Exploring the Usability of Pronounceable Passwords

Shing-hon Lau, Stephen Siena, Ashutosh Pandey, Sroaj Sosothikul, Lorrie Cranor, Blase Ur, Richard Shay
Carnegie Mellon University
Pittsburgh, PA, USA
{shinghol, ssiena, ashutosp, ssosothi} @andrew.cmu.edu, lorrie@cmu.edu, bur@cmu.edu, rshay@cmu.edu

1. INTRODUCTION

Text passwords are prevalent in computer systems, mainly due to their simplicity. While other forms of authentication such as graphical passwords, password tokens or biometrics exist, none have supplanted text passwords. However, one disadvantage of text passwords is that users are forced to remember many different passwords, often generated under different policy restrictions. To remember these passwords, users tend to rely on tricks and patterns that tend to result in easily guessable, and thus insecure, passwords.

One method of ensuring secure passwords is to use system-assigned passwords. However, these tend to be random strings of characters that are difficult for users to remember. A promising alternative is the use of pronounceable passwords. Shay et al. have previously demonstrated that pronounceable passphrases show promise compared to random passwords [3]. However, few researchers have examined pronounceable passwords in great depth.

In this preliminary work, we compare 3 non-pronounceable (heretofore called *random*) and 4 different pronounceable system-assigned passwords. We performed a 699 person, three part Mechanical Turk study to investigate the memorability and likability of these passwords. We explore two different non-pronounceable password generators: a slight modification of the existing Gasser algorithm [2] and our own novel pronounceable password generator.

2. RESEARCH QUESTION

In this work, we explore two different research questions:

1. Are pronounceable passwords more memorable than random passwords?
2. Do users like pronounceable passwords more than random passwords?

2.1 Password Conditions

We explored seven different password conditions. Three random conditions were controls and four conditions were pronounceable. Table 1 contains examples of each condition, as well as the level of entropy. The conditions are:

1. **rand-char-5:** 5 random characters chosen randomly from a 64-character set of lowercase and uppercase letters, digits, and symbols. Characters with ambiguous appearance (e.g. 'O' and '0') were removed. The set of characters is:
   `abcdefghjkmnpqrstuvwxyz`
   `ABCDEFGHJKLMNPQRSTUVWXYZ`
   `23456789@!$*#%-&_.`
2. **rand-char-7:** The same as rand-char-5, but with 7 characters.
3. **rand-low-7:** 7 random lowercase letters chosen from the entire alphabet.
4. **pro-gas:** 8 letter pronounceable strings generated by a slightly modified version of Gasser's algorithm [2]. We use an implementation[1] that makes all possible strings equi-probable, to close a vulnerability noted by Ganeson et al. [1].
5. **pro-pair-4:** 4 concatenated *atoms*; An *atom* consists of a consonant sound (1 or 2 letters) followed by a vowel. We use 40 consonant sounds; 21 single consonants (including 'qu'), and 19 two letter consonant combinations. The 40 consonant sounds are: `b,c,d,f,g,h,j,k,l,m,n,p,r,s,t,v,w,x,y,z bl,br,dr,ch,cr,dr,fl,fr,gr,gl,pl,pr,qu,sh,sl, sp,st,sw,th,tr`.
6. **pro-pair5:** The same as pro-pair-4, but with 5 atoms.
7. **pro-pair-3d2:** pro-pair-5 with 2 digits inserted after the 2nd or 3rd atom.

Condition	Entropy	Examples
pro-gas	30.2	cytuchva, oktesauc
pro-pair-4	30.6	rishespuhi, spesterudre
pro-pair-5	38.2	huthuslawoce, rehipuweya
pro-pair-3d2	45.2	sujohu46spucra, revi56yoguwa
rand-char-5	30	y_Qzw, a$_A3
rand-char-7	42	US$#-P5, L#%mwQg
rand-low-7	32.9	vfkmlqc, hrkvtuf

Table 1: Password conditions. Entropy (in bits) of the 7 password conditions, along with example passwords.

3. METHODOLOGY

3.1 Experimental Design

We conducted a 700-participant between-subjects experiment of the 7 password conditions using Amazon's Mechanical Turk. Subjects were assigned in a round-robin fashion to each of the 7 conditions. The study consisted of three surveys: 1) an initial survey, 2) a follow-up survey conducted 1

[1] `http://www.adel.nursat.kz/apg/` (visited 5/2012)

day after the initial survey, and 3) a follow-up survey conducted 1 week after the initial survey. The differing delays on the follow-up surveys allow us to test short-term and long-term memorability of the password conditions. Subjects were compensated 50 cents, 70 cents, and 80 cents, for the three surveys, respectively.

The initial survey consisted of subject consent, password assignment, a questionnaire about demographic information and sentiments towards the assigned password, followed by a password-recall task. Follow-up surveys consisted of a password recall task followed by a short questionnaire about password memorability. Our survey forms are largely derived from the forms used by Shay et al. [3]. However, we have two followup surveys to test both short-term and long-term memorability while they only had a single followup survey that took place 2 days after the initial survey.

3.2 Data Analysis

We used both omnibus tests and pairwise tests. For quantitative data, we used a Kruskal-Wallis omnibus test. For categorical data, we used a χ^2 omnibus test and a Fisher's Exact Test for pairwise comparisons. We used a Bonferroni-Holm correction for multiple testing. We used $\alpha = 0.05$ for all significance results presented.

4. RESULTS

4.1 Demographics

Our study was conducted entirely in April of 2014. A total of 754 subjects began the first survey and 699 completed it. 555 participants returned for the second survey and 450 also returned for the third survey. Each condition had between 62 and 69 subjects complete all three surveys. We focus our analysis only on the subjects that completed all three surveys.

Out of the 450 subjects that completed all three surveys, 174 (39%) were women and 276 (61%) were men. The mean age was 32 years, while the median was 29 years. The standard deviation was 10.5. The youngest subject was 18 years old and the oldest subject was 82 years old. 86 (19%) subjects reported that they had degrees or jobs in "computer science, computer engineering, information technology, or a related field", 360 (80%) reported that they did not, and 4 declined to answer. 280 (62%) of our subjects reported that they had an associate's degree, a bachelor's degree, or graduate degree, 166 (37%) did not, and 4 declined to answer.

We found no statistically significant differences between any of the conditions in terms of gender, age, degree type, or education. There were also no statistically significant differences in the dropout rate between conditions.

4.2 Storage Behavior

During the second and third surveys, we asked participants if they stored their password by writing it down on paper or electronically storing it.

There was no statistical difference, across conditions, for storage usage reported for the second or third studies. 251 (56%) of our subjects reported that they did not use storage, which comprise the *no-storage subjects*.

4.3 Recall

We asked our subjects to recall their password after the initial survey and before they began each of the two follow-up surveys. On all three occasions, there was no significant difference between conditions in the number of attempts required to enter the password or the fraction of subjects that successfully recalled the password. This was true when considering all subjects together and also when considering only the no-storage subjects.

4.4 User sentiment

Users were generally positive towards the pronounceable passwords, but we did encounter two specific issues.

First, users were often concerned about finding the "correct" pronunciation of a password. For example, one subject commented that "The first part, 'blay' can be 'bligh' or 'blaaay'. The 'rusho' at the end can be 'ruhsho' or 'roosho'."

Second, users were concerned that the pronounceable passwords may not offer enough security since they did not have uppercase letters, digits, or symbols, which are typically associated with Strong passwords. For example, one subject commented that "I would have made it more difficult by adding capital letters and numbers/symbols".

5. DISCUSSION

The most surprising result from our study is that we found absolutely no difference in ability to recall between any of the 7 conditions. The fact that subjects were able to recall long pronounceable passwords at the same rate as short, random passwords was quite surprising. This suggests that pronounceable passwords may be able to offer additional security without negatively affecting user ability to recall passwords.

The challenges with pronounceable passwords seem to be two-fold. First, users need to be taught how to take advantage of the benefits of pronounceable passwords. Second, users must be assured that they are secure, despite the deviation from what is generally considered a strong password.

6. CONCLUSION

In our preliminary exploration of pronounceable passwords, we have found them to be just as memorable as traditional strong passwords. However, pronounceable passwords face some challenges moving forward. Users need to be educated about how to take maximum advantage of pronounceable passwords and educated about how they can provide security while deviating from traditional strong passwords.

7. ACKNOWLEDGMENTS

Thanks to Richard Shay for his assistance in managing the online study.

8. REFERENCES

[1] R. Ganesan, C. Davies, and B. Atlantic. A new attack on random pronounceable password generators. In *Proceedings of the 17th {NIST}-{NCSC} National Computer Security Conference*, 1994.

[2] M. Gasser. A random word generator for pronounceable passwords. Technical report, DTIC Document, 1975.

[3] R. Shay, P. G. Kelley, S. Komanduri, M. L. Mazurek, B. Ur, T. Vidas, L. Bauer, N. Christin, and L. F. Cranor. Correct horse battery staple: Exploring the usability of system-assigned passphrases.