2005

Password Selection by End Users from an eCommerce Site: An Empirical Study

B. Dawn Medlin
Appalachian State University, medlinbd@appstate.edu

Joseph A. Crazier
Appalachian State University, cazierja@appstate.edu

Dinesh S. Dave
Appalachian State University, daveds@appstate.edu

Password Selection by End Users from an E-Commerce Site: An Empirical Study

B. Dawn Medlin
Appalachian State University
medlinbd@appstate.edu

Joseph A. Cazier
Appalachian State University
cazierja@appstate.edu

Dinesh S. Dave
Appalachian State University
daveds@appstate.edu

ABSTRACT

Individuals generally have the responsibility of creating their own passwords on an e-commerce site. As users attempt to create a password that they can remember, they often create one that others can easily guess. This situation can also create another paradox, where the user cannot remember their password in their quest to create an unpredictable one. This paper examines what passwords were created by users on an e-commerce site, their gender, into what categories they appear, and how their choices could be identified through a software cracking program. This paper also addresses the security of users' password choices in comparison to positive password actions suggested by security experts. The results of this study could assist both consumers and e-commerce sites in recognizing and recommending secure password choices and policies.

Keywords

Security, Passwords, User Behaviors, E-commerce

INTRODUCTION

The Internet continues to grow at an ever-increasing rate, with secure e-commerce transactions becoming a necessity for both consumers and businesses. Even though there have been advances in security technology and authentication applications, one aspect remains constant: passwords still play a central role in system security. The difficulty with passwords is that all too often they are the easiest security mechanism to defeat.

Network policy enforcement and electronic commerce both depend on a single concept: trust. The administrator of a network essentially acts as the gatekeeper, hopefully ensuring that trust is not misplaced. This task is not an easy one. According to the Gartner Group, within a few years half of today's more robust methods of customer authentication won't be strong enough to safeguard against malicious acts of network terrorism such as phishing, spamming, and spoofing.

But without secure passwords, e-commerce sites invite online criminals attempt fraudulent schemes that mimic the goods and services that legitimate e-commerce merchants offer. With increasing numbers of users on an increasing array of e-commerce sites, often requiring the use of passwords, users often choose to reuse the same simplistic password, and do so on multiple sites (Campbell, P., Calvert, B., and Boswell, S., 2003). For consumers, identity theft and fraud can be very costly, while for e-commerce sites it can undermine consumer confidence, resulting in lost revenue and profits.

An additional concern is that many of the deficiencies of password authentication systems arise from the limitations of human memory. If humans were not required to remember a password, a maximally secure password could be established that would consist of a string of numbers, characters, and symbols. It would also not be a password that the individual had previously used. Redundancy of the same passwords is common because of the short-term memory of humans.

The purpose of this paper is to examine user generated passwords and the effect of memory, creativity, and gender as they relate to specific categories. Examination of these passwords illustrates the connectivity between password choices and users' cognitive limitations.

LITERATURE REVIEW

Research in the area of password security has dramatically increased over the past twenty years (Ives and Walsh, 2004). Even though there has been increased awareness surrounding password protection techniques and password security software, password vulnerabilities remain significant.

Most of today's computer and network operating systems provide measures to secure access to data, applications and the operating system by granting permissions and end user's rights based upon their machine or network logon. In most cases, the additional key security mechanism is a password, also known as a single-factor method of authentication, as the password is known (theoretically) to the end user. This mechanism generally consists of a combination of characters, numbers, symbols, and alpha characters that the user enters, along with his or her account name, to verify that the account really belongs to the person who is logging into the system. The password, then, functions like the key to a lock; anyone who has it can get in.

For most computerized systems, passwords are the first line of defense against hackers or intruders. There have been numerous published articles that have created guidelines on how to create better or safer passwords. One of the first guidelines was published in 1985 by the Department of Defense (DOD) and is still relevant today. Their guidelines recommended the following: 1) passwords must be memorized, 2) passwords must be at least six characters long, 3) passwords must be replaced periodically, and 4) passwords must contain a mixture of letters (both upper- and lowercase), numbers, and punctuation characters.

Most networks administrators and security experts have concurred with all of the above DOD recommendations, and in fact have added other recommendations. These suggested guidelines include never using personally identifiable information, beginning the password with an eight or nine character word or phrase and randomly adding numbers and/or symbols, and never use words that are commonly found in a dictionary. The enforcement of most of the above recommended password actions are not always evident on e-commerce or internet sites. Generally, in this venue, once a user has created a password it is generally never changed, unless the customer has forgotten or lost his or her password.

Given enough time though, any password can be cracked. But the amount of time it takes to crack a password depends only on how good the password is. Password crackers use a variety of methods and tools that includes guessing, dictionary lists, or brute force attacks. Word or dictionary lists are an automated program that includes a text file of words. The program repeatedly attempts to log on to the target system, using a different word from the text file on each try. A brute force attack is a variation of the dictionary attacks, but it is designed to determine passwords that may not be included in the text file used in those attacks. In a brute force attack, the attacker uses an automated program that generates hashes or encrypted values for all possible passwords and compares them to the values in the password file. When an attacker successfully gains access to a resource, they have the same rights as that user whose account has been compromised to gain access to that resource. If the account has sufficient privileges, the attacker can create a "back door" for future access, without concern for any status and password changes to the compromised user account.

An additional issue surrounding password security is the limitation of a human's cognitive ability (Pond, Podd, Bunnell and Henderson, R., 2000). If humans were not required to remember a password, a maximally secure password would be one with maximum length that could consist of a string of numbers, characters, and symbols, and in a manner that provided no redundancy. In fact, the requirements to remember long and complicated passwords are contrary to a well-known property of human memory. First, the capacity of human memory in its capacity to remember a sequence of items is temporally limited, with a short-term capacity of around seven items plus or minus two (Kanaley, R., 2001). Second, when humans remember a sequence of items, those items cannot be drawn from an arbitrary and unfamiliar range but must be familiar 'chunks' such as words or familiar symbols. Third, the human memory thrives on redundancy. In fact, studies have shown that individuals' short term memory will retain a password for approximately 30 seconds thereby requiring individuals to attempt to immediately memorize their passwords (Atkinson, R.C. & Shiffrin, R.M., 1968). It has also been shown that if an individual is interrupted before they fully memorize the password; it will fall out of their working memory and most likely be lost.

Additionally, it was found that if an individual was in a hurry when the system demanded a new password, individuals sacrificed either the concentration of the critical task at hand or the recollection of the new password (Brostoff and Sasse, 2001). Also, related to this issue is having to create the content for this new quickly demanded password. The pressure to

choose creative and secure passwords quickly generally results in individuals failing in their attempt to memorize this new password, which can result in higher reset rates or systems having to remind end users of their selections (Brostoff and Sasse, 2001).

If though an individual adheres to most security experts' suggestions about password security, it usually involves a tradeoff. If a password is easy to create and remember, it is most likely that it is easy for others to guess or a hacker to crack. In contrast, some passwords are very secure against hacking because they include numbers, symbols, and upper and lower case letters, and/or characters, but these are the passwords that are also the ones that are the most difficult for individuals to remember. The irony of the latter example is that in trying to create a more secure password, the user is significantly less likely to remember it.

With the increase of Internet shopping, online banking, credit card reconciliation, and bill paying, individuals are required to manage a multitude of accounts, and generally each requiring a password. This increased need for password selection has revealed several interesting issues associated with users' difficulty in developing and remembering passwords (Jones, T. 2002).

In order to combat the issue of having to remember so many different passwords some users have resorted to the selecting familiar terms such as a pet or family name, their own name, their phone number, or other common terms that could be found in a dictionary. British psychologist Helen Petrie, Ph.D., a professor of human/computer interaction at City University in London analyzed the passwords of 1,200 British office workers who participated in a survey funded by CentralNic, an Internet domain-name company in 2001. She found that most individuals' choices of passwords fell into one of four distinct password "genres" or categories (Table 1).

The first group, which she labeled as "family", comprised nearly half of the respondents. These individuals selected their own name or nickname, the name of a child, partner or pet, birth date, or significant number such as a phone or social security number. Further, Dr. Petrie found that these individuals choose passwords that symbolized people or events with emotional value or ties.

Category #	***Name***	***Definition***
1	Family	Name or nickname, name of a child, partner or pet, birthday
2	Fan	Names of athletes, singers, movie stars, fictional characters or sports teams
3	Fantasists	Interest in sex is evident in passwords such as "sexy" "stud" and "goddess"
4	Cryptic	Unintelligible passwords or a random string of letters, numbers and symbols, such as Jxa+157

Table 1: Petrie's Category Definition

One third of the respondents were "fans," using the names of athletes, singers, movie stars, fictional characters or sports teams. Dr. Petrie also found that these individuals were generally young and wanted to align themselves with the lifestyle represented by or surrounded around a celebrity status. Two of the most popular names were Madonna and Homer Simpson.

Eleven percent of responses were "fantasists." Petrie found that their passwords were comprised of sexual terms or topics. Some of the examples included in this category were terms such as "sexy," "stud", and "goddess".

The final ten percent of participants were identified as "cryptics." These users were seemingly the most security-conscious, but it should also be noted that they were also the smallest of all of the four categories. These individuals selected unintelligible passwords that included a random string of letters, numerals and symbols, such as Jxa+157.

Password choice according to a certain category may be an important finding, but the gender of the user and their choice of a password may be just as, if not more, and interesting one. Men and women differ in the ways they use language and their choice of words. Though men and women technically speak the same language, some scholars have concluded that men and women use language differently. According to Lakoff (1975), she suggests that women's language makes more frequent use of emotionally intensive adverbs such as "so," "terribly," "awfully," and "quite." Similarly, Eakins and Eakins (1978) observed that men and women use different vocabularies. They suggest that women's language is more punctuated with adjectives and adverbs that "connote triviality or unimportance" such as "sweet," "dreadful," "precious," and "darling" (p.30). The findings of Eakins and Eakins are also supported by the work of Aries (1976) and Leet-Peregrini (1980) as cited in Tannen (1990). Tannen categorizes women's talk as "interdependent" and "cooperative," whereas male conversational patterns express "independence" and assertions of power. Based on the scholarship cited above, we would ordinarily expect

men to be less inclined than women to engage in emotional and relationship patterns of communication with evidence of this possibly being found in their choice of passwords.

In summary, the study reported here has several goals: to assess the prevalence or lack thereof risky password practices in an online e-commerce site, and to explore the categories and gender in relationship to password selection. The remainder of the paper presents the findings and their implications.

RESEARCH METHODOLOGY

Data Collection

This study examines an online Internet bookstore's user password choices, the security of those choices, and the relationship of gender to his or her password choice. The data set was collected from an e-business site over a one year period from the summer of 2003 until the summer of 2004. The bookstore sells primarily religious books and did not require a signed consent from the user as the information was obtained via the owner of the site. This data is not based on a survey or on a simulation; it is current data using e-commerce users' actual passwords and demographic information for the current year measuring current consumer behavior and security practices after September 11th and the increasing press coverage of identity theft and other Internet hazards.

The data set contains 520 customer profiles, with 499 of them having identifiable genders. Of those identified, 143 are males (29%) and 356 are females (71%). The majority of the respondents are from the western United States (US), where the company was founded; however there are also many from other parts of the US.

This particular website did not enforce any particular password rules or suggest password guidelines during the sample period. Because of a lack of rules, system created passwords, and/or guidelines, passwords collected are of particular interest, as they show current consumers' password security practices when left to their own device and without being forced by a network or other administrator to create a "strong" password.

Rating Instrument

To assess the security level of the passwords we reviewed current best practices for online password security (Ohio State 2004, University of New Mexico, 2004, Department of Defense 1985, and Security Stats 2004), and aggregated the guidelines into an instrument that was used to rate the security level of each password.

Using the 5-point human rating system gathered from a review of the literature, a panel of three individuals familiar with the current literature related to password security developed the categories presented in Appendix A. To provide a greater degree of granularity for analysis and in order to recommend specific guidelines, a series of dichotomous yes or no questions were developed around the relevant guidelines, both positive and negative. From these questions, a standardized scoring system was developed (Appendix B).

In the scoring system we had eight questions concerning the positive recommendation of what people should be doing according to best practices, and an additional eight concerning negative "what not to do" guidelines. We then developed an aggregate score for the positive and negative section by assigning a 1 or a 0 to the password for each question and summing the scores for each section. An overall score was assigned by subtracting the overall negative score from the overall positive score. Since we have 8 positive questions and 8 negative questions, this gives us a theoretical range of between -8 and +8 for the overall score. Additionally, this allows us to identify specific strengths and weaknesses in the sample on each guideline and also to compare males and females independently on each question. Our rating sheet and questions are displayed in Appendix B.

RESULTS

The review of Table 2 suggests that end users, regardless of their gender, are using good password practices that include numbers, symbols, and upper and lower case letters. Interestingly, this result differs from the study conducted by Petrie (2001), in that Cryptics were the smallest group of users in her study but were the largest in this study. The second largest group in this study was "Other", where consumers' selected common words found in an English dictionary such as "checkers," "onions," "cake," etc. Females appeared to use these passwords more often than males. The next highest category that appeared was Family. Both genders equally used this category in forming their password.

As noted earlier, Dr. Petrie's study identified four distinct categories in the order of Family, Fan, Fantasists, and Cryptics, while this study identified eight separate categories. It should be noted that in this study the classification of groups become more refined. Our additional categories are Other, Faith, Place, and Numbers, which appear to be more descriptive of consumers' passwords (Table 3). It could be argued that the category of "Other," "Faith," "Place," and "Numbers," should be included in for instance the "Family" category in that it may represent words that relate to individuals choices of favorite food, his or her religious choices, favorite places or places that have emotional ties, and numbers that are meaningful to them personally. It appears that individuals and males in particular, are becoming more careful in their password selection.

	Overall		Males		Females	
Category	N	%	N	%	N	%
1 – Family	101	19.3%	27	18.9%	70	19.7%
2 – Fan	11	2.1%	2	1.4%	9	2.5%
3 – Fantasists	2	.4%	0	0%	2	.6%
4 – Cryptic	201	38.5%	64	44.8%	127	35.7%
5 – Faith	30	5.7%	7	4.9%	21	5.9%
6 – Place	7	1.3%	2	1.4%	4	1.1%
7 – Other	143	27.4%	33	23.1%	107	30.1%
8 – Numbers	25	4.8%	8	5.6%	16	4.5%
Total	520	100%	143	100%	356	100%

Table 2: Descriptive Statistics

Category #	Name	Definition
5	Other	Common English dictionary terms that did not include religious terms or places.
6	Faith	Terms associated with religion or religious activities
7	Place	Names associated with towns or cities
8	Numbers	A string of all numbers

Table 3: Current Study Category Definition

The results in Table 4 present the results of an ANOVA procedure with comparison of means. For both the Dichotomous and 5-Point scale the F-statistic provided a significant value. For the dichotomous scale we see that the means tend to be slightly more towards the positive. There are, however, some very significant differences by category. The Fan category is statistically the lowest mean and is significantly different from all the other groups. Appearing next are the categories of Faith and Other. These categories are not significantly different from one another, but are significantly different from the other categories.

For the original 5 points human categorization scale, 1 being not secure, 5 being very secure, we see that categories 1, 5, and 7 are not statistically different from each other. However, Category 8 is significantly higher than this group, indicating greater security. The cryptic category is highest of all with a mean value of 3.36.

This suggests that the mean of each category are statistically different. Also, the comparison of means identified four groups for the Dichotomous scale: Faith and Other; Fan; Cryptic; and Numbers. The five point scale identified three groups as: Cryptic; Numbers; and Fan, Other, and Faith. It should be noted that because of the small sample size in the categories of Fan, Fantasist, and Place we will not able to test them.

The review of Table 4 and Table 5 indicates password strength. In further review of Table 5 it suggests that the means for both Dichotomous scale and 5-Point scale are statistically different by gender. Interestingly, the mean values for males are higher as compared to females for both of the aforementioned scales. This result indicates that males are more cautious in selecting a secure password than females. This finding concurs with the earlier result as presented in Table 4. Because of the low number of the male sample size in some of the categories, we were not able to analyze the comparison category by category, but as an overall difference.

Variable	F-Statistics (p-value)	Category 1 Fan	Category 4 Cryptic	Category 5 Faith	Category 7 Other	Category 8 Numbers
Dichotomous positive/negative scale (-8 to 8)	111.51* (0.0001)	0.44 [A]	2.78 [C]	1.00 [B]	0.91 [B]	1.88 [D]
5 point human categorization	78.68* (0.0001)	2.34 [E]	3.36 [G]	2.37 [E]	2.52 [E]	2.92 [H]
Sample size (n)	--	101	201	30	143	25
Means with the same letter indicate that they are not significantly different in each row *H_o: $\mu_1 = \mu_4 = \mu_5 = \mu_7 = \mu_8$ vs. H_1: At least one of them are not equal						

Table 4: Results of Linear Model Procedure for the Comparison of Means

Variable	F-Statistics (p-value)	Male	Female
Dichotomous positive/negative scale (-8 to 8)	5.30* (0.0217)	1.85 [A]	1.47 [B]
5 point human categorization	7.95* (0.0050)	2.92 [C]	2.76 [D]
Sample size (n)	--	143	356
Means with the same letter indicate that they are not significantly different in each raw *H_o: $\mu_{Male} = \mu_{Female}$ vs. H_1: : $\mu_{Male} \neq \mu_{Female}$			

Table 5: Comparison of Means by Gender

CONCLUSION

Ultimately, it is the goal of every network administrator and e-commerce site to encourage individuals to choose more secure passwords. Achieving that goal, however, can be a difficult one. The problem appears that, though humans may be creative, we are still very predictable in our choices. It is evident through our research that individuals are becoming more aware of the need to create passwords that are relatively secure. Our research subjects, though different from the prior study (Petrie, 2001) have apparently learned the necessity of secure password selection. Overall, this study indicates that there are some statistical differences between the categories, and that the category's password may give some clues about its security level.

This study found that the majority of the consumers on this e-commerce site, identified as Cryptics, have seemingly learned to avoid most of the negative password practices as found by security experts. It is not known whether that knowledge has been gained through education, direction from administrators, information regarding ID theft, nature of the demographic population, or possibly a post 9-11 paradigm shift, but individuals have apparently seen the need for more secure passwords that involve both length and complexity.

Extensions of this study are also needed. One such issue surrounds the time it takes to crack a password and the relationship between crack times and selected passwords. Policies could be created from this information that could guide companies in their efforts to enforce password best practices. Additionally, further research into the areas of gender and categories may lead system administrators and e-commerce sites to make specific recommendations and/or policies related to password creation based upon gender. A study could also be conducted to see what specific policies or educational activities have

affected individuals in their password choices and that information could also guide companies in their security efforts in relationship to password selection. Lastly, the study will be extended by assigning appropriate weights to enhance the current password categories.

REFERENCES

1. Adams, A., and Sasse, M. (1999) Users are not the enemy. Communications of the ACM, 42(12), 41-46.
2. Andrews, L.W. (2004) *Passwords Reveal Your Personality*. Retrieved February 1, 2005 from http://cms.psychologytoday.com/articles/pto-20020101-000006.html.
3. Atkinson, R.C. & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In Spence, K.W. & Spence, J.T. (Eds.), *The Psychology of Learning and Motivation*, New York: Academic Press.
4. Brostoff, S. and Sasse, M.A. (2000) Are Passfaces more usable than passwords? A field trial investigation. *In Proceedings of HCI 2000*, September 5-8, Sunderland, U.K., 405-424.
5. Campbell, P., Calvert, B., and Boswell, S. (2003) *Security+ Guide to Network Security Fundamentals.* Boston, MA: Course Technology.
   Eakins, B. W., and Eakins, R. G. (1978) *Gender Differences in Human Communication.* Boston, MA: Houghton Mifflin.
   Department of Defense. (1985) *Password Management Guideline*. Retrieved September 2004, from http://www.alw.nih.gov/Security/FIRST/papers/password/dodpwman.txt.
6. Jones, T. (2002) "Too many secrets? Password proliferation leads to user fatigue," in *Columbia News Service, Columbia University Graduate School of Journalism*, New York.
7. Kanaley, R. (2001) Login error trouble keeping track of all Your sign-ons? Here's a place to keep your electronic keys, but you better remember the password. *San Jose Mercury News*.
8. Lakoff, R. (1975) *Language and woman's place.* New York: Harper Colophon Books.
9. Leet-Peregrini, H. M. (1980) *Conversational dominance as a function of gender and expertise. In H. Giles, W. P. Robinson, and P. M. Smith (Eds.), Language: Social psychological perspectives.* Oxford: Pergamon. 97-104.
10. Ohio State University. (2004) Password Best Practices, Office of Information Technology, The Ohio State University, March 4, 2004.
11. Pond, R., Podd, J., Bunnell, J., and Henderson, R. (2000) "Word Association Computer Passwords: The Effect of Formulation Techniques on Recall and Guessing Rates," *Computers and Security*, Vol. 19, pp. 645-656.
12. Richardson, L. (1987) "Gender stereotyping in the English language." In L. Richardson and V. Taylor (Eds.), Feminist frontiers II. New York: Random House.
13. University of New Mexico. (2004) *Password Methodology: How to make, remember and change good passwords.* Retrieved October 10, 2004 from http://www.unm.edu/cirt/accts/psswrdmethodology.html.

Appendix A: Original 5 Point Human Password Rating System

Level	Description	Criteria
1	Obvious name or number.	1. Password contains all or part of a username, email, name, etc. 2. Or is somehow related to known information, such as a spouse or a child's name. We may not have enough information to always assess this. 3. It is easily guessable by someone that knows a little about the individual, based on publicly available knowledge.
2	Common word or number	1. Password is a repeating number (i.e. 99999) 2. Or it is a common word recognizable as either a name or word likely to be found in a dictionary?
3	Words with number at start or end or non-repeating numbers	1. Example John52 or 73horses, combining recognizable words or phrases with a number. 2. Or having a non-repeating number, such as 345820, that does not appear to be a social security number, phone number or street address. 3. Or adding a compound word, such as lovejohn or tomuchfun. 4. Or unrecognizable alpha characters that did not form a common word or name (i.e. mtlyfll).
4	Unrecognizable alpha and number combinations	1. Has a mixture of unrecognizable alpha characters with numbers thrown in (i.e.jds932). 2. May include compound words with numbers lovejohn99.
5	Special characters	1. Numbers and letters mixed with the inclusion of special characters like #, %, and, @ etc. 2. Spaces and dashes are included in this category.

Appendix B – Dichotomous Password Scoring Sheet

Question	Yes	No
Positive Questions		
1. Does the password have both upper and lower case letters?	1	0
2. Does the password have both upper and lower case throughout the password, not just the beginning?	1	0
3. Does the password have both letters and numbers?	1	0
4. Does the password have both letters and numbers throughout, not just at the beginning or end?	1	0
5. Does the password have any special characters?	1	0
6. Does the password have at least 6 characters?	1	0
7. Does the password have 8 or more characters?	1	0
8. Does the password appear to be random?	1	0
Negative Questions		
9. Is the password the same as the username, email or name?	1	0
10. Does the password resemble the username, email or name?	1	0
11. Does the password appear to be the name of a person (real or in book)?	1	0
12. Does the password appear to be the name of a place (real or in book)?	1	0
13. Does the password appear to be a word that could be found in an English dictionary?	1	0
14. Does the password appear to be a word in a foreign dictionary?	1	0
15. Does the password appear to have a discernible pattern to it? (i.e. 123321, 8888888 or aabbccbbaa)?	1	0
16. Does the password appear to be a date?	1	0