Available online at www.sciencedirect.com
ScienceDirect
Int. J. Human-Computer Studies 65 (2007) 744–757
International Journal of Human-Computer Studies
www.elsevier.com/locate/ijhcs

Improving password security and memorability to protect personal and organizational information

Kim-Phuong L. Vu[a,*], Robert W. Proctor[b], Abhilasha Bhargav-Spantzel[b], Bik-Lam (Belin) Tai[b], Joshua Cook[b], E. Eugene Schultz[c]

[a]*California State University Long Beach, Long Beach, CA, USA*
[b]*Purdue University, West Lafayette, IN, USA*
[c]*High Tower Software, Aliso Viejo, CA, USA*

Received 26 July 2006; received in revised form 19 December 2006; accepted 26 March 2007
Communicated by C.M. Karat
Available online 7 April 2007

Abstract

Personal information and organizational information need to be protected, which requires that only authorized users gain access to the information. The most commonly used method for authenticating users who attempt to access such information is through the use of username–password combinations. However, this is a weak method of authentication because users tend to generate passwords that are easy to remember but also easy to crack. Proactive password checking, for which passwords must satisfy certain criteria, is one method for improving the security of user-generated passwords. The present study evaluated the time and number of attempts needed to generate unique passwords satisfying different restrictions for multiple accounts, as well as the login time and accuracy for recalling those passwords. Imposing password restrictions alone did not necessarily lead to more secure passwords. However, the use of a technique for which the first letter of each word of a sentence was used coupled with a requirement to insert a special character and digit yielded more secure passwords that were more memorable.

Keywords: Authentication; Information security; Passwords

1. Introduction

The Internet and World Wide Web have become important parts of most people's lives. Users regularly conduct various transactions over the Web that involve personal information. These transactions include, among others, on-line banking, use of E-health services, and engaging in E-commerce (i.e., selling or purchasing products or services on line; Bidgoli, 2004). The organizations with which these transactions occur maintain personal information about users on their computers, as well as a variety of other types of sensitive information crucial to the organizations' success, requiring that this information be secured and access restricted to authorized individuals (Bidgoli, 2006). Moreover, the organizations should have policies in place to ensure that the users' privacy will be protected so that their personal information does not fall into the hands of people for whom it is not intended (Ackerman and Mainwaring, 2005).

Many methods can be used to restrict access of information to authorized users and personnel through a process of authentication (i.e., having an individual provide some kind of credentials to gain access to a system or network) and authorization (i.e., allowing the individual access to requested information after verification of the identity of that individual and use of access controls such as file permissions; Schultz et al., 2001). The most widely used authentication method is the username–password combination. For the username–password method to be effective, it is essential that users generate and use strong passwords that are resistant to guessing and cracking. However, there typically is a tradeoff between password memorability and security (Proctor et al., 2002). Passwords

*Corresponding author. Tel.: +562 985 5021; fax: +562 985 8004.
E-mail address: kvu8@csulb.edu (K.-P.L. Vu).

1071-5819/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijhcs.2007.03.007

that are easy to remember tend to be biographical information (birth date, names of significant others, favorite items, etc.) or simple words that can be guessed or cracked by other individuals or computer programs. Many studies have shown that users in fact tend to create passwords of these types that are easy to remember (e.g., Klein, 1990; Riddle et al., 1989). For example, Leyden (2003) reported that 12% of users used "password" as their password, and the three most common types of passwords included a user's own name, favorite football team, or date of birth. Another problem with passwords is that people tend to use the same password for multiple accounts (Ives et al., 2004). If the password for one account is exposed, then the security of all the other accounts is jeopardized.

The username–password method provides less security than other authentication methods such as biometric devices, smart cards, and token devices (e.g., De Luis-Garcia et al., 2003; Ives et al., 2004). However, for many Web sites that maintain personal information, the username–password combination is, and will continue to be, the primary method of identifying and authenticating users. The popularity of the method lies in its being accepted widely by users and being easy to implement (e.g., Pinkas and Sander, 2002; Viega, 2005). Thus, use of username–password combinations is not likely to be replaced any time soon because of convenience and practicality considerations (Pinkas and Sander, 2002). Consequently, it is important to determine methods for password generation that will yield passwords that provide adequate security but are also memorable (Proctor et al., 2002).

1.1. Memory and passwords

Memorable passwords will allow users to recall their passwords when needed so that users avoid engaging in bad security practices such as writing down their passwords or choosing an easy-to-guess password. In this section we review several factors known to influence human memory and their implications for password memorability.

1.1.1. The generation effect

Basic research on human memory has shown that memory for items is better when participants are required to generate those items instead of just reading them (see, e.g., de Winstanley and Bjork, 2004; Slamecka and Graf, 1978). This *generation effect* has an important implication for improving the memory of passwords: If users are allowed to generate their own passwords, then the act of generating a password itself should allow users to remember the password better than would be the case if passwords were just provided to them. As noted, though, a drawback with allowing users to generate their own password is that users tend to generate passwords that may not be secure, such as their own name (e.g., Klein, 1990; Leyden, 2003). One method that has been used to produce secure passwords is to have a computer generate strings of random alphanumeric characters and symbols. Though secure, the random nature of these strings tends to make them meaningless to users, causing the passwords to be difficult to remember and tempting users to write them down. If users are asked to generate passwords that include strings of "random" alphanumeric characters and symbols, though, they could do so in a way that is likely to have more meaning to the user than computer-generated passwords, making self-generated passwords more memorable.

1.1.2. Memory load

The ability to retrieve items in memory is dependent on memory load, the number of items that must be retained: As memory load increases, the number of forgotten items increases (Neath, 1998). With regard to passwords, people should be able to remember a few unique passwords without much difficulty; however, as the number of passwords that users have to remember increases, the likelihood of recalling a specific password decreases. The growth of E-commerce and E-services has made memory-load problems more evident because users must generate multiple passwords satisfying different criteria for a variety of Web sites. For example, some Web sites have no restrictions on a user's password, whereas others require a minimum length, a mixture of letters and digits, and so on. Some sites additionally require that a special character be included in the password, whereas others do not allow use of special characters.

A survey conducted by Safenet (2005), showed that 35% of users had 3–4 passwords, 18% had 5–6, 6% had 7–8, and 23% had 9 or more passwords. A press release by RSA Security (2005) also indicated that 30% of their survey respondents manage 6–12 passwords, with 28% indicating that they must keep track of more than 13 passwords. With these large numbers of passwords that must satisfy various restrictions, it is difficult for users to remember the password for any one site. This difficulty likely contributes to their engaging in behaviors that threaten security, such as writing down their passwords, sharing them with another person, or using the same password for multiple accounts.

In fact, a survey of 3050 Web users conducted by Rainbow Technologies (2003) found that 55% of the respondents admitted to writing down at least one password, with 8% indicating that they wrote down all of their passwords. In a follow-up to the Rainbow Technologies survey conducted by SafeNet (2005), 50% indicated having written down at least one password, and 10% said that they always wrote down their passwords. Furthermore, about 50% of the respondents specified that they needed to have their passwords reset on occasion due to forgetting them. The need to have passwords reset may cause users frustration and companies extra costs. For example, RSA Security (2005) reported that each call to a help desk to reset a password cost between $25 and $50.

1.1.3. Proactive interference

Proactive interference occurs when people cannot recall current items because of interference produced by previously learned items (e.g., Bunting, 2006). It will increase as the number of previously learned items and associations increase. Proactive interference has become a significant memory problem for password use. Because users tend to give out their passwords willingly when asked (see, e.g., social engineering, Viega, 2005), many sites now require that users change their passwords monthly. The idea behind requiring password changes is that access to an old password will not enable access to information stored in an account or system (Hilton, 2006). Although changing passwords frequently may be good for security, it is not good for usability: Proactive interference from older passwords may create difficulty for users trying to remembering their current passwords (Bunting, 2006).

1.1.4. Elaborative processing and mnemonics

The memory literature also shows that recall of items can be improved by elaborative processing of the material (e.g., Craik and Lockhart, 1972; Jacoby and Craik, 1979). Depth of processing, which refers to how "deeply", with respect to meaning, one encodes the information could improve recall of that information at a later time. For example, determining whether the words "computer" and "printer" are related requires a deeper level of processing than determining whether the two words are printed in the same font size. Research has shown that the deeper the information is processed, or the more it is elaborated (such as thinking about the meaning of the information or how it relates to other things), the better it can be recalled later (Craik and Lockhart, 1972; Jacoby and Craik, 1979; Parkin, 1984). The benefit in recall as a function of depth of processing can be explained by the fact that more connections are made between items, which can provide many retrieval paths for later recall (e.g., Craik and Tulving, 1975).

Another class of elaborative processing techniques known to improve recall of items is mnemonic techniques (Neath, 1998). These techniques allow for items that need to be remembered to be encoded in an organizational scheme that can aid their retrieval. For example, with first letter mnemonics, people can remember the colors of the rainbow by remembering the name, "Roy G. Biv" and using the first letters in the name to generate the colors: R = red; o = orange; y = yellow; G = green; B = blue; i = indigo, and v = violet. As with depth of processing, mnemonic techniques work because they provide connections between the items needing to be recalled with another well-established structure.

Elaborative techniques can aid the recall of passwords because they allow users to impose meaning to a random string of letters and characters. For example, if the password were YiA3p!, the user could construct a sentence that includes the letters and characters (e.g., first letters of the sentence, Yesterday i Ate 3 pizzas!).

1.2. Security of passwords

Several methods have been proposed for strengthening passwords. These aim at improving (a) the quality of the password generated initially, (b) the memorability of the password that is generated, or (c) the complexity of an encryption method that is used to store passwords within system files. Attempts at improving quality of passwords include imposing password restrictions (e.g., Bergadano et al., 1998), making people update their passwords regularly, and educating people about what makes a good password (e.g., Ives et al., 2004; Viega, 2005). Some researchers have explored the benefits for memory of using graphical passwords, for which a user denotes the password by clicking on certain areas of a scene in a particular sequence (e.g., Wiedenbeck et al., 2005) or on one of several faces (Tari et al., 2006). Encryption methods include, for example, use of hashing techniques that allow the stored representation of the password to be more crack resistant (e.g., Halderman et al., 2005).

The method on which we focus is that of password filtering, or proactive password checking (Bishop and Klein, 1995). The proactive technique of password generation allows users to generate their own passwords, but the system checks to determine whether the generated password adheres to specified restrictions (e.g., Proctor et al., 2002). The intent is to ensure that the generated passwords allowed by the system meet certain standards and thus are more secure than they would be if the restrictions were not imposed. The restrictions include, for example, that the password be of a minimum length and that it contain both letters and digits. An advantage of proactive password checking over computer-generated passwords is that the user is allowed to generate the password himself or herself, making it more memorable (Slamecka and Graf, 1978), while still incorporating the recommended characteristics of a good password (Blackwell et al., 2004).

The security provided by passwords can be assessed using one of several commercial cracking programs. Such programs provide measures of the proportions of passwords cracked during a certain period of time using various methods. One of the most effective cracking programs for Windows operating systems security is l0phftcrack 5 (lc5). It was used in the present study because it is a widely available tool that many security experts use to help identify weak passwords (Posey, 2003). lc5 is considered to be an intermediate strength password cracker, which is the range that should be most sensitive for detecting differences in password security for different generation conditions. The program uses two methods to crack passwords. The first combines a dictionary attack, in which whole words are tested against a set of words taken from a dictionary (e.g., Webster's dictionary), with a hybrid attack that adds numbers to dictionary and other words and uses different combinations of upper and lowercase letters. If this first method is not successful, then a brute force attack

is used where a very large number of combinations of letters, numbers, and characters are tested.

1.3. Proactive password checking

Proctor et al. (2002) showed that adding password restrictions increased the time taken by users to generate the passwords, but the restrictions did not affect the users' ability to recall the passwords later or the time it took to log in to the system. More important, for passwords of five characters, proactive password checking resulted in a 40% reduction of the number of passwords cracked by a cracking program. Furthermore, requiring users to generate passwords with a length requirement of eight characters was sufficient to reduce the number of passwords cracked to 17% compared to 75% for the 5-character passwords generated with no restrictions.

Thus, Proctor et al. (2002) study showed that proactive password checking improves security, and that this method results in the generation of passwords that are no harder to recall and do not require more time to enter when users login. However, in Proctor et al.'s study, participants generated only one password for a single account and then used the password to gain entry to the account after a short delay. Because each Web site requiring registration usually uses different username–password combinations, which are used after various delays, it is important to evaluate whether the use of proactive password checking is effective when users have multiple accounts and need to access those accounts after longer delays. The goal of the present study was to perform such an evaluation for proactive password checking in general and for specific mnemonic techniques that should be effective in improving both retention and security.

1.4. Present study

Experiment 1 provided an initial evaluation of the effectiveness of proactive password checking in generating passwords for three or five accounts, with each password having to satisfy seven restrictions. Memory for the passwords was tested after a 5-min delay and a week later by providing the name for each account and having the participant attempt to recall the associated password. Because many of the passwords generated in Experiment 1 were cracked, we conducted two additional experiments oriented toward evaluating use of mnemonic techniques to generate more secure but still memorable passwords. In Experiment 2, in addition to satisfying proactive password restrictions, the passwords were to be derived from the first letters of the words of a sentence. In Experiment 3, this method was compared to one in which characters in the password were based on whole words from a generated sentence. Experiment 3 also evaluated whether having users recall the passwords after 5 min, or reenter them immediately, benefited retention over the week interval. The experiments provide evidence that both memorability and security can be improved through careful consideration of human memory capabilities.

2. Experiment 1

Experiment 1 examined how well users could remember unique passwords generated for different accounts when seven restrictions were imposed. These restrictions were chosen because they conform to common recommendations for good passwords, such as "mixing special characters with numbers and letters" is better than mixing numbers and letters or using letters alone (Blackwell et al., 2004). Users were assigned to conditions in which they generated passwords satisfying the restrictions for either three or five accounts. These numbers of accounts were chosen because they fall within the range of 3–6 accounts held by most users (see, e.g., SafeNet, 2005).

2.1. Method

2.1.1. Participants

Thirty-two students from Purdue University participated for partial credit toward their Introductory Psychology course. All were experienced computer users and were familiar with generating passwords.

2.1.2. Apparatus

A program, written in Java, was used to present instructions to participants, record and check the generated passwords, and record the time to generate an acceptable password. Recall time was also measured after a short retention interval of 5 min and a long retention interval of 1 week. Generation and recall times were measured by using the Java code that gives the system time in milliseconds (System.currentTimeMillis), subtracting the time when the password prompt was presented from the time when the password was entered.

2.1.3. Procedure

All participants were tested individually in a quiet, well-lit room. Participants were informed that they would be asked to generate passwords for several different "fake" accounts. Sixteen of the participants generated passwords for three accounts, and the other 16 participants did so for five accounts. The experiment itself was divided into three parts. In the first part, the participant generated passwords for each account. The experimenter read the instructions for generating the passwords to the participant. For the 3-accounts group, the generic account names were: E-mail, bank, and eBay. For the 5-accounts group, the two additional accounts were labeled travel and books.

Seven password restrictions were imposed, namely that the password:

(1) Be at least six characters.
(2) Contain an uppercase letter.
(3) Contain a lowercase letter.

(4) Contain a digit.
(5) Contain a special character (e.g., ! or #).
(6) Be unique from the passwords generated for the other accounts.
(7) Not contain the person's username or any variant of it.

Participants were informed that they would be asked to recall the password for each account after generating all of them. Each participant was asked to enter a username for his/her file and afterwards was presented with a prompt to enter a password for one account. All of the password restrictions remained visible on the screen during this process. If the generated password met all conditions, the prompt for the next account was displayed. If the password did not meet one or more of the restrictions, a prompt to reenter the password was presented along with a list of the restrictions that were not met by the previous password entry. The generation time and number of attempts were recorded and sent directly to a log file. Participants were not given feedback about their generation times. Once the passwords for all accounts were generated and accepted, they were printed on the screen along with their corresponding account names for the participant to review. Participants were not allowed to write the passwords down.

The computer screen was cleared of any information relating to the first part of the experiment, and all participants took a 5-min break. During this time, they left the room and were encouraged to walk around and engage in activities such as getting a drink of water. At the end of this break, the second part of the experiment was started. In this part, participants were presented with a list of the account names for which they had generated passwords. They were informed that one of the account names would appear on the screen, and that they were to recall and enter the password for that account. The participants logged into each account four times in random order. They were also told that they would have a maximum of 10 attempts for each account occurrence. For each occurrence, the login time and number of incorrect attempts were recorded. Before leaving the experiment, the participants were instructed not to write down the passwords that they had generated during the study.

For the last part of the study, participants came back a week later to recall the passwords. The procedure was identical to the second part of the experiment. At the end of the experiment, the log files were compiled, and mean generation and login times for each condition for each participant were computed.

2.2. Results

2.2.1. Generation time and number of attempts

There were 16 participants assigned to each group. A one-way ANOVA, with group (three or five accounts) as a between-subjects factor, was conducted on two dependent measures: mean generation time and number of attempts. There was no significant difference in generation times or number of attempts to come up with an acceptable password for the 3- and 5-account groups. On average, participants took 29 s to generate a password for an account and arrived at an acceptable password after a single attempt ($M = 1.2$ attempts). The number of characters for each password ranged from 6 to 15, with the mean length of each password being 9.1 characters for the three accounts group and 8.5 characters for the five accounts group.

2.2.2. Login time and number of attempts

Mean login time and number of attempts for each participant were submitted to 2 (Group: three or five accounts) × 2 (Recall delay: 5-min or 1 week) repeated-measures ANOVAs with recall delay as a within-subjects factor and group as a between-subjects factor. There were no significant effects for login time. However, the main effect of number of accounts was significant for the mean number of login attempts, $F(1, 30) = 26.98$, $p < .025$. Participants in the 5-accounts condition ($M = 2.8$ attempts) made more errors in recalling the passwords than those in the 3-accounts condition ($M = 1.4$ attempts). There was no significant effect of recall delay, and it did not significantly interact with the number of accounts.

2.2.3. Forgetting

We also examined the number of participants who forgot the password for one or more accounts. For the 3-accounts group, one participant failed to recall a password during short-term recall, and only two participants were unable to recall the password for one account during long-term recall. In contrast, for the 5-accounts group, two participants did not recall the passwords for three of the five accounts, both at short and long-term recall. In addition, five participants were not able to remember a password for one account during short-term recall, and three participants for one account and one participant for two accounts at long-term recall.

2.2.4. Crackability of passwords

We limited the password cracking time for which the lc5 program was run to approximately 4 h. The lc5 program was able to crack a significantly larger percentage of the passwords from the 3-accounts group (60%; $N = 29$) than from the 5-accounts group (40%; $N = 32$), $\chi^2(1,128) = 5.01$, $p < .03$. The requirement to generate five unique passwords rather than just three apparently forced participants to be more creative in their password generation.

2.3. Discussion

Experiment 1 shows that imposing proactive password restrictions does not necessarily produce crack-resistant passwords: Generated passwords were required to satisfy seven password criteria, yet approximately half of the

passwords were cracked within 4 h. The dramatic increase in the number of passwords cracked in the present study compared to Proctor et al. (2002) study is likely due to differences in the cracking software used in the two studies. This study used lc5, which was run in Windows, whereas Proctor et al.'s study used John the Ripper 1.6, which was run in Linux.

Many of the passwords generated were common words starting with a capital letter, with a digit and special character appended at the beginning or end of the word to satisfy the restrictions. Because the generated password followed a simple pattern, it was easy to crack. If restrictions were made such that symbols or numbers were used more instead of letters, the strength of the passwords would be increased.

Many passwords generated were closely related to the type of account. For example, "4Money!" was generated for a bank account. This relational strategy increases memorability, for the password but at the expense of reduced security. A simple strategy for password generation like this would not be appropriate for a bank account because the risks associated with security violations are high, though it might be appropriate for a site at which financial risks are lower (e.g., using "4Email!" for an e-mail account). Therefore, effort should be made to generate passwords in accordance to the level of security required.

3. Experiment 2

The number of passwords that users had to remember affected their ability to recall the passwords when logging in. Most users were able to recall the password for each account after one or two attempts, but recall was more difficult for the 5-accounts group than for the 3-accounts group. In addition, not all participants were able to recall the correct passwords for all accounts. For the 5-accounts group, 11 users (69%) were unable to recall the password for at least one account, compared to three users (19%) for the 3-accounts group. Thus, an increase in memory load seems to lead to a decreased ability to recall unique passwords for different accounts. These findings suggest that generation of unique passwords for different accounts can increase security at little cost to memorability when the number of username–password combinations is small (e.g., 2–3 accounts), but not when it is larger.

The purpose of Experiment 2 was to examine the effectiveness of proactive password checking using a sentence-generation method to produce passwords that are more crack-resistant and memorable. This method is one that has been used more recently to generate passwords (see Blackwell et al., 2004). The sentence-generation method requires users to generate a password by formulating a real sentence and taking the first letter of each word in the sentence to make a password. The act of generating a sentence that makes sense should promote better recall because it requires both generation and deeper processing of information. The context of the sentence can also be used as cues to help the users recall the password. The sentence-generation method should also improve the security of the password because it will be composed of strings of letters rather than full words.

However, because of the power of modern password-cracking software, strings of letters alone can be easily cracked. To evaluate this possibility, we also included a condition in which users were required to embed a special character and digit into the sentence and resulting password, which has not been examined before. Comparison of the crackability and memorability of the passwords with and without the character and digit restriction should provide additional evidence regarding the effectiveness of the sentence generation method.

3.1. Method

Forty new students from Purdue University participated for partial credit toward their Introductory Psychology course requirement. The experiment was conducted using a similar Java program to that described in Experiment 1.

The procedure was identical to that of Experiment 1, except for the differences noted. All participants generated and recalled passwords for only three different accounts. To generate a password, participants were asked to create a sentence and then to combine the first letter of each word in the sentence to form a password.

Twenty participants were given the following three rules to apply to their sentence:

(1) The sentence should have at least six words.
(2) The sentence should make sense.
(3) The sentence (and password) for each account should be unique.

The remaining 20 participants were given those rules plus two additional rules:

(4) The sentence (and password) should contain a special character (e.g., !, @, or#).
(5) The sentence (and password) should contain a digit.

Participants in the latter condition were allowed to embed the special character and digit between words or in place of words in the sentence.

3.2. Results

For the 3-restrictions group, the number of characters for each password ranged from 6 to 10, with an average of 6.6 characters. For the 5-restrictions group, the passwords ranged from 6 to 12 characters, with mean of 7.6 characters.

3.2.1. Generation time and number of attempts

Twenty participants were assigned to each group. A one-way ANOVA, with group (three restrictions or five restrictions) as a between-subjects factor, was conducted on two dependent measures: mean generation time and

number of attempts. There was a difference in the time to generate an acceptable password for the two groups, $F(1, 38) = 18.65$, $p < .001$. Participants took 50.9 s to generate a password with three restrictions but 84.9 s with five restrictions. However, there was no significant difference in the number of attempts needed to generate an acceptable password between the two groups ($M = 1.3$ attempts for the 5-restrictions group and $M = 1.2$ attempts for the 3-restrictions group).

3.2.2. Login time and number of attempts

Mean login time and number of attempts for each participant were submitted to 2 (Group: three or five restrictions) × 2 (Recall delay: 5-min or 1 week) repeated measures ANOVAs with recall delay as a within-subjects factor and group as a between-subjects factor. There was a difference in the time to login for the two groups, $F(1, 38) = 6.06$, $p = .042$. It took twice as long for participants to login with a password satisfying five restrictions ($M = 26.4$ s) than with one satisfying three restrictions ($M = 13.2$ s). There was no significant effect of recall delay on login times, and delay did not significantly interact with the number of restrictions.

For the number of attempts measure, there was no significant main effect of group or recall delay. The interaction between group and recall delay was marginal, $F(1, 38) = 3.48$, $p = .07$: Participants in the 5-restrictions group tended to make more attempts to recall the password ($M = 2.9$ attempts) than those in the 3-restrictions group ($M = 2.0$ attempts).

3.2.3. Crackability of passwords

lc5 cracked 62% of the passwords from the 3-restrictions group but only 2% from the 5-restrictions group, $\chi^2(120) = 31.4$, $p < .001$. The high percentage of passwords cracked in the 3-restrictions group can be attributed to the absence of special characters and digits, both of which make successful dictionary/hybrid attacks much less likely.

3.2.4. Forgetting

For the 3-restrictions group, two people failed to recall one password and one person two passwords at short-term recall. At long-term recall, four people did not remember one password. For the 5-restrictions group, one person was unable to recall one password, two people two passwords, and two people all three passwords at short-term recall. At long-term recall, three people did not recollect one password, three people two passwords, and two people all three passwords.

3.3. Discussion

The sentence-generation method is effective at producing crack-resistant passwords when users are required to embed a special character and digit into the sentence. Only 2% of the passwords generated using the sentence-generation method with the character and digit requirement were cracked, compared to 62% of the passwords generated without the special character and digit. It is important to note, though, that the requirement of including a special character and digit in itself is not sufficient to yield crack-resistant passwords. In Experiment 1, approximately 50% of the passwords generated with a digit and special character were cracked with the same program and time-frame as those used in Experiment 2. Thus, the increased security of the password is a combined result of the sentence-generation method and embedding a character and digit within it.

Despite the improvement in security, the requirement of including a special character and digit caused participants to take longer to generate the sentences (and passwords). In addition, there was also a cost associated with the memorability of the passwords, at both short- and long-term recall. Participants in the group that generated passwords with the additional character and digit took two times longer to recall and login with the passwords, made almost twice as many errors before being able to recall the password, and completely forgot the password twice as often.

Because of this cost to memorability, it is important to determine what types of errors that users are prone to make in order to investigate techniques that will help minimize these errors. With the sentence-generation technique, we hypothesized that there are five likely reasons why participants had a harder time remembering the passwords when they had the additional requirement of including a number and special character (see Fig. 1). To determine the types of errors made by the participants, we classified the types of errors into five categories: Order errors, in which participants recalled all components, but in incorrect order; forgetting sentence errors, for which the sentence was recalled incorrectly; error of special character and/or digit, for which the sentence was correct, but the character or digit embedded in it was wrong; both sentence and digit/character error, for which both of the preceding error types were present; error of association, for which a correct password for a different account was recalled. Analysis of these error categories showed that the two most common errors were forgetting the sentence generated (38% of errors) or the special character and/or digit embedded into the sentence (25%) and the resulting password. Failure to remember the sentence was mainly due to participants' remembering the gist of the sentence but not its exact phrasing. For example, the original sentence was "I am going to lunch @ 5 today." It was recalled incorrectly on one occasion as "I will eat lunch @ 5 today" and on another as "I will have lunch @ 5 today." This type of error is consistent with memory studies showing that people attend more to the meanings that are being expressed, and less to the exact way the meanings are expressed (e.g., Brewer, 1977). Thus, the problem of not being able to recall the exact wording of the sentence is a limitation of the sentence-generation technique.

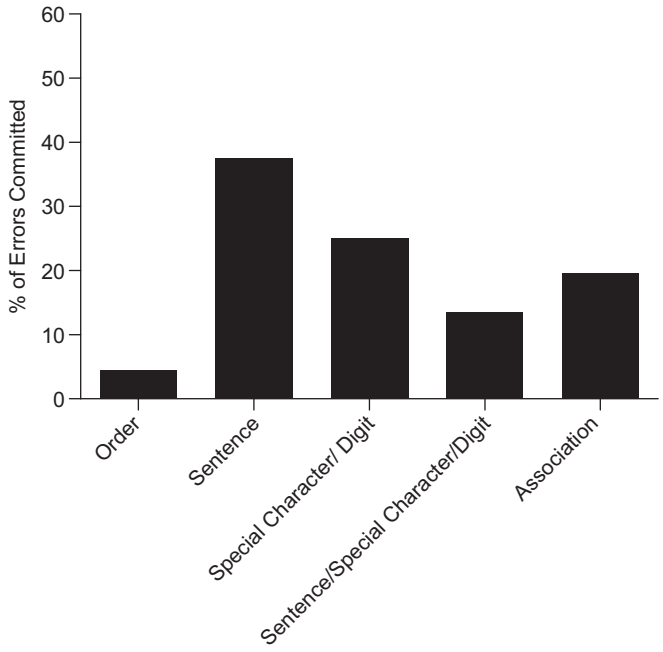

Fig. 1. Mean percentage of errors committed in Experiment 2 as a function of error type (order; sentence; special character/digit; sentence and special character/digit; association). An error of *order* occurs when participants remembered the sentence they generated as well as the special character and digit, but did not remember where the character and digit were embedded within the sentence. A *sentence* error occurred when participants forgot the sentence they generated or its exact wording. The *special character/digit* error occurred when participants forgot the special character and/or digit they used in the sentence. The *sentence and special character/digit* error occurred when participants forgot both the sentence generated and the special character and/or digit they used in the sentence. Finally, the error of association occurred when participants made a wrong association for a password of a different account.

4. Experiment 3A

Although Experiment 2 showed that passwords generated with the sentence-generation technique were secure when special characters and digits were used in conjunction with sentence generation, participants had difficulty remembering the digits and/or symbols that were used in the password. This difficulty could arise because, in many cases, the symbols and/or digits were not meaningfully related to the sentence. As a result, the sentence itself did not act as a cue for retrieving the character and/or digit. Experiment 3A used an entire-word generation technique to try to get users to use the special characters and digits in a more meaningful manner. With the entire-word method, participants constructed the password with entire words from their sentence rather than just the first letters of each word. They were told to substitute words, or portions of words, in the sentence with characters and digits that were phonetically similar to them.

We based the entire-word method on our observation that some currently used passwords appear to be random strings of digits and letters, but when inspected closely are really mnemonic sentences, phrases, or words (e.g., The password EyeD8tedM@ can be remembered as the sentence, I dated Matt; see also, Blackwell et al., 2004). Intuitively, this mnemonic technique for associating entire words within a sentence or phrase with different representations of the word or syllables seems to be a promising method for helping users to generate secure and memorable passwords. This is because meaningful items are easier for participants to remember than non-meaningful items (e.g., Richardson, 2003). Because other factors may offset this benefit of meaningfulness, Experiment 3A was conducted to assess the effectiveness of the entire-word technique and compare it to the first-letter sentence-generation technique examined in Experiment 2.

In the previous experiments, all users were tested for recall at both a short 5-min interval and a long 1-week interval. Upon debriefing participants at the end of Experiments 1 and 2, several reported that they thought that the passwords were easier to remember at the week delay. Some mentioned that they had more difficulty remembering the password when recall was tested 5-min later, and that they thought the login repetitions during the short-term recall phase of the study helped them remember the passwords better at the week delay. This intuition of some participants is in agreement with findings in the memory literature, which show that being tested over material enhances retention (Roediger and Karpicke, 2006). Although it is not common practice to have users login after a short delay, many Web sites do ask users to reenter their passwords for verification purposes after generation. In Experiment 3A we examined long-term retention of passwords with and without a prior recall test at a 5-min interval; in Experiment 3B, we examined whether immediate reentry of each password as it was generated had a similar influence on long-term retention to that of the short-term recall test. To increase memory load, the number of accounts was increased to five.

4.1. Method

4.1.1. Participants

Sixty new students from Purdue University's Introductory Psychology Subject Pool participated in the experiment. Fifteen were tested in each of four different conditions: entire-word generation/short- and-long-term recall; entire-word generation/long-term recall only; first-letter generation/short- and-long-term recall; first-letter generation/long-term recall only.

4.1.2. Procedure

Participants were instructed to create passwords for five fictitious on-line accounts. The accounts were distinguished by the names Amazon, eBay, Yahoo!, AOL, and Dell. In the first-letter password generation condition, participants were instructed to generate a sentence and then to use the first letter of each word in the sentence to form the password, as described in Experiment 2.

In the entire-word condition, participants were instructed to create a sentence and then to replace entire words of the sentence with a mnemonic equivalent to create the password. In other respects, the procedure and guidelines were the same as for the first-letter condition, with the additional requirement that the special character and digit should be embedded in the sentence in a way that it makes sense relative to the context of the sentence.

Participants in the entire-word condition were also given the following examples of how to transform phrases or words of their sentences into passwords. The examples were intended to illustrate different ways that specific words or phrases can be transformed into a mnemonic string of letters, digits, and special characters:

- Before I had coffee at work → "B4EyeH@CofE@w".
- I had four snakes → "EyeH@4$nake$".
- Later I ate tofu → "L8erEye8 + ofu".
- You got to go home at nine → "UG2Ghom@9".

Participants in both conditions were then prompted with the name of the first account and instructions to enter their sentence. After a sentence had been entered, a new prompt appeared and instructed the participant to create a password from her or his sentence using the restrictions provided by the instructions. If the password created from the sentence did not fit the criteria for the appropriate condition, a message was presented indicating which criteria were not met. The participant was then instructed to re-enter a sentence and password. The process was repeated until a password meeting the criteria was generated. Then, a prompt for the next account appeared, and the process was repeated for that account.

Half of the participants in each password generation condition were asked to recall the passwords for all accounts after a 5-min break (the short- and long-term condition), as in Experiments 1 and 2. The other half of the participants in each password generation condition did not engage in the 5-min memory test (the long-term-only condition). All participants returned a week later to perform the recall task.

4.2. Results

The number of characters for each password in the first letter condition ranged from 6 to 15, with a mean length of 7.5 characters. The passwords generated for the entire word condition were longer, ranging from 6 to 25 characters, and averaging 10.1 characters.

4.2.1. Generation time and number of attempts

Mean generation time and number of attempts needed for successful password generation were calculated for each user, and submitted to separate 2 (Generation technique: entire word or first letter) × 2 (Login condition: short-and-long-term or long-term-only) between-subject ANOVAs. For generation time, participants took approximately 12-s longer to generate passwords for the entire-word method (83 s) than for the first-letter method (95 s), though this difference was not significant. The main effect of login condition was not significant either, and it did not interact with the generation method. There were no significant effects for the number of attempts data: On average, users needed 1–2 attempts to generate an acceptable password.

4.2.2. Login time and number of attempts

For users who had to login at both short and long delays, the short-delay data were analyzed using a one-way ANOVA, with generation technique as a between-subjects factor. The mean login time was 33.7 s, and the mean number of attempts was 3.77. There was no significant effect of generation technique on login time or number of attempts, indicating that, for immediate recall, both generation techniques yielded passwords of similar memorability.

Because all four groups were tested for password retention at the 1-week interval, a 2 (Generation technique: entire word or first letter) × 2 (Login condition: short-and-long-term or long-term-only) ANOVA was performed (see Figs. 2 and 3). Only the main effect of login condition was significant, $F(1, 56) = 20.05$, $p < .001$. Users were able to login approximately 25-s faster at the long delay if they had also performed the login trials at the short delay than if they had not. A similar finding was obtained with the number of attempts data, $F(1, 56) = 7.96$, $p < .008$; participants who had logged in at the short delay needed fewer attempts than those who had not (Ms = 4 and 6 attempts, respectively).

4.2.3. Forgetting

The mean number of passwords forgotten at the week delay for each user was computed (range = 0–5 passwords) and submitted to a 2 (Generation technique: entire word or first letter) × 2 (Recall condition: short-and-long-term recall or long-term-recall only) between-subjects ANOVA.

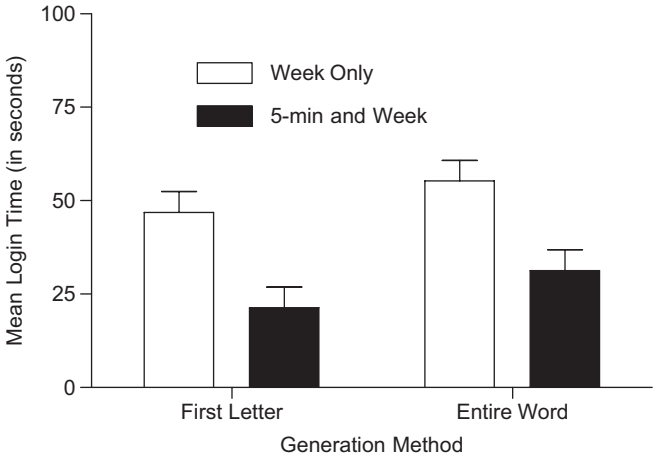

Fig. 2. Mean login time (in s) at the week delay as a function of the generation method (first letter or entire word) as a function of recall delay (week only or after 5-min and 1 week).

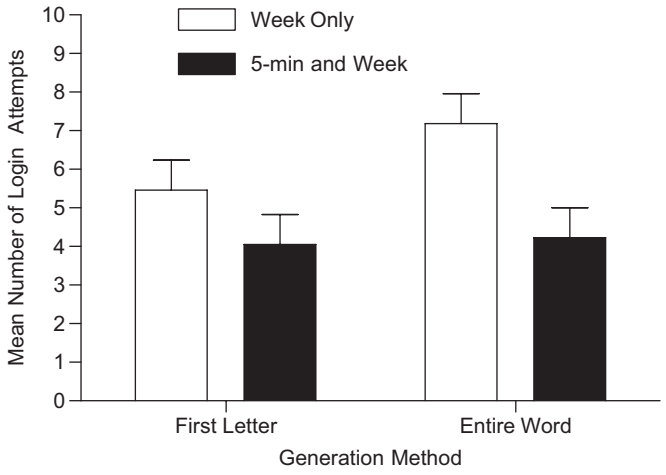

Fig. 3. Mean number of login attempts at the week delay as a function of the generation method (first letter or entire word) as a function of recall delay (week only or after 5-min and 1 week).

Participants who used the entire-word technique tended to forget more passwords ($M = 2.5$) than did participants who used the first-letter technique ($M = 1.7$), $F(1, 56) = 3.13$, $p = .082$. More importantly, there was a significant effect of login condition, $F(1, 56) = 6.29$, $p = .015$, and the interaction between login condition and generation method was not statistically significant. These results demonstrate that participants who received a short-term recall test (recall after 5 min) forgot fewer passwords than those who participated in only the 1-week recall test (Ms = 1.5 and 2.63 words, respectively).

4.2.4. Crackability of passwords

lc5 cracked 13% ($N = 20$) of the passwords from the first-letter generation group and 10% ($N = 15$) from the entire-word generation group, a difference that was nonsignificant.

4.3. Discussion

Participants were able to recall their passwords much more effectively at a 1-week retention interval when they were tested for their memory of the passwords 5 min after generating them than when they were not. This finding suggests that requiring users to recall their passwords shortly after generating them should help the users to remember the passwords.

The entire-word method did not improve recall of the passwords compared to the first-letter method. There are several possible reasons why the entire-word method was not as effective as expected. Participants remembered the sentence, but often forgot the exact transformation that had been used to convert it into the password. For example, participants would forget whether they spelled out a word completely or abbreviated it, and if they abbreviated the word, whether they used a symbol to take the place of a subset of letters in the word or to replace the whole word.

Participants used special characters to represent different parts of the password inconsistently for the five accounts and were confused at recall as to how they used the character for a particular password. For instance, participants tended to use the "$" sign to represent the letter "s", the word "money", or the word "dollar."

Below are some example sentences that participants created which had problems similar to the ones described above.

Inconsistent within the password:

- I ate breakfast at eight. > Eye8breakfast@ate (uses the number "8" for "ate" and then spells out the number "8" with "ate").

Inconsistent between one or more passwords:

- I had four cats > Ih@d4c@ts (uses "@" symbol for letter "a").
- I live at Maclur > Il@maclur (uses "@" symbol for the word "at").

The percentage of passwords cracked in the first letter condition was larger in the present experiment than in Experiment 2 (13% vs. 2%), even though the passwords generated in the two experiments were similar in range and average length. One reason for the higher crack rate is that about half of the passwords cracked from the present experiment were from only two participants (for both participants, four of the five passwords were cracked). For these participants, the common element in their password generation was placing the special character at the end of the password rather than embedded it within.

5. Experiment 3B

Because Experiment 3A showed that long-term retention of passwords was better when they were recalled after a 5-min delay, one might speculate that the immediate re-entry of a generated password required by many Web sites for verification purposes has an unintended beneficial effect of improving password retention. However, there are several differences between the short-term login method used in Experiment 3A and the re-entry password-generation method commonly used by Web sites. The most salient differences are that recall was performed four times for several accounts in Experiment 3A, after a 5-min delay, as opposed to a single time for one account, with no delay. Because immediate re-entry of a single password can be performed on the basis of the currently activated memory and requires little effort, it likely does not produce the benefit found in Experiment 3A. To assess the value of immediate re-entry for long-term retention, we conducted another experiment in which each password was re-entered immediately after initially being generated, with a retention

test similar to that used in Experiment 3A performed 1 week later.

5.1. Method

Fifteen new participants from the Introductory Psychology course at Purdue University participated in the experiment. As in Experiment 3A, users sat in front of a computer and were instructed to create passwords for five fictitious on-line accounts. Because the entire-word generation technique was not better than the first-letter generation technique in Experiment 3A, in the present experiment participants generated passwords with only the first-letter technique. The sentence and password had to satisfy the same restrictions for the first-letter technique as in Experiment 3A.

After a successful password was generated for an account, the password would disappear and a prompt would immediately appear on the screen instructing the participant to re-enter the password that he or she had just created. When the participant entered the password correctly for a second time, the prompt for the next account appeared and the process was repeated. As soon as the participant finished passwords for all five accounts, he or she left and returned 1 week later to recall the passwords.

5.2. Results

Passwords generated ranged from 6 to 13 characters in length. The average length was 7.9 characters.

5.2.1. Generation time and number of attempts

For the re-entry method used in this experiment, the generation time included the time it took the participant to generate the password and immediately re-enter it. To compare this generation time with that for the short-and-long-term and long-term-only generation conditions in Experiment 3A, mean generation time and number of attempts needed for successful password generation were submitted to one-way ANOVAs with recall condition (re-entry; short-and-long-term; long-term-only) as a between-subject factor. There was an effect of recall condition, $F(2, 42) = 6.85$, $p = .003$. Tukey post hoc analyses showed that the mean generation time of 114 s in Experiment 3B was significantly longer than the generation time for either of the first-letter conditions of Experiment 3A (Ms = 81 s for the short-and-long-term condition and 85 s for the long-term-only condition). This finding suggests that the process of having users re-enter a password takes about 30 s. The mean number of attempts needed to successfully generate and re-enter a valid password was 1.56, which was not significantly different from the means for the two first-letter conditions in Experiment 3A that did not require re-entry (Ms = 1.55 for the short-and-long-term condition and 1.51 for the long-term-only condition).

5.2.2. Login time and number of attempts

Performance was examined at the week-delay interval. Mean login time, number of attempts needed to enter the correct password, and number of passwords forgotten were calculated for each user. The login data for the short-and-long-term and the long-term-only condition using the first-letter generation technique of Experiment 3A were used as a comparison. The data were submitted to separate one-way ANOVAs, with login condition (immediate re-entry; short-and-long-term; long-term-only) as a between-subjects factor (see Fig. 4).

There was a significant effect of login condition on long-term recall of the passwords, $F(2, 42) = 6.56$, $p = .003$, with login times being shorter for users who performed both short- and long-term logins ($M = 21$ s) than for users who re-entered the password immediately after generation ($M = 45$ s) or logged in only at the week interval ($M = 47$ s). There were no significant effects for the number of attempts and forgetting data. Although the difference in number of password attempts was not significant, participants who performed both the short- and long-term logins took an average of 4.05 tries to enter the correct password at the long-term login, whereas participants in the long-term-only and the immediate re-entry conditions took averages of 5.84 and 5.46 tries, respectively.

lc5 cracked 19% ($N = 14$) of the generated passwords, which was not significantly different from the 13% of passwords cracked for the first-letter condition in Experiment 3A. Five of the cracked passwords were from one individual. Like those two subjects in Experiment 3A for whom the majority of their passwords were cracked, this participant placed the special character at the end of all passwords.

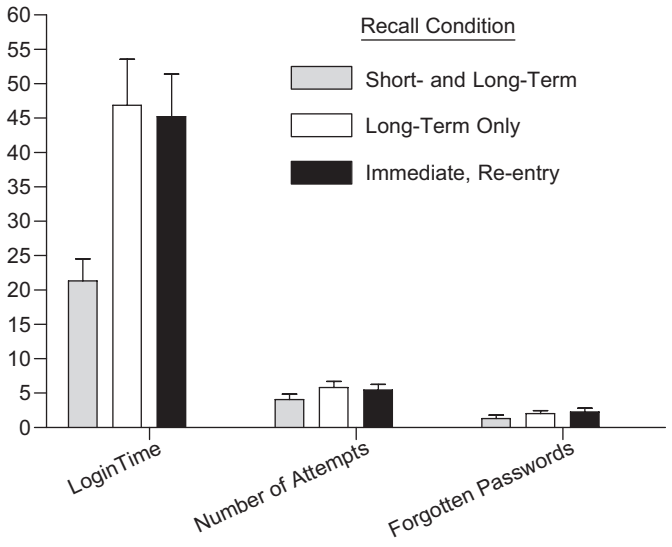

Fig. 4. Performance after a 1-week delay: mean login time (in s), number of attempts to recall passwords, and number of passwords forgotten as function of recall condition for users in the first-letter condition of Experiments 3A and 3B.

5.3. Discussion

The generation data showed that, although re-entering the password for verification purposes took more time, it did not increase the number of attempts needed for successful password generation because users re-entered the passwords correctly. The time to recall the passwords for users in the immediate re-entry condition was similar to that for the users who participated in the long-term only recall condition of Experiment 3A. Thus, having users re-enter a generated password is not sufficient to promote improved long-term retention of the passwords. Requiring users to engage in recall of passwords after a delay of several minutes is a more effective method for improving password retention.

6. General discussion

Having users adopt effective passwords is an important step in protecting personal and organizational information, and ensuring privacy. Examining the tradeoff between password security and memorability is important because there are growing numbers of on-line transactions being performed in which the username–password combination is used as the sole means for identification and authentication. Because of the unprecedented growth in identity theft resulting from unauthorized access to personal and financial information and because consumers have become increasingly concerned with Web privacy, the success of a Web site will be determined in part by how well the site can protect personal and organizational information. If the site cannot safeguard against unauthorized transactions, the organization will likely incur financial costs (e.g., refund of money for items not purchased by the user, personnel costs for investigating fraud, etc.) as well as lower customer satisfaction or loss of customers to competing organizations (see Schultz, 2005).

Experiment 1 showed that approximately half of the passwords created by users were cracked by lc5 within 4 h, even though seven restrictions were imposed on the passwords that were generated. These findings indicate that simple proactive password generation may not be sufficient to produce very secure passwords. Moreover, passwords were more difficult to recall when memory load was increased by increasing the number of passwords that participants had to remember.

Experiment 2 used a first-letter, proactive generation technique that required participants to create a sentence and then form the password from the first letters of the words of the sentence. This technique was more successful at promoting the generation of secure passwords by users than simply imposing restrictions. With the first-letter technique, only 2% of the generated passwords were cracked by lc5 when participants received the additional instruction to embed a special character and digit into the password. However, the security provided by these passwords was at the cost of memorability. Participants experienced some difficulty recalling the exact wording of the sentences that they generated to form the passwords and exactly how the special characters and digits had been embedded within the sentences.

Experiment 3A used a technique intended to provide a structure for embedding the special character and digit into the sentence, with the aim of improving the password's memorability. However, we found little difference between this new, entire-word generation technique and the first-letter technique used in Experiment 2 in terms of generation and login times, and the percentage of passwords cracked. Moreover, the first-letter generation technique showed less forgetting of passwords than did the entire-word generation technique. The entire-word method was slightly more successful at generating crack-resistant passwords than the first-letter method, but this was likely due to the fact that the passwords generated with the entire-word method were longer (10.1 vs. 7.5 characters). We previously have shown that the percentage of passwords cracked decreases as the length increases (Proctor et al., 2002).

One positive aspect of the results for Experiments 1 and 2 was that logins using the passwords a week later benefited from the participants having used the passwords to login 5 min after completing the generation phase. Experiment 3A confirmed this benefit, showing that for both the first-letter and entire-word techniques, long-term recall was better of the passwords was better when short-term recall of them had been performed 5 min after generation. Experiment 3B showed that immediate re-entry of a password, as required by many Web sites for verification purposes, has little benefit on long-term retention. This outcome is not surprising, though, because retrieval of active contents from working memory is relatively effortless and, consequently, has little impact on long-term retention (Neath, 1998).

Gehringer (2002) noted that one significant problem associated with having unique passwords for multiple accounts is that users tend to write these passwords down. This tendency to write down passwords is amplified when the system employs lockout techniques to avoid brute-force attacks on the system. Gehringer also noted that in many systems, the number of guesses prior to a lockout is three, and if the limit is increased, may reduce the users' incentives to write passwords down. To evaluate whether the "3 times and you're out" rule is justified, we looked at the average number of attempts for participants who were able to remember their passwords (did not meet the 10 maximum criterion for forgetting).

All three experiments showed that, if participants did not forget their passwords, many were able to recall their passwords within 3–6 attempts. When participants had to remember passwords for only three accounts, most were able to recall the password with three attempts or less at long-term recall. However, when participants had to remember passwords for five accounts, many required as many as six attempts before recalling the correct password.

Thus, it seems that the "3 times and you're out" rule is justified if the user must remember only a few passwords, but not when the user has many unique passwords for multiple accounts. Increasing the number of attempts allowed may reduce the incentive for users to write down passwords while still protecting the system against brute force attacks.

Future research should focus on methods that help participants remember which passwords are associated with which accounts, for instance, making the participants generate sentences that they associate with the account they are trying to access. As an example, a participant may generate a sentence such as I bought eight books at Amazon (Ib8b@A) or I sold books for $20 (Isbf$20). These would still be strong passwords, and they may help the user remember the password because they would have another line of association for recalling the password.

As a consequence of the tradeoff between security and memorability, users will continue to resort to writing down computer-generated passwords or generating passwords that are easy to remember and crack until a technique is established that promotes generation of memorable passwords that are secure. Some experts are skeptical about whether any method can allow generation of passwords that are both memorable and secure. This skepticism was noted by Sasse et al. (2001), who said, "Both security and usability experts have stated that recalling strong passwords is a humanly impossible task because strong passwords are non-meaningful items and hence inherently difficult to remember" (p. 126). However, Sasse et al. went on to indicate that they do not agree with this assessment, with their view being, "It is possible to create passwords that are strong and meaningful" (p. 126). Although designing methods that will generate secure and memorable passwords is difficult, the potential benefits from developing such methods justify their pursuit, especially because the username–password combination is likely to remain widely used in E-commerce and other on-line transactions for years to come.

7. Recommendations

Below we provide recommendations for properties of passwords that should enhance security and memorability.

(1) There should be a minimum length restriction. Passwords should be at least eight characters long. The longer the password, the less likely that it will be cracked.
(2) Inclusion of special characters and digits increase the security of passwords. However, if the digits and special characters are not placed in a meaningful manner, memorability of the password decreases.
(3) Avoid using simple patterns. For example, avoid placing the uppercase letter at the beginning of the password and/or special characters and digits at the end of the password. Simple patterns may assist recall of a password, but they also make the password easier to crack.
(4) Elaborative processing such as that required by the first-letter sentence generation technique can improve memory of passwords.
    (a) More secure passwords can be generated using the first-letter generation technique when a special character and digit are included in the sentence.
    (b) Memory for the special character and digit within the sentence is better when they are embedded in a way that is meaningful (e.g., I had 2 go home @ 5 o'clock = Ih2gh@5o).
(5) Administrators should use a lock-out procedure that results in an account lockout after a certain number of attempts. A "six-strikes-and-you're-out" rule is more justified than a "three-strikes-and-you're-out" rule. Due to the fact that users have many accounts for which they must generate passwords, several attempts may be required for the user to retrieve the unique password for a specific account. Allowing users multiple attempts to login will likely reduce their tendency to write down passwords, while at the same time minimizing the vulnerability to password hacking.
(6) Having users login multiple times after generating the password for an account increases its memorability. Simply re-entering the password immediately after its generation (e.g., for verification purposes) is not sufficient.

The recommendations above are mainly for alphanumeric passwords. However, graphical password systems provide an alternative (Wiedenbeck et al., 2005) to alphanumeric strings that should be explored more fully because they exploit people's ability to recognize pictures and faces with high accuracy.

Acknowledgements

Portions of Experiments 1 and 2 were presented at the 47th and 48th annual meetings of the Human Factors and Ergonomics Society, and portions of Experiments 3A and 3B at the Fifth Annual Security Conference, and are included in their proceedings. This research was supported in part by NSF ITR Cyber Trust grant #0430274 and a Scholarly and Creative Activity Committee Award from CSULB to the first author.

References

Ackerman, M.S., Mainwaring, S.D., 2005. Privacy issues and human–computer interaction. In: Cranor, L.F., Garfinkel, S. (Eds.), Security and Usability. O'Reilly, Sebastopol, CA, pp. 381–399.

Bergadano, F., Crispo, B., Ruffo, G., 1998. High dictionary compression for proactive password checking. ACM Transactions on Information and System Security 1, 3–25.

Bidgoli, H., 2004. Preface. In: Bidgoli, H. (Ed.), The Internet Encyclopedia, vol. 1. Wiley, Hoboken, NJ, pp. XXIII–XXV.

Bidgoli, H. (Ed.), 2006. Handbook of Information Security. Wiley, Hoboken, NJ.

Bishop, M., Klein, D.V., 1995. Improving system security via proactive password checking. Computers and Security 14, 233–249.

Blackwell, A., Anderson, R., Grant, A., 2004. Password memorability and security: empirical results. IEEE Security and Privacy, 25–31.

Brewer, W., 1977. Memory for the pragmatic implications of sentences. Memory and Cognition 5, 673–678.

Bunting, M., 2006. Proactive interference and item similarity in working memory. Journal of Experimental Psychology: Learning, Memory, and Cognition 32, 183–196.

Craik, F.I.M., Lockhart, R.S., 1972. Levels of processing: a framework for memory research. Journal of Verbal Learning and Verbal Behavior 11, 671–684.

Craik, F.I.M., Tulving, E., 1975. Depth of processing and the retention of words in episodic memory. Journal of Experimental Psychology: General 104, 269–294.

De Luis-Garcia, R., Alberola-Lopez, C., Aghzout, O., Ruiz-Alzola, J., 2003. Biometric identification systems. Signal Processing 83, 2539–2557.

de Winstanley, P.A., Bjork, E.L., 2004. Processing strategies and the generation effect: implications for making a better reader. Memory & Cognition 32, 945–955.

Gehringer, E.F., 2002. Choosing passwords: security and human factors. Proceedings of IEEE 2002, 369–373.

Halderman, J.A., Waters, B., Felten, E.W., 2005. Security through the eyes of users: a convenient method for securely managing passwords. In: Proceedings of the 14th International Conference on World Wide Web. ACM Press, New York.

Hilton, R., 2006. Password change policies. Retrieved from 〈http://blog.air0day.com/2006/04/27/password-change-policies/〉. Downloaded on May 23, 2006.

Ives, B., Walsh, K.R., Schneider, H., 2004. The domino effect of password reuse. Communications of the ACM 47, 75–78.

Jacoby, L.L., Craik, F.I.M., 1979. Effects of elaboration of processing at encoding and retrieval: trace distinctiveness and recovery of initial context. In: Cermak, L.S., Craik, F.I.M. (Eds.), Levels of Processing in Human Memory. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 1–21.

Klein, D., 1990. Foiling the cracker: a survey of, and improvements to, password security. Proceedings of the Second USENIX Security Workshop, 5–14.

Leyden, J., 2003. Office workers give away passwords for a cheap pen. The Register. Available on-line at 〈http://www.theregister.co.uk/2003/04/18/office_workers_give_away_passwords/〉. Downloaded May 21, 2006.

Neath, I., 1998. Human Memory: An Introduction to Research, Data, and Theory. Brooks/Cole, Pacific Grove, CA.

Parkin, A.J., 1984. Levels of processing, context, and facilitation of pronunciation. Acta Psychologia 55, 19–29.

Pinkas, B., Sander, T., 2002. Securing passwords against dictionary attacks. In: Proceedings of the Ninth ACM Conference on Computer and Communications Security. ACM, Washington, DC, pp. 161–170.

Posey, B.M., 2003. Recover lost passwords with these tricks and tools. Available online at 〈http://articles.techrepublic.com.com/5100-6329-5078287.html〉. Retrieved on December 10, 2006.

Proctor, R.W., Lien, M.C., Vu, K.-P.L., Schultz, E.E., Salvendy, G., 2002. Improving computer security for authentication of users: influence of proactive password restrictions. Behavior Research Methods, Instruments, & Computers 34, 163–169.

Rainbow Technologies, 2003. Password survey results (June 2003). Retrieved November 14, 2005, from 〈http://mktg.rainbow.com/mk/get/pwsurvey03〉.

Richardson, J.T.E., 2003. Dual coding versus relational processing in memory for concrete and abstract words. European Journal of Cognitive Psychology 15, 481–509.

Riddle, B.L., Miron, M.S., Semo, J.A., 1989. Passwords in use in a university timesharing environment. Computers and Security 8, 569–579.

Roediger Jr., H.L., Karpicke, J., 2006. The power of testing memory: basic research and implications for educational practice. Perspectives on Psychological Science 1, 181–210.

RSA Security, 2005. RSA security survey reveals multiple passwords creating security risks and end user frustration. Available on-line at 〈WWW.rsasecurity.com〉. Downloaded on May 7, 2006.

SafeNet, 2005. 2004 annual password survey results. Available on-line at 〈www.safenet-inc.com〉. Downloaded on May 21, 2006.

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

Schultz, E.E., Proctor, R.W., Lien, M.-C., Salvendy, G., 2001. Usability and security: an appraisal of usability issues in information security methods. Computers & Security 20, 620–634.

Schultz, E.E., 2005. Web security and privacy. In: Proctor, R.W., Vu, K.-P.L. (Eds.), Handbook of Human Factors in Web Design. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 613–625.

Slamecka, N.J., Graf, P., 1978. The generation effect: delineation of a phenomenon. Journal of Experimental Psychology: Human Learning and Memory 4, 592–604.

Tari, F., Ozok, A.A., Holden, S.H., 2006. Password management, mnemonics, and mother's maiden names: a comparison of perceived and real shoulder-surfing risks between alphanumeric and graphical passwords. In: Proceedings of the Second Symposium on Usable Privacy and Security SOUPS '06. ACM Press, New York.

Viega, J., 2005. Solutions to many of our security problems already exist, so why are we still so vulnerable? Queue, 41–50.

Wiedenbeck, F., Waters, J., Birget, J.-C., Brodskiy, A., Memon, N., 2005. PassPoints: design and longitudinal evaluation of a graphical password system. International Journal of Human–Computer Studies 63, 102–127.