ISSN: 0067-2904

User Authentication via Mouse Dynamics

Osama A. Salman, Sarab M. Hameed*
Department of Computer Science, College of Science, University of Baghdad, Baghdad, Iraq

Abstract

Nowadays, the development of internet communication and the significant increase of using computer lead in turn to increasing unauthorized access. The behavioral biometric namely mouse dynamics is one means of achieving biometric authentication to safeguard against unauthorized access. In this paper, user authentication models via mouse dynamics to distinguish users into genuine and imposter are proposed. The performance of the proposed models is evaluated using a public dataset consists of 48 users as an evaluation data, where the Accuracy (Acc), False Reject Rate (FRR), and False Accept Rate (FAR) as an evaluation metrics. The results of the proposed models outperform related model considered in the literature.

Keywords: behavioral biometric, mouse dynamics, user authentication

مصادقة المستخدم من خلال ديناميكية الفارة

سراب مجيد حميد، اسامة علي سلمان
قسم علوم الحاسوب، كلية العلوم، جامعة بغداد، بغداد، العراق

الخلاصة

تطور الاتصالات في الوقت الحاضر من خلال الانترنيت والزيادة الحاصلة في استخدام الحاسبه ادى الى زيادة الوصول غير الماذون به. الباميتري السلوكي وتحديدا ديناميكية الفارة هي احدى الطرق لتحقيق المصادقة الباميترية للحماية من الوصول غير الماذون به. هذا البحث يقدم نموذجين لتمييز الاشخاص الى حقيقين ومزورين باستخدام ديناميكية الفارة . تم تقييم اداء النماذج المقترحة بأستخدام مجموعة بيانات مكونة من 48 شخص و معايير تقييم الدقة (Acc) ومعدل القبول الكاذب (FAR) ومعدل الرفض الكاذب (FRR). تم تقييم اداء النماذج المقترحة والتي اثبتت تفوقها من خلال مقارنتها مع بحث منشور في الادبيات.

1. Introduction

With the growth of the Internet and the continuous increase in the availability of huge amount of information, the importance of user identification and authentication have rapidly increased. Some of this information is confidential and should not be accessed by an unauthorized user. The common method used to address this problem is using a user authentication [1]. Biometrics provide a number of advantages over other authentication methods that are cannot be forgotten and compromised. However, a basic restriction in biometric is the demand for specific hardware to get biometric data.

*Email: sarab_majeed@yahoo.com

Due to these restrictions, an increasing interest in the research on mouse dynamics that does not need any specific hardware to collect biometric data. Mouse dynamics include monitoring the user behavior through how he/she interacts with the mouse as a means for authentication [2]. Mouse dynamics has attracted more research attention over the last decade. Some of these are Ahmed and Traore in 2007proposed a behavioral biometrics via mouse dynamics using artificial neural network algorithm. The signature of mouse dynamics involved movement speed (MSD), movement direction average (MDA), movement direction histogram (MDH), action type average (ATA), traveled distance histogram (TDH) and movement elapsed time histogram (MTH). A dataset with twenty-two participants was used as an evaluation data. The results showed that the proposed model produced false acceptance rate (FAR) of 2.4649% and false rejection rate (FRR) of 2.4614% [3].

Ahmed and Traore in 2010 introduced the mouse dynamics biometric concepts and presented a detector together and process mouse movements. In addition, various factors adopted to form user signature were considered. Testing of the detector was performed on the dataset in [3] in addition to mouse data involved twenty-six users. The proposed detector result achieved FAR equal 2.6052% and FRR equal 2.506% [4].

Shen et al. in 2013 presented a simple approach for user authentication using a fixed mouse-operation task. To characterize a user's behavior, holistic features and new procedural features were extracted. Dynamic time warping (DTW) distance and Manhattan distance were used to obtain a distance vector. Then, kernel principal component analysis (KPCA) was applied to get the feature components of the original feature-distance vectors. Finally, one-class classification techniques including SVM, nearest neighbor and a neural network were used for conducting the user authentication. The evaluation on a dataset from thirty-seven users demonstrated the effectiveness of the proposed approach with FAR of 8.74% and FRR of 7.69% [5].

Shen et al. in 2014 investigated the performance of anomaly detection algorithms based on mouse dynamics. The evaluation was performed on a dataset containing 17,400 samples from fifty-five users and seventeen detectors were applied. The results show that the six top-performing detectors produced Error Equal rate (ERR) between 8.81% and 11.63% [6].

Mondal and Bours in 2015 presented a study regarding the performance of continuous authentication using mouse dynamics. They used weighted fusion scheme, score boost, static trust model and dynamic trust model for analyzing and SVM and ANN as a classifier. The evaluation was done on a dataset that includes the mouse dynamics data obtained from forty-nine users. The results showed significant improvement over the beforehand performance results on the same dataset [7].

Mondal and Bours in 2016 introduced a new a technique based on pairwise user coupling for identification and continuous user authentication. They build a dataset that contains a combination of the data behavior of keystroke and mouse dynamics. The accuracy result was 62.2% and the detection rate was 58.9% [8].

Lu et al. in 2017 proposed an authentication approach using mouse movement and eye movement tracking. Two neural networks were used for multi-class classification and binary classification. In addition, the regression model with fusion was used for classification purpose. The performance of the proposed approach was evaluated on a dataset collected from forty users. The results clarified that coupling eye tracking with the mouse dynamics are applicable for authentication [9].

The proposed work presented in this paper is similar to the work developed by Ahmed and Traore in 2010 in utilizing histograms of features extracted from the mouse actions. However, the proposed work differs in using different features that can characterize the user behavior more accurately. A Gaussian Naive Bayes is utilized for classification purpose.

The rest of this paper is organized as follows. The mouse dynamics is briefly described in section 2. Section 3 introduces the proposed mouse dynamics user authentication model. Section 4 presents the evaluation metrics that are used to evaluate the proposed model. The results of the proposed model are then evaluated in section 5. Section 6 presents the comparison results with other methods. Finally, section 7 provides the conclusions and some hints for future work.

2. Mouse Dynamics

Mouse dynamics is considered as an example of behavioral biometric that was presented at the University of Victoria in the research lab of information security and object technology (ISOT) in 2003.The key strengths of the mouse dynamics compared with other biometric technologies is that it enables monitoring the user dynamically and passively. Accordingly, it can be utilized to track

continuously genuine and imposter users during computing sessions. Mouse dynamics need the availability of a mouse that is low-priced. When a user uses the mouse device, the characteristics of the mouse actions are analyzed. In mouse dynamics, some features are extracted from user's mouse actions. Then these features are stored. When the user wants to access the system later, the system compares his/her actions with the stored one and decides if he/she is a genuine or an imposter [10].

The mouse actions can be classified as silence action that denotes no movement and movement activities. Movements of the mouse can involve movement type, movement speed, traveled distance and movement direction. Movement type contains Mouse-Move (MM) action, Drag-and-Drop (DD) action and Point-and-Click (PC) action [5]

3. The Proposed Mouse Dynamics Model

The main challenge in developing a mouse dynamics user authentication model is detecting distinguished features that characterize user behavior. The proposed model introduces new features constructed from the properties of mouse movement to observe the user behavior.

3.1 Dataset Preprocessing

In this paper, the dataset developed in [3] [4] is used. This dataset includes the mouse dynamics data of 998 sessions collected from 48 users. The collected data contains mouse activities. Each activity holds the characteristics of an intercepted mouse movement. The collected data contains four main mouse activities as described below:

1. Action type: the action type takes four values 1, 2, 3 and 4 for mouse move (MM), silence, point click (PC) and drag and drop (DD) respectively.
2. Traveled distance in pixels (d).
3. Elapsed time in seconds (t).
4. Movement direction takes eight values (1 to 8) according to the mouse movement.

The collected mouse raw data from each user have different ranges depending on an environment setting and the accuracy of mouse dynamics modeling can be affected by nature of the data. Two types of filtering are used in this paper. In the first filter, mouse data with distance value equal zero value and between 25 and 1200 are considered while in the second filter all users speed greater than 800 are eliminated.

To construct the feature set F that determining user's behavior, the mouse dynamics raw data for each user is divided into a number of mouse actions called sessions. In other words, the mouse dynamics raw data are organized into a session with a proper number of actions (S_{len}).

To characterize the behavior of a user of each session, the features are extracted and aggregated into histograms as in [4]

1. Movement direction histogram (MDH) feature denoted by eight values obtained by computing the proportion of actions performed per direction.
2. Action type histogram (ATH) feature denoted by three values. Each value is obtained as the relative frequency of the mouse actions in a session.

Statistical features can be extracted for user authentication via mouse dynamics. Some of the features used in [4] are utilized in this paper as follows:

1. Average movement speed per action type (ATA) feature denoted by three values. Each value is obtained as the average speed of carrying out the mouse actions (MM, PC and DD).
2. Average movement speed for each direction (MDA) feature represented by eight values. Each value is derived from the average speed (i.e., traveled distance in pixels/ time in seconds) over all the actions per direction.

Furthermore, backpropagation neural network as mentioned in [4] is utilized to extract a feature that defines the user behavior from mouse dynamics as a curve approximate to user-collected data. The trained backpropagation neural network is used to investigate the behavior of the user. To identify the behavior of the user, the minimum and maximum values of the speed and the distance $(s_{min}), (s_{max}),\ (t_{min})$ and (t_{max}) respectively should be found for each testing session. Then, twelve values of the speed and the distance are extracted.

Moreover, new features are extracted for user authentication via mouse dynamics as in what follow:

1. Average acceleration per movement direction (coined as AAD) feature denoted by eight values. Each value is obtained as the average acceleration (i.e., action speed divided by time) over all the actions per direction.

2. Average acceleration per actions type (coined as AAA) feature represented by three values. Each value is derived from the acceleration of carrying out the mouse actions (MM, PC and DD).

Integrating the above extracted features, two feature sets are introduced to characterize the behavior of the user authentication. The first feature set coined as MD#1 that contains MSD, MDA, MDH, ATA and ATH and the second one coined as MD#2 model that contains MSD, AAD, MDH, AAA and ATH features.

After extracting the features from mouse raw data that describe the user behavior, the values of extracted features have different ranges; therefore, the features are set in a uniform range to avoid some features' domination over others. The features are scaled linearly to the range $[-1, 1]$

3.2 User Classification

The role of classification stage in the proposed user authentication model is to categorize a user behavior as either genuine or an imposter. The extracted features resulted from preprocessing stage are used as the input to this stage. Gaussian Naïve Bayes classifier is utilized to show the ability of the proposed model to recognize user behavior as a genuine or an imposter. Gaussian NB classifier contains two stages: learning stage and testing stage. In learning stage, Gaussian NB is trained with features extracted from data preprocessing phase, given feature vectors extracted from preprocssing stages $\mathbb{F} = \{F_1, F_2, \dots, , F_n\}$ and their corresponding labels $C = \{0,1\}$, the prior probability $P(c_j)$, $c_j \in C, \forall j \in \{1,2\}$, is calculated as the frequency of user behavior belongs to c_j divided by the total number of user behavior in training dataset

Estimating the distribution of the feature of the given class is achieved by calculating the mean μ_j and variance σ_j^2 of feature F_i

In testing stage, the prior probability and mean and variance of each feature resulted from the learning phase are used as input to the classification phase. Then, for each of feature vector in testing dataset$\mathbb{F}' = \{F'_1, F'_2, \dots, , F'_{nt}\}$, the posterior probability of each class $c_j \in C, \forall j \in \{1,2\}$ is computed as in Equation 1.

$$P(c_j|F'_i) = P(c_j) \times \prod_{k=1}^{nt} PDF(F'_i) \quad (1)$$

Where

PDF is the probability density function that is computed as in Equation 2

$$PDF(F'_i) = \frac{1}{\sqrt{2\pi\sigma_{ij}^2}} e^{-\frac{1}{2}\left(\frac{f'ij - mij}{\sigma_{ij}^2}\right)^2} \quad (2)$$

4. Performance Metrics

The following metrics are used to evaluate the performance of biometric authentication system **[11]**

1. False Reject Rate (FRR) measures the ratio of genuine users that are misclassified as an imposter user) on the total number of genuine users (N_G).

$$FRR = \frac{Number\ of\ rejected\ \text{genuine users}}{N_G} \quad (3)$$

2. False Accept Rate (FAR) measures the ratio of imposter users that are misclassified as a genuine user on the total number of imposter users (N_I).

$$FAR = \frac{Number\ of\ acceptance\ \text{imposter users}}{N_I} \quad (4)$$

3. Accuracy ($\boldsymbol{Acc}$) measures the ratio of correctly classified users to the total number of users.

$$Acc = \frac{number\ of\ correctly\ classified\ users}{N_I + N_G} \quad (5)$$

5. Experimental Result

Dataset of mouse dynamics of 48 users is used to evaluate the performance of the proposed models. Each user has a different number of actions. In the proposed work, four different settings for the length of session $S_{len} = \{500, 1000, 1500, 2000\}$ are chosen to show the impact of session length on the ability of the proposed user authentication model in discriminate among users. The session length represents the number of actions required to complete a session.

Testing the proposed mouse dynamics models is performed by applying 3-fold cross-validation approach. Table-1 reports the evaluation of MD#1 and MD#2 models in terms of Acc,FRR and FAR**.**

Table-2 illustrates the performance of the proposed models compared with [4] in terms of average accuracy (Acc′), average false reject rate (FRR′) and average false accept rate (FAR′). MD#1 and MD#2 decreases the number of features from the complete set of 39 features to 34 features. The reported result in the table clearly points out that MD#1 model outperform [4] model. This belongs to the positive impact of excluding TDH and MTH features in the proposed MD#1 model. Furthermore, adding the new proposed AAD and AAA features have a positive impact on the performance of MD#2 model. This comes from the suitability of inclusion of AAD and AAA features to distinguish users. In addition, the results show that the session length effects on the performance of user authentication model. Increasing session length provides the proposed models with more information for constructing user signature from mouse dynamics. Accordingly, the performance of the user authentication models regarding the accuracy is increased and FAR and FRR are decreased.

Table 1-Acc,FRR and FAR of MD#1 and MD#2

		***MD*#1 Model**			***MD*#2 Model**		
S_{len}	**Fold #**	***Acc*%**	***FRR***	***FAR***	***Acc*%**	***FRR***	***FAR***
	1	90.368	0.3	0.084	91.103	0.338	0.073
500	**2**	91.103	0.329	0.074	91.176	0.342	0.073
	3	90.956	0.342	0.075	91.25	0.354	0.071
	1	91.801	0.162	0.077	90.776	0.27	0.082
1000	**2**	91.52	0.256	0.073	89.62	0.163	0.1
	3	90.643	0.233	0.084	91.667	0.302	0.069
	1	93.682	0.121	0.059	91.285	0.061	0.089
1500	**2**	91.068	0.2	0.082	92.593	0.167	0.068
	3	90.414	0.13	0.094	89.325	0.174	0.103
	1	93.931	0.042	0.062	93.353	0.083	0.065
2000	**2**	88.473	0.091	0.117	89.625	0.091	0.105
	3	92.775	0.095	0.071	93.642	0.19	0.055

Table 2-Comparison of the proposed model with [4]

S_{len}	**Model**	***Acc′***	***FRR′***	***FAR′***
	[4]	90.711	0.277	0.081
500	***MD*#1**	90.809	0.324	0.078
	***MD*#2**	91.176	0.345	0.072
	[4]	89.761	0.217	0.095
1000	***MD*#1**	91.322	0.217	0.078
	***MD*#2**	90.688	0.245	0.084
	[4]	90.632	0.117	0.092
1500	***MD*#1**	91.721	0.151	0.078
	***MD*#2**	91.068	0.134	0.087
	[4]	91.147	0.179	0.082
2000	***MD*#1**	91.726	0.076	0.083
	***MD*#2**	92.207	0.122	0.075

6. Conclusions

The results illustrate the capability of the proposed models to differentiate the genuine user from imposter one. The results show the proposed authentication models provide better accuracy, false reject rate and false accept rate than [4] despite the proposed models contain 34 features while [4] consists of 39 features. Also, the results illustrate the impact of session length. Increasing session length lets the models produce better results. Also, as a scope of further work, silence action can be considered that may improve the performance of the proposed models.

References

1. Khan, H.Z.U. **2010**. Comparative Study of Authentication Techniques. *International Journal of Video & Image Processing and Network Security -International Journals of Engineering & Sciences (IJVIPNS-IJENS)*, **10**(4): 9-13.
2. Zheng, N., Paloski, A. and Wang, H.M. **2011**. An Efficient User Verification System via Mouse Movements. *In Proceedings of the 18th ACM Conference on Computer and Communications Security*, October, Chicago, JL, USA.
3. Ahmed, A.A.E. and Traore, I. **2007**. A New Biometric Technology Based on Mouse Dynamics. *IEEE Transactions on Dependable and Secure Computing*, **4**(3): 165–179.
4. Ahmed, A. A. E. and Traore, I. **2010**. Mouse dynamics biometric technology. In *Behavioral Biometrics for Human Identification: Intelligent Applications*, K. Klinger and J. Snavely, Eds., , pp: 207-223.
5. Shen, C. Cai, Z. Guan, X. Du, Y. and Maxion, R. A. **2013**.User Authentication Through Mouse Dynamics. *IEEE transactions on information forensics and security*, **8**(1): 16-30.
6. Shen, C. Cai, Z. Guan, X. and Maxion, R. A. **2014**. Performance evaluation of anomaly-detection algorithms for mouse dynamics, *computer & security*, **45**: 156-171.
7. Mondal, S. and Bours, P. A. **2015.** computational approach to the continuous authentication biometric system. *Information Sciences*, **5**(21): 28 – 53.
8. Mondal S. Bours P. **2016**. Combining keystroke and mouse dynamics for continuous user authentication and identification. *IEEE International Conference on Identity, Security and Behavior Analysis* (ISBA), Japan.
9. Lu, H. Rose, J. Liu, Y. Awad A. and Hou L. **2017.** *Combining Mouse and Eye Movement Biometrics for User Authentication*. In: Traoré I., Awad A., Woungang I. (eds) Information Security Practices. Springer, Cham.
10. Nakkabi, Y. Traore. I. and Ahmed, A. A. E. **2010**. Improving mouse dynamics biometric performance using variance reduction via extractors with separate features. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans*, **40**(6): 1345-1353,.
11. El-Abed, Charrier, M. C. and Rosenberger, C. **2012**. Evaluation of biometric systems. In *New Trends and developments in biometrics*, Y. Jucheng and J. Shan, Eds., 2012, pp. 149-169.