5-899 / 17-500 Usable Privacy and Security
Text Passwords

Lecture by *Sasha Romanosky*
Scribe notes by *Ponnurangam K*
March 30, 2006

1 Topics covered

- Authentication and authorization
- Pass-sentences, pass-phrases and mnemonic passwords
- Randomly generated passwords
- User Study: Human selection of text passwords

1.1 Authentication and authorization

The management of multiple versions of user identities across multiple applications is difficult and one type of building blocks of identity management is[1]:

- Password reset
- Password synchronization
- Single sign-on
- Access management software

Identification, Authentication and Authorization are three components of Identity management. There are many ways of defining these terms. One of the way is :

- *Authorization* is a process by which an entity (such as a person or a computer system) determines whether another entity is who it claims

[1] http://infosecuritymag.techtarget.com/2002/apr/cover_casestudy.shtml

to be. And *authentication* ensures that a person is who he or she claims to be, but it says nothing about the access rights of the individual[2].

- *Identification* asks a person to identify himself by means of a token or an identification string such as an email address or account number; *authentication* is a process of providing the evidence of his or her identity; *authorization* allows an authenticated person to take a set of actions according to the permissions granted [1, pp. 104]. This is also discussed as “what you know, what you have and what you are” in [2].

Password management and passwords in general become an important topic in the identity management field. And passwords are one means by which the users authenticate themselves to the systems. There are many challenging topics in the area of passwords. There are three ways by which a person can remember an item: *uncued recall, cued recall* and *recognition* [1, pp. 108]. As part of improving the security systems developers should also improve the mechanism by which users can remember the password. For password recovery process most of the implementations use the mechanism by which the users are suppose to remember some unique word or phrase which can be used to obtain their password. These passwords can be in different forms: text based, graphical, etc. One of the questions could be, why is that the “challenging questions” asked to the users are during password recovery and not during the general authentication process also [1, chapter 6]. The challenge question posed to the users while password recovery have increased memorability and it is difficult for others to guess the answer for the challenge question. There are many ways by which the challenge questions are created: default list of questions available in the system, the questions can be created by users, etc.

1.2 Pass-sentences, pass-phrases and mnemonic passwords

A pass-phrase is a phrase that is more secure and probably easier for the users to recall and use them. Many studies have shown that users most of the time choose passwords that are from dictionary and so ends up being easy to make the dictionary attacks. A pass-phrase can be a complete sentence

[2] https://www.privacyassociation.org/images/stories/pdfs/CIPP_Privacy_Glossary_0306.pdf

and preferably a nonsensical one. It is believed that these sentences would be hard to break [3]. But the pass-phrase does not remove the possibility of vulnerability like social engineering and theft. To decrease the crackability of passwords, researchers designed pass-sentence using which they created the password. For example, “I ate my oatmeal today“ might become “I8my02-day” which is a pass-sentence password [3], [4]. Researchers have also tried designing a compliance score which is derived by comparing the actual phrase and the phrase that is given by the user to get authentication. For example, comparing the actual phrase “Jack bought a pizza at Martin’s for half a dollar” to the sentence given by the user “Jack got some food at Martins place.” But it is not clear about the efficiency and the effectiveness of using this compliance score mechanism. Analysis of such passwords requires natural language technologies which has inherent problems. *Mnemonic passwords* are passwords which are short form of the pass-sentence or pass-phrase which are used by the users to authenticate to the systems. The example mentioned above “I8my02-day” can also be considered as a mnemonic password.

1.3 Randomly generated passwords

Another way by which passwords can be created is randomly creating the characters of the password. There are many ways by which you can get these characters, one of the ways is using the syllables of a given language. Since these are syllables of a language it has a high pronounceability but one can create these random passwords which can be unpronounceable also. The unpronounceabilty can provide little bit more security to the passwords, but there are no studies showing such results. Since the characters are chosen from the same language this can have the affect of “smallest bucket” attacks, where the attackers can focus only on a small set of characters of a language.

1.4 Memorability

Any system which uses the password management system should implement the mechanism in which the users have a high level of memorability for the passwords. One of the question raised was, “On an average how many passwords have students used and do they remember all the passwords?” Most of the students in the class felt that they do not remember all the

[3] http://www.iusmentis.com/security/passphrasefaq/

passwords. Some students mentioned that they have about 6 - 12 passwords to manage. And they use different mechanism to memorize the passwords.

1.5 User Study: Human selection of text passwords

To understand whether users create stronger mnemonic or random passwords, researchers conducted a study for 2 weeks in February 2006. About 298 subjects took part in the study where the users were evenly distributed among the random and mnemonic password generation group. To study the strength of the passwords researchers did the following:

- Scored the strength of each password: created a scoring function using the "number of characters" which is the sum of characters in the search space, "length" which is the number of characters in the password. The function they used for the study is

$$passwordlength = log_{10}(numberofcharacters^{length}) \tag{1}$$

- Tried to crack each password: each password was tried to crack using dictionary attack and 62 hours of brute force attack.
- Frequency distribution: compared the frequency distribution of characters for both types of passwords.

1.5.1 Results

- Mnemonic passwords are generated from phrases from the Internet so they are susceptible to dictionary attacks
- Random generated passwords are less crackable than assumed

1.5.2 Questions / Comments

- What is the significance of the difference in the results with respect to User-generated score (15.7 and 17.2)?
- Since the study was conducted on the Internet with less control over the participants, how generalizable are the results?
- Since there was an ipod given for a winner, the participants had incentive to create a strong password, but it is not clear whether we have a similar situation in the real-world.

References

[1] Cranor, L. F., and Garfinkel, S. *Security and Usability: Designing Secure Systems that People Can Use.* O'Reilly., Aug, 2005.

[2] Lininger, R., and Vines, R. D. *Phishing: Cutting the Identity Theft Line.* Wiley, publishing Inc., 2005.

[3] Office of Information Technology. Passwords. Retrieved March 31, 2006, http://www.security.umd.edu/protection/passwords.html.

[4] Spector, Y., and Ginzberg, J. Pass-sentence a new approach to computer code. *Comput. Secur. 13*, 2 (1994), 145–160.