Password Based Shadow Attack: Quantitative Testing Analysis

A.Poorna Chandra Reddy
Department of Computer science and Engineering, Christu Jyothi Institute of Technology and Science, Jangaon

Abstract- **With the proliferation of websites, the security level of password-protected accounts is no longer purely determined by individual ones. Users may register multiple accounts on the same site or across multiple sites, and these passwords from the same users are likely to be the same or similar. As a result, an adversary can compromise the account of a user on a web forum, and then guess the accounts of the same user in sensitive accounts, e.g., online banking services, whose accounts could have the same or even stronger passwords. We name this attack as the shadow attack on passwords. To understand the situation, we examined the state of- the-art Intra-Site Password Reuses (ISPR) and Cross-Site Password Reuses (CSPR) based on the leaked passwords from the biggest Internet user group. With a collection of about 70 million real-world web passwords across four large websites in China, we obtained around 4.6 million distinct users who have multiple accounts on the same site or across different sites. We found that for the users with multiple accounts in a single website reused their passwords and for the users with multiple accounts on multiple websites reused their passwords across websites. For the users that have multiple accounts but different passwords, the set of passwords of the same user exhibits patterns that can help password guessing: a leaked weak password reveals partial information of a strong one, which degrades the strength of the strong one. Given the aforementioned findings, we conducted an experiment and achieved an improvement of guessing success rate with John the Ripper guessing tool. To the best of our knowledge, we are the first to provide a large-scale, empirical, and quantitative measurement of web password reuses, especially ISPR, and shed light on the severity of such threat in the real world.**

1. INTRODUCTION

Password-based authentication [1] is one of the most widely used methods to authenticate a user before granting accesses to secured websites. The wide adoption of password-based authentication is the result of its low cost and simplicity: a user can enter his or her passwords anywhere by a keyboard or a touch screen without any other extra devices. The popularity of passwords and the proliferation of websites, however, lead to a concern on password reuses between accounts on different websites [2] or even on the same websites. Moreover, the recent numerous high-profile password leakage events did not make the password situation better, and we ask the questions: What do password reuses mean to accounts between websites and even the ones within the same websites? What is the implication of a compromised website or account to others? How easy are shadow attacks, i.e., an adversary compromises an account utilizing the passwords of other accounts that are either on the same site or from other sites? To find out the answers, in this paper we analyze password reuses and shadow attacks empirically. It is well-known that passwords are usually reused by a user across different websites [2][3], yet little work has been devoted to understanding passwords being shared among multiple accounts of the same user on the same website. Since both password reuses within the same website and across multiple ones can enable shadow attacks, in this paper, we analyze the both scenarios:

A user creates accounts with the same password on the same websites, which we term as Intra-Site Password Reuses (ISPR), and a user creates accounts with the same password across different websites, which we term as Cross-Site Password Reuses (CSPR). While having the same passwords for multiple accounts is simple and convenient to users, it raises security concerns, e.g., if a password on one website is leaked, an adversary can have an enhanced chance to crack the other accounts of the same user, regardless of whether the accounts are on the same or

different websites. We note that account ownership can be identified by the registered email addresses. As a result, we argue that users' accounts with passwords of higher security level could be relatively easily compromised, given the knowledge of the passwords at a lower security level, e.g., web forums. Although the password reuses are known to researchers for years, a large-scale in-depth empirical analysis of password reuses is still absent so far. Das et al. [2] leverage 6,077 distinct accounts to answer the question of How often does a user reuse the same password across multiple sites? Our work is along the same line. Yet we conduct a first-of-its-kind in-depth empirical study on web password reuses (both ISPR and CSPR) at a much larger scale. We leverage a collection of more than 70million real-world leaked web passwords in clear text to investigate the fine-grained patterns and threats of password reuses. These leaked passwords are from four main-stream websites with millions of users in China: CSDN [4], Tianya [5], Duduniu [6], 7k7k [7]1. Luckily, two websites allow users to register multiple accounts using the same email address. This provides a valuable opportunity to study the ISPR, which has never been studied in the literature [2], to the best of our knowledge.

2. RELATED WORK

Password-guessing algorithms: This method assumes that you can retrieve the hash of the password to be guessed and that the hashing algorithm is the same between the rainbow table and the password.

Password cracking algorithms: Brute Force Password Cracking Algorithm trying to write a brute force password cracker in which tests all possible alphanumerical strings of length , then all possible strings of length

Attacks:

Phishing Attack: Phishing is the attempt to acquire sensitive information such as usernames, passwords, and credit card details (and sometimes, indirectly, money), often for malicious reasons, by masquerading as a trustworthy entity in an electronic communication.

Dictionary Attack: Vulnerability to password or decryption-key assaults can be reduced to near zero by limiting the number of attempts allowed within a given period of time, and by wisely choosing the password or key. For example, if only three attempts are allowed and then a period of 15 minutes must elapse before the next three attempts are allowed, and if the password or Key is a long, meaningless jumble of letters and numerals, a system can be rendered immune to dictionary attacks and practically immune to brute-force attacks.

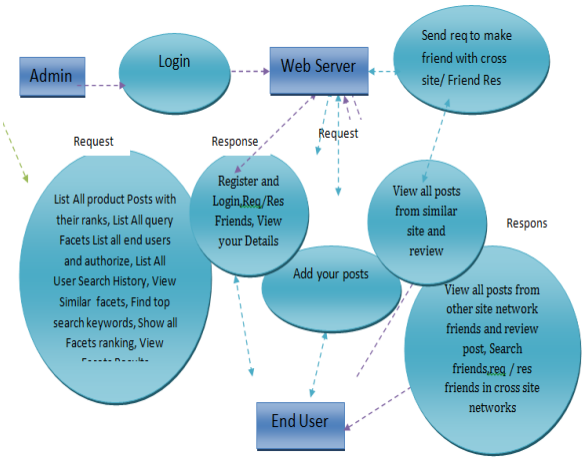
Brute Force Attack: Brute force (also known as brute force cracking) is a trial and error method used by application programs to decode encrypted data such as passwords or Data Encryption Standard (DES) keys, through exhaustive effort (using brute force) rather than employing intellectual strategies.

Password Guessing Attack: A common threat web developer's face is a password-guessing attack known as a brute force attack. A brute-force attack is an attempt to discover a password by systematically trying every possible combination of letters, numbers, and symbols until you discover the one correct combination that works.

3. BACKGROUND

Existing password schemes, many voices have called for password replacement or enhancement. Described many ancillary means to replace the current password-based authentication mechanism. Existing that a user should group their accounts when he or she has many different passwords. Proposed that a user should reuse their passwords in similar accounts, because they argue that it is impossible for a user to remember so many passwords, and input them in correct user interfaces. They proposed that each attack method has its strength in cracking passwords of certain strength. They also pointed out that the probability of guessing a correct password will decrease exponentially as the search space grows up, which is consistent with our experiment results. Proposed that a user should group their accounts when he or she has many different passwords.

4. SYSTEM FLOW



The so-called DFD is a bubble graph. The system refers to the system in terms of data entry, and a simple graphical tradition that can be used to address the implementation of various data from the output system data. Data Flow Diagram (DFD) is one of the most important modeling tools. The parts of the system used in modeling. These components include an external standard system and process flow in the system, and processes, and data into the method used by data exchange. DFD shows how to navigate through the information system and make changes in the changes it changes. The data from the input and output input will be applied as infectious and the flow of a graphic style is shown. DFD is a bubble chart. It can use DFD to represent system at any level of abstraction. These DFD levels are broken up into the increased flow of information and technological know-how.

5. EXPERIMENTAL RESULTS

6. CONCLUSION

According to the conclusion made in prior work that the number of valid email addresses on the same website should be smaller than 25 on averages, we removed all accounts whose email addresses had been used for more than 25 times on one website. Finally, we found that and CSDN contained some accounts with the same emails and usernames but different passwords. the passwords in Disport are stronger than those in Dips against online guessing attacks. The first two metrics indicate that the occurrence of the most frequent passwords in Disport is lower than the ones in Dips. The rate of CSPR is the lowest for users with education email addresses, and the number is smaller than the general rate . This result confirms our hypotheses that users in academic organizations are better educated with web security than common users and tend to use different passwords for accounts in different websites. Another reason may be that users incline to reuse passwords when registering with low-valued or easily replaceable email accounts. Academic emails, however, are difficult to be replaced.

REFERENCES

[1] R. Morris and K. Thompson, "Password security: A case history," Communications of the ACM, vol. 22(11), pp. 594–597, 1979.
[2] A. Das, J. Bonneau, M. Caesar, N. Borisov, and X. Wang, "The tangled web of password reuse," in NDSS'2014, 2014.
[3] D. Florencio and C. Herley, "A large-scale study of web password habits," in WWW'07 Proceedings of the 16th international conference on World Wide Web, 2007, pp. 657–666.
[4] CSDN,http://www.csdn.net/company/about.html
[5] Tianya,"http://help.tianya.cn/about/history/2011/06/02/ 166666.shtml."
[6] Duduniu,"http://baike.baidu.com/view/1557125.htm."
[7] 7k7k, "http://www.7k7k.com/html/about.htm."
[8] J. Bonneau, "The science of guessing: Analyzing an anonymized corpus of 70 million passwords," in 2012 IEEE Symposium on Security and Privacy (SP), 2012, pp. 538–552.
[9] J. Ma, W. Yang, M. Luo, and N. LI, "A study of probabilistic password models," in Proceedings

of IEEE Symposium on Security & Privacy, 2014.

[10] Z. Li, W. Han, and W. Xu, "A large-scale empirical analysis of chinese web passwords," in 23rd Usenix Security Symposium. San Diego: USENIX, 2014.

[11] D. Wang, H. Cheng, Q. Gu, and P. Wang, "Understanding passwords of chinese users: characteristics, security and implications," https://www.researchgate.net/, July 2014.

[12] D. Schweitzer, J. Boleng, C. Hughes, and L. Murphy, "Visualizing keyboard pattern passwords," in Visualization for Cyber Security, 2009. VizSec 2009. 6th International Workshop on. IEEE, 2009, pp. 69–73.

[13] Wikipedia, "Levenshtein distance," http://en.wikipedia.org/wiki/Levenshtein distance, May 2014.

[14] "Longest common subsequence problem," http://en.wikipedia.org/wiki/Longest common subsequence, May 2014.