A study of user password strategy for multiple accounts

Conference Paper · February 2013

DOI: 10.1145/2435349.2435373

CITATIONS
17

READS
443

3 authors:

S.M. Taiabul Haque
University of Texas at Arlington
9 PUBLICATIONS **56** CITATIONS
SEE PROFILE

Matthew Wright
Rochester Institute of Technology
95 PUBLICATIONS **2,281** CITATIONS
SEE PROFILE

Shannon Amerilda Scielzo
University of Texas Southwestern Medical Center
88 PUBLICATIONS **256** CITATIONS
SEE PROFILE

A Study of User Password Strategy for Multiple Accounts

S M Taiabul Haque
Department of CSE
University of Texas at Arlington
Arlington,TX USA 76019
eresh03@gmail.com

Matthew Wright
Department of CSE
University of Texas at Arlington
Arlington,TX USA 76019
mwright@cse.uta.edu

Shannon Scielzo
Department of Psychology
University of Texas at Arlington
Arlington,TX USA 76019
scielzo@uta.edu

ABSTRACT

Despite advances in biometrics and other technologies, passwords remain the most commonly used means of authentication in computer systems. Users maintain different security levels for different passwords. In this study, we examine the degree of similarity among passwords of different security levels of a user. We conducted a laboratory experiment with 80 students from a public university in the southern United States. We asked the subjects to construct new passwords for websites of different security levels. We collected the lower-level passwords constructed by the subjects, combined them with a comprehensive wordlist, and performed dictionary attacks on their constructed passwords from the higher-level sites. We could successfully crack almost one-third of their constructed passwords from the higher-level sites with this method. This suggests that, if a user's lower-level password is leaked, it can be used effectively by an attacker to crack some of the user's higher-level passwords.

Categories and Subject Descriptors

D.4.6 [**Security and Protection**]: Authentication; H.1.2 [**User/Machine Systems**]: Human Factors

General Terms

Security, Human Factors

Keywords

Security, usability, passwords, laboratory experiment

1. INTRODUCTION

Password-based authentication is considered to be the most popular method of user authentication, mainly because of its simplicity and cost effectiveness. However, it is by no means a panacea as long as usability is concerned. As formulated by Wiedenbeck et al., a good password, with its associated character of being an easy-to-remember and a hard-to-guess sequence of characters, presents a dilemma [6]. Naturally, words that are easy to recollect from memory are short, single words found in dictionaries, or slight variations of them. A tendency to choose such words as passwords makes them susceptible to dictionary attacks.

CODASPY'13, February 18–20, 2013, San Antonio, Texas, USA.
ACM 978-1-4503-1890-7/13/02.

This password management problem is aggravated by the fact that users need to maintain multiple accounts that require passwords. In a large-scale study, Florêncio et al. reported that Internet users, on average, maintain 25 password-protected accounts [2]. An average user is not expected to be sufficiently equipped on a cognitive level to deal with 25 different passwords. In fact, Adams and Sasse reported that a typical user can be expected to cope with at most four or five passwords effectively [1]. Due to this cognitive capacity constraint, users reuse passwords across different sites, with little or no modification.

Many studies have been conducted for understanding the password habits of users. Researchers from industry periodically gather large-scale data about passwords and publish many insightful lists[1]. Florêncio et al. from Microsoft Research did a landmark study that involved half a million users and revealed many interesting findings about different user password habits [2]. Academic researchers, on the other hand, use various novel methods and laboratory studies to observe more closely a particular password behavior (e.g., password reuse habit) of a sample population. Shay et al. capitalized on the opportunity of Carnegie Mellon University (CMU) password policy change [5], while Gaw et al. gathered feedbacks from users after they had made actual login attempts in different websites [3]. Although these papers have reported about password reuse, our work attempts to look at finer-grained aspects of reuse – how similar are passwords to one another across sites, and how do they vary with perceived security level.

Our study was inspired by the work of Notoatmodjo et al., where they confirmed that users mentally group their accounts and tend to make stronger passwords for accounts that they consider more important [4]. Users have different levels of incentive to protect their different accounts. In this study, we examined how vulnerable the higher-level (webmail or banking account passwords) passwords of a specific user would become, if the lower-level (online news account or weather portal passwords) passwords of that user could be compromised. Our results showed that, almost one-third of the higher-level passwords of the participants could be cracked by using the lower-level passwords and a comprehensive worldlist. This demonstrated that, the knowledge of a password of a lower risk account seems to increase the chance to crack higher risk account based on similarity to the lower risk password.

[1]SplashData recently published its annual list of the most common passwords used on the Internet. The list can be found at http://splashdata.com/press/PR121023.htm.

2. METHODOLOGY

We conducted a laboratory experiment with 80 students from a public university in the southern United States. Although a larger number of participants could have been drawn from an online study, we preferred a laboratory study because our pilot study (N=12) showed that a laboratory study would produce more consistent responses. Students were assigned partial course credits in exchange for their participation. The complete study was approved by the local Institutional Review Board (IRB).

In our study, we considered online banking accounts and accounts in all kinds of merchant sites like Amazon.com or Ebay.com as financial accounts. Webmail accounts and social networking accounts were considered as identity accounts. Users are always concerned about the security of financial accounts because it is always important for them to keep their hard-earned money safe. Users also have a lot of incentives to protect the security of identity accounts because they build long term reputations of trust in their professional and personal lives through identity accounts.

On the other hand, users create accounts in some websites only to customize the contents of those sites. No significant interaction with other users or financial transaction happens through these accounts. Online news websites and search portals belong to this category. Users do not have much incentive to protect the security of these content accounts.

It is unlikely that all of the password-protected accounts of an individual user belong to the category of well recognized identity, financial or content sites as mentioned above. In our study, we considered users' accounts in all kinds of little recognized websites as sketchy accounts. It includes unfamiliar sites that claim to have various kinds of deals, little known online forums or content provider sites. Users have the least incentive to protect the security of these accounts.

We designed a PHP script that prompted the users to create passwords for their new accounts in eight different websites of these four different categories:

- Financial website : Chase.com and Wellsfargo.com
- Identity website : Yahoomail.com and Facebook.com
- Content website : Nytimes.com and Weather.com
- Sketchy website : Dreamdeals.com and Justchill.com (hypothetically constructed sites)

We selected Chase.com and Wellsfargo.com as representatives of banking/financial websites because these two banks should be familiar to the participant students due to the prevalence of their ATMs on the campus. Facebook.com and Yahoomail.com were selected as identity websites, mainly because of their popularity as social networking site, and webmail site, respectively. For content websites, we carefully selected Nytimes.com and Weather.com because these two sites readily present a clear distinction between identity sites and content sites, without any requirement of explicitly labeling them as content sites.

We did not want to give the participants any clue about our experimental motive because we expected the participants to spontaneously construct new passwords, exactly in the same way as they do in real life. Therefore, for all the six real sites, we designed the interfaces in such ways that they would look similar to the original sites. For the two hypothetical unfamiliar sketchy sites, we gave their interfaces very much informal looks.

2.1 Password Construction

For ethical and security reasons, we explicitly told the participants not to provide any of their existing passwords. For each website, we provided a brief introduction and presented a real-life scenario. For example, for Weather.com, the participants were presented with the following scenario:

> Weather.com provides the latest weather forecasts, maps, and alerts. You want Weather.com to show weather for your local city when you go to the site. To do that, you need to register an account on Weather.com so that you can customize your location. Imagine that you are registering a new account on Weather.com. You have reached the final step of registering your new account, and you need to input a password. Proceed to the next page to input your new password.

As they proceeded, the password construction page for Weather.com appeared.

2.2 Password Policy

For all the six real websites, we enforced exactly the same password policies as they are enforced in those sites. For the two hypothetical sites, we ensured that the passwords provided by the participants should be at least five characters long. Like original sites, participants were also required to retype their passwords in a second box, which prevented them from typing some random characters as their passwords.

In this way, we implicitly tried to trigger the real life password creation mechanisms of users for websites of different security levels. In designing the interfaces and providing the introduction for each site, we were thoroughly careful about not revealing the participants that our main objective is to categorize their constructed passwords based on financial, identity, content or sketchy websites.

3. RESULTS

We collected the passwords that were constructed by the participants and grouped together the passwords of the same category. We analyzed each group separately to find out the frequency of using capital letters, digits and special characters. We also calculated the lengths of the passwords. The length and the frequency values decreased as the security levels of the sites decreased. Figures 1 and 2 summarize our analysis.

Next, we tried to crack the financial and identity (higher-level) passwords of a participant by using the participant's content and sketchy (lower-level) passwords. For cracking purposes, we used John The Ripper (JTR) password cracker. We combined the "wordlist" mode of JTR with the "single crack" mode.

The "wordlist" mode cracking is basically a dictionary attack where every word in a wordlist is tried against the candidate password until a match is found. If word mangling rules are enabled, each word in the wordlist is modified or mangled to generate other possible combinations. The "single crack" mode is the default cracking mode of JTR where a large number of word mangling rules are applied to a small dictionary to perform a dictionary attack.

As the default set of word mangling rules is very small in the "wordlist" mode, we modified the configuration file

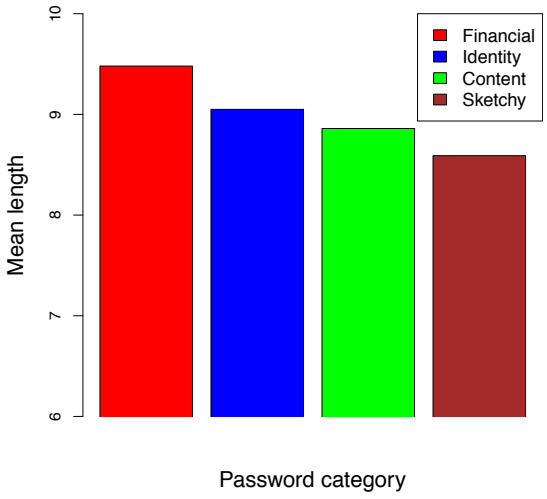

Figure 1: A comparison of mean lengths.

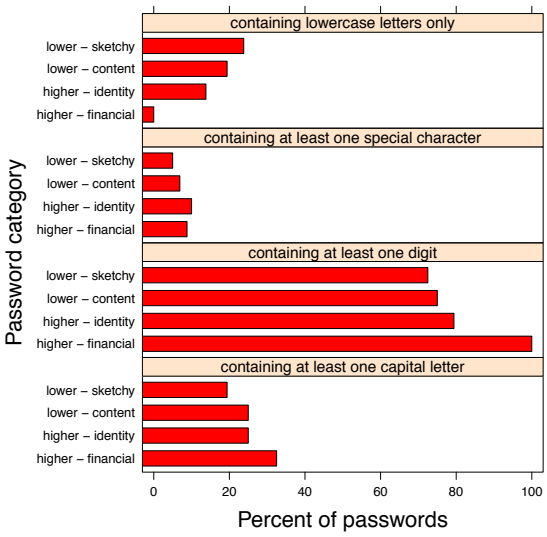

Figure 2: A comparison among passwords of different category.

of JTR so that it is possible to apply the large set of word mangling rules of the "single crack" mode while performing cracking in the "wordlist" mode.

For each participant, we combined the participant's four lower-level passwords with the Cain & Abel wordlist and tried to crack four higher-level passwords of that participant in our modified "wordlist" mode. Among the 320 higher-level passwords, we could successfully crack 33.1% (106) of passwords with this method.

We also tried to crack the four higher-level passwords of a participant by using the four lower-level passwords only, without using any wordlist. We could successfully crack 19.1% (61) of higher-level passwords this time. This demonstrated that, passwords used at a higher security level have a strong degree of syntactic similarity with the passwords used at a lower level.

4. DISCUSSION AND FUTURE WORK

We do not dispute the fact that our experimental setting was artificial. However, we tried our best to achieve realism in our experiment. For each website, we presented a real-life scenario to the participants and the scenario was created in such a way that it would resemble a real world application as much as possible. Although our sample size was not large, it can be considered as a reasonable one, compared to the sample sizes of [3], [4], and [6], which were also laboratory experiments among students. Finally, we note that the presence of an observer may, if anything, motivate users to create stronger passwords than they might otherwise.

Our cracking methodology through JTR relied only on syntactic similarity. Through word mangling rules, it modified the lower-level passwords in various ways in order to guess the higher-level passwords. The semantic similarity was not considered at all. For example, multiple passwords of a user can be inspired from a common source (e.g., music, film, sports etc.). If one of the passwords of a user is related with a personally meaningful word (e.g., name of the pet dog), it is probable that another password of that user is also inspired by a similar thing (e.g., color of the pet dog). Our cracking methodology did not leverage these kinds of semantic similarity. We believe that, by exploiting the semantic similarity, a larger percentage of higher level passwords can be cracked. We leave this as a future work.

In addition of asking the participants to construct new passwords, we had them answer a survey regarding their password behaviors for websites of different security levels. We plan to thoroughly review the responses of the survey and report in greater detail how users manage a range of passwords for websites of different security levels.

5. REFERENCES

[1] A. Adams and M. A. Sasse. Users are not the enemy. *Commun. ACM*, 42(12):40–46, 1999.

[2] D. Florêncio and C. Herley. A large-scale study of web password habits. In *Proceedings of the 16th international conference on World Wide Web*, pages 657–666, May 2007.

[3] S. Gaw and E. W. Felten. Password management strategies for online accounts. In *Proceedings of the second symposium on Usable privacy and security*, pages 44–55, July 2006.

[4] G. Notoatmodjo and C. Thomborson. Passwords and perceptions. In *Proceedings of the Seventh Australasian Conference on Information Security - Volume 98*, pages 71–78, January 2009.

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

[6] S. Wiedenbeck, J. Waters, J.-C. Birget, A. Brodskiy, and N. Memon. Authentication using graphical passwords: effects of tolerance and image choice. In *Proceedings of the 2005 symposium on Usable privacy and security*, pages 1–12, July 2005.