Examination of Incorrect Image Selection Method for Image based Authentication Using Browsing History of Web Pages

Yusuke Iizawa†, Yoshitaka Nakamura‡, and Hiroshi Inamura‡

†Graduate School of Systems Information Science, Future University Hakodate, Japan
‡School of Systems Information Science, Future University Hakodate, Japan
{g2118002, y-nakamr, inamura}@fun.ac.jp

Abstract - In recent years, with the increase in the frequency of use of smartphones,terminals require high security user authentication. In current personal authentication of smartphone, “What you know” and “What you are” are generally used. However, they have problems with memory load, operation load, and use of user-specific persistent data such as a fingerprint.

In this paper, we propose a new personal authentication method to reduce memory load, operation load without using user-specific persistent data. As an approach of this method, we consider combining image based authentication with excellent storage load and operation load and behavior history based authentication using continuous information with change. The proposed method uses screen shot of the web page that the user most closely watched as correct image, and screen shots of web page not seen by user as incorrect images. We examined two methods, selection method excluding web pages of the same domain as correct image and selection method based on semantic distance with search keyword as selection methods for incorrect images. We also verified its effectiveness by comparing the two methods. As a result, it was shown that the incorrect image selection method based on the semantic distance to the search keyword is effective.

Keywords: Smartphone, Personal authentication, Web browser history, Image-based authentication, Screenshot

1 INTRODUCTION

In recent years, smartphones have become widespread, and many people use them on a daily basis.

Due to the characteristics of smartphones, users may operate smartphones anywhere they can connect to the network. Therefore, the opportunity to use the smartphone terminal in the public place where many unspecified people exist increases, and the risk that the data in the terminal is accessed due to the theft of the terminal itself is also increased. Since the internal data of the terminal contains a large amount of personal information, there is a risk that the theft of the terminal may cause unauthorized withdrawal from the online bank account or leakage of the personal information. Therefore, in order to ensure security of data in the terminal even if the terminal is stolen, it is important to provide a user authentication method with high security when using the terminal.

As a personal authentication method for smartphones, two are generally used: “What you know” and “What you are”.

“What you know” refers to an authentication method in which the user arbitrarily sets and uses secret information such as PIN (Personal Identification Number), character string password, pattern and so on. This authentication method needs to set secret information that is complex enough not to be guessed by others. Although this improves security against attacks based on user behavior observation such as peep attacks and guess attacks, it requires increased memory load of the user and complicated authentication operations. On the other hand, setting simple secret information reduces the user's memory load, but makes the authentication vulnerable to the aforementioned attack methods. In addition, since there is a tendency to simply set the password of the PC[1], it is considered that the same tendency is also made to the smartphone.

“What you are” is called biometric authentication. It refers to an authentication method that sets physical features such as fingerprints, irises and faces as secret information.Fingerprint authentication method is installed in many smartphones because the input operation of secret information is very convenient with only the minimum operation of “touch”. On the other hand, there is a problem of spoofing by using permanent data of user's biometric data [3] [4].

Because of these problems, a secure authentication method is required without using permanent biometric data.

2 RELATED WORK

Image based authentication is a personal authentication method that can reduce memory load and operation load. Image based authentication presents a group of images in which a correct image set as secret information by the user and incorrect images are mixed, and performs personal authentication based on whether the user can select a correct image. The basic steps of image based authentication are shown in the Fig.1. Images have the following effects on human memory[5].

1. Easy to memorize in large quantities compared to text data
2. Easy to memorize for a long time compared to text data

The fact that the memory of an image is better than that of a text, as described above, is called the “Picture Superiority Effect(PSE)” [6]. Because of this effect, image based authentication is said to be superior to “What you know” in terms of memory load. “Dejavu”[7], which uses images of geometric patterns, and “Awase-E”[8], which uses photos taken with mobile terminals as images, have been devised. However, in

order to perform image authentication, it is necessary for the user to collect a large number of images and register them in advance, which places a heavy burden on the user.

As a personal authentication method using personal data other than biometric information, there are some researches that acquire the user's behavior history according to sensor data obtained by a smartphone and perform implicit personal authentication based on this information[9][10]. However, behavior based authentication may be highly dependent on the surrounding environment.

Therefore some researches were conducted to identify the user's subjective experience information such as yesterday's breakfast and the visited shops, and to identify individuals.

Authentication using the user's mail reception history[11] and "PassFrame" using video recorded from the user's viewpoint have been proposed[12].

But, behavior history is easily associated with information about the user's life and relationships, so it is necessary to use an appropriate behavior history that is not associated with such information.

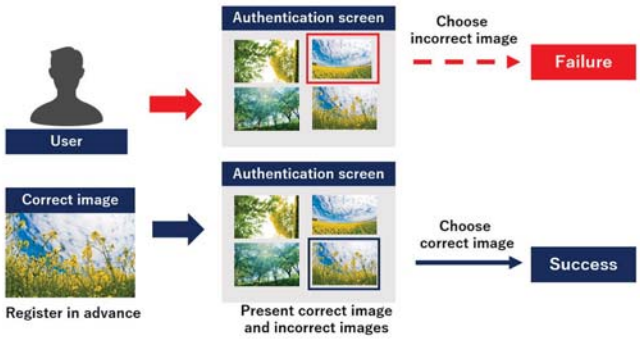

Figure 1: Basic steps of image authentication

3 Image based authentication using behavior history

3.1 Approach

By combining image based authentication and behavior history based authentication, the advantages of both methods can be obtained.

Web page is used as authentication behavior history information from application and video information in smartphone. Since many web pages are supposed to be published, deleting information on a specific personal web page makes it safe to use information on browsing history of web pages for authentication. The screen shot images that can be acquired when a web page browsing operation is performed using a standard browser on a smartphone have the following features.

1. Since browsing web pages by a browser is a user's voluntary behavior, the browsing history is information specialized for individual users

2. Since web pages are basically published to a large number of unspecified Internet users, it is difficult to identify the corresponding smartphone user only from the screen shot image of the web page

3. Services that use personal accounts tend to use dedicated applications, so browsers are more likely to perform simple browsing operations that do not link to personal information[13]

Since there are a large number of browsable web pages on the Internet, it is unlikely that the candidate images for authentication will be exhausted. Further, by using the URL recorded in the browsing history of the browser, the user's image registration work can be made unnecessary. In addition, since contents such as texts and images on web pages that the user has voluntarily browsed are stored in the memory of terminal, they are also effective for recall at authentication.

3.2 Research tasks

In image based authentication using behavior history, the selection method of correct image and incorrect images based on the selected information become problems.

The correct image used at authentication needs to be able to be recalled correctly from the memory by the user. In Ref.[7] [8], it is easy for the user to recall the correct image because users needs to register the correct image by themselves. However, in the case of image based authentication where the user does not register the authentication images, the user sees the correct image only after being presented on the authentication screen. Therefore, it is necessary to select the correct image which clearly remains in the user's memory.

On the other hand, incorrect images need to be easily distinguishable from the correct image for the user. In Ref. [7] [8], an image group other than the one registered by the user as a correct image is used as an incorrect images. When the user intentionally registers the correct image, it is easy to distinguish between the correct image and the incorrect images. However, if the incorrect images presented on the authentication screen is similar to the correct image, the authentication success rate decreases and the authentication operation time becomes longer. Therefore, it is necessary to prepare incorrect images looks significantly different to the correct image for the user himself. Also there is also a need to make few difference between the correct image and incorrect images for others.

4 Proposed method

4.1 Selection of correct image

There is an implicit method of acquiring information on the interest of each web page, without burdening the user, using the browsing time of the web page [14]. According to this method, the longer the web page is displayed, the more likely the user is interested in the page and the user is watching closely. It can be inferred that the user can easily remember the image of the web page that was watching closely. However, it is difficult to determine whether the user really watches closely at the screen, and an image not remaining in the memory of the user may be selected as the image of the

web page at which the user watched closely. Therefore, it is assumed that the user watches closely at the web page displayed on the screen of the smartphone while the operation of touching the screen with the thumb when browsing the web page.

$$\text{Gaze continuation rate} = \frac{\text{Touching time of the screen(ms)}}{\text{Browsing time of the web page(ms)}}$$

From the web page browsed by the user before closing the browser application, the web page with the highest gaze continuation rate is considered to be the most gazed, and the screen shot of that page is taken as the correct image.

4.2 Selection of incorrect images

4.2.1 Conditions to select incorrect images

It is important that the incorrect images are distinguishable from the correct image for the user and difficult to distinguish from the correct image for others. When a screen shot of a web page with low gaze continuation rate is selected as an incorrect image from browsed web pages, depending on the contents of the web page, the user may confuse the correct image with the incorrect image. Therefore, in order to prevent confusion of memory and to distinguish clearly from the correct image, screen shot images of the “unvisited” web page is selected as incorrect images.

4.2.2 Selection method excluding same websites as the correct image

As a first method, an incorrect images is selected from web page group obtained from another search result for the search keyword of the web page selected as the correct image. By selecting only an unvisited web page in this web page group as incorrect images, confusion of the user's memory is prevented, and it becomes easy for the user to distinguish between the correct image and the incorrect images. Figure 2 shows the flow of the selection method. First, a certain number of web page URL groups are extracted from the search result using the search keyword at the time of web page search selected as the correct image. The search result at this time may include the web page itself selected as the correct image, the viewed web pages, and web pages on the same domain name as those web pages. If these web page images are presented simultaneously at the time of image based authentication, the correct image may be confused with the incorrect images. Therefore, among the web page URL group obtained from the search result, URLs including the same domain name as the web page selected as the correct image and the browsed web page are excluded. And this selection method takes screen shots of the web pages from the URL of the remaining web pages and selects them as incorrect images.

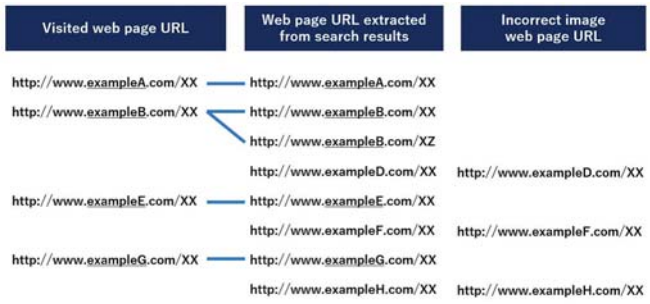

Figure 2: Selection method excluding the same website as the correct image

4.2.3 Selection method based on semantically distanced keywords

The second method is a method of extracting web pages that are semantically distant from web page that are correct images. By using the learning model using Word2Vec [15] as the search keyword of the web page URL selected as the correct image, it is possible to find a keyword that is semantically different from the search keyword of the correct image. Word2Vec takes a set of sentences as input and learns the vector representation of the word from other words that appear near the word. Since an incorrect images are selected using a search keywords different from the correct image, it is easy for the user to distinguish between the correct image and the incorrect images. Figure 3 shows the flow of the selection method. By calculating the cosine similarity with the search keyword of the web page selected as the correct image based on the language model learned by word2vec using the corpus in advance, a certain number of keyword groups with semantic distance apart are acquired. Screen shots of the web pages are randomly acquired one by one from the obtained search results of each keyword.

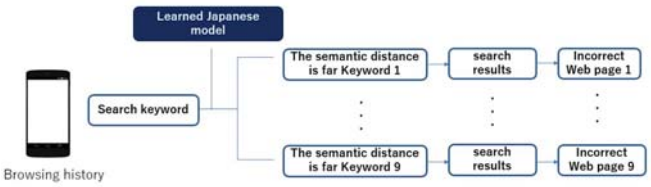

Figure 3: Selection method based on semantically distanced keywords

5 Evaluation experiment

The effectiveness of authentication and the validity of image selection are confirmed by evaluating the authentication success rate of the proposed system and the time taken for image selection at the time of authentication.

5.1 System configuration for experiment

In the evaluation experiment, a browser type data acquisition application for Android OS was developed and used as a system for the experiment. Table 1 shows the smartphones

used in the experiment, and Fig.4 shows the configuration of the experimental application. This application has a function of acquiring URLs, web page display time, screen contact time, gaze continuation rate, and search keywords. Each time the web page transitions, each data is sent to BrowserDB on the server. Google search engine was used for this search. In addition, the web page screenshots of the acquired URL are saved in the size of 320 x 568 pixels.

Table 1: Experimental device

Device name	VAIO Phone A VPA0511S
OS	Android 6.0.1
External dimensions	77.0 mm x 156.1 mm x 8.3 mm
Display size	5.5inch
Resolution	1080 x 1920

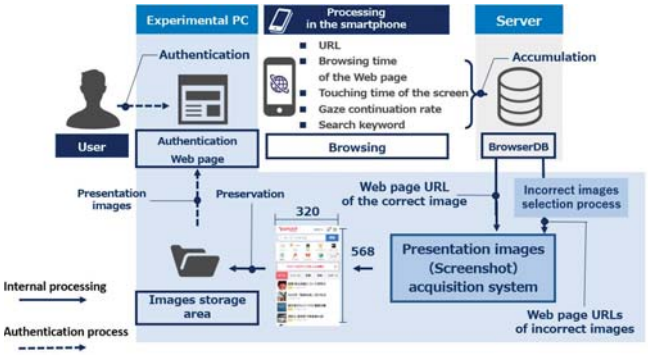

Figure 4: Experimental system

5.2 Experimental method

The subject browses the web page based on a search by a keyword designated in advance, and then performs an authentication operation of selecting a correct image from the authentication screen. On this authentication screen, 1 correct image selected by the correct image selection method and 9 incorrect images selected by the incorrect image selection method are displayed. In each trial of experience, the subject browses a web page for 10 minutes, takes 10 minutes break, and then performs an authentication operation. In order to measure the gaze continuation rate, and it is assumed that the thumb is always touching the screen when browsing web pages. In this method, the learned Japanese model[16] was used with Word2Vec. The search keywords to be specified are “カサブランカ (Casablanca)” and “端島 (Hashima)” in the trial using the incorrect image selection method excluding the same website as the correct image, and “土方歳三 (Toshizo Hijikata)” and “変身 (Henshin)” in the trial using the incorrect image selection method based on semantically distanced keywords between correct image. As keywords having a semantic distance from the search keyword at this time, the top 9 keywords with cosine similarity obtained by the learning model calculated by Word2Vec for each search keyword were used.

5.3 Result

In Fig.5 and Fig.6, the blue bar represents the result in the case of using Method A. The green bar represents the result in the case of using Method B. Figure 5 shows the average value of the authentication success rate in each method. When using Method A, the authentication success rate remains at 57.1%, whereas when using Method B, the authentication success rate has achieved 100.0%. Figure 6 shows the average image selection time in the authentication process for each method. The lines in the Fig.6 show the maximum and minimum values for each method. When using Method A, the average image selection time was 20.42 seconds, while when using Method B, it was significantly reduced to 9.04 seconds.

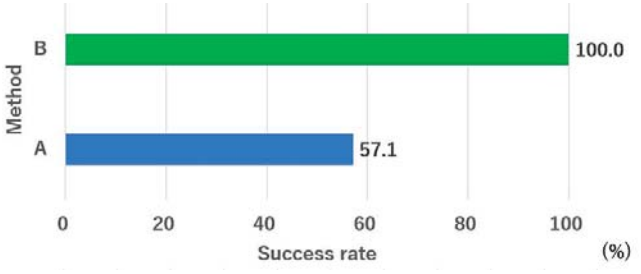

Figure 5: Authentication Success rate

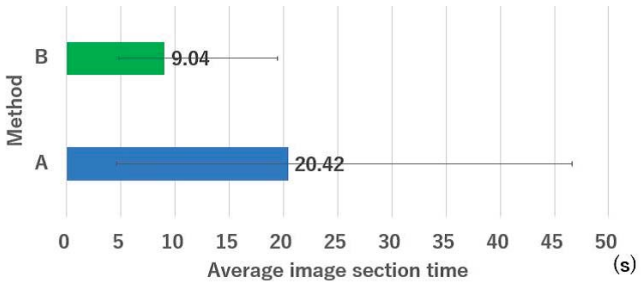

Figure 6: Average image section time

6 Discussion

Experimental results show that method B has a higher authentication success rate than method A. The cause of this is the content of the searched web pages, and the confusion of the user’s memory. Method A uses a search keyword when the correct image is browsed to select web pages that the user has not browsed as incorrect images. The content of the web pages of those incorrect images included the search keyword in the page, but the web pages of the content included in the search keyword were selected. For example, “カサブランカ (Casablanca)” is a web page such as an area or movie containing the keyword. For this reason, there are web pages that have similar searches but have similar contents, and I think that the user is confused. On the other hand, Method B uses the search keyword when viewing the correct image to acquire incorrect images using different keywords having a semantic distance. As a result, Web pages with content that was independent of search keywords were selected. For example, if it is “変身 (Henshin)”, these are web pages that contain

keywords such as “ロケット団’(Team Rocket)” and “カービィ(Kirby)”. For this reason, the user can easily distinguish the keyword from the content, and think that the correct image, which is the browsed web page, can be determined. This user’s memory confusion is considered to have an influence on the authentication speed, and it is considered that the authentication speed of Method B is faster than Method A, which easily causes memory confusion. Therefore, in this method, it can be said that by using an incorrect image selection method based on the semantic distance to the search keyword, it is possible to reduce user’s memory confusion and achieve a high authentication success rate.

7 Conclusion

In this paper, we proposed a personal authentication method that combines image based authentication, which is advantageous for reducing memory load and operation load, and behavior history authentication using personal data other than living body. There are problems are the selection of the correct image, and the selection of the incorrect images. The correct image used the screen shot image of the web page with the highest gaze continuation rate of the user. In selecting incorrect images, we proposed method A: selection method of incorrect images excluding same web page and method B: incorrect images selection method based on semantic distance to search keywords.As evaluation experiments, the effectiveness of two methods in choosing incorrect images was investigated. We performed experiments to select correct images from 1 correct image and 9 incorrect images by each method. In method A, the authentication success rate is 57.1 %, and the average image selection time is 20.42 seconds, but method B has 100.0% authentication success rate.The average image selection time was a result of 9.04 seconds. Therefore, it turned out that incorrect images selection method based on semantic distance to search keywords is effective.

As future tasks, we will investigate attack resistance in the Educated Guessing Attack, and discuss countermeasures for the vulnerable part.

REFERENCES

[1] splashdata, “The Top 50 Worst Passwords of 2018,” <https://www.teamsid.com/100-worst-passwords-top-50/> [Accessed June 6, 2019].

[2] statista, “Penetration of smartphones with fingerprint sensors worldwide from 2014 to 2018 ”, <https://www.statista.com/statistics/522058/global-smartphone-fingerprint-penetration/> [Accessed June 6, 2019].

[3] K. Cao A.K. Jain, “Hacking Mobole Phones Using 2D Printed Fingerprints,” MSU Technical Report, MSU-CSE-16-2, 2016.

[4] P. Bontrager, A. Roy, J. Togelius, N. D. Memon, and A. Ross, “DeepMasterPrints: Generating MasterPrints for Dictionary Attacks via Latent Variable Evolution,” Proceedings of the IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS2018), 2018.

[5] T. Takahashi, S. Kitagami, K. Miyashiro, E. T. Harada, and S. Suto, “Picture-authentication system(1):How users select registration pictures?,” Proceedings of the IEICE technical report. LOIS, Life intelligence and office information systems, Vol.112, No.35, pp.1-8, 2012(*in Japanese*).

[6] R. Niimi, A. Ueda, and K. Yokozawa, “Object perception, Series Integrated Perception,” Vol.2, pp.69-70, Keiso Shobo (2016). (*in Japanese*)

[7] R. Dhamija and A. Perrig, “Déjà Vu: A User Study Using Images for Authentication,” Proceedings of the 9th conference on USENIX Security Symposium(SSYM’00), pp.45-48, 2000.

[8] T. Takada and H. Koike, “Awase-E: the Method Enables an Image-based Authentication to be More Secure and Familiar for Users with Providing Image Registration and User Notification,” IPSJ Journal, Vol.44, No.8, pp.2002-2012, 2003(*in Japanese*).

[9] T. Ajioka, T. Umezawa, and N. Osawa, “Authentication Method for Mobile Device using Wrist Acceleration during Unlick Operation,” IPSJ SIG Technical Reports, Vol. 2016-MBL-81, No.22, pp.1-6,2016(*in Japanese*).

[10] K. Yamada, K. Notomi, and K. Saito, “Personal Authentication Method using Behavioral Feature Amounts of Smartphone Handling,” Proceedings of Journal of Biomedical Fuzzy Systems Association, Vol. 16, No.1, pp.41-48,2014(*in Japanese*).

[11] M. Nishigaki, and M. Koike, “A User Authentication Based On Personal History : A User Authentication System Using E-mail History,” IPSJ Journal, Vol.47, No.3, pp.945-956, 2006(*in Japanese*).

[12] N. Nguyen, and S. Sigg, “PassFrame: Generating image-based passwords from egocentric videos”, Proceedings of the 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), 2017.

[13] nielsen, ”Smartphone users who use the app and browser properly-Released “Digital Trends 2016 “ which summarized the trend of nielsen 2016-””, <https://www.netratings.co.jp/news_release/2017/03Newsrelease20170309.html/> [Accessed June 9, 2019](*in Japanese*).

[14] Y. Hijikata, “Exploiting Customer’s Preference - Leading Edge of User Profiling Technique- : Techniques of Preference Extraction for Information Recommendation,” IPSJ Magazine, Vol.48, No.9, pp.957-965, 2007(*in Japanese*).

[15] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J.Dean, “Distributed Representations of Words and Phrases andtheir Compositionality”, Neural Information Processing Systems2013, pp.3111–3119, 2013.

[16] AIAL, “Publish the learned Japanese model of word2vec,” <http://aial.shiroyagi.co.jp/2017/02/japanese-word2vec-model-builder/>(*in Japanese*) [Accessed October 29 2019].