International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

Shadow Attacks Based On Password Patterns Password Reuses

Bhavika Garse[1], N. D. Kale[2]

[1, 2]Computer Engineering Department, PVPIT Pune

Abstract: ***For secured websites, the password based authentication is frequently used approach for authenticating the end user before granting the access. The growing use of password based authentication approach at increasing websites leads to the important issue of possibility of password reuses among accounts of various websites or similar websites. Additionally, the recent study on numerous high profile password hacking claims that password situation is not better. Under such cases, there is huge possibility of shadow attacks in which an attacker can successfully compromise the account that reuses the password of other accounts those are from similar website or different websites. The reuse of passwords for different accounts under same website is called as Intra-Site Password Reuses (ISPR). The reuse of passwords for different accounts under different websites is called as Cross-Site Password Reuses (CSPR). Therefore in order to prevent such shadow attacks on passwords, first we need to understand and examine the both ISPR and CSPR based on publically available password datasets. However, there is no in-depth empirical study conducted in literature except the one very recently introduced on Chinese password datasets. However the problem with this method is that they are removing the duplicate profiles and passwords largely in their pre-processing step, this can reduces the scalability of password reuses.***

Keywords: Pre-processing, Dataset, ISPR, CSPR

1. Introduction

Password-based authentication is one in every of the foremost wide used methods to demonstrate a user before granting accesses to secured websites. The wide adoption of password-based authentication is that the results of its low value and simplicity: a user will enter his or her passwords anyplace by a keyboard or barely screen with none alternative additional devices. The popularity of passwords and therefore the proliferation of websites, however, result in a priority on password reuses between accounts on totally different websites or perhaps on identical websites. Moreover, the recent various high-profile password leakage events didn't build the password scenario higher, and that we raise the questions: What do password reuses mean to accounts between web sites and even those among identical websites? What's the implication of a compromised website or account to others? However simple are shadow attacks, i.e., an someone compromises an account utilizing the passwords of alternative accounts that are either on a similar web site or from alternative sites? To search out the answers, during this paper we tend to analyze password reuses and shadow attacks by trial and error.

It is well-known that passwords are usually reused by a user across different websites, 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: (i) a user creates accounts with the same password on the same websites, which we term as Intra-Site Password Reuses (ISPR), and (ii) 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.

The continued domination of passwords over all other methods of end-user authentication is a major embarrassment to security researchers. As web technology moves ahead by leaps and bounds in other areas, passwords stubbornly survive and reproduce with every new web site. Extensive discussions of alternative authentication schemes have produced no definitive answers. Over forty years of research have demonstrated that passwords are plagued by security problems and openly hated by users. We believe that, to make progress, the community must better systematize the knowledge that we have regarding both passwords and their alternatives. However, among other challenges, unbiased evaluation of password replacement schemes is complicated by the diverse interests of various communities. . In our experience, security experts focus more on security but less on usability and practical issues related to deployment; biometrics experts focus on analysis of false negatives and naturally-occurring false positives rather than on attacks by an intelligent, adaptive adversary; usability experts tend to be optimistic about security; and originators of a scheme, whatever their background, downplay or ignore benefits that their scheme doesn't attempt to provide, thus overlooking dimensions on which it fares poorly. As proponents assert the superiority of their schemes, their objective functions are often not explicitly stated and differ substantially from those of potential adopters. Targeting different authentication problems using different criteria, some address very specific environments and narrow scenarios; others silently seek generic solutions that fit all environments at once, assuming a single choice is mandatory. As such, consensus is unlikely. These and other factors have contributed to a longstanding lack of progress on how best to evaluate and compare authentication proposals intended for practical use. In response, we propose a standard benchmark and framework allowing schemes to be rated across a common, broad spectrum of criteria chosen objectively for relevance in wide ranging scenarios, without hidden agenda.

2. Literature Survey

In this section we are presenting the different methods those are presented to eliminate duplicate URLs.

Current research on duplicate URL detection can be classified in two different methodologies: content based and URL based. In content based methods, it is necessary to download all contents of the URL, inspect it and then match it. Thus this method consumes lots of resources and to avoid such a waste of resource several URL based methods has been proposed. In the paper we are focusing on URL based methods.

3. Related work

R. Morris and K. Thompson, "Password security: A case history," Communications of the ACM, vol. 22(11), pp. 594–597, 1979: This paper describes the history of the design of the password security scheme on a remotely accessed time-sharing system. The present design was the result of countering observed attempts to penetrate the system. The result is a compromise between extreme security and ease of use.

A. Das, J. Bonneau, M. Caesar, N. Borisov, and X. Wang, "The tangled web of password reuse," in NDSS'2014, 2014:We study several hundred thousand leaked passwords from eleven web sites and conduct a user survey on password reuse; we estimate that 43-51% of users reuse the same password across multiple sites. We further identify a few simple tricks users often employ to transform a basic password between sites which can be used by an attacker to make password guessing vastly easier. We develop the first cross-site password-guessing algorithm, which is able to guess 30% of transformed passwords within 100 attempts compared to just 14% for a standard password-guessing algorithm without cross-site password knowledge.

3.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.: Author report on the largest corpus of user-chosen passwords ever studied, consisting of anonymized password histograms representing almost 70 million Yahoo! users, mitigating privacy concerns while enabling analysis of dozens of subpopulations based on demographic factors and site usage characteristics.

4.J. Ma, W. Yang, M. Luo, and N. LI, "A study of probabilistic password models," in Proceedings of IEEE Symposium on Security & Privacy, 2014: A probabilistic password model assigns a probability value to each string. Such models are useful for research into understanding what makes users choose more (or less) secure passwords, and for constructing password strength meters and password cracking utilities.

A. Existing System Architecture

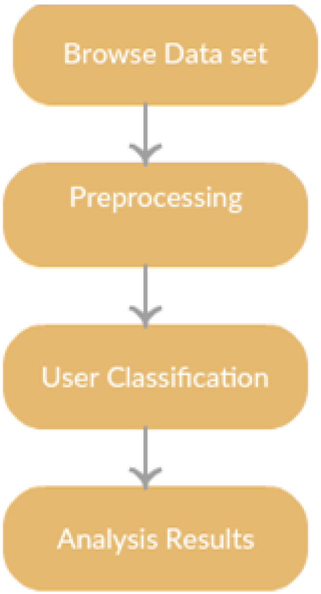


Figure 1: Existing System Architectural Diagram

For the users who have accounts on different websites, according to our research, 33.16 ± 8.91% of users use the same passwords across two sites (CSPR). This percentage is lower than the ones (4351%) reported in the literature of Das et al., because we took an conservative approach to process data and may have excluded a few reused passwords. For the users who have multiple accounts on a website, in our research, 59.72% of them reused their passwords (ISPR). This percentage is higher than the upper bound of the CSPR rate reported by Das et al. This suggests that users tend to reuse their passwords on the same websites than across multiple websites. We further investigate the security strength of the reused passwords in terms of how easily they can be guessed correctly by an adversary with dictionaries. With the same metrics (e.g., α-guesswork, α-workfactor) used by Bonneau, we find that the reused passwords across sites are stronger (i.e., harder to guess) against online password guessing attacks than all passwords, while intra-site reused passwords perform similarly to all passwords against online password guessing attacks. When we conducted offline password guessing attacks, all reused passwords perform weaker than all passwords. Even though some users use different passwords for their accounts across different websites, their passwords are sometimes created using the same building blocks. For example, among the users who use different passwords on the four websites, 15.36% of them add prefix to create passwords and 9.03% of them add suffix.

4. Proposed Approach Framework and Design

a) Problem Definition

For secured websites, the password based authentication is frequently used approach for authenticating the end user before granting the access. The growing use of password based authentication approach at increasing websites leads to the important issue of possibility of password reuses among accounts of various websites or similar websites. Additionally, the recent study on numerous high profile password hacking claims that password situation is not better. Under such cases, there is huge possibility of shadow attacks in which an attacker can successfully compromise the account that reuses the password of other accounts those are from similar website or different websites. The reuse of passwords for different accounts under same website is called as Intra-Site Password Reuses (ISPR). The reuse of passwords for different accounts under different websites is called as Cross-Site Password Reuses (CSPR). Therefore in order to prevent such shadow attacks on passwords, first we need to understand and examine the both ISPR and CSPR based on publically available password datasets. However, there is no in-depth empirical study conducted in literature except the one very recently introduced on Chinese password datasets. However the problem with this method is that they are removing the duplicate profiles and passwords largely in their pre-processing step, this can reduces the scalability of password reuses.

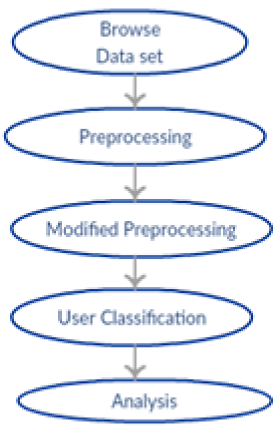
b) Aims and Objectives

The main aim of this project is to present the improved approach for dust detection and removal using genetic algorithm based multi sequence alignment method. Below are basic objectives of this method.

- To present literature review of different methods dust detection and removal.
- To present limitations of existing techniques.
- To present proposed algorithms and framework.
- To present practical analysis and performance evaluation.

c) Proposed System Architecture

As the number of websites is increasing, the security level of password enabled accounts is no longer secured. It may possible that end users can create the many accounts on either same website or multiple websites, hence the passwords for all accounts from the same user likely to similar or same. This leads the users account can compromised by attackers and then identify the same user passwords in sensitive accounts like banking accounts. This type of attacks on password is basically known as shadow attacks on password. In this project, we are presenting the framework to study the state of art of Cross Site Password Reused (CSPR) and Intra Site Password Reuses (ISPR) based on large scale password datasets in order to improve the password reuse success rate.



d) Mathematical Model

Terminologies: We call the passwords that are used more than once by a user (either ISPR or CSPR) in our dataset as reused passwords, and the pair of different passwords used by the same user as diverse password pairs.

Pre-processing: To ensure that all evaluated accounts are valid and map to real users, we pre-processed the leaked password data sets by removing rogue accounts before experiments.

User Classification: We extracted the following three types of accounts:

Users each of whom has at least two accounts on the same website.

Users each of whom has at least two accounts across different websites (cross-site users).

Users that belong to the intersection of the previous two data sets.

Min-entropy, $H\infty$

$H_{\infty} = -\log_2 (p)$

Marginal success rate or β-success rate,

$\tilde{\lambda}_{\beta} = \log_2 (\beta/\lambda_{\beta})$

Guesswork G

$\tilde{G} = \log_2 (2.G-1)$

α.-guesswork

$\tilde{G}_{\alpha} = \log_2 ((2.G_{\alpha}-1)/\lambda_{\mu_{\alpha}}) + \log_2 (1/(2-\lambda_{\mu_{\alpha}}))$.

5. Conclusion and Future Enhancement

To the best of our knowledge, this is the first empirical study on web password reuses by analyzing a large number of sample data. Although the web password reuses are known to researchers and Internet users, it is yet to perform a large-scale empirical study. We obtained 2,671,443 distinct users each of whom has at least two accounts from the same site, and 2,306,055 distinct users each of whom had at least two accounts from different websites. We also obtained 350,849 distinct users who has at least two accounts on the same site and across sites simultaneously. The quantitative answers shed lights on the serious threat of web password reuses, i.e., password shadow attacks, where an adversary may attack an account of a user using the same or similar passwords of his/her other less sensitive accounts. As a future direction, we would study CSPR from both adversaries' and defenders' points of view, leveraging the logs or activities that are available in the public domain. In addition, we will evaluate how the password policies affect CSPR after understanding the policies of these four websites. Last but not the least, we plan to study the impact of single sign-on tools on password reuses.

6. Acknowledgment

It gives me a great pleasure and immense satisfaction to present this paper of topic " **Shadow Attacks Based On**

Password Patterns Password Reuses" which is the result of unwavering support, expert guidance and focused direction of my guide **Prof.N.D.Kale** to whom I express my deep sense of gratitude and humble thanks, for his valuable guidance.

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. 657–66.
[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.