Survey of Services that Store Passwords in a Recoverable Manner

Kazutoshi Itoh and Akira Kanaoka(✉)

Toho University, Miyama 2-2-1, Funabashi, Chiba 274–8510, Japan
akira.kanaoka@is.sci.toho-u.ac.jp

Abstract. Passwords entered by users in web services and applications are essential and confidential information. Therefore, it is ideal for difficulty storing them to decipher in case of unauthorized intrusion from the outside. As a typical example, passwords are converted into hash values using the SHA2 algorithm and stored. However, not all web services and applications implement the ideal storage method. There have been many incidents in which personal information has been leaked. In some cases, the passwords were not stored correctly on the server-side but in plain text or encrypted in a reversible form. The passwords were leaked when there was an unauthorized intrusion or other damage. This research aims to clarify the actual situation of how services and applications store users' passwords in plaintext or reversible form on the server-side through external observation surveys. The method is to list the survey targets for each service or application and conduct the survey for each service or application. As a result of the survey, there were no services or apps that were confirmed to have implemented inappropriate storage methods in both the top sites in the Alexa ranking and the top apps in the Google Play ranking, and the survey revealed that there were not many services that returned plain text in general.

Keywords: Password · Plaintext · Web Technologies

1 Introduction

Passwords entered by users in web services and applications are essential confidential information, and therefore, it is ideal for keeping them in a complex state to decipher in case of unauthorized intrusion from outside. As a typical example, there is a method of storing passwords by converting them into hash values using a cryptographic hash function such as SHA-256. However, not all web services and applications implement the ideal storage method.

There have been several incidents where personal information has been leaked from web services and applications. Although there are various causes, there are cases where passwords themselves are leaked in the event of an unauthorized intrusion because the server did not store the passwords properly and encrypted them in plain text or reversible form.

Ideally, users should have different passwords for different services and applications, but in reality, it is known that users often tend to use the same password

A. Moallem (Ed.): HCII 2023, LNCS 14045, pp. 68–77, 2023.
https://doi.org/10.1007/978-3-031-35822-7_5

for multiple services and applications [1]. Given that it is difficult to change such users' behavioral characteristics, storing passwords in plaintext or reversible encryption should be avoided.

In this research, the current state of what services and applications store passwords in plaintext or reversible form on the server side is investigated, and the commonalities specific to these services are examined. In order to investigate whether or not services and applications store passwords on the server side in plain text or reversible form, observations are conducted on typical web services and applications.

A survey was conducted on services that store passwords in plaintext or reversible form. The results of a survey of Alexa's Top Japan Domains (#1 to 103) and Google Play's Top Free Android Apps (#1 to 28) did not reveal any web services or apps that can be determined to be storing passwords in plain text or reversible form. No web services or apps were found that could be determined to store passwords in plaintext or reversible form.

Forty services were surveyed, and seven services were found to store passwords in either plain text or reversible form. Based on these investigations, we analyzed the similarities among the services that store passwords in plain text or reversible form from two perspectives: the appearance of the web service or application and the HTML data provided by the server. When the HTML data were compared, no common points were found, and no common points that could be considered the cause were discovered.

The structure of this paper is as follows. First, in Sect. 2, several cases of password leaks are listed, and related studies are explained in Sect. 3. Section 4 presents the methodology, target of the survey, and results of the external observation survey of password storage methods conducted in this study. Section 5 discusses future issues, and Sect. 6 summarizes.

2 Examples of Plaintext Password Leaks

2.1 Largest Breach Cases

An investigation by UpGuard [2] reported that in October 2015, information from the Chinese service NetEase was leaked, and that the leaked information included hundreds of millions of plaintext passwords.

In the case of the Evite breach, 101 million records containing plaintext passwords were compromised in 2013, but it was not discovered until 2019.

In the case of Russia's VK in 2012, 93 million records were compromised, including plaintext passwords and email addresses.

The cases described here are the largest plaintext password leaks in history, but it is not hard to imagine that a massive number of plaintext password leaks, including small ones, occur frequently.

2.2 Facebook Password Breach

On March 21, 2019, Facebook announced its stance and efforts in response to the security incident in "Keeping Passwords Secure—Facebook," [3] which revealed that some passwords for Facebook and other applications were recorded in log files in plaintext format. On April 18 of the same year, Facebook announced its stance and efforts in response to the security incident. It was followed by an update to the release on April 18 of the same year. Facebook explained that it would notify the affected users.

2.3 Google's Plaintext Storage of Some G Suite Passwords

In a blog post on May 21, 2019, US time, Google informed customers of its G Suite service that some passwords were being stored on internal servers without encryption [4]. In the post, Suzanne Frey, vice president of engineering at Google Cloud Trust, said that the bug only affects enterprise users.

3 Related Works

3.1 Behavioral Tendencies of Developers

Naiakshina et al. have been conducting ongoing research on how software developers implement password storage. They surveyed students, freelancers, and corporate developers, and the results showed that a large percentage of the participants implemented password storage inappropriately [5–8].

These results suggest that it is generally difficult to implement proper storage, as participants in experiments who have not received proper lectures tend to store without encryption or hashing in the first place, and even when hashing, salt and stretching are not used.

3.2 Password Reuse

In today's world, where it is not uncommon to use multiple services, it is desirable to set different passwords for each service, but in reality, users tend to reuse passwords and share passwords across multiple services. While these were empirically known, In 2016, the first academic study was conducted by Wash et al. [1]. They studied the behavior of 134 people over six weeks and found that among the participants in the experiment, each person reused their password on 1.7–3.4 sites.

Based on the assumption that users are likely to reuse passwords, secure password storage becomes even more critical.

4 Observation Survey of Password Storage Methods

4.1 Methodology

The purpose of this research is to clarify the actual situation of what services and applications store users' passwords in plaintext or reversible form on the server-side. There are two approaches to this: external observation and internal observation. Internal observation is not realistic because it cannot be realized unless the confidential information on the server-side can be viewed. Therefore, in this research, we use external observation.

In this research, a survey is conducted on typical services and applications, and after discovering the services that store passwords in plain text or reversible form on the server-side, the method is to classify them and find common and similar points. First, typical services and applications try to reset their passwords and check if the service or application sends back the set password in its response. It is then checked whether the password is stored in plain text or reversible form.

Next, services and applications that store passwords in the reversible form are categorized, and similarities or commonalities are examined and analyzed. By taking these steps of investigation and consideration, we thought we could discover the root cause of implementing problematic storage methods. In this section, we describe these investigation targets, methods, and environments and the results of our investigation.

4.2 Website Survey of Alexa's Top Services

In order to investigate the status of password storage for typical services, a survey is conducted on the top Alexa sites.

The first step is to register as a user for each of the services on the top Alexa sites, log off, and then take action as "forgot password" when logging back in, and then check the response. There are two possible responses: one is to disclose the original password, and the other is not to disclose the original password. In this survey, the ones with the original password disclosure are searched.

Survey Target. Although there are many web services, including those operated by individuals and companies, this study decided to investigate the status of password storage for typical web services, and referring to "Japan Top Domains[1]" provided by Alexa, web services ranked from 1st to 103rd were targeted.

If the web services are used frequently by users and have a relatively high level of attention, storing passwords in plain text or reversible form suggests that the social impact of security threats is high.

Research Methods and Environment. In order to search for commonalities and similarities in what services store passwords in plain text or reversible form, items that can be used as indicators are listed. In order to list the survey targets, Alexa's Top Japan domain is accessed using a web browser, and the service name, URL of the registration page, and the existence and type of ID linking are listed. The listing was performed in one day. The Alexa rankings may change after a while, so the list should be compiled in a day. Although some services allow ID integration, each service should create its account without using ID integration.

[1] https://www.alexa.com/topsites/countries/JP.

After the list is completed, the survey is conducted one by one, starting from the first place. The survey method is described below.

- Access the registration page URL and actually register as a user.
- Save screenshots of the registration screen and any emails received during registration.
- After the user registration is completed, the user can take an action to the same service saying "I forgot my password" and record the pattern of the reaction, such as whether the service sends back the original password, and if not, what procedures are taken, using screenshots and memos.
  - If there are features specific to that service, such as how to register or reset passwords, record those as well, using screenshots or notes.

When registering for an account, some services read the past login information and log in automatically. Therefore, in order to avoid saving the login information and to proceed with the registration process smoothly, everything was performed in Google Chrome's incognito mode.

Result. The survey was conducted for the services ranked from 1st to 103rd. The detailed information of each service should not be presented from an ethical point of view.

A screenshot was used to save each transition through the registration screen of each service during the registration process. After the 51st place, the policy was to take screenshots only for those with the original password disclosure.

After listing and investigating Alexa's "Japan Top Domains" from No. 1 to No. 103, none were found to have the original password disclosure. Of the 103 cases, there were 9 cases where there was no user registration, 26 cases where registration was impossible during the survey experiment due to overseas services and the need to register a phone number, and 1 case where the service had been terminated.

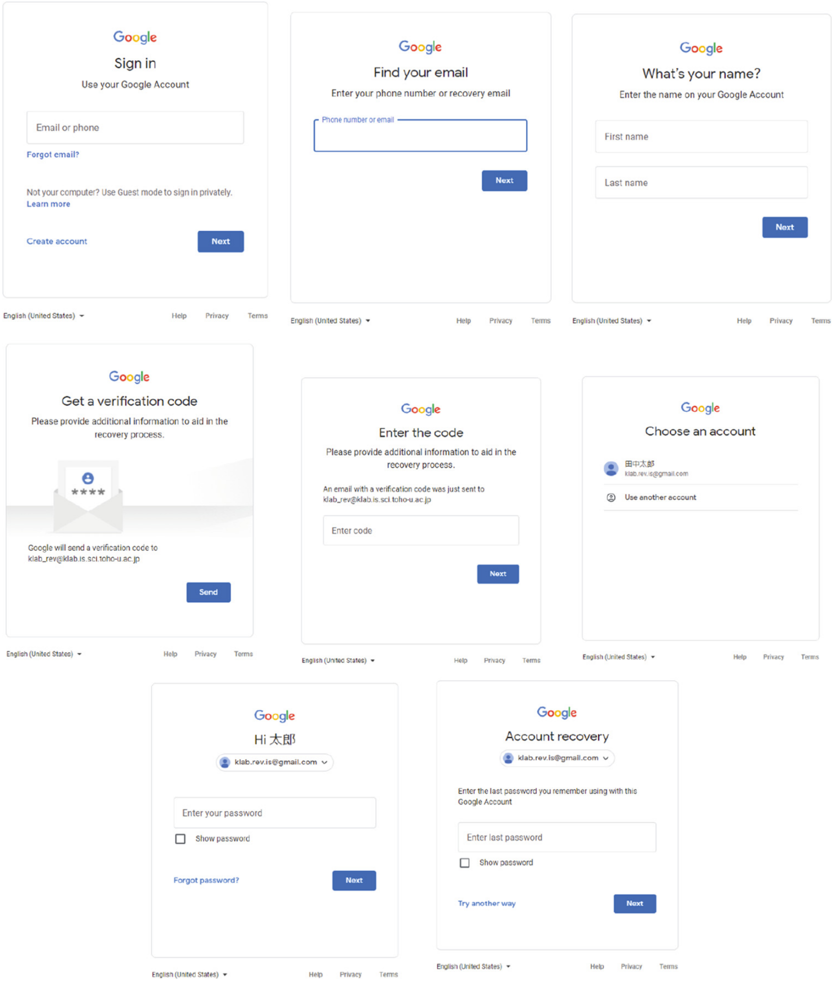
There were 67 sites that had completed membership registration, of which 48 were duplicates, excluding sites that had duplicate registration accounts, such as those dependent on Google.

The duplicates were Google account, Livedoor account, DMM account, Microsoft account, and Amazon account.

Figure 1 shows the registration screen of Google with Alexa rankings of 1, 2, 5, and 31, as well as the corresponding screen when you forget your password. Google has a unified authentication mechanism for its various services, and all logins are redirected to a page on the same domain.

4.3 Google Play Top Ranking Apps Survey

A "forgot password" action is taken on the top apps in Google Play to see how they respond. There are two possible responses: one is to disclose the original password, and the other is not to disclose the original password. This study will search for those with the original password disclosure.



Google
Sign in
Use your Google Account
Email or phone

Not your computer? Use Guest mode to sign in privately.

Google
Find your email
Enter your phone number or recovery email
Phone number or email

Google
What's your name?
Enter the name on your Google Account
First name
Last name

Google
Get a verification code
Please provide additional information to aid in the recovery process.
Google will send a verification code to klab_rev@klab.is.sci.toho-u.ac.jp

Google
Enter the code
Please provide additional information to aid in the recovery process.
An email with a verification code was just sent to klab_rev@klab.is.sci.toho-u.ac.jp
Enter code

Google
Choose an account
田中太郎
klab.rev.is@gmail.com

Google
Hi 太郎
klab.rev.is@gmail.com
Enter your password
□ Show password

Google
Account recovery
klab.rev.is@gmail.com
Enter the last password you remember using with this Google Account
Enter last password
□ Show password

Fig. 1. Google (Alexa Rank: 1, 2, 5, 31) registration screen and the response screen when you say "forgot password"

Survey Target. Although there are many applications distributed on Google Play, both paid and free, we decided to investigate the password storage status of representative applications in this study. We referred to Google Play's popular overall ranking "Free Top Android Apps[2]," and targeted applications ranked 1st to 28th. If the apps that are used frequently by users and have a relatively high profile store passwords in plain text or reversible form is considered to have a high social impact on security threats.

[2] https://play.google.com/store/apps/top?hl=ja.

Research Methods and Environment. In order to search for commonalities and similarities in what apps are storing passwords in plain text or reversible form, index items are listed. As a list of targets for the investigation, the free Top Android apps in Google Play's overall popularity ranking are installed from the top, and user registration is checked. Since the Google Play ranking may fluctuate over time, the listing is conducted in a single day. An Android device is used for the survey to install the apps.

The survey is conducted one by one, starting from the first place after the list-up. The survey method is described below.

- Register users for the installed application.
- Save screenshots of the registration screen and any emails received during registration.
- After the user registration is completed, the application is erased to return to the state before the user registration. After that, take action "I forgot my password," and record the pattern of the reaction, such as whether the app sends back the original password, and if not, what procedures are taken, using screenshots and notes. Use screenshots and notes to record the reaction.
    - If there are features specific to that app, such as how to register or reset passwords, record those as well, using screenshots or notes.

Result. The survey was conducted for the apps ranked from 1st to 28th. The detailed information of each app should not be presented from an ethical point of view.

A screenshot was used to save each transition of the registration screen of each application during the registration process. After 11th place, the policy was to take screenshots only for those with the original password disclosure.

After listing and investigating Google Play's popular overall ranking of "Free Top Android Apps" from No. 1 to No. 28, I could not find any with the original password disclosure. As a breakdown, out of 28, 6 apps were not registered, and 5 apps could not be registered. 16 apps completed registration, of which 3 were automatically linked to the device's Google account, and 12 were duplicates, excluding apps that depended on Google for the registered account. The duplicated ones were Google account and d-account.

4.4 Survey of Services Where Information on Inappropriate Password Storage Was Obtained Through Web or SNS Searches

Alexa's survey of the top 103 Japanese Top Domains and the top 28 Free Top Android Apps in Google Play's overall popularity ranking did not reveal any services that store passwords in plain text or reversible form. Therefore, as an additional survey, a survey was conducted not only on typical sites and apps, but also on web and SNS services that reportedly store passwords in reversible form.

Survey Target. In order to discover services that store passwords in plain text or reversible form, this survey will be conducted on services that have been informed that their original passwords are returned through web searches and SNS searches.

Research Methods and Environment. Since this survey will be conducted on the services that were informed when the original password was sent back, the survey method and environment is the same as the method and environment for Alexa top sites.

The list of survey targets was made based on web search results and Twitter search results. The search was conducted using the following free words, and then the listed targets were further scrutinized.

- "password" "plain text" (in Japanese)
- "password" "e-mail" (in Japanese)

Result. The number of cases actually investigated after listing was 40. As a result of our survey, a total of seven services that had disclosed the original password was found. In addition, we found three services that did not disclose the original password but sent a new or temporary password to the email after the user said, "I forgot my password.

The survey was limited to services informed that the original passwords would be sent back, but the survey results showed that few services sent back the original passwords. Among the services that were informed that the original passwords were sent back, it can be inferred that there were several cases where the passwords were stored in plain text when the information was provided but was later modified.

4.5 Analyze the Common Elements Between Services that Store Passwords in Plaintext or Reversible Form

The survey discovered services that store passwords in either plain text or reversible form. An analysis of the commonalities among services that store passwords in either plaintext or reversible form are conducted. The purpose is to understand how this situation is driven by a specific implementation or a specific operator.

Analysis Method. The method used is to analyze the similarities between services that store passwords in plain text or reversible form in terms of appearance and HTML data provided by the server. We have found several services that store passwords in either plain text or reversible form, and the unique characteristics of these services are recorded. These features are compared from two perspectives, the appearance and the HTML data provided by the server, and the common parts are searched for.

Result. The results of the analysis show a similar format in appearance between the two services. We avoid presenting detailed information about each service for ethical reasons.

We then compared the HTML data provided by the servers of the two services (let us call them Service A and Service B). The HTML for Service A had "action="index.aspx"" and the HTML for Service B had "action="/7cn-webapp/mobile/WMShinkiTorokuNyuryoku.do?&timestamp=20200129170029". There was a description. It is a description related to the location where the data entered by the user is sent, but there were no similarities because the architectures of the services were different, with Service A being .aspx and Service B being .do. The other services were also compared from the two perspectives, but no common points could be found that could be considered the cause.

5 Future Works

In this study, the password management status of famous websites and applications was investigated, the similarities between services that store passwords in plain text or reversible form were searched for, and the causes were analyzed. In this study, an approach based on external observation was adopted. However, since there is a limit to the information that can be obtained in an experiment based on external observation, we believe that the accuracy and efficiency of the investigation can be improved by using internal observation or a similar method. It is also necessary to increase the number of services compared and analyze the common parts of services that store passwords in plaintext or reversible form. One of the future tasks is to identify the root cause of the services that store passwords in plaintext or reversible form based on the analysis results.

6 Conclusion

Passwords entered by users in web services and applications are considered to be important confidential information, so it is ideal to store them in a manner that is difficult to decipher in case of unauthorized intrusion from the outside, but there have been incidents of personal information leaks in web services and applications. However, there have been incidents of personal information leaks in web services and applications, partly due to storing passwords in plain text or reversible form.

The purpose of this study is to find out what kind of services and applications store passwords in plain text or reversible form on the server side. As an approach to this, a survey was conducted from typical web services and applications. as a result of surveying web services ranked from No. 1 to No. 103 in Alexa's "Japan Top Domains" and from No. 1 to No. 28 in Google Play's "Free Top Android Apps" ranking of overall popularity, it was found that passwords were stored in plain text or reversible form. As a result, no web service or app was found that could be determined to be storing passwords in plaintext or reversible form. Therefore, a separate survey was conducted on services that were reported to store passwords in plain text or reversible form. 40 services were surveyed, and 7 services were found to store passwords in plain text or reversible form. Based on these investigations, we analyzed the similarities among the services that store passwords in plaintext or reversible form from two perspectives: the appearance of the web service or application and the HTML data provided by the server. When the HTML data was compared, no common points were found, and we were not able to discover any common points that could be considered the cause. Future tasks include increasing the number of services to be compared, analyzing the common parts of services that store passwords in plaintext or reversible form, and improving the accuracy and efficiency of the survey by using internal observation or similar methods if possible. The results can then be used to identify the root cause of the service storing passwords in plaintext or reversible form.

The fact that we did not find any services storing passwords in plaintext or reversible form on Alexa's top sites or Google Play's top-ranked apps indicates that the risk is not urgent. However, the services that have been reported still store passwords in plain text or reversible form, and some of these services have a large number of users. This study shows that the risk itself continues to exist, and the reality of the risk has been clarified.

Acknowledgements. This work was supported by JSPS KAKENHI Grant Number JP22K12035.

References

1. Wash, R., et al.: Understanding password choices: how frequently entered passwords are re-used across websites. In: Twelfth Symposium on Usable Privacy and Security (SOUPS 2016) (2016)
2. Tunggal, A.T.: The 62 Biggest Data Breaches (Updated for January 2022). UpGuard Blog (2022) https://www.upguard.com/blog/biggest-data-breaches. Accessed 11 Feb 2022
3. Keeping password secure—Facebook. https://about.fb.com/news/2019/03/keeping-passwords-secure/. Accessed 08 Oct 2019
4. Notifying administrators about unhashed password storage. https://cloud.google.com/blog/products/g-suite/notifying-administrators-about-unhashed-password-storage. Accessed 08 Oct 2019
5. Alena, N., et al.: Why do developers get password storage wrong? a qualitative usability study. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (2017)
6. Alena, N., et al.: Deception task design in developer password studies: exploring a student sample. In: Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018) (2018)
7. Alena, N., et al.: If you want, I can store the encrypted password a password-storage field study with freelance developers. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (2019)
8. Alena, N., et al.: On conducting security developer studies with CS students: examining a password-storage study with CS students, freelancers, and company developers. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (2020)