be used to answer the research questions this paper will aspire to answer one is an interview and the other one is an experiment.

The interview will consist of several different password related questions to get a better understanding of what users know about password related issues and what they consider strong passwords. The interviewee subjects are of varying age, education and have varying habits surrounding passwords. Then the information collected during the interviews was used to design the experiment part of the report.

The experiment was used to test if a user's passwords could be strengthened by giving them a lecture on the matter. For starters three different websites were created for the users to register accounts on, these websites were a bank, a pizzeria and an email service. Then from the information gathered during the interviews and during the background part of the report a lecture was created to educate users on password related matters. The lecture contained information on what a strong password is, how to manage your passwords, what kinds of passwords users use and how to create a password using a passphrase. 40 students at the University of Skövde conducted the experiment, 20 of these users got the lecture and the other 20 functioned as a control group to see if the education had any effect. After the users registered accounts on the three websites each password was studied individually to see if there was any difference between the groups. The group given the lecture had longer average password, there was also less users in this group that reused the same password over several services. Users were also able to check a box on the first website if they had an IT education to see if there was any difference in passwords for users with the IT education. It was found that these users had a slightly higher average password length than other users.