Password Multiplier Evaluation

XXX

May 23, 2006

1 Password Multiplier Evaluation

There are two orthogonal goals to the study. First, we are surveying novice users to learn about password management behavior. Second, we are using the same sample population to evaluate the Password Multiplier (PM) Firefox plug-in.

An abstract of PM is described in XXX. PM is a stateless password manager where the user has a single-master password. When logging in, the user need only double-click on the password field, type in their master password, and hit the OK button to have their website-specific password generated into the correct field of the website. The master password can be cached for a session to speed up the process.

1.1 Procedure

We should recruit 15 to 20 undergraduates participate in the study, where groups of 3-5 per session. We will use our department computer labs and machines will have Firefox already installed. Participants will be asked to bring any information that helps them remember their passwords (notebooks, bank statements, PIN cards, printouts of files, etc.) The study will then have four parts:

1. 15 minutes. Signing Consent forms and Getting Orientation.
2. 50 minutes. Logging into websites with personal information.
3. 25 minutes. Installing and using PM.
4. 20 minutes. Answering an online questionnaire.

The remaining 10 minutes will be slack time to transition between parts of the study.

In the second part of the study, the participants will login to websites with their own usernames and passwords. To reduce the time spent recalling commonly used websites, I have collected a list of approximately 100 websites with logins (XXX). In the study, we will ask the participants which websites they use. Then, the users will be asked to login to each of their websites with accounts using the Firefox browser, indicating whether or not they were successful:

- I was able to login.
- I do not have some login information. Check all that apply:
  - Username

- E-mail address
- Password
- Other (please specify)

For successful logins, we will ask participants to write down their password for the website. Loading and logging into the New York Times Online takes approximately one minute. In the 50 minute section, participants should be able to login to 20-30 websites.

In the third part of the study, participants will install the PM plugin to Firefox. We will provide users with the login information (username, e-mail, password) for dummy accounts on 5-10 websites. We will avoid personal websites such as Mail.com and Orkut where users may be confused with their own accounts. We may also provide a master password for PM. The task is a bit undefined. We want users to use PM to generate unique passwords for their websites, but we also want the users to spend time logging in and websurfing. Perhaps we will have them use PM to change the password of one or two websites, but have them login using this password and read one portion of the website before moving on to the next one.

In the fourth part of the study, participants will answer questions about their satisfaction with PM, rating it on a Likert scale (1 = Strongly Disagree, 3 = Neutral, 5 = Strongly Agree):

1. It was easy to create passwords with Password Multiplier.
2. It was difficult to login with a password created by Password Multiplier.
3. It is easier to use Password Multiplier than my current method of creating passwords for websites.
4. It is more difficult to use Password Multiplier than to use my current method of creating passwords for logging into websites.
5. For logging into multiple websites: it is easier to use Password Multiplier for website login than using a single, universal password.
6. For logging into multiple websites: it is more difficult to use Password Multiplier for website login than using a file or piece of paper to lookup my password.
7. After learning how to use Password Multiplier, I prefer my current method of creating passwords for websites.
8. After learning how to use Password Multiplier, I prefer Password Multiplier to my current method of storing passwords for websites.
9. After learning how to use Password Multiplier, I prefer my current method for recalling passwords for websites.

To learn more about password selection, participants will be asked to review their password sheets and count how many unique passwords they have, count how many times they reuse their passwords, and count how many permutations of their passwords they use.

Finally, we will want to ask a few open ended questions about password selection and management.

1. When was the last time you needed to create a new password for logging into a website?
   - How did you create your new password?
   - How did you store your new password?
   - What's another another method for creating or storing a password?

2. Are there two websites where you use different passwords?
    - If no: Why not?
    - If yes: Why do you have different passwords for these sites?
3. Are there two websites where you use the same password?
    - If no: Why not?
    - If yes: Why do you have the same passwords for these sites?
4. Is there a website where you have memorized your password for logging in?
    - If no: Why not?
    - If yes: What are some reasons why you have this password memorized?
5. When logging into different websites, some people use the same password for more than one website.
    - Why do you think someone would do this?
    - What benefits does someone have doing this?
    - What problems does someone have doing this?
    - How do you think you can get around these problems?
6. When logging into different websites, some people use one password for one website and another password for another website.
    - Why do you think someone would do this?
    - What kind of benefits do you think you might have doing this?
    - What kind of problems do you think there might be doing this?
    - How do you think you can get around these problems?
7. How many passwords do you think one person should have for the websites they visit? Do you have this many passwords? Why or why not?
8. Describe the last time you logged into a website and couldn't remember the password. How long ago was this? How did you login to the website?

After answering the questionnaire, we will shred the password logs kept by the participants during the study. Thus, we will not have access to personal password information.

1.2 Analysis

While the sample of websites is not a random sample of all websites, it will present qualitative data on what users remember about websites where they have accounts. We can report the number of tasks completed (answering if they have an account), the fraction of websites where people report having accounts, and the fraction of websites where people report remembering password information.

For the PM evaluation questionnaire, we plan on running chi-squared analysis with the expected value of 3 (Neutral).

As the password information is not a manipulated condition, we will only report a summary of fraction of recalled logins using unique passwords, reused passwords, and similar passwords. Unfortunately, this measure is inherently biased as we use self-reported measures (people may have incentive to lie about

the security of their online behavior) and it is biased as the participants only provide information about passwords they remember.

We will also provide summaries of responses on password selection and management, although I have not devised a coding scheme for answers yet. Perhaps the questions should be fixed responses instead of open ended.

1.3 Side notes

The login tasks in the second portion of the study have high variability in difficulty. Participants may remember login information for some websites more easily than others. Secondly, there is no consistency in the layout of websites for logins. Some participants may have fewer online accounts than others. It will be difficult to report summary statistics if there is a high variability in the number of logins completed.

We would prefer subjects used their own accounts when using PM, however, this biases our task; websites where participants have the login information are the least likely accounts where PM would be useful for recalling and storing passwords. Additionally, we would need to provide time for participants to revert their passwords to their original ones. We use dummy accounts for the third part of the study instead. We assume that using PM demonstrates how easy it would be to generate and recall unique passwords for every website and that participants will generalize this application to all of their websites with login accounts. Since participants also indicate which websites they do not remember their login information, we assume that subjects will be interested in PM for all website logins, not just the dummy account logins.

Designing the questionnaire presents a particular difficulty because of possible bias in the language of the question, yet we want to collect preference data: "It was easy to create passwords with PM." The Woodrow Wilson Survey Research Center may help with removing bias from the questionnaire.

We may also want to survey current methods for password management, reporting how many participants indicated they store their passwords on paper/file, for example. Since these may be known as "bad practices", the questions in the survey tended to ask "how could you get around these problems" for using unique passwords for each website. More direct questions might be added though.

I am particularly concerned that we will not have much data. Will 20-30 websites be enough? 50 minutes is already a long period to ask people to perform a task without breaks.

While we would ideally compute the information about unique, reused, and similar passwords, we run into the difficulty of actually attempting to compute these on the raw data. In the current design, participants will never show their password information to the investigators and the data are easy to report.