Proc. of the Intl. Conf. on Advances in Computing and Information Technology-- ACIT 2014

ISBN: 978-981-07-8859-9

The Significant of Character's Location in the Authentication Process

Kranogwan Krasaesat, Pattarasinee Bhattarakosol

Abstract— **Presently, the security of information significantly becomes an important issue for all users over the Internet. Since a single password is insufficient to protect attack from attackers as proved by various researchers, various biometrics have been applied as new alternatives for classifying and identifying users. Keystroke dynamics is a behavioral biometrics with individual characteristic pattern of each person. This unique pattern might be the result from typing skill. In addition, a research had proved that the eye vision can be counted as a biometric which can identify an individual person with higher accuracy when combining with keystroke dynamics. Nevertheless, the character's position on the keyboard has been analyzed in this research that it is a factor of time differences based keystroke dynamics and eye vision mechanisms. Therefore, this research has objective to propose that the vision speed and the location of the typing character should be integrated in the authentication mechanism to gain a better authentication result.**

Keywords— **Authentication, biometric, behavioral biometrics, character's location, eye vision, keystroke dynamic, multi-biometric.**

I. Introduction

Biometrics authentication is growing and counted as a controversial field in which civil liberties groups express their concern over the privacy and identity issues. Currently, biometric laws and regulations are in process and biometrics industry standards are being tested. Since, face recognition biometrics has not reached the prevalent level of fingerprinting, but with constant technological pushes and with threats of terrorisms, researchers and biometric developers will stimulate these technologies for twenty-first century. As a result, biometric characteristics can be divided in two main classes: physiological (fingerprint, face recognition, etc.), and behavioral (keystroke dynamics, Voice, etc.)[1].

Keystroke dynamics is a class of behavioral biometrics that captures the typing styles of users. The typing style includes such factors as the length of time taken when typing the login name or password, the interleave time when a user presses over the keyboard, and the pressing time per key of each user [2]. According to [3], the eye vision has proved that there are some impacts to keyboard typing which related to the keystroke dynamics concept.

In the past, many researchers have been studied the human eye movements in the eye vision because it can identified the character's style and become a behavioral for classifying the person [7-9].

Using password was the original authentication system that can secure the personal information from unauthorized users for their illegal benefits. Although, the computer technologies have rapidly changed to support users'tasks, unfortunately, these flexibilities also support intruders for breaking to the system in a short period of time, such as the brute force attack, dictionary attack, and etc. Therefore, using only a password is insufficient solution to protect the personal information although its length has been increase. Thus, the authentication system has to be continuously developed.

Based on the research proposed by [3], the combination of keystroke dynamics and speed of eye vision had been applied to the identification process. However, the typing patterns of users have not been considered. Therefore, this research has an aim to extend the idea of [3] by considering the users' typing patterns based on the position of characters on the keyboard into the classification model under the assumption that each user should have individual typing pattern.

The remaining part of this paper is that Section 2 is the relative works. Then, the proposed method is described in Section 3. Section 4 is data gathering and analyzing method; Section 5 mainly describes the behavioral classification system. Discussion and conclusion with future works are elaborated in Section 6, Section 7 and Section 8, respectively.

II. Related Work

Generally, the authentication system is a common process that every user must be validated. This authentication system has an aim to protect unauthorized users in accessing the resources over the network, especially accesses through web-applications. Although passwords are used to protect the system before users are enable to access the required files or CPU, this password mechanism is too weak to protect intruders. Thus, biometrics are applied and implemented to the authentication mechanism so the real users can be identified; these biometric are such as fingerprint, face recognition, iris, speaker recognition, keystroke dynamics, etc.

Various studies have indicated that using the biometrics in the authentication system can easily and accurately identify persons since human characteristics are much difficult to be forged. As a result, authenticated person can be distinguished from unauthorized users [1]. One basic technique that is implemented in various systems is keystroke dynamics.

Keystroke dynamics is a type of behavioral biometrics that captures characteristics of users, mainly related to the time

Kranogwan Krasaesat, Pattarasinee Bhattarakosol

Innovative Network and Software Engineering Technology Laboratory
Department of Mathematics and Computer Science, Faculty of Science
Chulalongkorn University, Bangkok, Thailand

capturing in various aspects. A time value can be measured in different criteria, such as a dwell time which is the pressing time over a key on the keyboard. Another time value is the interleave time that is the time measured the hand's movement from one key to the next key on the keyboard. In addition, the correctness of the used of this method can be indicated using one of these values: False Rejection Rate (FRR), False Acceptance Rate (FAR), and Equal Error Rate (ERR). FRR is the percentage of authorized users is identified as imposters; FAR is the percentage of imposters is identified as a valid users; and ERR is the crossover point at which FRR equals FAR [4], [5].

Although the keystroke dynamics is an efficient method in identifying authenticated users, all measurement indexes mentioned previously still show some mistaken identification. Therefore, many recent researchers proposed the combination of keystroke dynamics and another biometrics value to increase accuracy of the authentication process.

According to Obaidat and Sadoun [6], they studied the differences between statistical-based and neural network-based classification methods with keystroke dynamics. This study concluded that neural network-based methods gave better results as compared with statistical methods in keystroke patterns classification. In other words, the neural networks were trained in advance not only using legitimate users' sample, but also intruders' samples. Hence the classifier is expected to produce better results with low FAR.

In addition, the researches of [10], [11] also studied the typing pattern and discovered that the combining of keystroke pattern and users' passwords will lead to a high accuracy result. However, using password as a combining factor may not secure enough since the password can be captured easily by bots or imposters. Thus, using multi-biometrics in the authentication process should be a better alternative.

Base on the study of Nonsrichai, and Bhattarakosol [3], the biometrics under the eye vision had been proven that there are some impacts to keyboard typing which related keystroke dynamics. Meanwhile, merging between eye vision and skills of keystroke dynamics will create an individual pattern to identify persons who are genuine or imposter users.

Over the years, researches in keystroke biometrics applied many existing machine learning and classification techniques. All those researches meant to find the best classifying method in the highest accuracy of result and the lowest error rate. Therefore, this study will use multi-biometric based eye vision with keystroke dynamics to form characteristic patterns in the identification process.

III. Proposed Method

Currently, the eye vision ability can improve performance of the classification process when combining with the keystroke dynamics [3]. Moreover, the correctness of using multi-biometrics is higher than the use of individual value. Most applications prefer the use of keystroke dynamics based on the time capturing. Similarly, this research will also apply the timing concept to build up a new detection system by considering typing patterns of users based on positions of characters on the keyboard. This factor, character positioning, might affect the decision and hand's movement. The next section will elaborate the data collection method which is applied from the method of [3] and the user's typing pattern will be displayed.

IV. Data Gathering and Analyzing Method

This part states the source of sampling data for this research; those are used to classify and identify genuine and imposter users. The explanation will be divided into two parts: data collecting, and data analyzing. Details of each part are explained as follows:

A. *Data Collecting*

The sampling is performed in a class with 42 students, including staffs from private sections. These samples must use computer daily under their willingness. Moreover, each sample, with the age between 18-30 years, must use the same keyboard and computer system during the data collection period so the bias from unfamiliar equipment of users will be eliminated.

Each participant must join the experimental system for 30 days. Moreover, each volunteer has to type three times a day, once in the morning, once in the afternoon, and once in the evening before 6 pm. As a consequence, these time differences might be able to indicate the typing pattern of each user in different time lines.

In order to control the experimental environment, especially the used keyboard, in this research will uses only the QWERTY keyboard for collecting data and use for screen pattern for eye testing. Therefore, the character's location and distance between characters on the keyboard which related to the typing patterns can be easily defined.

Since this research is a continuing process from [3], therefore, the similar display screen of [3] is deployed, as shown in Figure 1. The keyboard is separated into 9 areas, as shown in Figure 2. Nevertheless, this research will consider the position of each character on the used keyboard with the integration of counting times.

1	2	3
4	5	6
7	8	9

Figure 1. Screen pattern for eye testing

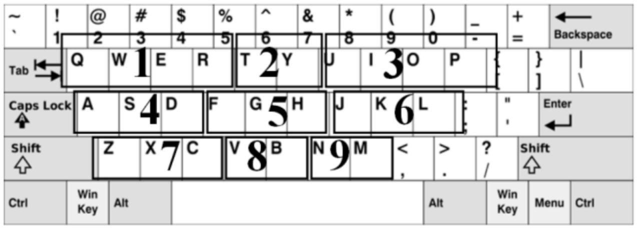


Figure 2. Separation groups of characters

According to Figure 1, the data collection system will random a character and display over those 9 segments for capturing time of keystroke. These times include the dwell time, interleave time, start time, and total time. These values are used for testing the eye vision of users. The random character will be displayed in different areas without replication.

Referring to groups in Figure 2, each character in each group will be randomly selected and displayed to one of the 9 segments. Thus, all volunteers will not be able to predict the next displayed character and position on the screen.

B. *Data Analyzing*

After data from 42 samples were collected, the next step is to analyze data. First, the position of all characters in the password, QLAMWXTV is drawn as a directed graph in Figure 3. This shows that all samples must move their hands or figures to type the full password.

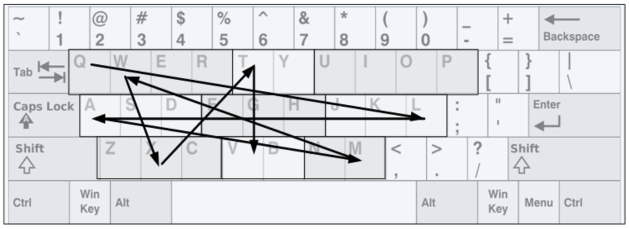


Figure 3. Positions of password's characters

According to positions presented above, the differences of dwell times of users' typing are considered. Therefore, the line graph that demonstrates the dwell times from 5 samples is presented in Figure 4.

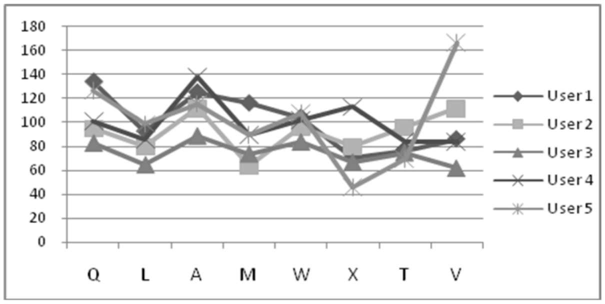


Figure 4. Example of dwell times from 5 samples

Figure 4 shows the sample of typing patterns of fixed password (QLAMWXTV) related to the dwell time when a user presses the character key. It is clear that samples have different patterns. In such cases, there is an opportunity that the differences between dwell times are related to the position of characters on the keyboard. Thus, consideration of the interleave time which is depended on the hand moving to each area of the keyboard is drawn in Figure 5.

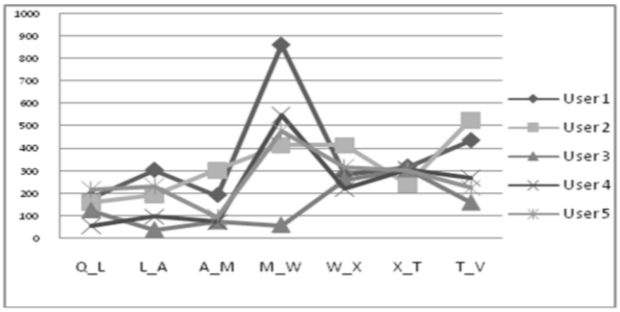


Figure 5. The pattern of interleave time

Referring to Figure 5, it is obvious that the character's location has influences to the interleave time of each user. Since every user used both hands for typing, therefore, when they typed Q to L, L to A, A to M, both hands were used and the figure's position normally put closed to the characters because characters are located on the boundary of the keyboard. Thus, small gaps of interleave times occurs. Similarly, when characters are located in the nearby areas, the interleave times are also small. Nevertheless, all these interleave times show that there is an impact when types different characters in difference areas. So, in order to prove this assumption, the statistical analysis using Analysis of Variance (ANOVA) of randomize completed block design (RCBD) is performed under the significant level (α) of 0.05. These tests will consider the significant differences of means of dwell times (MDT) and means of interleave times (MIT) from all users. Hypothesis of these statistical tests are defined below.

H_{01}: There is no significant difference among MDT of users when characters' locations are changed.

H_{11}: There is at least one significant difference between MDT of users when characters' locations are changed.

H_{02}: There is no significant difference among MIT of users when characters' locations are changed.

H_{12}: There is at least one significant difference between MIT of users when characters' locations are changed.

The result from ANOVA shows that there is at least one significant difference between means of dwell times of users when characters' locations are changed, *p*-value = 0.00 < 0.05(α). Thus, the multiple comparisons among MDT are performed using Scheffe analysis by considering only the differences related to the area of characters. Referring to the results, these confirm that the location of each character has impact to the MDT.

As same as the previous analysis of dwell times, the result

ISBN: 978-981-07-8859-9

from ANOVA also indicates that there is at least one significant difference between MIT of users when characters' locations are changed, p-value = 0.00 < 0.05(α). When performing the multiple comparisons among MIT values using Scheffe analysis. Based on the results, it is clear that locations of characters cause significant differences in the MITs.

Referring to results from MDT and MIT, it can conclude that models of keyboards have significant impact to the biometrics measurement. Therefore, not only the consideration in times based keystroke dynamic method but also the types of keyboards must be included.

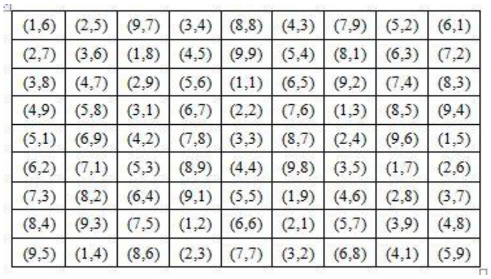
Although the research of [3] had proposed that the eye vision can be applied to the authentication system to increase the accuracy of the identification process, this research did not consider the impact from the position of the typing character. Thus, the results presented above can determine that the consideration of character's position should be counted as an important factor in the identification mechanism. The following section proposes the Behavioral Classification System (BCS) that applies the character positioning into the system proposed by [3].

V. Behavioral Classification System

Since the proposed solution of [3] considered only two factors, those are keystroke and vision time. However, the preliminary result of this research found that the position of the typing character might have impact to the typing time. Therefore, the assumption of this research is that the dwell time and the interleave time of the vision measurement should combine with the position of the character. Thus, the accuracy to classify a person will be obtained from combination of three factors: keystroke dynamics, speed of eye vision, and character's position.

This section proposes the concept to prove the impact of character position towards the typing time of the vision testing period. The BCS is the system that responsible for classifying a person via Key Positioning Collector System (KPCS) using users' profiles, typing area, character area, and time capturing.

Since there are 9 key areas and 9 presented areas, the combinations of these factors must be completely matched to obtain a complete set of data. So the Graeco-Latin Square is applied for this experiment, as shown in Figure 6. Thus, the obtained data will cover all factors without unsuitable replicated data set.



(1,6)	(2,5)	(9,7)	(3,4)	(8,8)	(4,3)	(7,9)	(5,2)	(6,1)
(2,7)	(3,6)	(1,8)	(4,5)	(9,9)	(5,4)	(8,1)	(6,3)	(7,2)
(3,8)	(4,7)	(2,9)	(5,6)	(1,1)	(6,5)	(9,2)	(7,4)	(8,3)
(4,9)	(5,8)	(3,1)	(6,7)	(2,2)	(7,6)	(1,3)	(8,5)	(9,4)
(5,1)	(6,9)	(4,2)	(7,8)	(3,3)	(8,7)	(2,4)	(9,6)	(1,5)
(6,2)	(7,1)	(5,3)	(8,9)	(4,4)	(9,8)	(3,5)	(1,7)	(2,6)
(7,3)	(8,2)	(6,4)	(9,1)	(5,5)	(1,9)	(4,6)	(2,8)	(3,7)
(8,4)	(9,3)	(7,5)	(1,2)	(6,6)	(2,1)	(5,7)	(3,9)	(4,8)
(9,5)	(1,4)	(8,6)	(2,3)	(7,7)	(3,2)	(6,8)	(4,1)	(5,9)

Figure 6. Design experiment of 9*9 Graeco-Latin Square.

According to Figure 6, the matching between character's location and presented areas on the screen is represented as an order pair (*x*, *y*) where *x* represents the character's position and *y* represents the presented area. For example, (1, 6) means a character within the keyboard area#1 will be random and presented in the screen area#6. All processes of the KPCS are drawn in Figure 7.

In each round, all testing patterns will be stored in the PatternDatabase (PattnDB); finally, there are 3780 records and all collected data will cover all necessary patterns as designed. Figure 8 shows collected attributes in the PattnDB.

After obtaining required data, the statistical analysis to determine relationship between eye speed and character positioning will be performed using ANOVA Greco-Latin under 95% confident interval; the software is SPSS v.17.0 of Chulalongkorn University. Then, the neural network mechanism will be applied to confirm the precision of the personal identification mechanism based on the combination of factors mentioned above. The accuracy of the identification process will also be measured using FRR and FAR values.

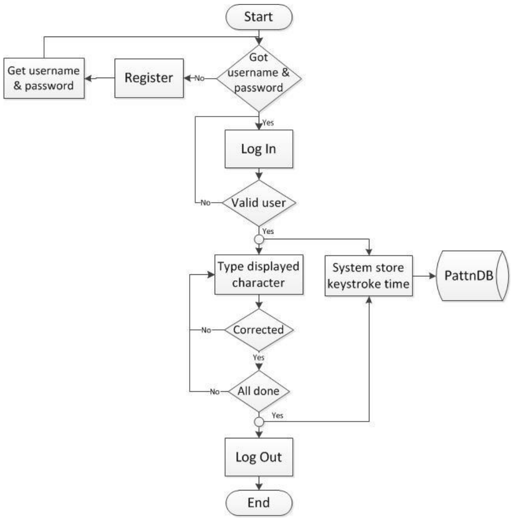


Figure 7. Flow diagram of the KPCS

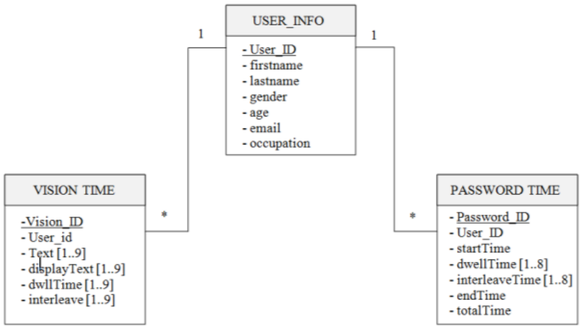


Figure 8. Attributes in the PattnDB

Based on the research of [3] and the preliminary study mentioned in the previous section, there is a possibility that the accuracy of the identification process should be higher than the combination between keystroke dynamics and the vision speed. Moreover, the result might lead to the optimized solution of using either keystroke with the character's position or eye vision with the character's position.

VI. Discussion

Since computers are installed over the Internet, various intruders try to attack the existing system in tremendous ways. One direct technique is to steal the user name and password to gain access from the system. Therefore, many protection mechanisms are proposed and implemented. However, the most popular technique in the present world is the use of biometrics that can be obtained only from individual person. These biometric are such as fingerprint, iris, face recognition, speaker recognition, and keystroke dynamics. Unfortunately, the only one metric cannot completely guarantee the correctness of the identification system. Therefore, multi-biometrics is applied to gain higher accuracy. Many researches combined the technique of keystroke dynamics to other biometrics values such as face recognition. However, these techniques need special equipment in the classifying process.

Even though the eye vision was proposed by [3] to combine with the keystroke technique, this research has found that the mean time capturing of the same password from various samples is significantly different. As a result, the suitable parameters to be deployed to the authentication process should include the location of typing character as well as consider the typing time when the character is presented. In addition, this technique needs no extra equipment and also can be applied to every keyboard system, including the touch screen technology.

VII. Conclusion

Currently, people access the Internet to perform their daily tasks. Unfortunately, there are unlawful users over the Internet and all resources on the network are in the high risk. Thus, password was implemented to protect the valuable system, sadly that the password was easily hacked. Therefore, the implementation of biometrics in the authentication system was proposed and implemented years ago. Nevertheless, the authentication process using only single biometrics cannot fully protect unauthorized users. As a consequence, the multi-biometrics was proposed to increase the accuracy of identification processes, including the eye vision ability.

This paper had found a new factor that affects to the typing times in various aspects; this factor is the character's position that appeared on the keyboard. Thus, the same user uses different keyboard styles might have different pressing times. Consequently, locations of characters should be included in the authentication process as same as another biometrics factor.

VIII. Future Work

The next step of this research is to prove that the vision speed is also relied on the character's position. Moreover, the classification mechanism using keystroke dynamics, eye vision, and characters' locations will be proposed, implemented, and test.

References

[1] Debnath Bhattacharyya, Rahul Ranjan, FarkhodAlisherov A., and Minkyu Choi, "Biometric Authentication: A Review," Int. J. u- and e-Service, Science and Technology, Vol. 2, No. 3, September, 2009.

[2] Kenneth Revett, Florin Gorunescu, Marina Gorunescu and Marius Ene. "A machine learning approach to keystroke dynamics based user authentication," Int. J. Electronic Security and Digital Forensics, Vol. 1, No. 1, 2007.

[3] K. Nonsrichai, and P. Bhattarakosol. "A New Alternative of an Authentication System using the Eye Vision Ability," Computing and Convergence Technology (ICCCT), 2012 7th International Conference on IEEE, 2012.

[4] Salil P. Banerjee, and Damon L. Woodard. "Biometric Authentication and Identification using Keystroke Dynamics: A Survey," Int. J. Journal of Pattern Recognition Research 7, 2012.

[5] F. Bergadano et al. "User Authentication through Keystroke Dynamics," Int. J. ACM Transactions on Information and System Security, Vol. 5, No. 4, November 2002.

[6] Obaidat, M.S., Sadoun, "Verification of Computer Users using Keystroke Dynamics," in IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics 27(2), 1997.

[7] G.C. De Silva, J.M. Lyons, S. Kawato and N. Tetsutani, "Human Factors Evaluation of a Vision-Based Facial Gesture Interface," in Conference on Computer Vision and Pattern Recognition, 2003, pp. 52.

[8] W. Kienzle, F.A. Wichmann, B. Schölkopf, and M.O. Franz, "Learning an Interest Operator from Human Eye Movements," in Conference on Computer Vision and Pattern Recognition, 2006.

[9] G. Shin, J. Chun, "Vision-based Multimodal Human Computer Interface based on Parallel Tracking of Eye and Hand Motion," in IEEE (ICCIT), 2007.

[10] S. Giroux, R.W. Smolikova, and Mark P. Wanchowiak, "Keypress Interval Timing Ratios as Behavioral Biometrics for Authentication in Computer Security," in IEEE, 2009.

[11] W. Chang, "Improving Hidden Markov Models with a Similarity Histogram for Typing Pattern Biometrics," in IEEE, 2005.

About Author (s):

Miss Kranogwan Krasaesat is a student under Master of Science, Computer Science and Information Technology of Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University.

Dr. Pattarasinee Bhattarakosol is an Assistant Professor in the Computer Science Program, Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University. Her research is mainly in the area of computer network. Currently, she has many international publications.