Shadow Attacks based on Password Reuses: A Quantitative Empirical Analysis

**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, 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 (i.e., 668 million members in China). 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, 59:72% reused their passwords and for the users with multiple accounts on multiple websites, 33:16 \_ 8:91% 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 a 39.38% 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.

**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.

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.

It is well-known that passwords are usually reused by a user across different websites 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.

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. The definitions of prefix and suffix patterns are described in Section 3.2.4. \_ Utilizing our findings to facilitate password guessing, we achieve a 39.38% improvement of password guessing success rate based on the JtR (John the Ripper) tool. By cracking a user’s weaker passwords first, an adversary greatly improves his chances of successfully guessing a stronger password of the same user. This suggests that the strength of a user’s passwords are somehow determined by the weakest one. Thus, shadow attacks can undermine the strength of relatively strong passwords.