A Survey of User Authentication Based on Mouse Dynamics

Kenneth Revett[1], Hamid Jahankhani[2], Sérgio Tenreiro de Magalhães[3], and Henrique M.D. Santos[3]

[1] Harrow School of Computer Science, University of Westminster, London, UK
revettk@westminster.ac.uk
[2] University of East London
Hamid.jahankhani@uel.ac.uk
[3] Universidade do Minho Department of Information SystemsCampus de Azurem
4800-058 Guimaraes, Portugal
{psmagalhaes,hsantos} @dsi.uminho.pt

Abstract. This work surveys biometric based authentication systems that deploy mouse movements. Typically, timing and movement direction, along with clicking actions are used to build a profile of a user, which is then used for authentication purposes. Most system relies on a continuous monitoring process, or require the user to interact with a program (such as a game) in order to derive sufficient statistical information regarding their mouse dynamics. In this work, a novel graphical authentication system dubbed Mouse-lock is presented. This system deploys the analogy of a safe, and the password is entered via the mouse in a graphical equivalent of combination lock. The question is whether this approach elicits sufficient discriminatory information from a relatively minimalist degree of interaction from the user. The preliminary results from a study with six subjects indicates, based on FAR/FRR values, that this is a viable approach.

Keywords: Accot-Zhai steering law, biometrics, Fitts' law, Hick's law, mouse dynamics, mouse-lock.

1 Introduction

This paper describes a relatively new approach to behavioral biometrics that relies on the way a user interacts with their computer using a standard mouse. In most graphical applications, the mouse is the method of choice for program interaction. A substantial amount of human computer interaction literature exists which explores how to arrange the graphical user interface (GUI) such that the user's interaction with the system in maximized with respect to some parameter(s) [1]. One common parameter is the interaction speed – how quickly can a user navigate through the GUI based application? This is the essence of the field of study in experimental psychology termed *interaction ergonomics* [2]. For instance, how quickly can a user position the mouse and click on an application icon? If there is a series of menus that must be navigated, what is the best way to arrange them for maximal throughput? These are fundamental questions in computer ergonomics, which has resulted in the formulation of two

H. Jahankhani, K. Revett, and D. Palmer-Brown (Eds.): ICGeS 2008, CCIS 12, pp. 210–219, 2008.

"laws" within experimental psychology: Fitts's Law and Hick's Law. Essentially, Fitts' law relates the length of time it takes to perform a task with a pointing device such as a mouse. For instance, how long does it take to move the mouse cursor to a particular position on the screen. The original formulation was derived from a 1D perspective, in that the vertical dimension was assumed to be infinite, and only the horizontal dimension was considered in the timing of the task. A 2D version of this model was produced by Accot & Zhai, termed the Accot-Zhai steering law [3]. These laws provided a quantitative estimate of time for task performance, and was used a metric for GUI layout optimization. The findings from these and related studies have by now become an integral part of the way GUI applications are designed.

The HCI research has also provided another law that relates the amount of time for a decision to be made with respect to selection of an entity (such as an icon) using a pointing device, termed Hick's law [4]. This law (sometimes referred to as the Hick-Hyman law) is a model that describes the amount of time it takes for a user to make a decision as a function of the amount of choices one has available to them. This law is applicable to purely graphical systems (as is Fitts's law), and is mentioned here for completeness – though it belongs in both places. In combination with Fitts's law, these HCI based models provide quantitative information regarding the thinking time and the motion time for users to interact with a system using a pointing device. To this author's knowledge, there is no comprehensive study employing these laws in the context of graphical authentication systems.

The application of Fitts's law can be a useful attribute when examining the dynamical aspects of mouse usage, especially in the context of the process of selecting (via a pointing device) the elements of a graphical password. What other attributes are available, akin to those obtainable from keystroke dynamics? Awad and colleagues proposed the average speed against the distance traveled and the average speed against the movement direction [5]. The attributes reported in the literature include clicking (left, middle, or center button) and mouse wheel movements. Gamboa and colleagues refer to the concept of a *stroke*, as the movement of the mouse between two successive clicks [6]. In addition, higher order features such as curvature, angle, and deviation can be acquired and used to capture the dynamical aspects of mouse movements [7]. In an interesting approach to user authentication via mouse movements, Syukri and colleagues asked users to write their signature using a mouse [8]. The next section provides a brief survey of mouse based biometrics.

1.1 Literature Review

Hashia and colleagues published a paper which examined the use of mouse movements as a method for user authentication [9]. The users enrolled in a static fashion, which was immediately followed by an authentication phase. For the enrollment process, the users were required to position the mouse pointer over a set of 10 dots that appeared in random positions on the screen (though for each enrollment trial, the dots appeared at the same position). The initial data that was captured was the movement of the mouse as the user positioned the mouse pointer over the dots (target areas). The x-y coordinate position of the mouse was captured every 50 ms, and stored for subsequent on-line analysis. The data was analyzed with respect to speed, deviation from a straight line, and the angle of movement. Deviation was measured as the

perpendicular distance from the point where the mouse is currently positioned to the line formed between the two points between which the mouse is moving. The angle was calculated by the angle between three points, which was either positive (0 to 180 degrees) or negative (0 to –180 degrees). The user was required to move 20 times over the 10 points to complete their registration/enrollment. Between each pair of points, the average, min, max, and standard deviation of the four attributes were recorded. This yields a total of 16 values for each point-pair, and a total of 144 attributes measures (16 attributes for each of the 9 pairs of points), which was stored as the BIR for each user. Note that all the data was normalized before being stored in the BIR. For the verification step, the same attributes are extracted and compared to the BIR for that user. As the user moves through the collection of dots, the system checks to see if the attribute values are within 1.5 standard deviations from the BIR values. If so, a counter is incremented, resulting in a score that is used for authentication purposes.

Hashia and colleagues tested their system on a set of 15 students, aged 22-30, in a university setting, employing the same mouse and mouse pad (associated with the same PC). The authors reported an EER of 20%. The authors also investigated using the average and standard deviation of the resulting counter values (instead of the range), which reduced the EER to 15%. These results are encouraging, but much higher than the EER obtained from keystroke dynamics, and graphical password authentication results. In addition, these authors report a continuous monitoring system, to validate the currently logged in user. During the enrollment process, mouse movements through a background process for a 15-minute block. The process calculates where the highest density of mouse movements are, and draws a convex hull around them. These are called states, and the system keeps track of movements between states, much like in the static enrollment, motion between the dots is recorded. The system calculates the speed (average and standard deviation), the average and the standard deviation of wavering. In addition, they store the transition state (each assigned a unique number), a count of how many times each state is visited in the 15-minute enrollment time, average speed, and wavering. During user verification, data is collected every two minutes and compared with the BIR data. If the speed and wavering is within 1.5 SD units found in the BIR, the user is positively verified. Otherwise, the user can be locked out of the system.

Gamboa and colleagues have produced some interesting results deploying the use of mouse movements for user authentication [10]. The approach produces a large number of interesting attributes, which are tuned to each user by a greedy search algorithm, and deploys a unimodal parametric model for authentication. The authors employ the use of a web based data acquisition system, built around the WIDAM (Web Interaction Display and Monitoring) system, that records all user interactions and stores the data in a file. The interactions employed in this system are related to mouse movements produced during the interaction with a graphical interactive program (a memory game). The raw data that is recorded is the <x,y> coordinates of the mouse pointer position, mouse clicks, and timing information associated with these activities. The authors use the concept of a stroke (analogous to that used in the Draw-a-Secret scheme), to represent the information (the <x,y> coordinates) contained between two successive mouse clicks. The authors employ a multi-stage processing schema in order to generate the data that will be used by the classifier. First, the data is cleansed by

applying a cubic spline, which smooths out inconsistencies/irregularities. Next, they extract spatial and temporal information (the input vector), and lastly, they apply a statistical model to extract salient features from the data.

With a 63-dimensional input vector available for each stroke, the feature vector was reduced based on how each attribute influenced the EER. Essentially, the authors employed a greedy search algorithm (Sequential Forward Search) algorithm, to examine sequentially which attributes produced the lowest EER. This is performed in a bootstrap like fashion, where each attribute was examined in isolation, the one producing the lowest EER was selected, and then the other attributes were examined sequentially, selecting those that also produced the lowest value for the EER. This process was repeated for all attributes for each user, producing a local set of attributes that best characterized each user in terms of a minimized value for EER.

The results indicate that the classification accuracy (as measured by the EER) was dependent upon the number of strokes included in the decision model. For instance, when including 50 strokes, the EER was 1.3%, very comparable to results obtained from physiological based techniques such as fingerprints and hand geometry [11]. Further, the authors also examined the EER as a function of the surveillance time. The results indicated that a 90-second surveillance time provides an EER on the order of 1:200, comparable to physiological based biometrics such as hand geometry and related technologies [12]. Whether these results can be improved upon is a matter for future research.

In a study by Pusara and Brodley, cursor movements and mouse dynamics was examined in order to determine whether these attributes would be suitable for user re-authentication [6]. Re-authentication is a technique suitable for the detection of a hijacking scenario (e.g. someone has replaced the originally logged in user). The data was collected from eighteen subjects working within Internet Explorer. After data collection, a detailed user profile was created for each user, tailored to the way each interacted with the application. Subsequent to model development, a supervised learning approach was used for the purpose of classifying data into authentic or imposter users.

In order to build a profile of user identity based on mouse dynamics, attributes representing discriminatory behavior were collected during an enrollment/training phase. The attributes collected were cursor movements, mouse wheel movements, and clicks (left, middle, and center). The two-dimensional coordinates of the current mouse position were recorded at 100 msec intervals. From this raw data, secondary attributes such as *distance*, *angle*, and *speed* between pairs of points were recorded. Note that the pairs may be consecutive or separate by some number of data points, k, termed by the authors to be a frequency measure. The mean, standard deviations, and third moment values were calculated over a window N of data points. In addition, all data points are time stamped. The data is constructed into a hierarchical form, in order to create a template onto which a user profile can be generated. At the top of this hierarchy is the sequence of mouse events for a given user. Next are the clicks, non-click moves, and mouse wheel events, followed by single or double click events. The same statistical measures are again applied to the N data points. This results in a feature set that represents the ensemble behavior of a windowed version of the original raw data.

From this hierarchical model of raw data, a set of features is extracted and used for the authentication/identification task. For each category in the hierarchy, there are six

features, corresponding to each of the categories in the hierarchy (e.g. wheel movements, clicks, etc.). The mean, standard deviation, and third moment of distance, angle, and speed between pairs of points are measured, resulting in 63 features. An additional 42 features were derived from statistics on the X and Y coordinates of cursor movement data. With the feature vector obtained from the user data, the authors consider performing a supervised versus an unsupervised classification strategy. The authors opted to employ a supervised approach to classification in this study. In this approach, the profile obtained for the current user must be matched to one of the user's models contained within the BIR database. The classification system is applied to a windowed dataset, and each point in the window is evaluated and raises an alarm if the value of the datapoint is not consistent with the user's profile. If a threshold number of alarms are indicated, the user is flagged as an imposter.

The author's evaluated their system on a set of eighteen student volunteers, whom provided data that consisted of 10,000 cursor locations. The data was collected from the use of Internet Explorer from a Windows based PC (this resulted in 7,635 unique cursor locations for the IE interactions alone). The data for each user was split into quarters, the first two used for training purposes, the 3rd quarter used for parameter selection, and the remaining was used for testing purposes. The authors employed a decision tree classification approach (C5.0). Essentially, the algorithm must find splits within the data based on an information gain ratio – which is an entropy measuring approach. In their first experiment, the authors attempted to build a decision tree classifier that could distinguish between pairs of users. The results indicated that the classification accuracy (based on minimizing EER) was highly dependent upon the characteristics of the user. For instance, if there were a significant number of events recorded within the window, the accuracy was increased. The overall results from this study indicate that the system was able to distinguish between pairs of users with considerable accuracy (with an accuracy of 100% for some pairs of users). In their next study, they examined whether the system could distinguish each user from all other users of the system. The results generated an average false negative rate was 3.06% for all 18 users. The corresponding false positive rate was 27.5%.

Ahmed and colleagues have employed mouse dynamics as an approach for intruder based detection [13], [5]. The authors measure several attributes with respect to the user's usage when interacting with a graphical based application such as general mouse movement, drag and drop behavior, point and click behavior, and silence. Using a variety of machine learning techniques, the authors develop a *mouse dynamics signature* (MDS) for each user. In this study, mouse usage data was collected from 22 participants over a 9-week period. The data that was collected was used in an off-line approach to evaluate their detection system. The detection system relied on the use of average mouse speed (against overall speed and direction), drag and drop statistics, mouse movement statistics, and point and click statistics. The users were separated into two groups, where one group (consisting of ten participants) represented authorized users and the remaining participants acted as imposters. These features were used to train a neural network to classify users of the system into genuine and imposters. To calculate the FRR, the first half of the sessions were used as reference values, and the remainder were used to test the system. The FRR was approximately 1.3% for this study. The FAR was calculated by allowing the other users of the system to act as

imposters, an yielded an average FAR of 0.65%. This approach – though applied to intruder detection, could just as easily been used to authenticate a user – or perform continuous user authentication. The next section provides the methodology employed in a novel mouse driven authentication method, which is based on the combination safe analogy.

2 Methods

In this paper, a novel graphical authentication system that is based on the analogy of entering a combination to a safe, dubbed *Mouse lock*. In this system, the numbers of the safe dial are replaced with graphic thumbnail images. The system is depicted in Figure 1. A with opening a typical safe, the user is required to select the correct position (indicated by the appropriate graphical element) and move it to the top dial indicator in the appropriate direction. A password is a combination of images and associated with each image is the direction to move direction to move the image to the top. For instance: <image 8>L, <image 2>R, <image 5> L, and <image 12>R would constitute a user's password. The direction of movement (L = left or R = right) is indicated by clicking the left and right mouse button respectively. The user clicks on the appropriate image, then clicks the correct direction button, which then places the image under the dial. Note that this process actually moves the images along the circular dial. It could also be implemented without moving the images – but in this study, the images were moved during password entry – whether the correct entry was made or not. An audible sound is produced when the user enters the correct combination (password). In the enrollment process, audible feedback is provided indicating whether or not the correct entry was made – but not when users were attempting to authenticate in the testing phase.

After a user selected a login ID (keyboard based entry), the users were required to select a password that consisted of five images. The direction in which the entries were to be moved to line up with the top fiduciary mark was randomly selected by the system. During enrollment – the password was displayed on screen for the user – and they were allowed to practise until they had entered their password correctly five times (not necessarily contiguously). During authentication, no feedback (audible nor on-screen) was provided to the user, and they were allowed a maximum of three failed attempts before the entry was considered invalid. The study group consisted of six subjects (university students aged 20-28, mean age 22). The graphical "lock" consisted of 24 small iconic images (each 15x15 mm) arranged in a circular fashion (see Figure 1), with the position indicator located at the top of the circle. After successfully enrolling into the system (note that the FTE was 0%), each user logged into their accounts 100 times, in order to generate FRR data, and each logged into the other accounts 20 times, generating 100 samples to generate FAR data. The FAR/FRR results from the six subjects are presented in Table 1.

Reference values (derived from the enrollment) from the test subjects were stored within the BIR for the selected user ID. When a user attempts to log into the system, first the accuracy of the "combination" is compared against the information associated with the specified login ID. If this is a match, then the timing information is acquired.

Initially, only the "digraphs" - in this case, the time between selecting successive selections, and the total time to enter the elements of the password were used for authentication purposes. If the measurements of these timing factors matched those stored in the BIR for that account, the user was authenticated. A distance metric based on the mean +/- 1.5 SD was used to determine whether each measure was successful. All timing metrics (including the total time) must pass the distance metric to be considered a successful attempt. Also note that with practise, the total time (and hence one or more of the digraphs) was reduced relative to the initial authentication attempt (the first attempt after the enrollment period). This is reflected in the Accot-Zhai steering law effect, an extension of Fitts' law of practise. This trend was detected for all users in this study, and probably reflects the familiarity associated with repeated performance of the same task.

In this study, the category effect of the thumbnail icons was also examined. In figure 1, the category selected by the users (as was the case for the initial results reported here) was constant: cars. We also experimented with various categories such as animals, people's faces, and outdoor scenes. In addition, heterogeneous thumbnails were also employed, with randomly selected mixtures from each of the four categories. The purpose was to determine whether the composition of the images would affect the fidelity and speed with which users would enter their combination. Table 2 presents the total time taken for each of the four categories, reported as the average (and SD). With respect to the accuracy of selection, table 3 presents the FAR/FRR from 100 trials.

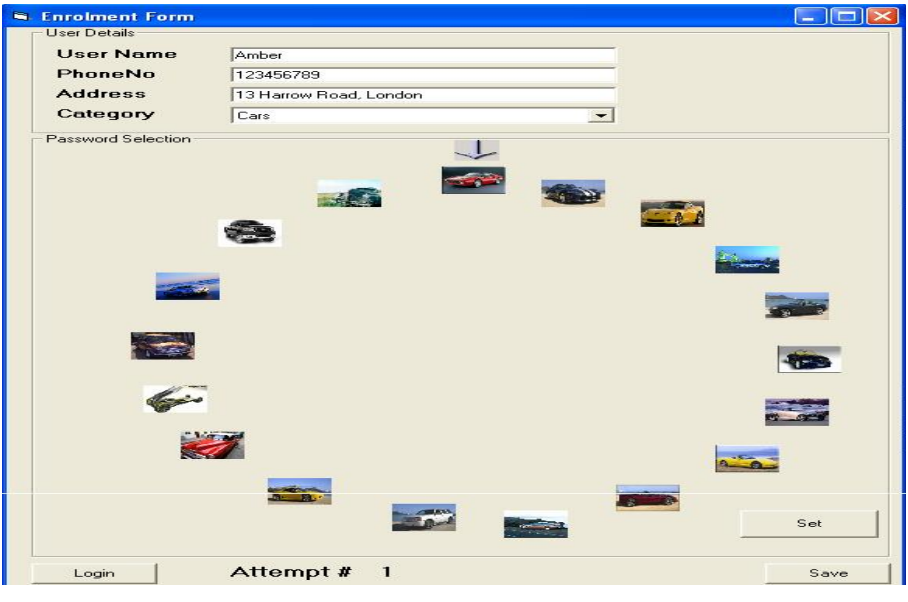

Fig. 1. The *Mouse-lock* system, waiting for user input. This version deploys a collection of vehicles (cars) as the graphical thumbnail elements of the password.

3 Results

Table 1. FRR and FAR results from each of the six users of the Mouse lock system. The FRR data was generated from 100 self-login trials, and the FAR results were acquired from 100 trials (20 attempts from each of the other five subjects). The results are reported as percentages, for each user separately (rounded to whole numbers).

Subject Number	FAR	FRR
1	4%	2%
2	6%	7%
3	2%	6%
4	3%	8%
5	2%	1%
6	4%	0%

Table 2 Total input time as a function of thumbnail categories. The users were to enter a five-image combination in all cases, and were allowed to practise until successful on five attempts (non-contiguous was allowed). These results are derived from an average of 100 attempts per user for each of the five categories, and the parenthetical values are the standard deviations.

Category	Total input time (seconds)
Cars	6.7 (2.8)
Faces	7.1 (3.9)
Animals	8.8 (4.2)
Outdoor scenes	7.8 (3.5)
Random mixture	9.3 (4.6)

Table 3. Summary of the FAR/FRR results according to image category (the same categories used in generating the data presented in table 2). The FRR data was generated from 100 self-login trials, and the FAR results were acquired from 100 trials (20 attempts from each of the other five subjects). The results are reported as percentages, for each user separately (rounded to whole numbers). Note that these results were derived from the entire subject cohort (six users).

Category	FAR	FRR
Cars	3%	2%
Faces	1%	1%
Animals	2%	3%
Outdoor scenes	4%	2%
Random mixture	1%	3%

4 Discussion

The results from the studies presented in this chapter indicate first and foremost that mouse dynamics can be a very effective means of authentication. With FAR and FRR values of approximately 2-5%, the data has provided evidence that *the way* users

interact with computer systems contains sufficient discriminatory power to distinguish authentic users from imposters. What features of a user's interaction with a pointing device – essentially we have mouse movements and clicks? The studies of Gamboa and Pusara indicate that a substantial number of secondary attributes can be extracted. These attributes include the speed, distance, and angles made when moving the mouse within some given time interval. Depending on the temporal resolution, a substantial amount of data can be extracted for classification purposes. In addition statistical information is typically extracted from the raw data to provide higher order feature vectors to enhance the discriminatory capability of the classifier(s) employed.

In this study, only the time between selecting the appropriate images and total selection times were used during the authentication process. Of course the application was very different from that of other studies examined in this paper. The reason for the low FAR/FRR results in this study may reflect the use of the safe analogy, which was quite familiar to all subjects employed in this study. The limited scope of the task involved in authentication in mouse-lock may limit the number of attributes that can be extracted for authentication purposes. One mechanism for exploring the attribute space – and augmenting the password space is to increase the number of thumbnail icons – though there is an upper limit, which is influenced by the device used for authentication (e.g. mobile phone versus a desktop PC). In addition, a larger password could be imposed – though the upper limit should be approximately eight thumbnail images as is the case for textual based passwords. With an eight-image password, the number of combinations approaches 735,471 - a considerable but not insurmountable password space. As with most graphical based authentication systems, it is a difficult task to emulate thumbnail entries remotely - and this value is well above the space offered by more conventional graphical authentication systems.

References

1. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology 47(6), 381–391 (1954)
2. Meyer, D.E., Smith, J.E.K., Kornblum, S., Abrams, R.A., Wright, C.E.: Speed-accuracy tradeoffs in aimed movements: Toward a theory of rapid voluntary action. In: Jeannerod, M. (ed.) Attention and performance XIII, pp. 173–226. Lawrence Erlbaum, Hillsdale (1990), http://www.umich.edu/~bcalab/Meyer_Bibliography.html
3. Accot, J., Zhai, S.: Refining Fitts law models for bivariate pointing. In: Proceedings of ACM CHI 2003 Conference on Human Factors in Computing Systems, pp. 193–200 (2003)
4. Hick, W.E.: On the rate of gain of information. Quarterly Journal of Experimental Psychology 4, 11–26 (1952)
5. Ahmed, A.E., Traore, I.: A New Biometric Technology Based on Mouse Dynamics. IEEE Transactions on Dependable and Secure Computing 4(3), 165–179 (2007)
6. Pusara, M., Brodley, C.E.: User re-authentication via mouse movements (2003)
7. Gamboa, H., Fred, A.: An identity authentication system based on human computer interaction behaviour. In: Proceedings of the 3rd International Workshop on pattern recognition on Information Systems, pp. 46–55 (2003)

8. Syukri, A.F., Okamoto, E., Mambo, M.: A User identification System using Signature Written with Mouse. In: Boyd, C., Dawson, E. (eds.) ACISP 1998. LNCS, vol. 1438, pp. 403–414. Springer, Heidelberg (1998)
9. Hashia, S., Pollett, C., Stamp, M.: On using mouse movements as a biometric. In: Perales, F.J., Campilho, A.C., Pérez, N., Sanfeliu, A. (eds.) IbPRIA 2003. LNCS, vol. 2652, pp. 246–254. Springer, Heidelberg (2003)
10. Gamboa, H., Ferreira, V.: WIDAM – Web Interaction Display and Monitoring. In: Quinlan, R. (ed.) The Proceedings of the 5th International Conference on Enterprise Information Systems. Data mining tools See5 and C5.0 (2003)
11. Jain, R.B., Pankanti, S.: Introduction to Biometrics. In: Jain, A., Bolle, R., Pankanti, S. (eds.) Biometrics. Personal Identification in Networked Society, pp. 1–41. Kluwer Academic Publishers, Dordrecht (2003)
12. Maio, D., Maltoni, D., Cappelli, R., Wayman, J.L., Jain, A.K.: FVC 2004: Third fingerprint verification competition. In: Proceedings of International Conference on Biometric Authentication, Hong Kong, China, pp. 1–7 (July 2004)
13. Ahmed, A.A.E., Traore, I.: Anomaly Intrusion Detection based on Biometrics. In: Proceedings of the 2005 IEEE Workshop on Information Assurance and Security, United States Military Academy, West Point, NY, pp. 452–453 (2005)