Poster: Waiting Makes the Heart Grow Fonder and the Password Grow Stronger

Experiments in Nudging Users to Create Stronger Passwords

Nathan Malkin, Shriram Krishnamurthi, David H. Laidlaw
Brown University, Department of Computer Science
Providence, RI
{nmalkin, sk, dhl}@cs.brown.edu

1. INTRODUCTION

Despite passwords' long history and present ubiquity, the general population exhibits password habits that are widely regarded as poor. Reuse is rampant, and users consistently choose weak (i.e., easy to crack) passwords, which is especially dangerous in light of increasingly sophisticated attack methodologies and exponentially faster modern hardware.

This paper contributes the design and evaluation of novel techniques for nudging users to create stronger passwords through the use of waiting periods. Users who experienced voluntary or forced timeouts in the process of password creation chose stronger passwords. The improvements were found to match or exceed those from other techniques, such as strength meters and monetary incentives. Additionally, we find further evidence that time spent on password creation correlates with password strength.

2. RELATED WORK

Pushing for Stronger Passwords

System administrators have employed a variety of approaches to combat the problem of weak passwords. One traditional technique has been the use of password composition policies: rigid requirements that force passwords to include capital letters, digits, special characters, etc. However, users find these policies annoying [3], and they are not necessarily more effective than simply requiring longer passwords [2].

Another popular[1] tactic is the use of password strength meters. These meters do not force any changes to the passwords users create but provide visual feedback about their strength, serving as a way of educating users about their habits. They have been found to result in longer passwords, but only stringent meters increase resistance to password-cracking algorithms [4].

Time and Security

In their work on strength meters, Ur et al. [4] also found evidence of a relation between password creation and time. Participants who received feedback from more stringent meters took significantly longer to create their passwords.

Using time as an independent variable, Egelman et al. [1] studied the question of user tolerance of security delays, finding that users exhibited limited tolerance for delays in software and were significantly more likely to cheat on a task in the absence of a detailed, security-minded explanation about the reasons for the delay.

No work to date, however, explores time as an independent variable in the password creation process.

[1]Ur et al. found them used by 70 of the top 100 websites, as ranked by Alexa [4].

3. EXPERIMENTAL FRAMEWORK

Password creation was evaluated under laboratory conditions. Subjects participated through Mechanical Turk, where the task was advertised as "a short survey about yourself and your browsing habits" in order to avoid priming participants by focusing their attention on passwords.

As part of the task, participants had to create a password that they would use when returning to finish the survey. They were paid $0.05 and $0.10 for completing the two parts of the survey, respectively. Additionally, 50% were randomly chosen to receive a bonus of $0.05.

While creating their passwords, users saw a strength meter that scored their password according to the `comprehensive8` heuristic [2,4]. The meter displayed a numeric score in addition to traditional feedback mechanisms such as color and coarse ratings (e.g., "poor," "great," etc.).

The strength meter's score was also used as the dependent measure for comparing password strengths across conditions. While this is not the most accurate measure of the password's susceptibility to real-world attacks, we are examining decisions by users to create more or less secure passwords, and their intent is best measured by the score they see on-screen.

4. CONDITIONS AND RESULTS

Participants

A total of 146 subjects participated in our study, approximately $n = 20$ per condition. Of these, 49% reported themselves as male, and 51% as female. 69% resided in the United States, while 31% were from India. 21% said they worked in a field related to Information Technology, while 75% said they did not. The median age was 29.

Voluntary Timeouts

In our first condition, users were able to opt in to timeouts by pressing a button and waiting for 5 seconds at a time (up to a total of 20 seconds). Since there is no intrinsic motivation for users to commit themselves to waiting, we attached a small monetary incentive to this action: opting in to a 5-second timeout increased the probability of the randomly determined bonus by 5%.

Password	A	B	C
Terrible	20 seconds	20 seconds	20 seconds
Bad	20 seconds	15 seconds	20 seconds
Average	20 seconds	10 seconds	20 seconds
Good	20 seconds	5 seconds	20 seconds
Great	20 seconds	no wait	no wait

A: universal waiting period

B: graduated waiting period

C: graduated waiting period with stronger incentives

Table 1: Mandatory Timeouts

Results from this condition demonstrate majority participation: 60% opted into timeouts. Of these, two thirds waited for only one period, and one third waited four periods (20 seconds) for the maximum expected payoff.

Interestingly, despite no incentives for stronger passwords, participants in this condition created significantly stronger passwords ($p < 0.05$)[2] than those in a control condition (without timeouts).

Forced Timeout

One possible explanation for stronger passwords in the presence of timeouts is that, while participants were waiting to accrue the extra probability (which they chose to do to increase their expected payoff), they were able to dedicate more time to creating a stronger password.

To test this hypothesis, we introduced a new condition with a mandatory waiting period. Participants were asked to stay on the password creation screen for 20 seconds, with no explanation for the delay and, this time, no monetary incentives (Table 1A).

The password scores that resulted were again significantly higher ($p < 0.05$) than those in the control. However, despite an increased total waiting time, they were not stronger than those in the voluntary condition.

Longer Timeout

We have seen that waiting resulted in stronger passwords. Does waiting even longer result in even better passwords? To address this question, we repeated the experiment with the mandatory waiting period extended to 40 seconds. Though the median password score was slightly higher, we saw no significant improvement.

Graduated Timeout

Our findings so far suggest that it may be possible to nudge users into creating stronger passwords by making them wait. However, websites typically wish to minimize the amount of time users have to spend on tasks like account creation, and the waiting period penalizes those who create strong passwords quickly: there is no reason to keep them waiting.

In light of this, a graduated waiting period may be more appropriate: the weaker the password, the longer the wait (Table 1B). As a result, those inclined to choose weak passwords would be motivated to make them stronger.

Though we expected performance with this mechanism to be on par with previous results, the average password in this condition was weaker. Furthermore, we were unable to reject the null hypothesis that this condition and the control were drawn from the same population ($p = 0.28$).

[2]We use the Least Significant Difference post hoc procedure for a one-way ANOVA as the significance measure.

Stronger Incentives

One explanation for the diminished strength is that the intermediate "penalties" were not strong enough: a 5-second timeout, for example, is not worth the effort of changing one's password.

To test if this is the case, we introduced harsher penalties: users with the best passwords still had no waiting period, but all others were required to wait for 20 seconds (Table 1C).

Contrary to expectations, the harsher penalties did not improve results. The median password strength increased only slightly, with still no statistically significant difference from the control ($p = 0.23$).

Conclusions

Our findings show that, for workers on Mechanical Turk, waiting periods — under some, but not all, circumstances — can result in stronger passwords. Less ambiguously, the results from our study as a whole support time as an important explanatory variable. A multiple regression of password strength on a variety of factors — including demographics, recall technique, and the presence of priming, incentives, and strength meters — found that time spent creating a password was a significant predictor of strength ($p < 0.001$), accounting for a greater portion of a password's strength score than all other variables.

5. REFERENCES

[1]: S. Egelman, D. Molnar, N. Christin, A. Acquisti, C. Herley, S. Krishnamurthi. 2010. Please Continue to Hold: An Empirical Study of User Tolerance of Security Delays. *Workshop on the Economics of Information Security* (WEIS '10).

[2] P. G. Kelley, S. Komanduri, M. L. Mazurek, R. Shay, T. Vidas, L. Bauer, N. Christin, L. F. Cranor, and J. Lopez. 2012. Guess Again (and Again and Again): Measuring Password Strength by Simulating Password-Cracking Algorithms. In *Proceedings of the 2012 IEEE Symposium on Security and Privacy* (SP '12).

[3]: R. Shay, S. Komanduri, P. G. Kelley, P. G. Leon, M. L. Mazurek, L. Bauer, N. Christin, and L. F. Cranor. 2010. Encountering stronger password requirements: user attitudes and behaviors. In *Proceedings of the Sixth Symposium on Usable Privacy and Security* (SOUPS '10).

[4]: B. Ur, P. G. Kelley, S. Komanduri, J. Lee, M. Maass, M. L. Mazurek, T. Passaro, R. Shay, T. Vidas, L. Bauer, N. Christin, and L. F. Cranor. 2012. How does your password measure up? the effect of strength meters on password creation. In *Proceedings of the 21st USENIX conference on Security symposium* (Security '12).