Scientific Research Publishing

Journal of Computer and Communications, 2016, 4, 26-36
http://www.scirp.org/journal/jcc
ISSN Online: 2327-5227
ISSN Print: 2327-5219

Online Accounts Management Method Using Risk-Based Approach

Yoshio Kakizaki*, Takumi Akiyama, Kazuya Otani, Ryoichi Sasaki

Department of Science and Technology for Future Life, Tokyo Denki University, Tokyo, Japan
Email: *kakizaki@im.dendai.ac.jp

How to cite this paper: Kakizaki, Y., Akiyama, T., Otani, K. and Sasaki, R. (2016) Online Accounts Management Method Using Risk-Based Approach. *Journal of Computer and Communications*, **4**, 26-36.
http://dx.doi.org/10.4236/jcc.2016.414003

Received: October 20, 2016
Accepted: November 11, 2016
Published: November 14, 2016

Abstract

In this paper, we propose an online accounts management method that uses a risk-based approach to minimize any risk of personal information leakage, while permitting users to reuse passwords. The proposed method clusters some accounts to reduce the number of passwords a user has to remember. This is achieved without increasing the risk value of any leaked personal information because the clustered accounts have similar attributes. Further, the proposed method suppresses any increase in risk by optimizing the cluster even if the password is shared by numerous accounts, and can reduce the number of passwords to approximately 30% - 40%. In evaluations conducted, 91% of participants accepted our method, indicating that it balances usability with security.

Keywords

Password, Authentication, Usability, Security, Risk Management

1. Introduction

In their 2014 annual security report, Cisco stated that, "although brute-force login attempts are by no means a new tactic for cybercriminals, their use increased threefold in the first half of 2013" [1]. This was said about list-based attacks, suggesting that list-based attacks have become a serious threat. If attackers succeed in logging into a user's web account, they can login to that user's other web accounts if they have the same password. This is critical as 59% of web users have in fact admitted to reusing passwords because remembering multiple different passwords is too difficult [2].

Countermeasures to list-based attacks include setting different IDs and passwords for different websites and installing Single Sign-On (SSO) or two-factor authentication (2FA). However, because SSO or 2FA can only be installed on the server, these methods

are not list-based attack countermeasures that users can adopt directly. One countermeasure that users can adopt on their end is the use of password manager software (such as LastPass, 1Password, and KeePass) installed on their computer. Such software allows users to manage several different passwords so that users do not need to reuse the same passwords for different websites.

On the other hand, only 20% of web users prefer to use SSO instead of conventional login methods [3]. In addition, both the usage rate and the support rate for 2FA are low. The most common method users utilize to manage passwords is simply their memory (63%). Only 8% of users actually use password manager software [2]. This is because many users cannot master these solutions; therefore, they continue to entrust password management to their memory. The number of such low literacy users who do not use these solutions is also significant. Thus, an easy-to-use and acceptable method for managing reusable passwords is required.

In this paper, we propose a method for managing online accounts for such low literacy users that uses a risk-based approach. The proposed method facilitates ease of use and reduces the risk of personal information leakage by allowing users to reuse passwords on multiple websites. The objective of this method is to minimize personal information leakage by regulating the risk of a list-based attack.

2. Related Works

Various password-related studies have been conducted. For example, Inglesant and Sasse [4] and Al Fayyadh *et al.* [5] conducted studies on password policies. Al Fayyadh *et al.* pointed out that users often adopt strategies to simplify password management, such as selecting weak passwords and reusing passwords across multiple accounts which, unfortunately, can cause security vulnerabilities [5]. Password management strategies have also been studied by Jeslet *et al.* [6], Choong [7], and Gaw and Felten [8]. Various password management applications have also been analyzed [9]-[14]. Karole *et al.* [9] conducted a comparative usability study of three popular password managers: Last Pass, Kee Pass, and Roboform2Go. Gasti and Rasmussen [11] analyzed storage formats used by popular password managers. Bojinov *et al.* [15] and McCarney *et al.* [16] proposed password managers. Ruiz *et al.* [17] researched personal behaviors regarding privacy that allows the leakage of information.

Gaw *et al.* [8] discussed how current systems support poor password practices from the perspective of user behavior. They also proposed changes to website authentication systems and password managers. To determine the current situation with regard to password reuse by users, they surveyed how often users reused passwords and their reasons for doing so. Consequently, they demonstrated that people relied on their memory. Alas, whereas current systems do not help users to recall their passwords, they do enable passwords to be reused.

Florêncio *et al.* [18] found that many currently available methods for managing passwords are ineffective at ensuring the safety of web services and establishing a basis for password management. Thus, they surveyed password durability and analyzed the

basis for strong password durability. They postulated that web accounts that do not contain important information about users did not need to be managed, and that only web accounts that do possess important information about users should be managed. Indeed, the following are commonly held opinions regarding safe passwords [19]:

1) Passwords should be random and strong.

2) Passwords should not be reused in all web accounts.

However, users frequently fail to adopt this advice. Crucially, there is no reliable way to remember strong passwords. To minimize the user's burden when managing passwords, and to minimize damage from personal information leakage, Florêncio *et al.* proposed grouping web accounts [19]. They investigated the influence of grouped web accounts that allow users to reuse the same password and found that reusing a password actually makes it stronger, whereas restricting the reuse of a password weakens it. There is thus a trade-off between reusing a password and the strength of that password. Furthermore, grouping web accounts and allowing passwords to be reused in web- account groups increases the probability that personal information will be leaked, but decreases the expected damage from such leakage. Florêncio *et al.* stated that it is actually difficult to manage several randomized passwords, and that methods for managing passwords must address weak passwords and their reusability.

3. Online Accounts Management Method

3.1. Concept

Not all web users understand password managers such as SSO and 2FA. Moreover, some users also do not understand account management. Such users reuse passwords, and are therefore exposed to the threat of list-based attacks. This research is geared towards these "low literacy" users.

It is necessary to accord password management with the importance of the account, as pointed out by Florêncio *et al.* [18]. We consider the similarities of the attributes of accounts. Some accounts retain similar attributes to other accounts. We believe that the risk when such accounts, which retain similar attributes, are attacked by list-based attacks is the same even if information is leaked.

Our proposed method is geared towards online account management from the viewpoint of usability and security. The method reduces the number of passwords that a user has to remember by clustering some accounts. However, because the clustered accounts all have similar attributes, the risk value of any leaked personal information does not increase. Thus, both the passwords that the user remembers and the risk value of any leaked personal information are minimized.

3.2. Preliminaries

3.2.1. Survey of Account Information

To determine what information can be leaked as a result of a malicious login, we analyzed popular websites from the Alexa Top 500 (January 5th, 2015), and targeted the 50 most accessed websites. Then we categorized account information into two groups: Re-

vealed information and unrevealed information.

Revealed information is information that can be confirmed after logging into personal pages such as "My Page". For example, on personal pages, there is information that can be seen, such as the username. **Unrevealed information** is information that cannot be confirmed even after has logged into personal pages such as "My Page". For our purposes, only the information that can be ascertained after logging in is considered **Revealed information**.

Table 1 lists information from a few of the websites we surveyed. In the table, "Y" denotes **Revealed information** and "N" denotes **Unrevealed information**. In this paper, both required and **Unrevealed information** are referred to as attributes.

3.2.2. Questionnaire Survey

We obtained 88 answers to our questionnaire given to university students. These answers were used to decide on an index for evaluating the proposed method's utility.

The list of questions asked and an explanation of each are provided below.

PQ1: What online accounts do you have? We permitted participants to indicate all of the web accounts that they have registered to among the surveyed websites described in Section 3.2.1.

PQ2: What information do you not want leaked? We asked participants to choose the information from the attributes that they did not want leaked. In this paper, we refer to this information as confidential information.

The breakdown of the answers to PQ1 and PQ2 are shown in Figure 1 and Figure 2, respectively. The "possession rate" in Figure 1 refers to the ratio of people who registered for the web account. Figure 2 represents the ratio of people who do not want the respective attribute information leaked. There were multiple answers to this question.

3.3. Method

To find web accounts with similar attributes, we cluster web accounts by calculating their similarity to each other. To calculate this similarity, we use the Revealed information. Clustering is a method for grouping data according to the similarity of its features. Grouped datasets are called "clusters". By clustering web accounts, we can generate sets of web accounts that have similar attributes.

To group web accounts according to the similarity of their attributes, we cluster web accounts using cosine similarity and k-means clustering.

Table 1. Sample of surveyed web sites.

Site	User name	Real name	E-mail	Phone number	Country	Sex
Google	Y	Y	Y	Y	Y	N
Rakuten	Y	Y	Y			
Amazon	Y	Y	Y			
Mixi	Y	Y	Y			Y

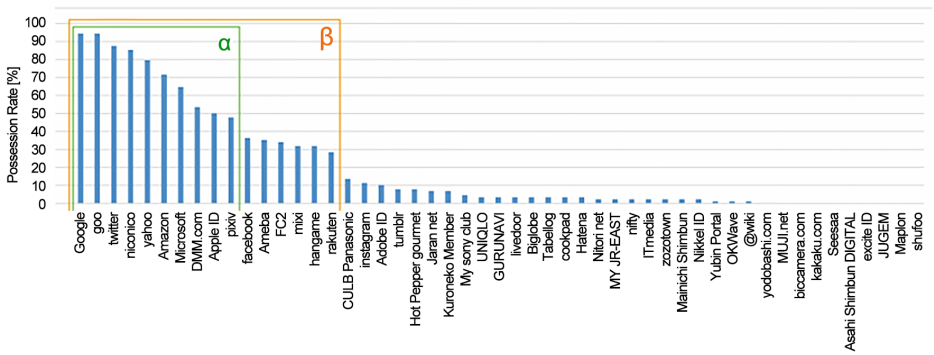

Figure 1. Possesion rate of commonly registered wen accounts.

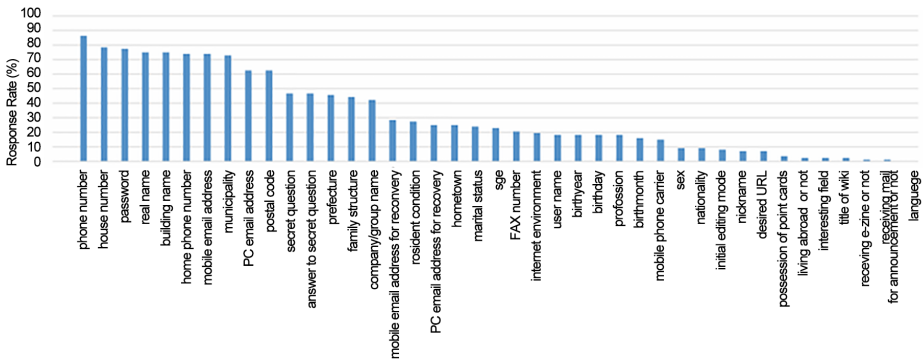

Figure 2. Information considered confidential by survey respondents.

To calculate the information leakage risk value, we regard the percentage of replies for each attribute as the risk value of leaked personal information in Figure 2. We then calculate the risk value of leaked personal information for each cluster. Specifically, the sum of the risk value for each attribute included in a cluster is regarded as the risk value for that cluster.

The number of clusters k is the number of passwords that must be remembered. The number of accounts that shares the same password increases when k is small, which increases the risk value. In our method, k and the risk value are minimized according to the accounts that the user is using. Thus, our method provides the account group that can reuse the password based on the risk value.

3.4. Example Execution

We present an example execution of our method in this section. To cluster the web accounts, we selected accounts based on the responses to our survey (Section 3.2.2). In this example, we selected the following web accounts for clustering (see Figure 1):

1) Web accounts registered by more than 40% of participants (the 10 web accounts between Google and Pixiv).

2) Web accounts registered by more than 20% of participants (the 16 web accounts between Google and Rakuten).

The results for Group A and Group B are shown in Table 2 and Table 3, respectively. The k = 1 case refers to the situation of a single password being used for all web accounts. Similarly, the k = 10 case in Table 2 and the k = 16 case in Table 3 refer to the use of a different password for each web account.

In Table 2, k = 4 is optimum when the result for k = 1 and that for k = 9 are removed because k and the risk value are both minimum. Likewise, in Table 3, k = 5 is optimum. Details of the result of k = 4 in Group A are shown in Table 4. Cluster 4 contains the most attributes; thus, its risk value is the maximum. Our method adopts the maximum risk value in the clustering result.

Our method minimizes the number of passwords and the information leakage risk, and can reduce the number of passwords to approximately 30% - 40%. For more details, please see [20].

Table 2. Risk value for Group A.

k	Risk Value	k	Risk Value
1	**5.275**	6	4.215
2	4.275	7	4.215
3	4.375	8	4.215
4	**4.215**	9	4.215
5	4.215	**10**	**2.784**

Table 3. Risk value of Group B.

k	Risk Value	k	Risk Value
1	**5.852**	9	4.215
2	4.602	10	4.215
3	4.784	11	4.215
4	4.784	12	4.215
5	**4.215**	13	4.215
6	4.215	14	4.215
7	4.215	15	4.215
8	4.215	**16**	**2.784**

Table 4. Details of k = 4 clustering results for Group A.

Cluster #	Accounts	Risk Value
1	Pixiv, niconico	1.644
2	Twitter, DMM.com, Amazon	1.624
3	Yahoo!, Microsoft, goo	2.871
4	Apple ID, Google	**4.215**

4. Evaluation

In evaluating the proposed method, we sought to determine how useful it is for web service users. For it to be useful, the proposed method should satisfy the following criteria:

1) Minimize attribute leakage resulting from list-based attacks.

2) Minimize the number of passwords that must be remembered by users.

4.1. Evaluation Procedure

To evaluate Criteria 1 and 2, we applied our method to 74 participants (different to those in Section 3.2.2). In this case, we requested that the participants indicate the web service they were using. We showed the participants the clustering result of our method, and asked whether it was acceptable. For those who stated that it was not, we asked for an explanation. In addition, we asked the participants the following:

EQ1: How many ID/password pairs do you remember?

EQ2: Do you use any password manager?

EQ3: Do you use two-factor authentication?

EQ4: Can you remember the number of passwords indicated by our method?

4.2. Evaluation Results

Table 5 shows the results of statistical analyses. The number of accounts is A; number of clusters is C, and reduction rate (RR) of the passwords is calculated from A/C. The risk value (RV) uses the value in Section 3.3.

To clarify the relation between the reduction rate in the number of passwords and the risk value, we sort by the number of accounts, and present the reduction rate and the risk value by 4-quantiles in Table 6. Table 7 shows the results for EQ2, EQ3, and EQ4.

Table 5. Statical analyses of the results.

Item	Average	Median	Mode	Min	Max	SD
EQ1	3.94	3	3	2	10	1.99
A	10.04	10	6	3	26	4.90
C	4.04	4	4	2	7	0.98
RR	0.47	0.48	0.50	0.18	0.80	0.17
RV	3.92	4.22	4.22	1.10	7.26	1.54

Table 6. Analyses by 4-quartiles.

Item	Q1/4	Q2/4	Q3/4	Q4/4
RR	0.67	0.53	0.41	0.27
RV	3.85	3.75	4.36	3.70

Table 7. Result of EQ2, EQ3, and EQ4.

Item	Yes	Maybe	No
EQ2	30	-	44
EQ3	18	-	56
EQ4	29	39	3

5. Discussion

According to Florêncio *et al.*, web service users should protect web accounts that contain confidential information, and they should reduce the costs from protecting web accounts that do not possess such information [18].

In this evaluation, Table 5 clarifies that the participants used 10 accounts on average, with the maximum being 26 accounts.

Table 5 shows that many of the participants could remember about three or four ID/password pairs (EQ1). This indicates that users cannot remember more than four passwords, even when our method reduces the number of passwords. Therefore, it is preferable that the clustering result of our method is approximately four.

In Table 5, our clustering result C is 4.04 on the average, and the standard deviation (SD) is 0.98. As a result, many results can optimize the number of passwords, and can also minimize the risk value. The risk value RV is 3.92 on average, and the reduction rate RR is 48%. Each result also is no different from the result in Section 3.4, which demonstrates the effectiveness of our method.

Table 6 shows the relation between the reduction rate of the number of passwords and the risk value. The risk value RV is virtually the same from Q1/4 to Q4/4. This result indicates that our method optimizes the combination of accounts and minimizes the risk value. In fact, criterion 1 is satisfied. The reduction rate is 67% at most because the number of accounts for Q1/4 is the lowest. Conversely, the reduction rate is 27% because the number of accounts for Q4/4 is the highest. Our method can suppress any increase in the risk value from these results by optimizing the cluster even if the password is shared by numerous accounts.

Figure 3 shows the relation between the number of accounts and the reduction rate for all participants. The reduction rate is also low when the number of accounts is small. Conversely, the reduction rate is high when the number of accounts is large. This result shows that many online accounts contain similar attributes. In other words, the similarity of the account is high.

From EQ2 and EQ3 in Table 7, most users use neither a password manager nor two-factor authentication. We examined whether participants with numerous accounts used a password manager or two-factor authentication. However, there was no significant difference. The result in EQ4 shows that our method was accepted by 91% of the participants. As described with EQ1, it is clear that our clustering result approximates the number of ID/password pairs that users can memorize. Thus, criterion 2 is satisfied.

Finally, our online accounts management method satisfies all criteria.

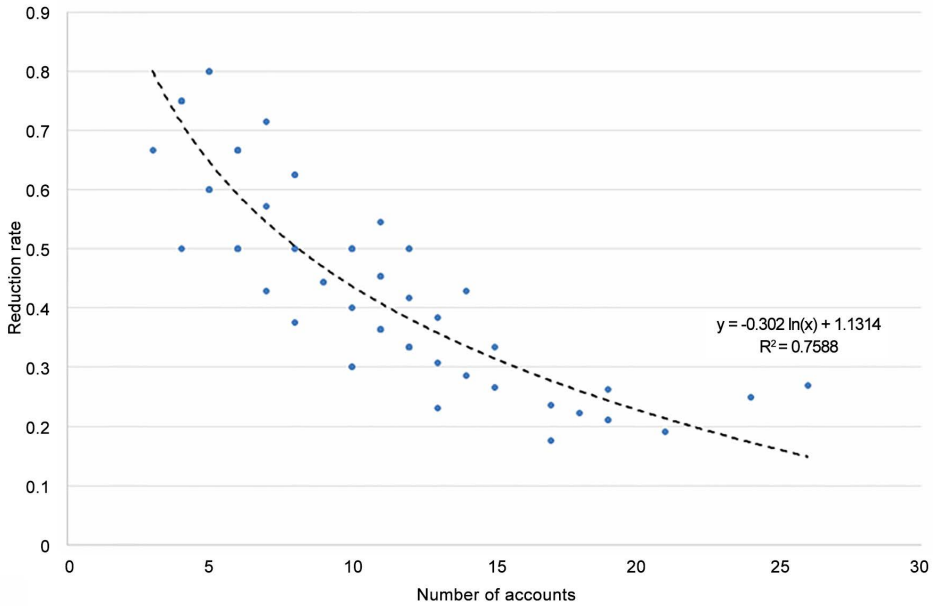

Figure 3. Relation between the number of accounts and the reduction rate.

6. Conclusions

In this study, we analyzed online accounts and proposed an online accounts management method that uses a risk-based approach. Multiple different passwords should be used for multiple different online accounts. Password managers, single sign-on, and two-factor authentication mitigate this issue. However, some users do not understand these solutions. Such users reuse passwords, and are therefore exposed to the threat of list-based attacks.

Our method reduces the number of passwords that a user has to remember. Further, the risk value of leaked personal information does not increase because the clustered accounts have similar attributes. Thus, both the passwords that the user remembers and the risk value of any leaked personal information are minimized.

The evaluations conducted of our method indicate that it can suppress any increase in the risk value by optimizing the clusters even if the password is shared by numerous accounts. The evaluation results also show that many online accounts contain similar attributes. In other words, the similarity of the accounts is high. Finally, our method was accepted by 91% of participants.

In future work, we will consider the influence of dynamic attributes. In this study, we targeted only static attributes. However, it is necessary to consider dynamic attributes such as historical data and activity logs. This will facilitate more accurate risk analysis of online accounts.

References

[1] Cisco. (2014) Cisco 2014 Annual Security Report. https://www.cisco.com/web/offer/gist_ty2_asset/Cisco_2014_ASR.pdf

[2] Password Boss. (2015) The State of Consumer Password Habits.

https://www.passwordboss.com/password-habits-survey-part-1/

[3] AddShoppers. (2015) 2015 On-Site Marketing Breakdown. http://learn.addshoppers.com/2015/

[4] Inglesant, P.G. and Sasse, M.A. (2010) The True Cost of Unusable Password Policies: Password Use in the Wild. *Proceedings of CHI* 2010, Atlanta, 10-15 April 2010, 383-392. http://dx.doi.org/10.1145/1753326.1753384

[5] AlFayyadh, B., Thorsheim, P., Jøsang, A. and Klevjer, H. (2012) Improving Usability of Password Management with Standardized Password Policies. *Proceedings of SAR-SSI* 2012, Cabourg, 22-25 May 2012.

[6] Santhi Jeslet, D., Sivaraman, G., Uma, M., Thangadurai, K. and Punithavalli, M. (2010) Survey on Awareness and Security Issues in Password Management Strategies. *International Journal of Computer Science and Network Solutions*, **10**, 19-23.

[7] Choong, Y.-Y. (2014) A Cognitive-Behavioral Framework of User Password Management Lifecycle. *HAS* 2014 *Second International Conference*, Heraklion, 22-27 June 2014, 127-137.

[8] Gaw, S. and Felten, E.W. (2006) Password Management Strategies for Online Accounts. *Proceedings of SOUPS*'06, Pittsburgh, 12-14 July 2006, 44-55. http://dx.doi.org/10.1145/1143120.1143127

[9] Karole, A., Saxena, N. and Christin, N. (2011) A Comparative Usability Evaluation of Traditional Password Managers. *Proceedings of ICISC* 2010, Seoul, 1-3 December 2010, 233-251.

[10] Ciampa, M. (2011) Are Password Management Applications Viable? An Analysis of User Training and Reactions. *ISEDJ*, **9**, 4-13.

[11] Gasti, P. and Rasmussen, K.B. (2012) On the Security of Password Manager Database Formats. 17*th European Symposium on Research in Computer Security*, Pisa, 10-12 September 2012, 770-787.

[12] Alkaldi, N. and Renaud, K. (2016) Why Do People Adopt, or Reject, Smartphone Password Managers? 1*st European Workshop on Usable Security*, Darmstadt, 18 July 2016, 1-14.

[13] Li, Z., He, W., Akhawe, D. and Song, D. (2014) The Emperor's New Password Manager: Security Analysis of Web-Based Password Managers. *Usenix Conference on Security Symposium*, San Diego, 20-22 August 2014, 465-479.

[14] Liou, J.-C. (2016) Performance Measures for Evaluating the Dynamic Authentication Techniques. *International Journal of Cyber-Security and Digital Forensics*, **5**, 83-93. http://dx.doi.org/10.17781/P002052

[15] Bojinov, H., Bursztein, E., Boyen, X. and Kamouflage, D.B. (2010) Loss-Resistant Password Management. *European Symposium on Research in Computer Security*, Athens, 20-22 September 2010, 286-302.

[16] McCarney, D., Barrera, D., Clark, J., Chiasson, S. and van Oorschot, P.C. (2012) Tapas: Design, Implementation, and Usability Evaluation of a Password Manager. *Proceedings of the Annual Computer Security Applications Conference*, Orlando, 3-7 December 2012, 89-98. http://dx.doi.org/10.1145/2420950.2420964

[17] Ruiz, R., Amatte, F.P., Park, K.J.B. and Winter, R. (2015) Overconfidence: Personal Behavious Regarding Privacy That Allows the Leakage of Information in Private Browsing Mode. *International Journal of Cyber-Security and Digital Forensics*, **4**, 404-416. http://dx.doi.org/10.17781/P001619

[18] Florêncio, D., Herley, C. and van Oorschot, P.C. (2014) An Administrator's Guide to

Internet Password Research. *Usenix Conference on Security Symposium*, San Diego, 20-22 August 2014, 44-61.

[19] Florêncio, D., Herley, C. and van Oorschot, P.C. (2014) Password Portfolios and the Finite-Effort User: Sustainably Managing Large Numbers of Accounts. *Usenix Conference on Security Symposium*, San Diego, 20-22 August 2014, 575–590.

[20] Akiyama, T., Otani, K., Kakizaki, Y. and Sasaki, R. (2015) Evaluation of a Risk-Based Management Method for Online Accounts. 4*th International Conference on Cyber Security, Cyber Warfare, and Digital Forensic*, Jakarta, 29-31 October 2015, 52-57.