Continuous user verification based on behavioral biometrics using mouse dynamics

Conference Paper · June 2013

DOI: 10.2498/iti.2013.0505

CITATIONS
4

READS
94

1 author:

Mirko Stanić
University of Zagreb
1 PUBLICATION **4** CITATIONS

SEE PROFILE

Continuous User Verification Based on Behavioral Biometrics Using Mouse Dynamics

Mirko Stanić
Agency for Science and Higher Education
Zagreb, Croatia
E-mail: mirko.stanic@azvo.hr

Abstract. *This paper provides overview of the current methods of identifying users based on their interactions with a computer keyboard, mouse or a touchscreen and argues that in their current state of development none of them are capable of establishing the users identity within the time it takes for a user to input a password.*

The paper proposes the application of behavioral biometrics as a supplement to regular password based user authentication as a safeguard against unauthorized users gaining access to a computer that is already running an authenticated session e.g. unattended computers in offices.

Keywords. *behavioral biometrics, mouse dynamics, verification*

1. Introduction

Today most computer systems identify users by means of secret phrases known as passwords. However this authentication system does nothing to protect the computer from unauthorized access once the user has started an active session. User authentication at sign on secures the workstation only against unauthorized access while the workstation is powered down or a user is logged off and even then only if the attacker does not know a valid user's password.

Furthermore it is shown by [14] that most users write down their passwords, pick weak passwords or are willing to tell them to a complete stranger in exchange for chocolate. Unattended computers with an active session present a much larger security threat. In offices it is common for people to step away from their desks be it to speak to a colleague in the same office, attend a meeting or just go on a break. Users which are not tech savvy will frequently leave their computers unlocked and with an active session.

It is already established by [12] most attacks originate from the inside the organization that is being attacked, be it on purpose, possibly by a disgruntled employee, or by accident by a user with privileges that are higher than it is actually required for his position. This allows for three types of attacks. A user of lower clearance can gain access to a terminal with higher clearance and access files or functions of the network to which he is not supposed to have access to or a user with the same or higher clearance can conceal his identity by performing malicious actions under the guise of a coworker. Lastly a person who is not affiliated with the company in anyway and is simply visiting can gain access to the internal network.

These limitations of password based authentication lead to the introduction of authentication techniques based on biometrics. We differentiate two types of biometrics: physiological biometrics and behavioral biometrics. Physiological biometrics is based on measuring physical human features that are relatively unique to each individual such as fingerprint, face, palm, iris, voice etc. Because of this physiological biometrics require specialized hardware such as fingerprint scanners which increase the cost of devices they are implemented in. On the other hand behavioral biometrics is based on a behavioral trait of an individual such as signature speech pattern, keystrokes or mouse movements. While they are more susceptible to change depending on the time of day when they are captured or subjects state of relaxation the benefit is that they do not require specialized hardware for acquiring them.

At its very core a biometric-based verification system is a pattern recognition system that acquire a persons biometric data, extracts a feature set and constructs a verification model. Said systems include the following elements: feature extraction which captures the data generated by standard input devices such as a mouse or a keyboard, feature extraction module that constructs the signature which characterizes a user based on his behavioral biometrics, a classifier that is used to construct a user verification model and a signature database

consisting of behavioral signatures of registered users. Fig 1. taken from [3] shows an example of behavioral biometric identification system architecture.

2. Existing methods

Behavioral biometrics on desktop computers is commonly based on keystroke dynamics and mouse dynamics. Performance of behavioral biometrics is measured by "False Acceptance Rate" (FAR), the ratio at which an attack is erroneously characterized as a valid user, and False Rejection rate (FRR), the ratio at which a login attempt by a genuine user is erroneously characterized as an attack. We also define an Equal Error Rate (EER) which is the point at which both FAR and FRR are equal. If FAR is high the system will be less likely to recognize a legitimate user as an attacker but there is also a higher chance that an attacker will be recognized as a legitimate user. On the other hand if FRR is high the system will become much more intrusive on the part of legitimate users by frequently erroneously login them out of the system but it will be much less likely to recognize an attacker as a legitimate user. In real life applications the desired objective is to keep FAR and FRR at approximately the same level.

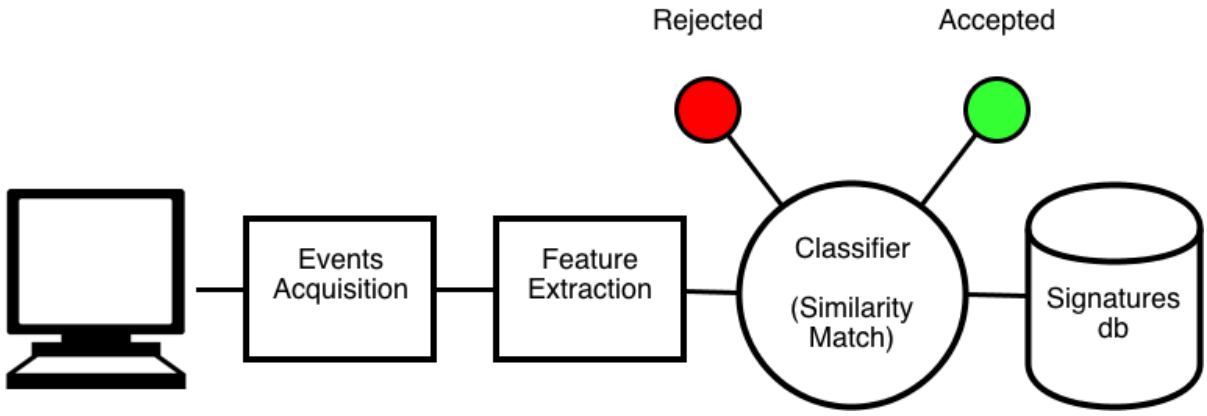

Figure 1. A typical framework of a behavioral biometric identification system [3]

2.1. Keystroke based methods

Gains et al. [4] were one of the first to propose a method for user identification via keyboard dynamics in 1980. The research was conducted on a small group consisting of 7 professional typists. Their work established that there is a "signature" to human typing which was in their case used to distinguish left handed typists from right handed ones.

Joyce and Gupta [7] developed classification techniques based on latencies between the time the user presses a key and releases it as well as the time that passes between to keystrokes. The method requires users to type a structured text and is not suitable for continuous user verification.

Monrose and Rubin [8] considered using multiple classifiers such as Euclidean distance measure, probabilistic measure and a third one which was an optimized version of the second classifier with the addition of weighted scores. The method was tested in an uncontrolled setting in order to better simulate a real life environment. Their method resulted in a FAR rate of 10%.

Yu and Cho [16] implemented a support vector machine (SVM) based classifier in their method as opposed to a neural network used in many earlier methods. The verification was done by recording the users keystrokes while they were typing in a password and resulted in a FAR rate of 0% and FRR of 3.69%.

Table 1. Comparison of existing user verification method by [17]

Source	FRR	FAR	Data required	Settings	Notes
[1]	2.4549%	2.4614%	2000 mouse actions	Continuous	Free mouse movements
[9]	0%	0.36%	2000 mouse actions	Continuous	Free mouse movements
[5]	2%	2%	50 mouse strokes	Static	Mouse movements from a game
[10]	1.75%	0.43%	Not specified	Continuous	Applies to a certain application
[13]	11.2%	11.2%	3600 mouse actions	Continuous	Free mouse movements
[11]	4%	3.5%	Not specified	Static	Mouse movements from a game
[3]	9.5%	17.66%	30 mouse actions	Continuous	Free mouse movements
[17]	1.3%	1.3%	20 mouse actions	Continuous	Free mouse movements

2.2. Mouse based methods

Gamboa and Fred [5] envisioned mouse based biometrics as a substitute for text based passwords. Their method required the user to identify matching pairs of images on tiles and verification was performed based on the characteristics of the user's mouse movements from one tile to the other. The system was tested on a sample of 50 users and produced EER of 0.7% for 100 mouse strokes which lasted 1 second each. That puts the detection time under 2 minutes.

Pusara and Bordley [10] proposed a web based verification method which recorded participants mouse movements while they were browsing a web site. Users were classified using the C5.0 decision tree algorithm. The method resulted in FAR of 0.46% and FRR of 1.75% with a highly variable detection time between 1 and 14.5 minutes.

Ahmed and Traore [1] developed a method which monitored user's interaction with a mouse throughout the whole session and extracted certain features which were then aggregated into histograms that were used to determine the identity of each user. A binary neural network was used as a classifier and the method achieved a FAR of 4.6% and FRR of 24% for a user session lasting about 4 minutes.

Revett et al. [11] based their work around using a GUI in which the correct sequence of elements was arranged using a mouse. The system was envisioned as a replacement for text based passwords. The system achieved FAR of 3.5% and FRR of 4.0%.

Bours and Fullu [2] tracked the users while they navigated the mouse through an onscreen maze. The method was tested on 28 users and resulted in a relatively high EER of 27%.

Zheng et al. [17] presented a method which could identify a user with as few as 20 mouse clicks and which is computationally not as intense as some earlier methods due to the use of a SVM as a classifying engine instead of a neural network. Their method produced FAR of 1.3% and FRR of 1.3%.

Feher et al. [3] proposed a novel verification method based on observing each individual mouse action performed by the user. The mouse actions were atomized to the point of a single click or movement and more complex actions were then defined using these basic actions. The basic actions are left click, right click, mouse move sequence and drag-and-drop action. This reduced the amount of time required to identify a user compared to other histogram-based methods. The method produced EER of 8.53%.

3. Situation today

Current methods of user identity verification based on mouse movements are not efficient enough to achieve the European Standard for Access Control Systems requirements, which are a FRR of less than 1% and FAR of under 0.001%. Table 1, taken from [17], shows the effectiveness of current mouse based user verification methods. To this end todays behavioral biometrics systems employ both mouse based behavioral biometrics as well as ones that are keystroke based. The problem with such systems is not of a technical nature but more of a social one because for a keystroke identification system to function it has to record all of the users input. Systems that identify users through the dynamics of their keystrokes in essence fall into the category of key loggers and a user has to trust that the system will not record his passwords or private messages and relay

them to a malicious third party. Because of this it is important to achieve a fully functional biometric identification system whilst solely relying on the data collected by observing the movements of the mouse. Up until recently mouse based verification systems were implemented using neural networks which are computationally very heavy and therefore degraded the performance of the machine they were running on and consumed more resources then a background process normally should.

The method demonstrated by Zheng et al. [17] uses support vector machines (SVM) [6] that have already been successfully used in a method for recognizing handwritten digits [15], which also fall under the category of behavioral biometrics. One of the key features of it is the small number of user actions that are required in order to identify a user.

4. Future direction

In the method proposed in [17] authors chose to track only mouse movements which ended with a mouse click and discarded all the others, although they also ran tests which recorded partial movements as well. This approach limited their ability to collect enough data in a short time and they have stated that the average time to collect 20 clicks was 15 minutes. If the method was extended to include more complex user actions in line with the method presented by Feher et al. [3]. Table 2 shows the basic mouse actions proposed by [3] and the number of features that are used to characterize each action. Table 3. shows higher level user actions composed of two or more basic mouse actions which are used in the process of user verification. Both tables are taken from [3].

Table 2. Basic mouse actions as defined by [3]

Basic action	Description
Mouse-move Event (m)	occurs when the user moves the mouse
Mouse Left Button Down Event (ld)	occurs when the left button is clicked
Mouse Right Button Down Event (rd)	occurs when the right button is clicked
Mouse Left Button Up Event (lu)	occurs after the left button is released
Mouse Right Button Up Event (ru)	occurs after the right button is released

Table 3. Proposed user actions used for verification by [3]

Action	Number of features
Left Click (LC)	2
Right Click (RC)	2
Drag and Drop (DD)	66
Double Click (DC)	6
Mouse Move and Left or Right Click Action (MM_LC)	70
Mouse Move and Double Click Action (MM_DC)	74
Mouse Move and Drag and Drop Action (MM_DD)	134

The method proposed by [17] records only Mouse-move Event followed by Right Button Down Event and does not differentiate between left click and right click. Authors also conducted tests which included Mouse-move Events as well, which increased the methods effectiveness. However they discouraged using that parameter since Mouse-move Events are frequently generated by a user idly moving a mouse to stop a screensaver from appearing, moving the mouse out of their way or moving the mouse by accident. Most Mouse-move Events have do not carry a definitive decision to act as opposed to Mouse-move Events followed by Mouse Left/Right Button Down events and would therefore introduce “noise” into the decision making system.

The most obvious mouse actions that could be added to the list of recorded parameters would be double click speed which consists of 4 basic

mouse actions conducted in quick succession. Research into user identification based on keystroke dynamics has already shown [3] that users can be identify by measuring the time between keystrokes and that can also apply to the time between two mouse clicks. It is also possible to further atomize the actions and measure the time between pressing the mouse button and releasing it. With continuous increase in computing power it can be expected that adding these additional parameters will not significantly increase load on the system but could significantly increase the effectiveness by lowering FAR and FRR.

5. Conclusion

This paper presents the current methods of verifying user identity by monitoring his mouse movements. The paper focuses on mouse verification methods which would be used in conjunction with password based identification methods to provide an extra layer of security especially in office environments where there are commonly a large number of workstations which can be left unattended for various reasons. One of today's problems is the time it takes to verify the identity of a user which, on average, exceeds 2 minutes. This problem stems from the fact that current methods require a large number of mouse actions to make a decision about the identity of the user. The paper focuses on two recently presented methods [3, 17], and proposes a merger of techniques employed by them in order to achieve greater effectiveness. The method proposed by [17] reduces the number of mouse actions that are required to identify the user but at the same time records only a single type of action which is mouse movement followed by a single click. If the method were to be extended to include other, more finely granulated mouse actions such as the ones proposed by [3] the sufficient number of mouse actions could be acquired in a much shorter time. Further research in that direction could result in further lowering the FAR and FRR even though it is unlikely that it would alone be enough to achieve the European Standard for Access Control Systems requirements.

6. Acknowledgements

This work was supported by the Croatian Agency for Science and Higher Education. This publication reflects the author's views.

7. References

[1] Ahmed A.A.E, Traore I, A new biometric technology based on mouse dynamics, IEEE Transactions on Dependable and Secure Computing 4, (2007), p. 165-179.

[2] Bours P, Fullu C.J, A login system using mouse dynamics, Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, (2009), p. 1072-1077.

[3] Feher C, Elovici Y, Moskovitch R, Rokach L, Schclar A, user Identity verification via mouse dynamics, Information Sciences, Volume 201, (2012), p. 19-36.

[4] Gaines R, Lisowski W, Press S, Shapiro N, Authentication by keystroke timing: some preliminary results, Rand Corporation, (1980).

[5] Gamboa H, Fred A, A Behavioral biometric system based on human computer interaction, Proceedings of SPIE 5404, (2004), p. 381-392.

[6] Joachims T, Text categorization with support vector machines: Learning with many relevant features. Proceedings of European Conference on Machine Learning, (1998), p. 137-142.

[7] Joyce R, Gupta G, Identity authorization based on keystroke latencies, ACM 33 (2), (1990), p. 168-176.

[8] Monrose F, Rubin A, Authentication via Keystroke Dynamics, Proceedings of the 4th ACM conference on Computer and communications security, (1997), p. 48-56.

[9] Nakkabi Y, Traore I, Ahmed A.A.E, Improving mouse dynamics biometric performance using variance reduction via extractors with separate features, IEEE Transactions on Systems, Man, and Cybernetics, (2010), p. 1345-1353.

[10] Pusara M, Brodley C.E, User re-authentication via mouse movements. Proceedings of the 2004 ACM, workshop on Visualization and data mining for computer security, (2004), p. 1-8.

[11] Revett K, Jahankhani H, de Magalhes S.T, Santos H.M.D, A survey of user authentication based on mouse dynamics, in: Proceedings of 4th International Conference on Global E-Security, Communications in Computer and Information Science, Volume 12, London, United Kingdom, (2008) London, p. 210-219.

[12] Schneier B, Secrets and Lies: Digital Security in a Networked World, John Wiley & Sons, Inc., (2000).
[13] Schulz D, Mouse curve biometrics, Biometric Consortium Conference, Biometrics Symposium, (2006), p. 1-6.
[14] Summers W.C, Bosworth E, WISICT '04 Proceedings of the winter international symposium on Information and communication technologies, (2004), p. 1-6
[15] Vapnik V, Statistical Learning Theory, Wiley, (2004).
[16] Yu E, Cho S, Keystroke dynamics identity verification—its problems and practical solutions, Computers & Security Volume 23, (2004), p. 428-440.
[17] Zheng N, Paloski A, Wang H, An efficient user verification system via mouse movements, 18th ACM Conference on Computer and Communications Security, (2011), p. 139-150.