Reuse and Recycle: Online Password Management

Shirley Gaw
Dept of Computer Science
Princeton University
35 Olden Street
Princeton, New Jersey 08540
sgaw@cs.princeton.edu

Edward Felten
Dept of Computer Science
Princeton University
35 Olden Street
Princeton, New Jersey 08540
felten@cs.princeton.edu

ABSTRACT

While password login is the dominate method for online authentication, the effect on users' creation and (re-)use of passwords is unclear. We can assume users visit many websites and have many online accounts, but the number of sites and the number of passwords may be much smaller than previously indicated [3]. This study attempts to quantify how many passwords people have and how often passwords are reused. Additionally, we were interested in who respondents thought could compromise their passwords and whether they understood how they could create stronger passwords.

1. INTRODUCTION

When asked why they use the same password across different, unrelated websites, respondents indicated human memory limitations:

> *Female, Public and International Affairs*: Because I can't/don't want to keep track of dozens of different passwords for dozens of different websites (or which passwords I used for which websites)
>
> *Female, English*: Because I can't keep track of a million different passwords.
>
> *Female, Molecular Biology*: I don't have the ability to remember 50 different passwords for all the sites I go to.
>
> *Male, Economics*: I have about 40 websites where I have a username and password...*(edited)*

In fact, the above statements also indicate that respondents would be overwhelmed if they had to manage the *large collections* created by having a unique password for each website. The respondents could be exaggerating the problem, however. How many passwords do people actually use? How many websites do they visit that require online authentication?

Even though password login is the most common authentication method on the the Internet, the literature indicates that users do not follow standard advice on password management. Prior work has studied password choice [2] and factors influencing password management [4, 1].

2. OVERVIEW

The goal of our survey was to collect usage statistics on how people reuse and manage their passwords for online authentication. Participants completed a questionnaire followed by a laboratory exercise. This was a multi-part survey, where topics included the tools used for password management; the use of the same password without changing it; the reuse of same password for different online accounts; the perceived likelihood of threat scenarios; and the meaning of tips for generating new passwords. Finally, there was a laboratory exercise on quantifying password reuse. Due to space constraints, we will focus on the last three parts of the study.

Participants were recruited through on-campus flyers and snowball sampling. 44 undergraduate and one graduate student (10 male, 35 female) have participated in the survey so far. Time to complete the survey ranged from 45 minutes to 90 minutes. Some participants left parts of the survey incomplete, but participants were paid $10 for completing the entire survey. As the survey is ongoing, we expect to complete further analysis in a future paper. Here we present from the data collected so far.

3. PERCEIVED THREAT QUESTIONNAIRE

The online questionnaire asked participants to rank types of people by their likelihood to compromise passwords. While we would have preferred to present categories with all combinations of three independent characteristics (personal relationship, computer expertise, affiliation), we believe this would have been too many choices for meaningful rankings. Instead, the population was partitioned unevenly: someone you know well, someone without computer expertise that you have met, someone with computer expertise that you have met, someone from your organization that you do not know, someone from a competing organization that you do not know, and someone that is unaffiliated that you do not know.

When asked to rank people by their ability, their motivation, and their likelihood to "access information without permission", respondents indicated highly varied threat models. The 45 respondents were instructed to consider only ability of the attacker. From the respondent rankings, those that were considered to be the most able attackers were either someone they knew well (53%, $N = 43$) or an acquaintance with computer expertise (26%). The least able attackers were an acquaintance without computer expertise (35%, $N = 42$) or an unaffiliated stranger (33%).

Especially interesting are rankings that considered only the motivation of the attacker. The people considered most motivated to attack were a stranger from a competing organization (40%, $N = 43$), an unaffiliated stranger (30%), or someone they knew well (26%). At the same time, the least motivated attackers were also someone they knew well (35%), an unaffiliated stranger (33%), or an acquaintance without computer expertise (23%). In other words, 13 respondents thought an unaffiliated stranger was most motivated to attack, and, at the same time, 14 respondents thought this person was least motivated. 11 saw someone they knew well as being most motivated, but 15 saw this person as the least motivated.

This same contrast is found when looking at the overall likelihood of being attacked. Those considered most likely to attack were an acquaintance with computer expertise (30%), an unaffiliated stranger (30%), or someone they knew well (23%). Yet, those ranked least likely to attack were acquaintances without technical expertise (40%) and also unaffiliated strangers (33%) or people they knew well (21%). Here again, someone they knew well was both "most" (10) and "least" (9) likely to attack. Unaffiliated strangers were also both "most" (13) and "least" (14) likely to attack.

4. PASSWORD LOGIN TASK

In the laboratory exercise, we were interested in quantifying password reuse. We expected participants would have difficulty estimating how many passwords they have and how often they reused their passwords. Instead of asking for estimates directly, we instructed participants to first complete a login task. We presented them with a list of 139 websites which used login authentication. We asked the participants to indicate which websites they used; participants attempted to login (or re-login) to these websites and wrote down their passwords. Participants were instructed to keep the list hidden from the researchers' view. The lists were destroyed with a shredder at the end of the exercise.

Using the password lists, participants answered questions about how many passwords they had and specifically how many semantically-related passwords and structurally-related passwords they had. Structurally-related passwords included reusing a password along with reusing a password after transforming it (character capitalization, addition, and/or removal). One question specifically asked participants to quantify how many times they reused a password without transformation. Participants first used the password list created by the login exercise to answer these questions, but afterwards, they were instructed to answer the questions again using a list of all passwords they could recall.

We were unable to find much literature quantifying how many passwords people had. [3] indicated respondents had an average of 19 passwords (passwords were not limited to online accounts). The respondents in our survey had far fewer passwords. The mean estimated number of recalled passwords for online accounts was 6.1 ($SD = 3.6$, $N = 35$) and the mean number of unique passwords was 3.2 ($SD = 1.9$). While mean number of times passwords were reused without transformation was 3.7 ($SD = 4.2$), this includes one outlier who repeated one of his password across 25 accounts. Without this response, the mean number of times participants reused a password across online accounts was 3.1 ($SD = 2.2$, $N = 34$).

The small number of unique passwords is consistent with respondents indicating they agreed with the statement "If I reuse a password, it is easier for me to remember it" ($M = 4.7$, $SD = .6$, $N = 39$) where 1 = "Strongly Disagree" and 5 = "Strongly Agree". Respondents also tended to agree that "It is easier to reuse a password than create a new one" ($M = 4.6$, $SD = .9$); unfortunately, respondents also disagreed with the statement "If a password is generated for me by a website then I use this password instead of one of my normal passwords" ($M = 1.7$, $SD = 1.3$). There are multiple interpretations of this response; one possibility is that respondents have trouble understanding why passwords are generated for them; we are interested in what further analysis of their responses to the password generation quiz indicate.

5. PASSWORD GENERATION QUIZ

The final portion of the study quizzed participants about their understanding of suggestions for generating new passwords (ie, "Use uppercase and lowercase letters", "Drop letters from a familiar phrase"). Participants were presented with the following scenario:

> Many websites have tips and rules for creating strong passwords. Pretend your friend Eve Jones (evjones@princeton.edu) is also a student at Princeton and she is having trouble understanding these rules. For each rule or tip, she's provided three example passwords with an explanation of how she created her password. Help her learn what makes a strong password by ranking her examples from strongest to weakest and explaining your ranking.

There were 13 password generation tips. Explanations of how the fictional Eve created her passwords described how the passwords related to her personal information and how the passwords were transformed to the displayed password. For example, "I spelled 'eve's password' backwards. I then removed the apostrophe." Future publications will describe the results in more detail.

6. REFERENCES

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

[2] H. Petrie. Password clues. `http://www.centralnic.com/news/research`. Accessed 2 May 2005.

[3] M. A. Sasse, S. Brostoff, and D. Weirich. Transforming the 'weakest link' a human/computer interaction approach to usable and effective security. *BT Technology Journal*, 19(3):122–131, 2001.

[4] D. Weirich and M. A. Sasse. Persuasive password security. In *CHI '01: CHI '01 extended abstracts on Human factors in computing systems*, pages 139–140, New York, NY, USA, 2001. ACM Press.