A Diary Study of Password Usage in Daily Life

Eiji Hayashi and Jason I. Hong

October 6, 2010

CMU-CyLab-10-016

CyLab
Carnegie Mellon University
Pittsburgh, PA 15213

A Diary Study of Password Usage in Daily Life

Eiji Hayashi
ehayashi@cs.cmu.edu

Jason I. Hong
jasonh@cs.cmu.edu

Carnegie Mellon University
Pittsburgh, PA

ABSTRACT
While past work has examined password usage on a specific computer, web site, or organization, there is little work examining overall password usage in daily life. Through a diary study, we examine all usage of passwords, and offer some new findings based on quantitative analyses regarding how often people log in, where they log in, and how frequently people use foreign computers. Our analysis also confirms or updates existing statistics about password usage patterns. We also discuss some implications for design as well as security education.

Author Keywords
Password, User Authentication, Diary Study

ACM Classification Keywords
H5.3. Information interfaces and presentation: Miscellaneous.

INTRODUCTION
Text-based passwords are the most commonly used authentication system today. There have been multiple studies investigating password usage, including people's selection of passwords [8], strength and memorability of user chosen passwords [9,11,12], and the number of passwords and accounts users have [3,4]. There are also studies investigating password usage in companies [5], as well as the effects of password policies on users' practices [6]. However, a relatively small amount of work [10] has investigated password usage in our daily lives.

In this paper, we present the results of a two-week diary study examining how participants used passwords in their everyday lives, spanning the entire day (as opposed to just work settings, as in [5]), as well as across all computers and services (as opposed to a single computer, as in [4]). This paper also contributes to our understanding of password usage by providing quantitative data on the number of logins and the number of accounts accessed. Finally, our analysis also includes aspects of password usage that have not been previously investigated, such as where people log into their accounts, what type of computer they used (e.g. users' computers or foreign computer), and what types of password aids they used to manage their passwords. These findings can help in the design of systems that facilitate password management.

RELATED WORK
Many studies have investigated how people use passwords. Adam and Sasse conducted a study focusing on people's attitudes towards password authentication systems. They found that if the authentication system did not mesh well with people's work practices, people tended to circumvent the authentication system, in effect undermining the security of the authentication system [1].

Inglesant and Sasse found that even if organizations enforced strict password policies on users, the policies did not guarantee security for certain attacks (such as key loggers) while also frustrating users [5]. Shay et al., through a survey of 470 students, showed that the students were annoyed when university adopted a new password policy requiring more complex passwords, but at the same time, the students felt more secure [6]. These studies imply that choosing appropriate password policies is crucial to making password authentication systems more secure and usable.

There have been other studies including password usage outside of organizations. Singh et al. conducted a qualitative study about how people shared banking passwords with spouses or significant others [7]. Gaw and Feltman interviewed 49 undergraduate students and found that the students had 7.8 accounts on average, with three or less passwords [3]. Florencio and Herley deployed software to more than a half million clients over three months. They reported that they observed about 25 online accounts per *client*. [4]. The difference in the number of online accounts between Gaw et al.'s study and Florencio et al.'s could be because multiple users shared one client.

In this paper, we present the results of a diary study investigating in what contexts people use passwords in their daily lives, examining password usage across all computers, services, and settings. Our analysis also provides novel data, such as where people log into their accounts and how frequently they use foreign computers. Furthermore, our

analysis confirms or updates some password statistics, such as the number of accounts people have.

METHOD

We provided small diaries to participants and asked them to carry the diaries throughout their day. We asked participants to record each *password event* in their diaries when they log into their accounts using desktop computers or laptops. Password events included typing passwords to log into online accounts and computers, unlocking screensavers, and logging into applications (e.g., email clients). Even if an authentication system automatically filled passwords, we instructed the participants to record the password event, as long as the participants had to click some button (e.g., login or OK). In contrast, if a system automatically logged into an account, we asked them not to record the password event since some participants would have difficulty distinguishing if they actually went through an authentication process. On the first day of the study, we asked participants to clear cookies in their browser to log out from all of their online accounts.

At the password events, the participants recorded contexts, including the participant's location, the purpose of the password event, type of computer they were using (e.g., personal or public computer), and whether they used a password aide (e.g., a sheet of paper with a list of passwords, or a piece of software to help with passwords).

The study lasted for two weeks from July 1st to July 14th. At the end of the study, we asked participants to complete a post-survey. We compensated participants $20 USD.

PARTICIPANTS

We recruited 20 participants using a university recruitment web site. Nine participants were male and 11 participants were female. Twelve participants were university students, two participants were university staff, and six participants were domestic residents. Their ages ranged from 21 to 59 with a median age of 29. In the survey, we examined the participants' expertise levels by asking whether participants agree or disagree with various statements (see Table 1). In general, the participants were comfortable using computers and estimated their expertise as average.

Table 1 Estimate of the participants' computer expertise (1 = strongly disagree and 5 = strongly agree). For the bottom row, one stands for novice and five stands for expert.

Sentence	Mean (SD)
I'm comfortable with using email	4.6 (0.49)
I'm comfortable with using web browsers	4.6 (0.49)
I'm comfortable with purchasing products on-line	3.8 (1.09)
I'm comfortable with configuring computers	3.3 (1.00)
How do you estimate your computer expertise?	3.1 (0.89)

ANALYSIS

We start with a descriptive analysis of the data. We collected exactly 1,500 password events. Figure 1 shows the distribution of password events per participant observed in the study period. The numbers ranged from 11 to 169 password events, with a mean of 75 (σ=35.3). The most common purpose of these events was to log into online services (75.6%), followed by to log into computers (20.3%), to use applications on computers (7.4%), and to unlock screensavers (3.3%). The small ratio of "unlocking screensavers" implied that a small number of participants were using passwords to unlock screensavers. In our post-survey, we also found that only three participants had screensavers that required passwords to be unlocked. These results implied that there are opportunities to design better user authentication systems for screensavers to facilitate its adoption.

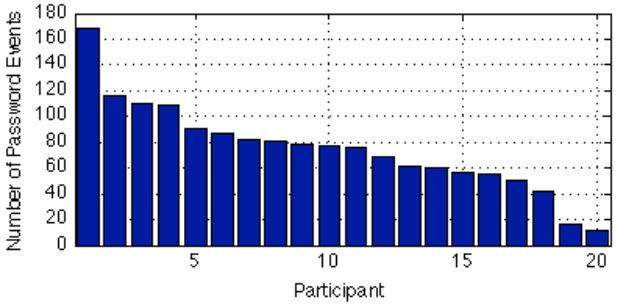

Figure 1 Distribution of password events across 20 users, sorted from most events to least. Most users accessed their accounts 40 to 110 times over a two-week study period.

ONLINE ACCOUNTS

In the study period, we observed 172 online accounts in total. The numbers ranged from 3 to 16 with a mean of 8.6 accounts (see Figure 2).

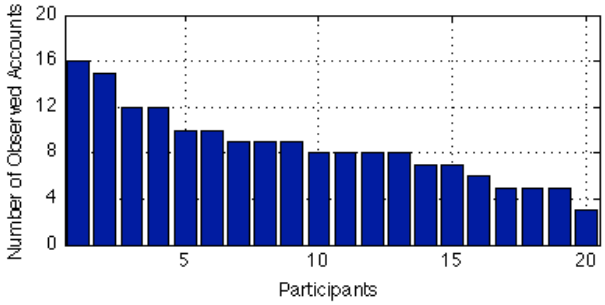

Figure 2 Distribution of the number of the observed accounts for each participant. The participants were sorted according to the number of the accounts, so this x-axis does not correspond to the x-axis in Figure 1.

Figure 3 shows the relationship between the number of days and the average number of accounts per participant observed. The dashed lines show one standard deviation. In the first two days, we observed five accounts. After that, the number increased to 8.6 constantly. Florencio et al. [4] reported that it took 60 days for this number to be saturated. They also reported that they observed about 70% of the online accounts in the first 14 days. Thus, we estimated that participants had about 11.4 online accounts. This estimated number of online accounts is slightly larger than in Gaw et al.'s study conducted in 2006 (7.8 accounts per a user) [3].

One implication here is that, while password aids do need to scale for users with vastly more accounts, systems that

can help people with about a dozen accounts would still be valuable. Another implication is that, for novel authentication systems, they should be tested for interference and memorability with roughly this number of accounts, rather than just for one (which is typically done).

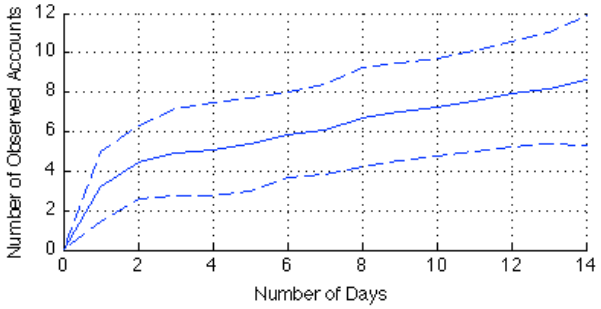

Figure 3 Cumulative number of the online accounts observed by day. The dashed lines stand for one standard deviation.

To facilitate analysis, we first categorized each account according to Google's categorization [13]. For the web sites not included in the Google's list, we manually categorized them using the same scheme. Then, we coded these categories into eight broader categories as shown in Table 2. "Email/Messaging" denotes webmail services, such as Gmail, or messaging services, such as Twitter. "Online Community" includes social networking sites or online forums. "University / Company" denotes web pages specific to universities or companies, such as online course registrations or work hour management system. "Portal" denotes pages such as MSN or Yahoo top pages. "Application" denotes online applications provided on web pages, such as Google Docs or Doodle.

Table 2. This table shows the number of accounts, number of the password events, and mean number of events per account for each category. While email/messaging consisted of 19% of the accounts, it consisted of 40% of the password events.

Category	# of Accounts	# of Events	Mean
Email/Messaging	33	418	12.7
Online Community	29	165	5.7
University/Company	17	128	7.5
E-commerce	35	95	2.7
Portals	10	73	7.3
Applications	16	69	4.3
Finance	14	37	2.6
Others	19	49	2.6
Total	173	1034	6.0

Table 2 shows the number of accounts in each category, as well as the number of password events per category. Email and messaging had the largest number of password events, with 40% of all password events. Note that there were 33 accounts in this category, as some participants had multiple email/messaging accounts.

Email/messaging, university/company, and portals were the three most frequently used categories, consisting of 34.7% of the total number of the accounts while covering 59.9% of the total number of the password events.

LOCATIONS AND COMPUTERS

We also let participants record their locations as well as the kind of computers used at password events. Table 3 shows the locations and the number of password events observed at those locations. 84.3% of the events were observed at either home or office. In contrast, only 6.9% of the events were observed in public places, such as libraries. Even if we include school as a public place, the total is 13.1%. Among the 20 participants, nine participants accessed their accounts only from home or office.

Table 3. Categorization of locations where participants accessed their accounts. 84.3% of the password events occurred either in home or office.

Place	# of Events
Home	889 (59.2%)
Office	377 (25.1%)
Public Places	104 (6.9%)
School	93 (6.2%)
Others	37 (2.4%)

Table 4. Categorization of computers that the participant used at password events. 93.9% of the time, the participant used either their personal computers or work computers.

Computer	# of Events
Personal	996 (66.4%)
Work	413 (27.5%)
Public	60 (4.0%)
Friends'	17 (0.9%)
Others	14 (0.9%)

Table 4 shows the type of computers that the participants used. We defined "personal computers" as computers primarily used by the participants for personal purposes, and "work computers" as computers primarily used by the participants for work-related purposes. Public computers were computers that anyone can access, such as those in libraries or in computer labs at universities. Friends' computers were computers owned by participants' friends.

We observe that 93.9% of password events occurred on either personal or work computers. There were 91 (5.8%) accesses from foreign computers (a public, a friend's, or other computer). Two participants accessed their accounts from foreign computers 45 times in total (accounting for half of the data) and nine participants never accessed their accounts from foreign computers. Naturally, those participants overlapped with those who accessed their accounts only from home or office.

Given that the vast majority of our participants only use their personal or work computers, and close to half of our participants do not login in public places at all, these findings suggest that if we can make the login process easier just for users' work and home computers, it can provide considerable benefit to a large number of users.

Furthermore, with the growing diffusion of location-based services, these findings suggest that we may be able to use one's current location at home or work as an additional factor in authentication. For instance, it is possible to build a screensaver that does not require a password to be

unlocked when a laptop is at home or work, but would require a password or perhaps additional authentication in other places. Additionally, since many individuals access their accounts in similar contexts (e.g., the same locations, on the same computers, with the same printers and devices nearby), authentication systems could utilize these contexts to modulate the level of authentication required. These approaches could potentially improve the security of an authentication system without adding burden to users.

PASSWORD AIDS

In the post-survey, we asked participants what password aids they used to manage the accounts observed over the study period. We also asked them to self evaluate how concerned they would be if someone obtained access to that account. We missed three accounts due of lack of data in the survey. Thus, we only had 169 accounts in this analysis.

Surprisingly, Table 5 shows that for 60.3% of the accounts, participants did not use any password aids. This carries two implications. First, according to the survey, all participants except one reused their passwords for multiple accounts. Given that people chose not to use any passwords aids for important accounts, this suggests that people realized that writing down important passwords is risky, but did not realize that reusing passwords is also risky. Although we do not know which passwords were reused (i.e., important vs. not important accounts), educating users about these risks seems prudent based on our data. Second, the low rate of adoption of password aids suggests that there is still a lot of room for helping people, and examining barriers to adoption may be a fruitful approach to improving security.

Table 5 Rows denote types of password aids. Columns denote participants' self-evaluation of how concerned if someone obtains accesses to these account, 5 denotes *very concerned* and 1 denotes *not concerned at all*.

	5	4	3	2	1	Total
No using password aids	46	27	22	6	1	102 (60.3%)
Browsers' auto-fill features	21	5	12	10	2	50 (30.0%)
Writing down on paper	5	1	0	0	1	7 (4.1%)
Dedicated password manager	0	0	0	0	0	0 (0.0%)
Others	4	2	0	3	1	10 (5.9%)
Total	76	35	34	19	5	169

LIMITATIONS

One of the biggest limitations in our study was participants' demographics. Although our participants involved university staffs and domestic residences, 60% of the participants were university students. Thus, our participants may not represent the general population.

Another limitation is that our study did not capture auto-logins using cookies or session information. As a result, while our data helps estimating users' actual workload, it under-estimate the frequency of user authentication.

Moreover, our study was limited to password events using computers. As other forms of computers, such as smart phones or tablet computers, people would have to use their passwords in wide variety of contexts. Further investigation would be necessary for password usages on these devices.

Finally, our study period could be short for some analyses. In the analysis of the number of accounts, the number did not saturate in the study period. Similarly, in the analysis of password aids, we may have observed larger number of infrequently used accounts, for which the participants might use different types of password aids.

CONCLUSION

Through a diary study, we collected 1,500 password events, which illustrated how participants used passwords in their everyday lives. The analyses of the data provided several implications about user authentication systems. We hope that this paper contributes to further investigation and development of user authentication systems.

REFERENCES

1. A. Adams, M. A. Sasse. Users are not the enemy. Comm. of the ACM (1999)
2. K.R. Walsh, B. Ives, H. Schneider. The Domino Effect of Password Reuse. Comm. of the ACM (2004)
3. S. Gaw, E. W. Felten. Password management strategies for online accounts. In Proc. of SOUPS (2006)
4. D. Florencio, C. Herley. A large-scale study of web password habits. In Proc. of WWW (2007)
5. P. Inglesant and M. A. Sasse. The true cost of unusable password policies: password use in the wild. CHI (2010)
6. R. Shay, S. Komanduri, P. G. Kelley, P.G. Leon, M. L. Mazurek, L. Bauer, N. Christin, L. F. Cranor Encountering stronger password requirements: user attitudes and behaviors. In Porc. of SOUPS (2010)
7. S. Singh, A. Cabraal, C. Demosthenous G. Astbrink M. Furlong. Password sharing: implications for security design based on social practice. In Proc. of CHI (2007)
8. E. F. Gehringer. Choosing passwords: security and human factors. In Proc. ISTAS (2002)
9. J. Yan, A. Blackwell, R. Anderson, A. Grant. Password memorability and security: Empirical results. IEEE Security & privacy (2004) vol. 2 (5) pp. 25-31
10. P. Dourish, E. Grinter, J. D. Flor, M. Joseph. Security in the wild: user strategies for managing security as an everyday, practical problem. Personal and Ubiquitous Computing (2004) vol. 8 (6)
11. C. Kuo, S. Romanosky, L. Cranor. Human selection of mnemonic phrase-based passwords. SOUPS (2006)
12. D. Florencio, C. Herley, B. Coskun. Do strong web passwords accomplish anything? In Proc. USENIX Hot Topics in Security (2007)
13. Google 1000 most-visited sites on the web. http://www.google.com/adplanner/static/top1000/