

Personal Identification Techniques Based on Operational Habit of Cellular Phone

Takamasa Isohara¹, Keisuke Takemori², and Iwao Sasase¹

¹ Graduate School of Science and Technology, Keio University,
3-14-1, Hiyoshi, Kohoku-ku, Yokohama, 223-8522, Japan
{isohara, sasase}@sasase.ics.keio.ac.jp
<http://www.sasase.ics.keio.ac.jp/>

² KDDI R&D Laboratories,
2-1-15 Ohara, Fujimino, Saitama, 356-8502 Japan
takemori@kddilabs.jp
<http://www.kddilabs.jp/>

Abstract. Biometrics authentication which identifies a legitimate user of a cellular phone has been adopted. However, it is not used so actively, because control procedure is complicated, and a user has a resistance to provide own biological information to the communication terminal. In this paper, we propose personal identification techniques based on operational habit of cellular phone, which authenticate the legitimate user by collecting typing histories of cellular phone in a background process. The typing histories that are recorded in the ring buffer are divided into two profiles. One is “short term profile” which is histories from latest typing, and the other is “long term profile” which are older histories than the short term records. Also, we implement two authentication algorithms; first algorithm retrieves a remarkable key of the short term profile from the long term profile. Second algorithm computes weight values of keys of the short term profile by using long term profile. We evaluate the FAR (False Acceptance Rate) and FRR (False Reject Rate), and we show that our method can apply to personal identification on the cellular phone.

Key words: Biometrics Authentication, Cellular Phone, Operational Habit

1 Introduction

Mobile communication terminals such as a cellular phone or a PDA (Personal Digital Assistant) are widely spread in our life. In these terminals, many kinds of personal information are installed, like an address book, instant messages, schedule and so on. In addition, the user accesses to e-commerce sites or on-line banking that require personal authentication. Therefore, a personal identification is needed to protect the personal information and to prevent malicious accesses when the terminal is stolen or lost. Therefore, biometrics authentications have been actively applied for the cellular phone [4, 1, 6]. There are two major methods, such as fingerprint authentication and face authentication [2, 3]. However, to use the biometrics authentication, there are following problems.

- To use the fingerprint authentication or face authentication, it is necessary to implement a sensor on the mobile communication terminal. It causes the cost and enlargement increase.
- Biometrics authentication needs complex calculation such as statistics analysis or neural network. However, complex calculation can not apply to the mobile terminal because of the low CPU power and few amount of memory.
- Calling the authentication procedure is cumbersome to be used. Moreover, there is a great deal of resistance to enter the one’s physical feature. Thus people are not always willing to use the biometrics authentication.
- When the physical feature is leaked, there is no alternative of the physical information.

This paper addresses the personal identification technique based on operational habit of mobile communication terminals. We design this method in consideration of the complexity of calculation and simplicity of usage. We focus attention on the ease of discrimination of personal difference in operational habit on the mobile communication terminal. To authenticate the users, we use the presence and frequency of keystroke entry. We make a prototype model and evaluate the difference of operational habit among a legal user and illegal users. Additionally, we examine the possibility of the application to the personal authentication.

The paper is organized as follows. Section II presents related works. Section III introduces the methodology of our technique. In Section IV, we show the results of experiment. Section V shows conclusions and future work.

2 Related Work

2.1 Classification of Authentication Method

Authentication in computer security is classified into “Client Authentication” and “User Authentication”. Client Authentication is an identification of the client machine that connects to the server machine. User Authentication identifies the user who accesses the secret information or operates the machine.

Here, we focus on the User Authentication because our goal is personal identification which is operated in the mobile communication terminal. User Authentication can be classified into three cases:

objects based It uses the hardware such as IC card or USB memory key to authenticate.

knowledge based In this case, the user is authenticated by knowledge such as password PIN number, etc.

physiology and actions based Examples of this cases are fingerprint, face and keystroke pattern. This authentication method is called “Biometrics Authentication.”

Recently, biometrics authentication is widely used because spoofing of this method is more difficult than other two methods.

2.2 Major Biometrics Authentication

Biometrics Authentication on Mobile Communication Terminal The most popular biometrics authentications on the mobile communication terminal are fingerprint authentication and face authentication [1, 6]. In the case of fingerprint authentication, we compare the fingerprint image of legal user with the examinee’s one. This method has an advantage that has high verification accuracy compared with the other biometrics authentication because the fingerprint does not change easily.

Face authentication is another spread authentication method on the mobile communication terminal. In this case, image of face is used to identify the user as well as fingerprint authentication. Authentication system compares the image which is registered beforehand with examinee’s one.

These methods are used to cut out the need of the password entry or to achieve higher security by combining with a password authentication method.

Keystroke Dynamics Keystroke dynamics are used on the personal computer environment. They are also known as typing biometrics. These methods analyze the way a user types at a personal computer by monitoring the keyboard inputs in attempt to identify users base on their typing rhythm pattern. There is less sense of resistance than fingerprint or face authentication because they are based on action characteristics. Moreover, it is easy to implement these methods on the many types of communication terminals which have a keyboard. The work in [4] shows an FRR of 16.67 % and an FAR of 0.25 %.

2.3 Problems

To use the above authentication techniques, there are following problems. First, to use the fingerprint authentication or face authentication, it is necessary to implement a fingerprint sensor or camera on the mobile communication terminal. It causes the cost increase and the enlargement of the communication terminal. Second, biometrics authentication especially in keystroke dynamics needs complex calculation such as statistics analysis or neural network and, thus, can not apply to the mobile communication terminal directly because the communication terminal has a low CPU power and few amount of memory. Third, calling the authentication procedure is cumbersome to be used. Moreover, there is a great deal of resistance to enter the one’s physical feature. Thus people are not always willing to use the biometrics authentication. Lastly, when the physical feature is leaked, there is no alternative of one’s information.

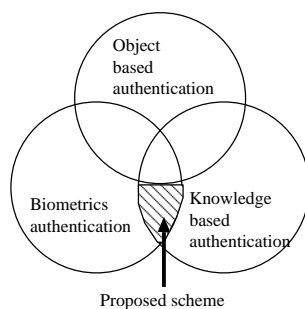


Fig. 1. Position of proposed scheme.

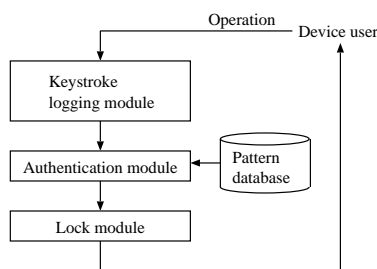


Fig. 2. Authentication model of proposed scheme.

3 Proposed Personal Identification Technique

We propose a personal identification scheme based on operational habit of mobile communication terminals. Fig. 1 shows a position of proposed scheme. This method is based on the operational knowledge of the communication terminal and the typing habitual pattern. Fig. 2 shows an authentication model of the proposed scheme. A detail of each process is shown as follows.

3.1 Collecting the Keystroke Record

In our method, a keystroke record is collected in the background of the communication terminal operation and it is used as an personal characteristic. Collecting method of the keystroke record can be classified into two forms. One is that all of keystroke records on operation of communication terminal are collected. Another is that only keystroke records of some application are collected. Additionally, the method of collecting all keys from beginning to end of operation and the method of collecting specific keys to several steps from beginning of operation are considered in each case. What method is used depends on the authentication algorithm.

3.2 Authentication Algorithm

Authentication algorithm is used to analyze collected keystroke records. We develop two algorithms as follows.

Presence Based Algorithm In this algorithm, compare long term profile to short term profile by focusing on the presence of key. Fig. 3 shows the definition of long term profile and short term profile. As a personal characteristic of operational habit, numbers of keys recorded in both long term profile and short term profile are used in this algorithm. To identify the legal user, matching score is generally adopted. Fig. 4 shows a concept of Matching Score. Here, we take the Fig. 3 as an example to calculate the matching score. In Fig. 3, long term profile has 5 types of keys in 7 entries. And short term profile has 3 types of keys in 4 entries. When to compare long term profile with short term profile, 2 types of keys are matched. Thus matching score is 2.

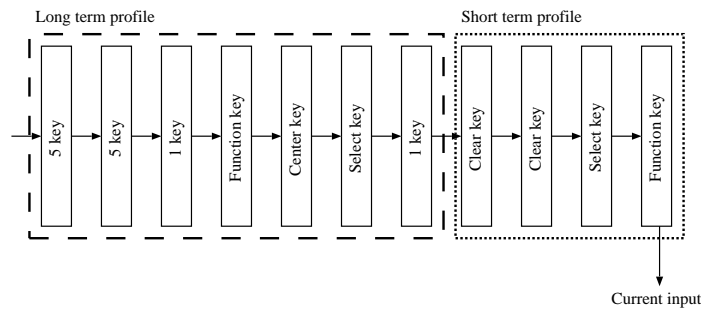


Fig. 3. Definition of long term profile and short term profile.

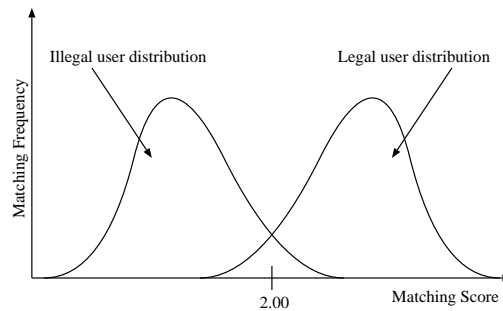


Fig. 4. Concept of Matching Score.

Frequency Based Algorithm In addition to the above algorithm, this method considers a frequency of appearance in long term profile. The frequency of appearance is defined as the proportion of some key to length of long term profile. Therefore, matching score is calculated by following expression

$$P = \sum_{k=1}^m F_k \cdot N_k \tag{1}$$

Where P is a matching score, m is number of all keys, F_k is frequency appearance of key k , and N_k is number of appearance in short term profile of key k . Table 1 shows an example of matching score in frequency based algorithm.

Table 1. An example of matching score in frequency based algorithm.

Key	Frequency in long term profile	Number of appearance in Short term record	
Call key	0.11		Matching score = 0.34+0.05+0.10*2 = 0.59
Upper arrow key	0.22		
Right arrow key	0.04		
Select key	0.34	1	
Down arrow key	0.03		
Function key	0.05	1	
Clear key	0.10	2	
1 key	0.08		

Lock mechanism When the illegal use is detected, a lock mechanism such as password authentication works.

4 Experiments

To evaluate the difference of operational habit among a legal user and illegal users in our scheme, we have some experiments. We make a prototype system build on BREW emulator[5]. BREW is one of the most popular application platform in mobile application.

10 examinees who are graduate students of our laboratory attend in this experiment. We divide them into a 1 legal user and 9 illegal users. A legal user is an well operator of emulator and an illegal user is a user who operates an emulator for the first time. Table 2 shows an outline and operation method of applications for evaluation. By comparing the operation of these applications, we are able to find the differences between the applications that have different operation behaviors. Testers freely operate an application. In BREW emulator, it is only able to collect a keystroke record for operation of application. So, we collect all operation steps in one application.

Table 2. Outline and operation method of applications for evaluation.

Application	Outline	Operation type
Media player	Playing the music	Cursor keys only
Cashbook	Holding the purse	Cursor and ten keys

4.1 Feature of Keystroke Records

Here, we show the features of the keystroke records. In a case of Media player, maximum length of sequence is 364, minimum length is 105, and average length is 242. On the other hand, in Cashbook, maximum length is 810, minimum length is 172, and average length is 455. Here, we focus in the “session” that is the length of sequence from beginning to end in operation of application. In the Media player, maximum session length is 218, minimum length is 1 and average length is 65. In the Cashbook, maximum length is 472, minimum length is 1 and average length is 162.

Next, we show that the difference of the kind of key between Media player and Cashbook. Fig. 5 shows a histogram in the case of Media player. Fig. 6 shows a histogram in the case of Cashbook. In these figures, white box means the record of legal user, and black box means the record of illegal user. When comparing both graphs, it is clear that the distribution of keys are different. From this result, we conduce that it is effective to optimize the attention of key to apply an authentication algorithm to this record.

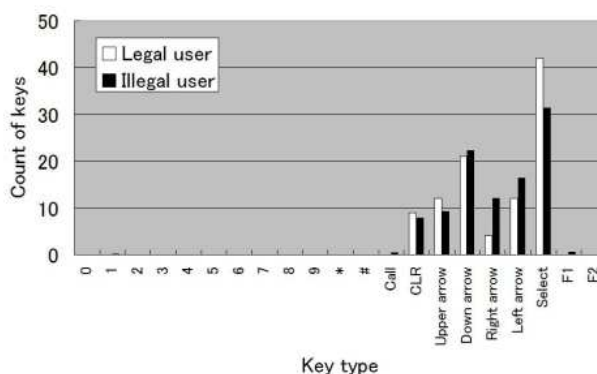


Fig. 5. Operational habit of Media player.

4.2 Authentication Accuracy of Proposed Algorithms

Using the keystroke records from experiments, we calculate a FRR and a FAR of proposed algorithms. FRR is the percentage of incorrectly rejected legal users, and FAR is the percentage of impostors incorrectly matched to a legal user’s biometric. Based on the experiment results, the length of long term

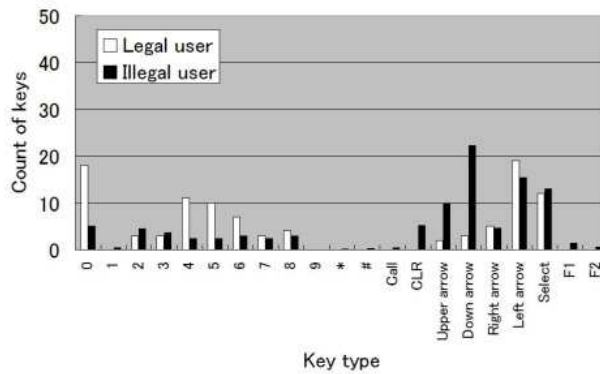


Fig. 6. Operational habit of Cashbook.

profile and short term profile are set to 100 and 10 respectively. Fig. 7 and Fig. 8 show EER (Equal-Error Rate) curve of presence based algorithm and frequency based algorithm, respectively. EER is the point at which the percentage of FAR is equal to percentage of a FRR. X axis shows the decision threshold and y axis shows the error rate. Decision threshold is equal to matching score in Fig. 4. When EER of Media player in both Figures are compared, EER is about 50 % in either Fig. 7 or Fig. 8. On the other hand, in the case of Cashbook, EER in Fig. 7 is about 60 %, but EER in Fig. 8 is about 30 %. Therefore, by attempting frequency based algorithm to Chashbook application, EER has been improved about 50 %. This result is caused because a personal characteristics is taken sharp by frequency information of keys in the application that is operated with many types of keys like cashbook.

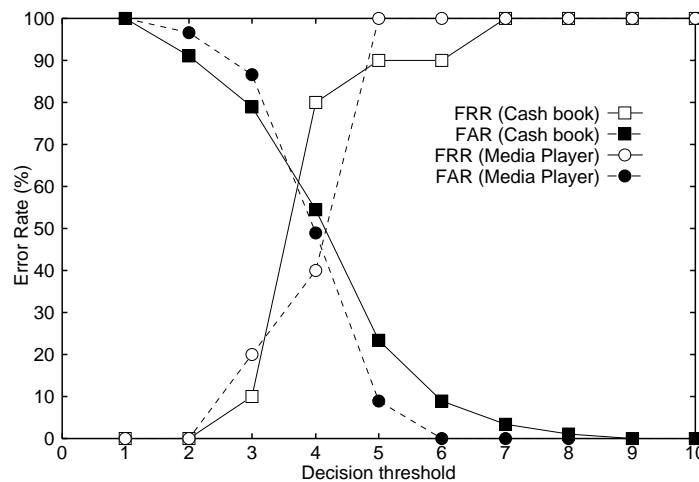


Fig. 7. EER of presence based algorithm.

5 Conclusion and Future Work

The personal identification scheme based on operational habit of cellular phone has been proposed. By getting the keystroke records in the background of communication terminal operation, we can achieve the operational simplicity and low sense of resistance. We design an authentication algorithm focused attention on presence and frequency of keystroke. By evaluation, we showed that the frequency based algorithm improve the EER on the case of application which is operated by many types of keys.

References

1. R. Bole A. Jain and S. Panakanti. *Biometrics: Personal Identification in Network Society*. Kluwer Academic Publishers, 1999.

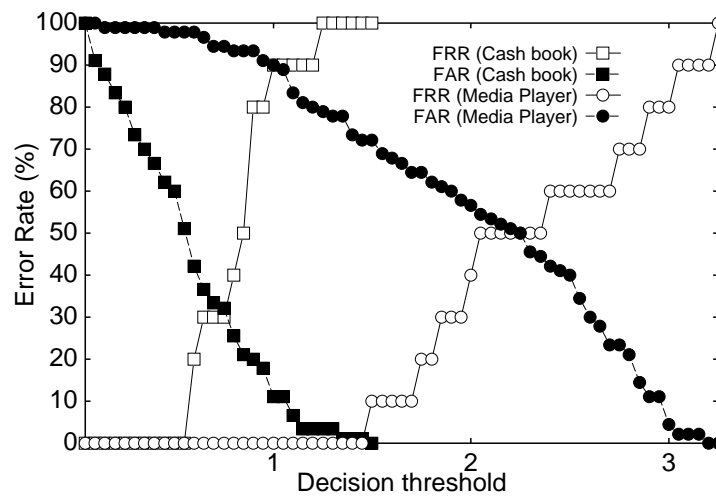


Fig. 8. EER of frequency based algorithm.

2. FOMA F901iS. http://www.nttdocomo.co.jp/product/foma/901i/f901is/topics_01.html 2006.
3. FOMA P901iS. http://www.nttdocomo.co.jp/product/foma/901i/p901is/topics_03.html 2006.
4. Rick Joyce and Gopal Gupta. *Identity authentication based on keystroke latencies*, *Comm. ACM*, **33**(2):168–176, Feb. 1990.
5. Qualcomm BREW Home. <http://brew.qualcomm.com/brew/en/> 2006.
6. Takeo Kanade Ying-li Tian and Jeffrey F. Cohn. *Recognizing action units for facial expression analysis*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **23**(2): 97–115, Feb. 2001.