Randomized Mouse Movement for Behavioral Biometric Identification

Nazirah Abd Hamid, Suhailan Safei, Siti Dhalila Mohd Satar, Suriayati Chuprat and Rabiah Ahmad

Abstract—A biometric system consists of pattern-recognition system that recognizes a person based on either physiological or behavioral characteristics. This system can be used either to identify or verify a user. In this paper, we proposed a behavioral biometric system that used random mouse movement to identify a user. We developed an application and tested it with five users. The experiment produced 14 matching or equal to 46.67% of successful matching. The experiments are done to observe the human actions under unpredictable situation because in real life, human acts are depending on their mood, stress or the surrounding environment.

Index Terms— behavioral, biometric, identification, mouse movement

1 INTRODUCTION

The evolutions of biometric systems are progressing in such significant ways that nowadays, the biometric systems can be found almost everywhere. A biometric system is basically a pattern-recognition system that consists of [1]:

1. Representation: indicate the input data that measure the pattern's characteristics.
2. Extraction: involve the process of pattern's characteristics to become measurement vectors.
3. Classification and identification: these processes able to produce positive identification within a specific range.

This pattern-recognition system able to recognize a person based on either physiological or behavioral characteristics [2]. Physiological biometrics identify a user based on physiological characteristics, such as fingerprints and eye retina/iris scanning, whereas behavioral biometrics depend on detecting the behavioral features of the user, such as signature, voice, keystroke dynamics and mouse movements [3].

Biometric system can be used as an authentication method and it can be recognized into two different modes. The first mode, identification can be defined as a process to identify a user by goes through all the users in database for a match. The second mode namely verification is a process where the system tries to accept or to reject a user's legit identification and usually comes with other identifier such as id or password.

For behavioral biometric identification system, to obtain the required data that are significant for identification can be difficult. Thus, a suitable application or Graphical User Interface (GUI) to acquire this data is important especially in determining the accuracy of the system. A suitable GUI needs human-computer interaction (HCI). Human-computer interaction involves the interaction between people and computers to achieve a goal. A successful interaction between a people and computers include the processes and actions under any given environment.

In this research, an application was designed to capture data of a user when the user moved a mouse to follow a set of random buttons. This activity needed for the user to move the mouse and clicked on the buttons according to where the buttons appeared. The random environment is the main contribution because we want to see the behavior of the user when he/she interacted with the random buttons. Then this behavior could be translated into profiles that could be used for identification process to identify a user.

The paper is organized as follows. Section 2 describes some related research on mouse biometric system. Section 3 explains the experimental design and implementation. We present the result analysis and discussion in Section 4 and lastly we conclude this work in Section 5.

- *Nazirah Abd Hamid is with the Faculty of Informatics, Universiti Sultan Zainal Abidin (UniSZa). E-mail: nazirah@unisza.edu.my*
- *Suhailan Safei is with the Faculty of Informatics, Universiti Sultan Zainal Abidin (UniSZa). E-mail: suhailanh@unisza.edu.my*
- *Siti Dhalila Mohd Satar is with the Faculty of Informatics, Universiti Sultan Zainal Abidin (UniSZa). E-mail: sitidhalila@unisza.edu.my*
- *Suriayati Chuprat is with the Advanced Informatics School (AIS), Universiti Teknologi Malaysia (UTM) Kuala Lumpur. E-mail: suria@ic.utm.my*
- *Rabiah Ahmad is with the Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM) Melaka. E-mail: rabiah@utem.edu.my*

2 RELATED WORKS

Behavior biometric can be defined as knowledge that evaluates human behavior, actions or skills [4]. It is important to remember that behavioral biometric systems are involving human actions that base on human skills, knowledge, style and motor-skills. Table 1 illustrates behavioral biometric system's measurements and features that can be extracted [5].

TABLE 1. BEHAVIORAL BIOMETRICS: MEASUREMENTS AND FEATURES

Behavioural Biometric	Measures	Features
Car driving style	Skill	Pressure from accelerator pedal and brake pedal, vehicle speed, steering angle
E-mail behaviour	Style	Length of the e-mails, time of the day the mail is sent, how frequently inbox is emptied, the recipients' addresses
Haptic	Style	3D world location of the pen, average speed, mean velocity, mean standard deviation, navigation style, angular turns and rounded turns
Mouse dynamics	Style	x and y coordinates of the mouse, horizontal velocity, vertical velocity, tangential velocity, tangential acceleration, tangential jerk and angular velocity

As any biometric system, mouse movement biometric system basically is adopting a basic biometric system model. A simple mouse biometric system usually consists of three modules. The three modules are:

1. Data Capture Module consists of an application that can collect raw data regarding the mouse behavior of a user when he or she is interacting with a GUI.
2. Feature Extraction Module purpose is to analyze the raw data to generate user feature vectors that can be used to distinguish each user behavior through their mouse movements.
3. Classifier Module where the extraction feature will be used to identify or verify a user.

Among the earliest research on mouse dynamic was to propose a scheme to identify the identification of a user by using a mouse. Hayashi, Okamoto and Mambo [6] conducted two sets of experiment to capture data of mouse movements. On the first experiment, the users were asked to draw a circle inside a given circle on the screen while the second experiment allowed the users to draw any sort of shape (i.e. triangle) but still inside the given circle. From the experiment, the authors recorded the coordinate of (x, y) and the elapsed time.

Then during the feature extraction module, the coordinate of (X, Y) and the elapsed time that was in a file were processed and put in a database with new added data; the length from the coordinate to the center of the circle. Lastly in classifier module, authors used comparing and verifying method. Whenever an input data was presented, it would be compared with the database [6].

At the end, Match-rate is calculated using a formula, and the results showed that the authors managed to get only 7% of False Acceptance Rate (FAR). FAR was the probability that the system incorrectly matches the input data to a data in the database. This research shows that identification process has big potential to be done through mouse movement [6].

In a research to actively authenticate a user by mouse movements, Aksari and Artuner [7] had created an application to capture data. The users were required to click on the first square, and then the second square would emerge in another location of the screen. The random order of the path was created so that the users could not guess the next square, hence a comparable dataset could be collected each time the users used the application. As for the raw data, this application would record the coordination of cursor and the time when the event had happened.

Then, after the raw data were acquired, a set of features were obtained. The features vectors were speed, deviation from a straight line, angle and positive acceleration using statistical method. To be able to differentiate each user, the authors further analyse the features vectors by calculating the average, standard deviation, maximum and minimum of them (path features) and normalized them so that better comparison could be done [7].

Aksari and Artuner [7] did training phase before the verification phase as the classification methods. In the training phase, the authors extracted path features from the mouse movement attributes. Then the authors calculated the differences of the path features. The outputs were average of difference and standard deviation of difference. To verify, after the users used the application, the application would extracted the path features as in the training phase. Then for each user, counter values were calculated and if the user counter values were high it was meant that he or she was the intended user. This experiment managed to get even lower results of FAR; 5.9%.

Recent research trends in behavioral biometric identification are focusing on the implementation in mobile devices. According to Wolff [8], the used of continuous authentication such as identification process would improve the security of the devices. The author used different sensors from mobile phones to collect data namely accelerometer, touch screen, and keyboard by analyzing the user's interaction with the device. Then an analysis engine was applied as a feature extraction module to extract data that could differentiate users. This research is important because the results can be a preliminary result of observing the ability of mobile devices to identify users based on their behavior.

3 SYSTEM DESIGN AND IMPLEMENTATION

This section describes the overall structure of Randomize Mouse Movement for Behavioural Biometric Identification. The structure can be illustrated in Fig. 1 as it shows every process that involved in the system design.

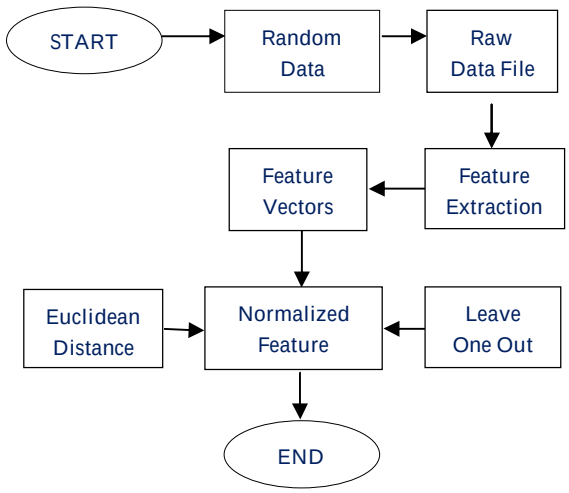

Fig. 1. Randomize mouse movement for behavioral biometric identification structure.

3.1 Randomized Data Capture Application

The Data Capture application was well hidden from the user and only could be functioning when the user started to interact with an experiment. The experiment involved the user to perform an activity called "follow the button". The user needed to move the mouse and clicked on the buttons according to where the buttons appeared. The buttons were arranged in random pattern so that the user could not predict the co-ordinate of the next button. This was a very important feature because if the user could not predict the pattern of the buttons, the users could produce a data size that is not uniform and varied.

When the user clicked on the first button as in Fig. 2, the second button would appear randomly as shown in Fig. 3. This user had to click until the 20th button to finish as illustrated in Fig. 4 without any constraint of time.

For this research, the experiment was conducted by using the user's own laptop and mouse within seven days to collect six different data. This user was asked to perform the experiment six times to establish the user personal profile consisting of user behaviours and characteristics.

.

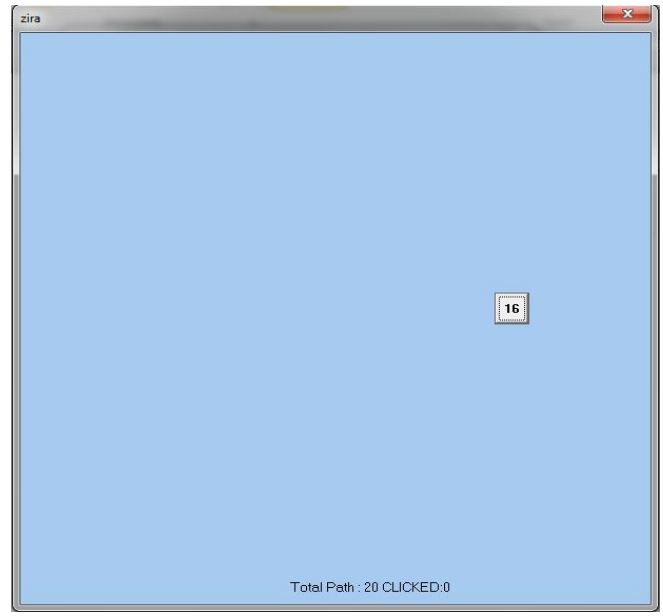

Fig. 2. Random Button GUI - Start.

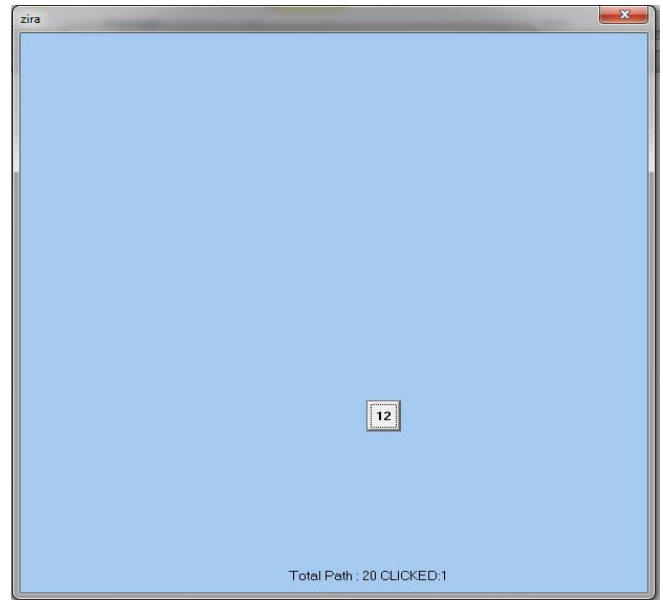

Fig. 3. Random Button GUI.

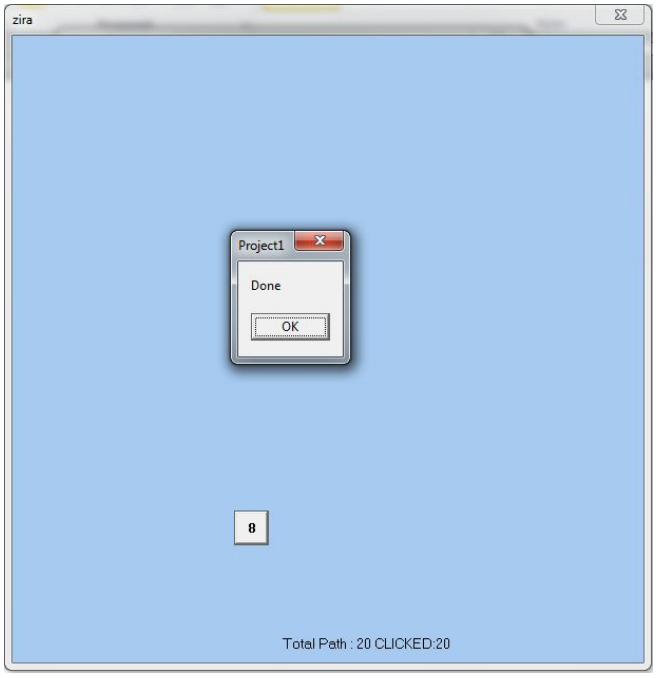

Fig. 4. Random Button GUI – End.

3.2 Raw Data

During the interaction of the user with the random buttons, the application would record three main attributes that would be used in the feature extraction module. These attributes or what we called as raw data consisted of coordinate X and Y as the user clicked on a button that set in a certain grids. The other attribute that involved in the process was time in milliseconds.

3.3 Feature Vectors Extraction

The purposes of this phase were to process and analyze the raw data and generated user features vectors. These features vectors were uniquely different for each person and can be generated into a signature of each user.

1. Mouse Movement Profile
   The raw data bring no significance values and did not represent any meaning about a user's behaviors or characteristics. To get mouse movement profile for each user, the raw data were applied by with calculations. The formulas were time, speed, acceleration, deviation and angle of deviation.
2. Mouse Movement Profile Measurement
   After the mouse movement profiles for each user were created, there existed some counts of the mouse movement points that could differentiate each user. So to find the nearest value from those counts that could distinguish a user to another user, the values of average and standard deviation were calculated. These values would be the best way to describe a user and could be known as mouse movement profile measurement.

The following steps summarized the whole steps:

- The raw data, x and y coordinates and t (ms) from a data file was applied with five different formulas to create a mouse profile.
- Then the average and standard deviation of each result would be determined as the mouse profile measurement.

3.4 Classifier Process

The main function of this process was to verify the validity of values from previous processes and classify the patterns that could identify a user. Reference [9] stated that classifier program could be functioning into two ways namely authorization technique or identification technique. In this research the classifier process was used for identification process.

1. Normalization of Data
   The experiment required the user to test the application for six times. It was important to establish a user's profile/signature. All the values would go through normalization before the identification process begun. This normalization process would set the values within 0-1 which was a common number range. This experiment would use Euclidean distance "(1)," as a method to find the similarity values, it is important that the values are within 0-1 values [9].

$$Distance(A,B) = \sqrt{\sum_{i=1}^{N}(a_i - b_i)^2} \quad (1)$$

2. Leave One Out Method
   The Leave One Out Method was used in the identification process. This method could be done by comparing or testing a test file value against all the file's values in the training data set by using Euclidean Distance formula.
   The classification was considered as successful attempt if the Euclidean Distance values between the test data and the data in the database were near with each other. In respect of the data were from the same user.
   For example, four users were involved in the experiment. Each user had two files. The training data set was Saida = {Saida1, Saida2}, Linie = {Linie1, Linie2}, Mala = {Mala1, Mala2} and Zalehah = {Zalehah1, Zalehah2}.

   Training Set = {Saida1, Saida2, Linie1, Linie2, Mala1, Mala2, Zalehah1, Zalehah2}

- Saida1 file was taken and compared it to the rest of the files in the training data set.
- We used Euclidean distance to compare the value of Saida1 to all the values of the files in the training data set.
- A match was successful when, the value of Saida2 was the closest to Saida1.
- Then the process would continue on until the last value in the training data set was tested. All the successful or positive matching would be recorded.

Following section illustrated the steps that included in classifier module:

- The average and standard deviation values were normalized so that all the data values were in a normalized manner.
- Identification process was implemented by using Leave One Out Method.
- To compare both data, Euclidean distance was chosen as to find the distance between the test data and the data in the training data set. If from the same user and the distance between the data were close, so the identification process is successful.
- The values/results from above process would be analysed to produce success percentages.

4 EXPERIMENTAL RESULT AND ANALYSIS

The results of Leave One Out Method for identification can be summarized as in Table 2. The results in Table 2 show that the successful matching is 14 matching from 30 data to produce 46.67% of success percentages.

TABLE 2. RESULTS OF IDENTIFICATION MATCHING

Case Description	Total Data	Matching	Percentage
Successful matching	30	14	46.67%

The experiment on Randomize Mouse Movement for Behavioural Biometric Identification manages to identify 14 identifications of users. This experiment is conducted in an uncontrollable environment whereby the experiment is tested by using the user own laptop and mouse. The results can be a preliminary result in observing the behaviour of a user when the user was interacting with a random application by using a mouse. Moreover, we can conclude that it is quite difficult to get a match for a user or to identify a user in a random environment.

5 CONCLUSION

The core contribution of our research is to create a random environment for identification process and test it against several users. This behavioral biometric identification system is designed to identify a user by capturing the data of a user when a user moved the mouse to follow a set of random buttons. The experiments are done to observe the human actions under unpredictable situation because in real life, human acts are depending on their mood, stress or the surrounding environment. From the experimental results, it can be concluded that to identify a user is quite challenging task especially when a user has to deal with random situation or environment.

REFERENCES

[1] F Monrose, and A.D Rubin, "Keystroke Dynamics as a Biometric for Authentication," Future Generations Computing System, 16(4), 2000, pp. 351-359.

[2] A.K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," IEEE Transactions on Circuits and Systems for Video Technology, 14(1), 2004, pp. 4- 20.

[3] A.A.E. Ahmed and I. Traore, "Detecting computer intrusions using behavioural biometrics," in 3rd Annual Conference on Provacy, Security and Trust, New Brunswick, Canada: ISOT, 12-14 October 2005, pp. 91-98.

[4] A. Yannopoulos, V. Andronikou, and T. Varvarigou, "Behavioural biometric profiling and ambient intelligence," in M. Hildebrandt and S. Gutwirth, (Ed.), Profiling the European Citizen, Springer Netherlands, 2008, pp. 89-109.

[5] R.V. Yampolskiy and V. Govindaraju, "Behavioural biometrics: A survey and classification," International Journal of Biometrics. 1(1), 2008, pp. 81-113.

[6] K. Hayashi, E. Okamoto, and M. Mambo, "Proposal of User Identification Scheme Using Mouse," in Information and Communications Security, Springer Berlin, 1997, pp. 144-148.

[7] Y. Aksari and H. Artuner, "Active authentication by mouse movements," in 24th International Symposium on Computer and Information Sciences, 14-16 September 2009, North Cyprus:IEEE, pp. 571-574.

[8] M. Wolff, "Behavioral biometric identification on mobile devices," in Foundations of Augmented Cognition. Springer Berlin Heidelberg, 2013, pp. 783-791.

[9] P. Kasprowski, "Human Identification Using Eye Movements," Doctoral Thesis, 2004, Silesian University of Technology, Poland.

Nazirah Abd Hamid is a lecturer in University Sultan Zainal Abidin, Terengganu. She holds a degree in Bachelor of Information Technology from University Utara Malaysia (UUM), in 2004 and M. Sc. Com. (Information Security) from University Teknologi Malaysia (UTM), Malaysia. Her research interests are Information Security and Human Computer Interaction (HCI).

Suhailan Safei is currently working as Deputy Dean, Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Malaysia. He has 5 years of teaching experience and 7 years of working in real time software development. He has obtained his B.Sc. in Computer Science and M.Sc. Computer Science (Real Time Software Engineering) from Universiti Teknologi Malaysia. His research interests include real time programming and e-learning**.**

Siti Dhalila Mohd Satar is currently works as a lecturer at Faculty of Informatics and Computing, University Sultan Zainal Abidin, Terengganu. She holds a B.S. degree in Information Technology from University Kebangsaan Malaysia (UKM), in 2008 and MSc in Computer Science (Information Security) from University Teknologi Malaysia (UTM), Malaysia. Her research interests are Information Security, and Information Quality.

Suriayati Chuprat is a senior lecturer at Advanced Informatics School (AIS), UTM International Campus, Kuala Lumpur.

Rabiah Ahmad is an Associate Professor at the Faculty of Information Technology and Communication, University Technical Malaysia Melaka (UTeM), Malaysia. She received her PhD in Information Studies (health informatics) from University of Sheffield, UK, and MSc. (information security) from Royal Holloway University of London, UK. Her research interests include healthcare system security and information security architecture. She delivered papers at various health informatics and information security conferences on national as well as international level. She has also published papers in accredited national/international journals. Besides that, she also serves as a reviewer for various conferences and journals.