5-2018

Cybersecurity: Personal Information and Password Setup

Yizhi Ma
Missouri University of Science and Technology, Ymcy2@mst.edu

Nathan W. Twyman
Missouri University of Science and Technology, nathantwyman@mst.edu

Cybersecurity: Personal Information and Password Setup

Yizhi Ma
Missouri University of Science and Technology
Ymcy2@mst.edu

Nathan W. Twyman
Missouri University of Science and Technology
nathantwyman@mst.edu

ABSTRACT

This study is designed to explore what are the most common and most suggested way to create passwords with better security, especially for the group of people who set their personal information as a part of their passwords. This study was conducted under survey methods, the participants were randomly collected from the students at Missouri S&T as well as online crowdsourcing internet, Amazon Mechanical Turk. The results of this study illustrated that more than 50% of the participants are still using putting personal information in their password regardless how insecure and unsuggested this move is. With the concern of this phenomenon, 150 results collected from the survey were interpreted and analyzed to provide a better guide for password setup and password security.

Keywords

Cybersecurity, Personal Information, Password Strength, Password

INTRODUCTION

Cybersecurity is defined as the collections of skills, security services, risk analysis and management means, security activities and precautions as well as its assurance and technologies that are utilized to protect the cyber assets and property rights (Craigen, Diakun-Thibault, Purse, 2014). It has been a heated topic for its implements in both academic and practice field. Meanwhile, password security, which has been a continuous problem in cybersecurity, which has not been explored enough under the background of cybersecurity (Schaffer, 2013). With the essential role internet plays in people's daily lives and activities, the password is one of the most important parts that secure and authenticate all these activities, such as online banking, online communication, E-commerce among one user to another through the internet (Bidgoli, 2004), that makes the importance of password security should not be overlooked.

Password has been used to grant access to secure properties since ancient times. Even today, it is not only used to secure physical exist properties in a traditional sense but can also secure digital assets through digital devices or internet. It is the only security that could enforce the unauthorized access towards an application or software, but grant the authorized access to access the application or software (Chanda, 2016).

A good password needs to be both secure and memorable to make it harder to break. It should be complexed enough itself firstly and avoiding being known by other people, while most of the people's passwords only focused on the first character but not the second (Vu, Proctor, Bhargav-Spantzel, Tai, Cook, Schultz, 2007). For instance, even it is not recommended, there are still a lot of internet users intend on putting their personal information such as name, birthday or phone number etc. in their password. Many studies have indicated the fact that there are still a large number of people putting their personal information as part of their passwords for better memorability, with knowing that this move is not suggested for security reason at all (Bishop, & Klein, 1995; Riddle, Miron, Semo, 1989; Leyden et al., 2003; Li, Wang, & Sun, 2017). The study was done by Li et al. (2017) concludes that as long as password memorability is still concerned as one of the most crucial aspects while creating a password, the fact that including personal information in user passwords will keep remaining, regardless of different languages, cultures, and background etc. Another study (Sharma, Nah, Cheng, Siau, Hall, 2017) found out a positive relationship between users rick adverse actions with confident regarding cybersecurity, and that also leads to greater cybersecurity risks. However, the specific use of personal information in passwords and its security implications have barely been studied under this context.

This study intended to focus on explaining two questions related to password security. Firstly, this study was focused on giving a clear analysis that showed what do people preferably to do while creating strong passwords even with their personal information in the password. Also, this study aimed to verify if there is any relationship between password strength and personal information contained.

This study was conducted using survey methodology with pre-designed questions through an online survey software, Qualtrics. There were 202 participants took this survey including 52 students from Missouri S&T and 150 random participants found from Amazon Mechanical Turk, after excluded all the invalid and uncompleted data, only 150 results were valid for further analysis.

This study ended with a discussion of empirical results, theoretical and practical implications, limitations and future research directions, and a section that concludes all the research findings and its applications.

METHODOLOGY

In this research study, a survey approach with pre-designed questions was proposed to collect the data. The survey designed for this research was majorly based on a detail-focused questionnaire on various aspect of cybersecurity under the context of password strength and personal information.

The survey for this study was built on a free web-based survey software, Qualtrics. The main reason for choosing Qualtics for this research is that it could secure the information the participants and make sure they are anonymous since this study is related to cybersecurity, we want to make sure that our participants will feel secure while filling out the survey. Meanwhile, it provides a user-friendly interface in all different operating systems, making it convenient for the participants to accomplish their test without any possible restrictions or technological difficulties. Qualtrics enables us to set up a different type of survey measurement to make sure the survey questions will be conveyed in the same way as we designed. It also allows us to view the results of this survey in reports or download them for further analysis.

The survey tasks were generated based on the testing objectives and the goal of this survey to gather both qualitative and quantitative information. The survey questions are majorly focused on determining what and how much methods are the participants mostly using their password while creating a strong password.

This survey was conducted within the participants from both Missouri S&T and on Amazon Mechanical Turk, fifty-two valid results were collected from the students at Missouri S&T and others were collected from more than one hundred fifty random testers filtered by Amazon Mechanical Turk. The survey was designed into two flows to separate the people based on different answers they have as Figure 1 indicates below.

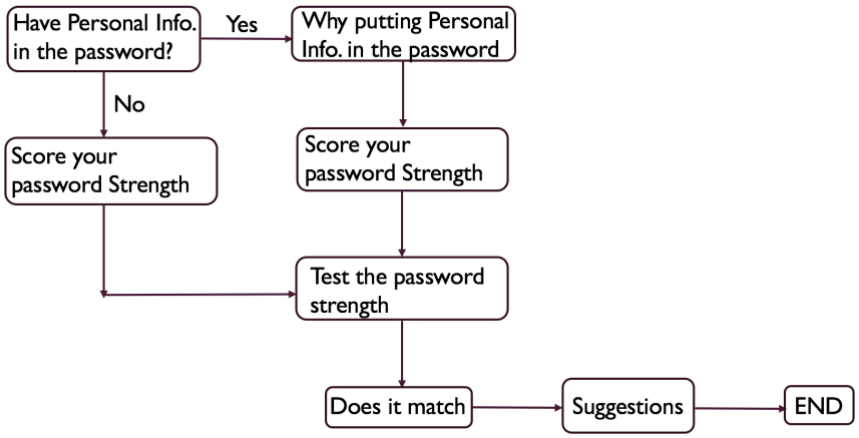

Figure 1. Survey Flow

DATA ANALYSIS

This study was conducted under a storyboard, it intends to picture a specific context for the participants, so as to make sure the collected data could be as close to reality as possible. There were 202 participates submitted data originally, but after considering all the results carefully and eliminate the invalid and uncompleted data, only 150 results were valid and qualified to move forward with further analysis.

The primary results from data analysis indicates 56% participates chose "Yes" as they included their personal information in their password, comparing there were 44% participates who chose "No", which they do not include personal information in their password. This result illustrates that there are more than half of the people are still containing their personal information in their passwords, this gives us a more persuasive knowledge about why this study is necessary and important to be done.

To further analysis, four sets of regression and correlation analysis were done by SAS software between PS. info Contain (if the password contains personal information or not) and Strength Evaluation (personal evaluation for their password strength), PS. info Contain (if there is personal information contains the password or not) and PW Tested Strength (the password strength test result from given resource), PS. info Contain (if there is personal information contains the password or not) and Results Matching (if their evaluated results match tested results or not), and PS. info Contain and Awareness Evaluation (self-evaluation of cybersecurity awareness).

PS. info Contain and Strength Evaluation

As for the regression analysis between PS. info Contain and Strength Evaluation, we found that there is a relationship Personal information containing and password strength evaluation with the P-value = 0.0002 while the R-square = 0.087, which is very low, indicates a weak correlation between PS. info Contain and Strength Evaluation. Even though the relationship is not as strong as expected, it still indicates that people put personal information in their password are more likely to evaluate their password to be relatively weaker comparing to people who don't.

PS. info Contain and PW Tested Strength

Surprisingly, there is no observed relationship between PS. info Contain and PW Tested Strength in this study with the P-value = 0.9361, R-square = 0.0. From this analysis, we can say that there is no relationship between personal information input situation and the password tested-strength.

PS. info Contain and Results Matching

From the regression analysis between PS. info Contain and Results Matching, there is no relationship between PS. info Contain and Results Matching since the P-value = 0.0089 (>.001), R-square = 0.0454, which indicates that there is a weak correlation between personal information input situation and password strength matching rate.

PS. info Contain and Awareness Evaluation

From the regression analysis between PS. info Contain and Awareness Evaluation, there is a relationship between PS. info Contain and Awareness Evaluation with the P-value = 1.27845E-10, R-square = 0.245853213, which indicates that there is a not strong correlation between personal information input situation and personal cybersecurity awareness.

Common ways of creating strong password

Given all the qualitative results collected by the survey, I classified all of them into ten following classifications, including:

- Special Characters (Letter Substitution)
- Extend Password Length
- Special Number
- Randomness
- Mix cases
- Make it very private
- Password Manager
- Pick letters from a sentence
- Other Information

The data was visualized in Figure 2., 27% of suggestions including putting more special characters along with letter substitution while they consider creating a strong password, which was the most suggested one. Whereas only 1% of the suggestions suggest on using a password manager or pick up letters in a certain sentence while creating a password, which was the least two suggestions while creating a strong password.

Interestingly, 3% of the suggestions suggested people put more private information in terms of the concern for both cybersecurity and password memorability.

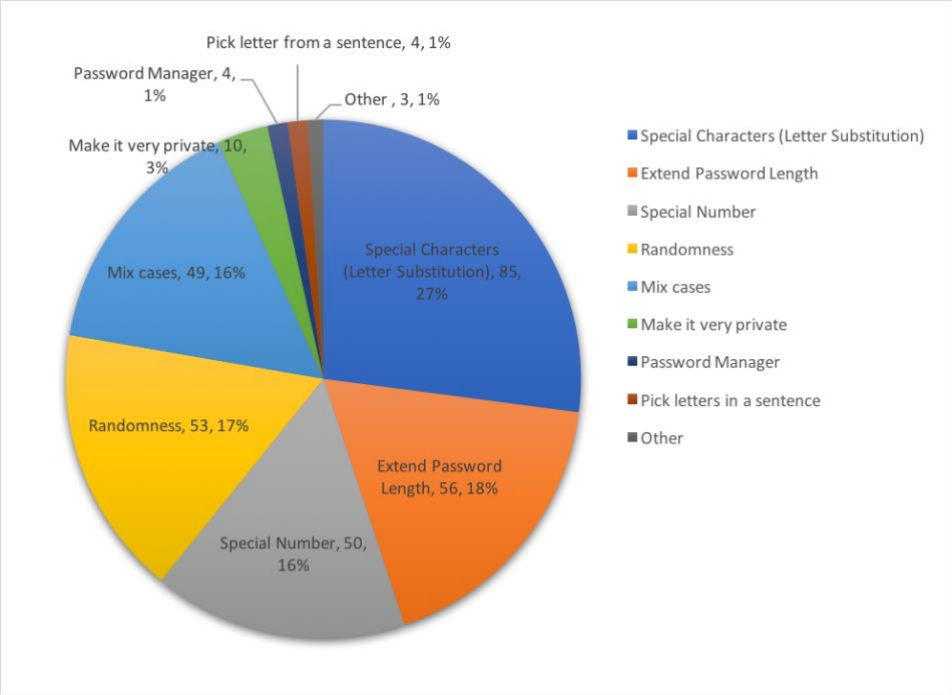

Figure 2. Common Ways for Creating Strong Password

DISCUSSION

From this study, the results indicate that there are more than 50% people are still using personal information in their password even it is not a suggested move for the concern of password security. With these results, the wide usage of personal information in the password is still a severe concern regarding cybersecurity issues. Hence, there should be more relative studies working in this area to provide a better understanding and guidance for people to increase people's password strength and cybersecurity awareness.

The regression analysis of PS. info Contain and PW Tested Strength turned out to be different than expected as containing personal information in a password can increase the risk regarding cybersecurity. In that case, people who chose to set personal information in their password intend to have relatively weaker cybersecurity awareness than the people who don't. But the tested password strength did not show a severe strength difference between people who put personal information in their password and people who don't.

For the analysis between PS. info Contain and Results Matching, this result turned out to be different than expected either, since people with personal information in their password are more likely to have lower cybersecurity awareness, so the results matching rate was expected to be lower. But the result indicates there is no found relationship between them, which might because of the password strength testing algorithm only evaluates the password strength based on password design itself, it did not take user's personal information into concern for the evaluation. In that case, the tested results of might not be accurate enough for this study.

For the analysis between PS. info Contain and Awareness Evaluation, the result surprisingly indicates that people with personal information in their password intend to have higher self-evaluation for cybersecurity awareness rather than lower, which might because of they are more cautious while doing the evaluation.

In order to provide a clearer guidance for people to create stronger and solid passwords with the concern of password security and personal information, the most suggested way to use special characters in the password with a letter substitution. Even putting personal information in the password can increase the possibility for one password being cracked, but with the combinations of using various special characters and letter, substitutions will make one password way harder to break it due to the uncertainties and irregularities.

Meanwhile, the usages of special numbers, extending password length, mix cases, and increase the randomness of a password was also widely suggested for creating better strength password. As all these three ways share the same idea while creating a

strong password, which is to increase the irregularities of a password. Extending password length could increase the difficulties and time of decode the password.

Picking up letters in a sentence is suggested few times for creating password also.

Using other information was suggested few times also regarding cybersecurity issues. Other information is still in the sense of personal information, but more like irregular personal information to make it harder to guess what information is put in the password, e.g. favorite song, favorite food, pet's name. This is the same idea as making the input personal information more private, which is aiming to prevent the leak of password information in the first place.

LIMITATIONS

As this study was designed under a storyboard with the preset password for a school account for the concern of generability, while school account password might not be able to represent all the possible situations related to password security. This might have led to relatively biased results regarding to the passwords for other accounts, the findings from this research study might not be able to apply well to other accounts perceive with less value, such as Netflix account, online shopping account etc.

Furthermore, the original research design did not collect any demographic inforamtion and password itself, the collected data was not able to indicate any other relationship regarding demographic information.

Hopefully, this study can provide a prior study within this field in order to better design further studies.

The algorithm used to test the password strength only focus on the password setup itself, but not consider the szainformation input related to user information, which made the test results exclude the effect of containing personal information. In that case, the tested results for the password with or without personal information might decrease the accuracy of password strength. At the same time, there is no best way to determine password strength to some extent after all, this fact might make the final results not explanatory enough to the real situation.

CONCLUSIONS

This study found out that people put personal who information in their passwords are more likely to evaluate their password to be relatively weaker comparing to people who don't put their personal information in their passwords

The results indicate that special characters including letter substitutions are the most used and most suggested way while creating a strong password.

Meanwhile, this study is proved to be consistent with the previous findings, which highlighted again that memorability issue should be focused on the most if aiming to resolve password security issue under cybersecurity context.

Moreover, the results of this study can be applied into relative password security education program to provide a more practical and persuasive instructions regarding cybersecurity.

This study provides us a relatively more clear and explanatory view about what should people do to increase their password security and how should they do to make it possible. Even though it would have performed a better analysis if collected more information including demographic information and passwords information from the participants, so we could get to observe and analyze more aspects from the password regarding personal information setup.

REFERENCE

1. Bishop, M., Klein, D.V., 1995. Improving system security via proactive password checking. Computers and Security 14, 233–249.
2. Bonneau, J. (2012, May). The science of guessing: analyzing an anonymized corpus of 70 million passwords. In Security and Privacy (SP), 2012 IEEE Symposium on (pp. 538-552). IEEE.
3. Bonneau, J., Herley, C., Van Oorschot, P. C., & Stajano, F. (2012, May). The quest to replace passwords: A framework for comparative evaluation of web authentication schemes. In Security and Privacy (SP), 2012 IEEE Symposium on (pp. 553-567). IEEE.
4. Chanda, K. (2016). Password Security: An Analysis of Password Strengths and Vulnerabilities. International Journal of Computer Network and Information Security, 8(7), 23.

5. Craigen, D., Diakun-Thibault, N., & Purse, R. (2014). Defining cybersecurity. Technology Innovation Management Review, 4(10).
6. Hossein Bidgoli. (2004). The Internet Encyclopedia (Vol. 3). John Wiley & Sons.
7. Li, Y., Wang, H., & Sun, K. (2017). Personal information in passwords and its security implications. IEEE Transactions on Information Forensics and Security, 12(10), 2320-2333.
8. Lu, H., & Ali, A. M. (2010). Making smart cards truly portable. IEEE Security & Privacy, 8(2), 28-34.
9. Riddle, B.L., Miron, M.S., Semo, J.A., 1989. Passwords in use in a university timesharing environment. Computers and Security 8, 569–579.
10. Schaffer, K. (2013). Passwords, Privacy, and Policies: Can They Do Business Together?. Computer, 46(12), 76-79.
11. Sharma, K., Nah, F., Cheng, M., Siau, K., Hall, R. (2017), Impact of Framing and Priming on Users' Behavior in Cybersecurity, WITS Feeder Conference on Security and Risk in SMAC (Social, Mobile, Analytics, and Cloud) and Cyber-Physical Systems, May 25-27.
12. Vu, K. P. L., Proctor, R. W., Bhargav-Spantzel, A., Tai, B. L. B., Cook, J., & Schultz, E. E. (2007). Improving password security and memorability to protect personal and organizational information. International Journal of Human-Computer Studies, 65(8), 744-757.