Is your password sexist? A gamification-based analysis of the cultural context of leaked passwords

Daniel Mølmark-O'Connor[1] and Emmanouil Vasilomanolakis[2]

[1] Aalborg University, Denmark
{doconn17}@student.aau.dk
[2] Technical University of Denmark, Denmark
{emmva}@dtu.dk

Abstract. Passwords are still the most common authentication method for various digital services. The majority of the research into passwords is focused on technical concerns rather than the human elements of password construction. In this paper, we aim at studying cultural aspects of leaked passwords with the usage of an online game. In particular, we introduce a novel web-based data collection tool utilizing gamification elements that benefits from appealing aesthetics and implemented narrative elements to engage users into prolonged play. The player's role is to label presented passwords with available descriptive tags. Our goal is to collect a large number of gaming data to identify prevalent tag choices through consensus, and as such, assign perceived meaning to the passwords through the tags. An initial field test of the prototype returned a high number of responses that were determined to be valid when assessed via internal controls.

1 Introduction

Passwords play a pivotal role in modern computing. They are still regarded as the most common form of user authentication in digital services. Despite this, user generated passwords remain a weak and exploitable security measure resulting from users generally creating passwords that are easy to crack [13, 6]. Research into passwords tends to favour investigations into the storage and security of passwords, such as research into the effectiveness and effect of password strength meters [12], or the guess-ability of a password [4].

The human-side of passwords, in the form of improper password creation practices, is a sizeable vulnerability in the authentication scheme, yet there remains little research in the area of the sociological factors of password creation. It poses the question: *What are users thinking when creating passwords?*. A survey of 470 Carnegie Mellon University computer users collected data on the behaviours and practices related to password use and creation when faced with new stricter password policies [10]. Results showed 19% of users had difficulty remembering passwords, and that over half the users reused or slightly modified

old passwords, often with the inclusion of special characters, as they were a new policy requirement. Almost 80% of users created passwords based on a word or name. Overall, the paper suggested that users create passwords that just meet the minimum technical requirements, and that common use of words or names is to assist memorability.

Further studies into password creation habits showed that users are conscious of the strength of their passwords, believed in incorrect security practices, and over estimated the privacy of their personal information [13]. Regarding how cultural factors play into password construction, a 2018 study of a metadata rich leak from a Middle Eastern bank showed that there were identifiable trends present that separated individuals from different cultures [1]. The aforementioned research suggests that there is a human context to password creation beyond just technical restraints.

In this paper, a solution to the lack of research into the sociological meaning of passwords is presented in the form of a game. Our prototype[3] takes real passwords from various data-leaks and assigns tags with some form of context or meaning to the password. With this data, it is possible to identify trends occurring across different data-leaks, such as establishing a prevalence of sexist or explicit passwords in one social media website when compared to another. These findings can assist in the research of password creation and further the understanding of not only *why* humans make weak passwords, but *how* they make weak passwords. Studies have shown that humans are great problem solvers and are motivated by assisting scientists [3, 11]. Our game takes this approach by being freely accessible to all interested parties who wish to take part in its quick and easy game-play through their web-browser. The game engages with the player and leverages the power of human reasoning to interpret the presented leaked passwords.

2 System design

Our prototype is split into a front-end, which is the game a user experiences, and a back-end, which is the database of passwords used and the storage of the user responses. The front-end is hosted on two indie game platforms and can be played in-browser or downloaded to play locally. The back-end database is hosted online through a web hosting service called Hostinger [8]. This MySQL database is populated with real passwords taken from 10 different sources, most of which are data-leaks from well known sources such as LinkedIn and NordVPN. The interaction between the front and back-end comes when a user boots up the game. The front-end contacts the database and returns a password at random to present on screen to the user. To ensure uniformity of the data, there is a check to ensure all passwords are labelled an equal number of times, or as close as possible. Once a user has assigned labels to the password presented, the database updates that password with this data and returns a new password to the user, and the loop continues.

[3] https://simmer.io/@AAUUser/password-labeller

The front-end is the game experience a user interacts with and can be seen in Figure 1. It was designed with the idea of promoting user engagement to prolong playtime. To do this, it leaned into game design theory [7], colour theory [5] and interaction design [9]. It also leaned into the gestalt principles of design for improved usability [2].

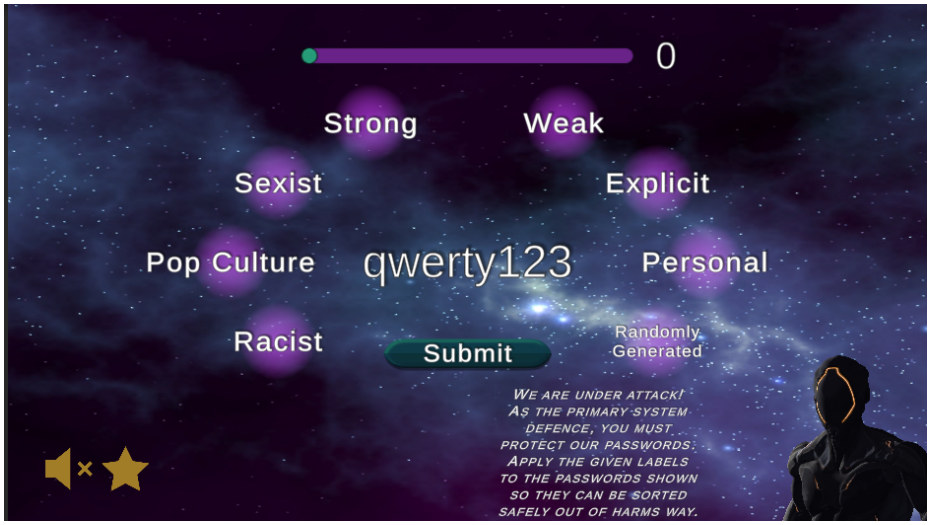

Fig. 1. Prototype gameplay

As depicted in Figure 1, the game consists of a (real world leaked) password placed in the centre of the screen surrounded by a collection of buttons with labels that can be chosen. The core game-play loop is for the user to interpret the password and chose the labels they feel appropriate before submitting and receiving a new password. A robot character provides encouragement and asserts a narrative in which the user is under attack and must label passwords to protect them. There are also points in the game-play where the user is presented with multiple choices associated with the narrative that results in a branching tree style story-line with 7 potential endings. Upon completing the game, the user is congratulated and given the option to play again.

3 Preliminary Results

The prototype was loaded with 1829 passwords taken from 10 different sources. Eight of these sources were legitimate passwords taken from data-leaks. In particular the data leaks were from LinkeDin, NordVPN, YouPorn (porn website), Ashley Maddison (online dating service), Hotmail, 000Webhost (web-hosting service), Muslim Match (Muslim dating website) and Faith Writers (Christian

focused website). The heterogeneity of the data sources allows us to experiment and analyze how the cultural context of passwords is altered on (very) different digital services. The two remaining sources were controls for validation, one being the top 20 most common passwords seen in 2022[4], and the other a list of 40 randomly generated passwords.

Eight labels were chosen to be presented to the users for the experiment. Three labels were negatively charged (i.e., Sexist, Racist, Explicit) to see the distribution of this sort of label across the sources. Two labels (i.e., Personal and Pop Culture) were chosen to examine how the passwords were perceived to connect to the creator and their surrounding culture. Finally, three labels (i.e., Randomly generated, Secure, Insecure) were chosen both to assist validation control, but also to gauge the users' perception of password strength in combination with the other labels.

This setup has been active for 26 days before the responses were analysed. There were 3074 instances of passwords being labelled with a total of 4012 labels across the total of 1829 passwords. The distribution can be seen in Figure 2.

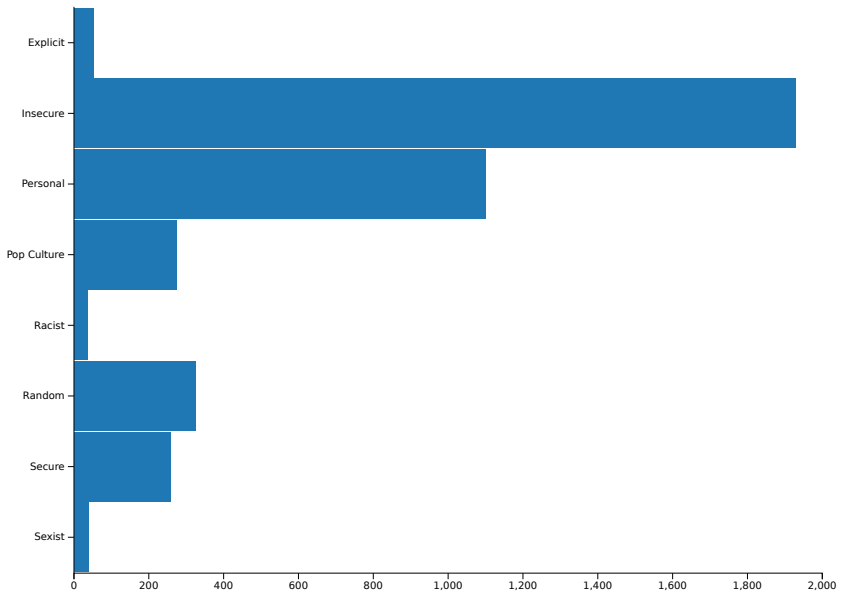

Fig. 2. Preliminary results of leaked password labels

The most frequently occurring labels across all sources were the *Insecure* and *Personal* labels. This is in line with related research which states users make passwords that are weak and based in common words and names to improve memorability. They were in fact the highest occurring labels for all sources

[4] https://www.tomsguide.com/news/worst-passwords-2022

other than the two control groups. Further results showed the *Pop Culture* label occurring twice as often in the Faith Writers (Christian religion related) website over a Muslim religion related website, which could be explained by the likelihood that the prototype was played primarily by western culture users. The negatively charged labels were most commonly seen in sex-related websites but also the Faith Writer website. On the one hand, it could be expected to see this representation in the sex-related websites as research shows users tend to be careless with passwords for these kinds of accounts. On the other hand, the presence of the Christian religion website could be explained by the potential negativity of the players towards religion or a false positive understanding of the meaning of the password.

Regarding the validity of the results, there are three factors that we took into account. First, none of the top 20 most common passwords of 2022 were labelled as secure. Second, there were twelve cases of a password appearing in multiple sources (between 2 and 4 times). It was seen that each of these incidences, that password was labelled identically. Finally, the randomly generated passwords were labelled as Random, Secure and Insecure in concurrence with what was expected based on the strength of their construction. These three results, on top of the expected result of seeing the high frequency of the Insecure and Personal labels, give validation to the experiment by conforming to expected outcomes, and attest that the data is reflecting the perception of the population regarding the meaning of the passwords.

4 Conclusion

In this work, we attempt a first look on the cultural context of leaked passwords. We design a prototype that utilizes gamification techniques and feeds with real world leaked passwords for users to examine and label passwords. Our preliminary results utilizing data from heterogeneous sources suggest: *i)* large portions of insecure passwords, *ii)* personal context being dominant in password creation and *iii)* negative labels (e.g., sexist and explicit) being dominant on porn or dating services.

The validation of the experiment shows that our prototype is getting accurate results, and the accumulation of 3074 user entries over 26 days shows the potential for much larger data gathering cycles. The experiment is considered a successful test run and gives motivation for the design and implementation of a second, larger experiment. This new experiment would benefit from a much larger experimental run-time and appropriate advertising. Furthermore, the labels and password sources could be refined to answer new questions. As the experiment runs and each password gets more and more instances of labelling, a picture will be drawn of a reflection of the populations perception of the interpreted meaning of passwords, and give context to passwords in a way we have never had before.

References

1. AlSabah, M., Oligeri, G., Riley, R.: Your culture is in your password: An analysis of a demographically-diverse password dataset. Computers & security **77**, 427–441 (2018)
2. Chang, D., Nesbitt, K.V.: Identifying commonly-used gestalt principles as a design framework for multi-sensory displays. In: 2006 IEEE International Conference on Systems, Man and Cybernetics. vol. 3, pp. 2452–2457. IEEE (2006)
3. Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D., Popović, Z., et al.: Predicting protein structures with a multiplayer online game. Nature **466**(7307), 756–760 (2010)
4. Dell'Amico, M., Michiardi, P., Roudier, Y.: Password strength: An empirical analysis. In: 2010 Proceedings IEEE INFOCOM. pp. 1–9. IEEE (2010)
5. Ferris, K., Zhang, S.: A framework for selecting and optimizing color scheme in web design. In: 2016 49th Hawaii International Conference on System Sciences (HICSS). pp. 532–541. IEEE (2016)
6. Florencio, D., Herley, C.: A large-scale study of web password habits. In: Proceedings of the 16th international conference on World Wide Web. pp. 657–666 (2007)
7. Fullerton, T.: Game design workshop: a playcentric approach to creating innovative games. CRC press (2014)
8. Hostinger: Hostinger web hosting service, `https://fold.it/`, last Accessed: 30/05/22
9. Preece, J., sharp, H., Rogers, Y.: Interaction Design: Beyond Human-Computer Interaction. John Wiley & Sons Inc (2015)
10. Shay, R., Komanduri, S., Kelley, P.G., Leon, P.G., Mazurek, M.L., Bauer, L., Christin, N., Cranor, L.F.: Encountering stronger password requirements: user attitudes and behaviors. In: Proceedings of the sixth symposium on usable privacy and security. pp. 1–20 (2010)
11. Sullivan, D.P., Winsnes, C.F., Åkesson, L., Hjelmare, M., Wiking, M., Schutten, R., Campbell, L., Leifsson, H., Rhodes, S., Nordgren, A., et al.: Deep learning is combined with massive-scale citizen science to improve large-scale image classification. Nature biotechnology **36**(9), 820–828 (2018)
12. Ur, B., Kelley, P.G., Komanduri, S., Lee, J., Maass, M., Mazurek, M.L., Passaro, T., Shay, R., Vidas, T., Bauer, L., et al.: How does your password measure up? the effect of strength meters on password creation. In: 21st USENIX security symposium (USENIX Security 12). pp. 65–80 (2012)
13. Ur, B., Noma, F., Bees, J., Segreti, S.M., Shay, R., Bauer, L., Christin, N., Cranor, L.F.: " i added'!'at the end to make it secure": Observing password creation in the lab. In: Eleventh Symposium On Usable Privacy and Security (SOUPS 2015). pp. 123–140 (2015)