CSEP564 : Computer Security : Reading X

Karuna Sagar Krishna

November 18, 2024

Paper Title

Of Passwords and People: Measuring the Effect of Password-Composition Policies

Paper Authors

Saranga Komanduri, Richard Shay, Patrick Gage Kelley, Michelle L. Mazurek, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, and Serge Egelman

Problem

The authors aim to understand the relationship between password composition policies and the strength of the resulting passwords. They also aim to understand the effect of such policies on user behavior on how they create or reuse passwords, how they remember or store passwords and how they cope with failures during password creation and hence user sentiments.

Approach

At high level, the authors approach is to conduct a large scale study (5000 participants) to gather empirical and survey data. The authors analyze this data and find various interesting insights that shed lights to questions that they set out to answer.

The authors provide the background of previous work and note that passwords are becoming more complex over time due to wide spread use of technology across various parts of human society. The authors explain their methodology on how they conducted their study. In short, this was a 2 part study; in first part users were provided a scenario along with password composition policy and users were asked to create passwords and fill in a survey. In second part which occurred couple of days later asked the users to remember their passwords and fill in another survey. The paper describes various details of the study including the different conditions and scenarios - 5 conditions spanning 2 scenarios. Scenarios provided the background to the users - one described need for a password to protect low value account and the second was for a high value account. The various conditions combined composition policies with scenarios leading to 5 conditions - basic8survey, basic8, dictionary8, basic16, comprehensive8.

The paper provides various details about analysis starting with understanding the demographics of the users/participants in this study. To have a rationale empirical understanding we need a formal password strength metric and the authors use a variation of Shannon's method for information entropy. Entropy is calculated for various elements of the password and summed up; this additive entropy has improved accuracy with smaller sample size than the traditional approach. The authors use various statistical methods like chi-square test,

Fisher's Exact Test and Permutation testing to determine the statistical significance of their observations and analyze the differences between entropy estimates.

Some of the interesting finds are - though NIST estimates similar entropy for basic16 and comprehensive8, the authors find that basic16 has more entropy; numbers provided lot more entropy than characters and symbols; symbols were used less frequently than numbers and median length consistently exceeded requirements. Though adding dictionary checks did not result in entropy increase, it did strengthen password against heuristic guessing used by popular password cracking tools. Comprehensive8 was least user friendly since more users failed to create it successfully or drop out of the study. Dictionary checks made it hard for users to create passwords. Storing passwords led to passwords with higher entropy.

Conclusions

From their study, basic16 is the clear winner. The authors point out that there is a tradeoff between the empirical strength of passwords and the usability of passwords under various password composition policies. The paper concludes by identifying several misconceptions - adding number significantly increases the entropy, dictionary checks provide less entropy though they are effective against heuristic cracking and users typically create passwords that exceed requirements.

New Ideas

The authors idea of using additive entropy seems interesting since we can understand the entropy contribution by various parts of a passwords i.e. entropy contribution by password length, use of numbers, symbols, characters and their positions, etc. This makes entropy composable metric.

Passwords are created and used by humans. So the idea of weaving empirical data with user behavior and sentiments is important and interesting.

Improvements

The paper feels incomplete without talking about using different languages in passwords. The online study could have been easily extended to include users from different countries and languages.

The paper doesn't clarify why certain composition policies were chosen and doesn't mention other policies if there are any.

New Directions

A new direction to explore - given the increased use of multi-factor authentication, password managers and passkeys, are password composition policies still important and add value?

As noted in the paper, entropy alone is not sufficient metric. So, what are other metrics would be useful to measure password strength. And in general, what metrics would be needed to measure security of authentication mechanism in general.