A User Authentication Technic Using a Web Interaction Monitoring System

Hugo Gamboa[1] and Ana Fred[2]

[1] Escola Superior de Tecnologia de Setúbal,
Campo do IPS, Estefanilha, 2914-508 Setúbal, Portugal
hgamboa@est.ips.pt
[2] Instituto de Telecomunicações Instituto Superior Técnico
IST - Torre Norte, Piso 10,
Av. Rovisco Pais 1049-001 Lisboa Portugal
afred@lx.it.pt

Abstract. User authentication based on biometrics has explored both physiological and behavioral characteristics. We present a system, called Web Interaction Display and Monitoring (WIDAM), that captures an user interaction on the web via a pointing device. This forms the basis of a new authentication system that uses behavioral information extracted from these interaction signals. The user interaction logs produced by WIDAM are presented to a sequential classifier, that applies statistical pattern recognition techniques to ascertain the identity of an individual - authentication system. The overall performance of the combined acquisition / authentication systems is measured by the global equal error rate, estimated from a test set. Preliminary results show that the new technique is a promising tool for user authentication, exhibiting comparable performances to other behavioural biometric techniques. Exploring standard human-computer interaction devices, and enabling remote access to behavioural information, this system constitutes an inexpensive and practical approach to user authentication through the world wide web.

1 Introduction

Personal identification / authentication plays an important role in current security and personalization systems. As opposed to traditional security systems, that based authentication on something *one has* or on something *one knows* (magnetic card, keys, etc. in the first case and passwords or personal identification numbers in the second), recent methodologies explore biometric characteristics. These methods are based on something *one is*, leading to increased reliability and immunity to authorization theft, loss or lent.

We can divide the biometric systems in two types [9]: (1) Identity verification (or authentication) occurs when a user claims who he is and the system accepts (or declines) his claim; (2) Identity identification (sometimes called search) occurs when the system establishes a subject identity (or fails to do it) without any prior claim.

Biometric techniques can also be classified according to the type of characteristics explored : (1) physiological — a physiological trait tends to be a stable physical characteristic, such as finger print, hand silhouette, blood vessel pattern in the hand, face or back of the eye. (2) behavioural — a behavioural characteristic is a reflection of an individual's psychology. Because of the variability over time of most behavioural characteristics, behavioural biometric systems need to be designed to be more dynamic and accept some degree of variability. On the other hand, behavioural biometrics are associated with less intrusive systems, leading to better acceptability by the users. Two examples of behavioural biometric techniques presently used are handwritten signature verification [6] and speaker recognition via voice prints [2].

The evaluation of a biometric technique requires the definition of metrics that can be used for the comparison of performance among different techniques [10], typically: False rejection rate (FRR) — rate of accesses where a legitim user is rejected by the system; False acceptance rate — rate of accesses where an impostor is accepted by the system; Equal error rate (EER) — the value at which FAR and FRR are equal.

In this paper we propose both a web based user interaction monitoring system called Web Interaction Display and Monitoring, WIDAM, and a new behavioural biometric technique based on web interaction via a pointing device, typically a mouse pointer. The normal interaction through this device is analyzed for extraction of behavioural information in order to link an identification claim to an individual.

In the following section we present the user interaction acquisition system, WIDAM. In section 3 we describe the authentication system, focusing on the sequential classifier. Section 4 presents experimental results obtained using the collected data. Conclusions are presented in section 5.

2 The Acquisition System

The acquisition system, WIDAM, (this system is presented with more detail in [4]) enables the user interaction monitoring, analysis and display on web pages. The system can be called as a remote display system that enables the synchronous and asynchronous observation of the user interaction, offering four different services : (1) Synchronous Monitoring Service — real-time monitoring of the user interaction; (2) Synchronous Display Service — real-time observation by other users; (3) Recording Service — storage of the user interaction information in the server database; (4) Playback Service — retrieval and playback of a stored monitored interaction.

WIDAM allows the usage of an interaction recording system directly over a web page, based on the Document Object Model [7] (DOM) of the web page. The system works in a normal web browser with java and javascript capabilities, without the need of any software installation. WIDAM is a light weight networked application using low bandwidth comparatively to image based remote display systems.

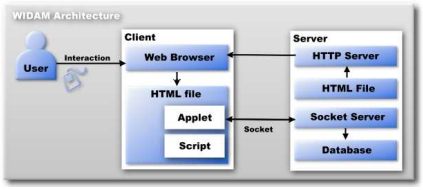

Fig. 1. The WIDAM Architecture.

The WIDAM Architecture is composed by a client and server applications, as depicted in figure 1. The user accesses the WIDAM application via a web browser that connects to the server. Then, the server sends back to the user a web page that is capable of monitoring and displaying the user interaction. This page creates a connection to the server and selects one of the services provided by WIDAM. Then the client and the server exchange messages using a specific protocol.

The client works in any web browser capable of executing Javascript code and Java Applets, independent of the operating system. When the users enters into a page of the WIDAM system, an applet is launched. This applet creates a socket connection that enables the message passing from, and to the server. The client loads the html page and sends an handshaking message through the open socket, specifying which type of service is requested.

In the case of a Recording Service or Synchronous Monitoring Service, the script sends a request to the browser, asking for notification of the user interface events (a sub set of the events from the Document Object Model Events [11] listed in table 1).

In the case of a Synchronous Display Service or Playback Service, the web browser creates a virtual mouse pointer and waits for messages from the server specifying which event should be emulated in the web browser.

ID	Event handler	Event cause
0	onMouseMove	The user moves the cursor.
1	onMouseDown	The user presses a mouse button.
2	onKeyPress	The user presses a key.
3	onUnload	The user exits a document.
4	onMove	The window is moved.
5	onSelect	The user selects some text.
6	onResize	The window is resized.
7	onBlur	The window loses focus.
8	onFocus	The window receives focus.

Table 1. DOM events captured by WIDAM.

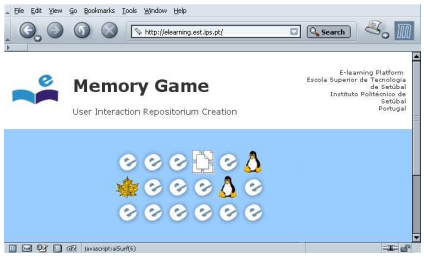

Fig. 2. Interaction test page: the memory game.

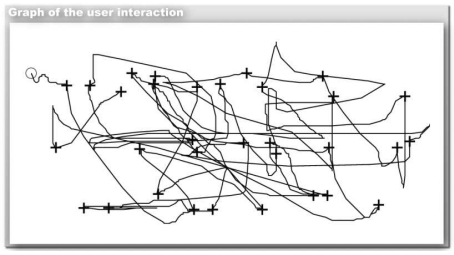

Fig. 3. Graph of the user interaction in the memory game.

For the purpose of the authentication technique being developed, the WIDAM system operated in the Recording Service mode, over a web page with the memory game: a grid of tiles, each tile having associated a hidden pattern, which is shown for a brief period of time upon clicking on it; the purpose of the game is to identify the matching tiles. The WIDAM system presents a web page to the user, asking for his identification (name, and a personal number). Then the system presents an *interaction acquisition page* with the memory game (that could be any html web page), depicted in figure 2. This page is monitored by the WIDAM application that records all the user interaction in a file stored in the web server. Figure 3 shows a graph of a user interaction while playing an entire memory game. The graph is produced by joining every sequential mouse movement with lines and using a cross mark to indicate a mouse click.

3 The Authentication System

An experimental system — the authentication system — was developed to verify the possibility of discriminating between users using their computer interaction information, specifically based on mouse movements performed between successive clicks, which we will call a *stroke* (see figure 4).

Figure 5 presents the acquisition and recognition systems and its respective building blocks. The acquisition system was addressed in the previous section. The recognition system comprises the following modules: (a) feature extraction; (b) feature selection; (c) parametrical learning; (d) statistical sequential classifier.

The recognition system reads the interaction data from the stored data files produced by the acquisition system. The interaction data passes a feature extraction procedure, creating a 63-dimensional vector, exploring both spatial (related to angle and curvature) and temporal (related to duration, position, velocity and acceleration) characteristics of the strokes. More details can be found in [5].

The system has an enrolment phase, where the global set of extracted features are used in an algorithm that selects a set of "best" features for each user, using the equal error rate as performance measure (feature selection block in figure

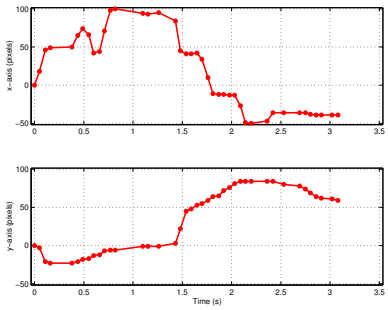

Fig. 4. Example of a stroke — input signals generated by the mouse move events between successive mouse clicks.

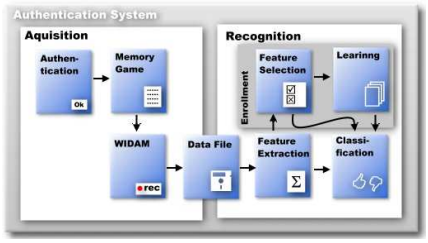

Fig. 5. Authentication system architecture.

5), using the Sequential Forward Selection (SFS) [8] that selects the best single feature and then adds one feature at time to a the vector of previously selected features. The algorithm stops when the equal error rate does not decrease.

The classification rule assumes a statistical model for the feature vectors. The learning phase consists of the estimation of the probability density functions, $p(X)$ (where X is the feature vector of a stroke), from each user's data. Considering that each user constitutes a recognition class, and assuming statistical independence between features, $p(X)$ factorizes into $p(X|\ user) = \prod p(x_i|\ user)$. We use as parametrical model for $p(x_i|\ user)$ the *weibull* [1] distribution ($p(x|a,b) = abx^{(b-1)}e^{(-ax^b)}$). Given the data from one user and one feature, maximum likelihood estimates of the parameters a and b are obtained.

The classifier's purpose is to decide if a user is who he claims to be, based on the patterns of interaction with the computer. We consider that the i^{th} user is denoted by the class w_i, $\ i = 1, \ldots, L$, and L is the number of classes. As defined before, a feature vector is associated with one stroke. Given a sequence of n_s consecutive strokes executed by the user, w_i, interaction information is summarized in the vector $\mathbf{X} = X^1 ... X^{n_s}$, consisting of the concatenation of the feature vectors associated with each stroke. $X^j = x_1^j ... x_{n_{f_i}}^j$, the feature vector representing the jth stroke, has n_{f_i} elements, n_{f_i} being the number of features identified for user w_i in the feature selection phase.

Considering each stroke at a time, and assuming statistical independence between features, we can write $p(X_j|w_i) = \prod_{l=1}^{n_f} p(x_l^j|w_i)$. Considering stroke independence we can further write $p(\mathbf{X}|w_i) = \prod_{j=1}^{n_s} p(X_j|w_i)$.

The classifier will decide to accept or reject the claimed identity based on two distributions: the genuine distribution $p(\mathbf{X}|w_i)$, and the impostor distribution $p(\mathbf{X}|\ \overline{w_i})$ that is based on a mixture of distributions (weibull distributions), one for each other user not equal to i, expressed as $p(\mathbf{X}|\overline{w_i}) = \sum_{j \neq i} p(\mathbf{X}|w_i)\frac{1}{L}$. In the previous equation we assume that the classes are equiprobable, $p(w_i) =$

$1/L \quad i = 1...L$. We can therefore express the posterior probability function as $p(w_i|\mathbf{X}) = \frac{p(\mathbf{X}|w_i)}{\sum_{k=1}^{L} p(\mathbf{X}|w_k)} = 1 - p(\overline{w_i}|\mathbf{X})$.

Since $p(w_i|X_j)$ represents an estimate of the probability of the classification being correct, we establish a *limit*, λ, to select one of the decisions, using the decision rule in equation 1. To present result about the classifier performance we adjust λ to operate in the equal error rate point.

$$Accept(\mathbf{X} \in w_i) = \begin{cases} true & \text{if } p(w_i|\mathbf{X}) > \lambda \\ false & \text{otherwise} \end{cases} \tag{1}$$

4 Results

We asked 25 volunteers (engineering students) to use the developed system, playing several memory games during about 10-15 minutes. This way, we created an interaction repository of approximately 5 hours of interaction, providing more than 180 strokes per user. The acquisition system monitors the pointing device with a sample rate of 50 times per second, producing messages form the client to the server that require approximately 1 Kbytes/s (950 bytes per second) as the maximum bandwidth. For instance, the five hours of interaction occupies 18 Mbytes of disk space.

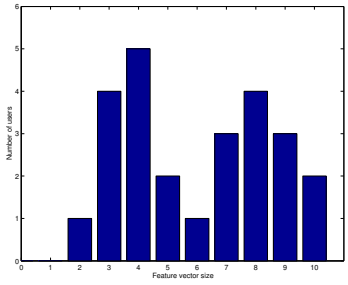

Fig. 6. User feature vectors size histogram

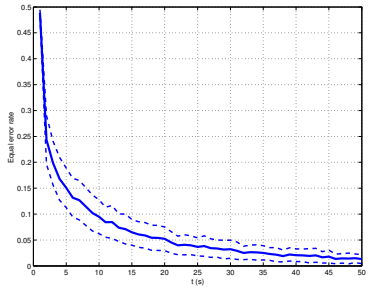

Fig. 7. Equal error rate results of the verification system. The solid line is the mean of the equal error rate of all users. The dashed lines are the mean plus and minus half standard deviation.

In order to use the same number of strokes per user in the tests performed, we randomly selected 180 strokes from each user. The set of strokes was divided into two equal parts, one for the training phase and other for the testing phase. Using the training set we learnt the parametrical distribution $p(x_i|user)$ for each user and each feature. Feature selection used the same data set and was tuned

l	eer	std
1	0.489	0.01
2	0.243	0.09
5	0.151	0.07
10	0.095	0.06
20	0.052	0.04
50	0.013	0.02
100	0.005	0.001

Table 2. Mean equal error rate (eer) and respective standard deviation (std) for different stroke sequence lengths (l).

Biometric technique	Equal error rate
Retinal Scan	1:10 000 000
Iris Scan	1:131 000
Fingerprints	1:500
Hand Geometry	1:500
Signature Dynamics	1:50
Voice Dynamics	1:50
30s of User Interaction	1:50
60s of User Interaction	1:100
90s of User Interaction	1:200

Table 3. Comparison between several biometric techniques

for each user, based on the performance of the system using sequences of 10 strokes. Figure 6 presents the histogram of the feature vector sizes for all the users; the average size of the feature vector is 6.

When testing the system for one user, we consider an imposter as one of the other users. The test function returns the equal error rate given N sequences of strokes of length l using the classifier tuned for user i. The input sequence of strokes of a test is composed of $N/2$ strokes randomly sampled from the testing set of the user, and $N/2$ strokes randomly sampled from the testing sets of all the other users.

One of the free variables of the system is the number of strokes that the system will use in the verification task. Bootstrap [3] estimates of the system performance as a function of the sequence stroke length (from 1 to 100 strokes) was obtained using 10000 bootstrap samples from the test set. The mean duration of a stroke is approximately 1 second. In table 2 we present the mean results of the equal error rate for all 25 users for several stroke sequence lengths. A graphical display of these results is shown in figure 7. As shown, the mean value and the standard deviation of the EER progressively tends to zero as more strokes are added to the decision rule. This illustrates the refinement of the performance obtained by the sequential classifier.

Table 3 presents EER values reported in the literature for several biometric techniques [12]. Preliminary results show that the proposed user authentication system, based on behavioural information extracted from the interaction with the computer, can achieve comparable performances with other biometric techniques.

5 Conclusion

We have explored the human computer interaction behavioural information to create a novel user behavioural biometric verification technique. For collecting the user interaction through the pointing device movements and clicks in a web

page, we developed a system, WIDAM, working on the world wide web. This system comprises a user interaction acquisition module, responsible for the collection of user interaction data that is capable of synchronous and asynchronous recording and playback of web user interaction activity. The biometric technique is implemented in the authentication system that produces the user classification and estimates of the performance of the decision rule.

This authentication system is based on a sequential statistical classifier that receives the sequential data produced along the user interaction. A large set of features were initially extracted from the collected data, using both time domain related and spatial information from the mouse movement patterns. A feature selection procedure reduced this initial set to a small number of features, using a greedy search, and taking the classifier performance, measured by the EER, as objective function.

The results of the tests with 25 users and a total of 5 hours of interaction showed that this technique is a promising tool for user authentication, considering that the performance results are comparable to some of the behavioural biometric techniques and that it is an inexpensive technique that operates remotely using the human-computer interaction behaviour.

References

1. Robert B. Abernethy. *The New Weibull Handbook*. Robert B. Abernethy, 2000.
2. Mohamed F. BenZeghiba, Hervé Bourlard, and Johnny Mariethoz. Speaker verification based on user-customized password. Technical Report IDIAP-RR 01-13, Institut Dalle Molle d'Intelligence Artificial Perceprive, 2001.
3. Bradley Efron and Robert J. Tibshirani. *An Introduction to the Bootstrap*. Chapman & Hall, 1993.
4. Hugo Gamboa and Vasco Ferreira. WIDAM - Web Interaction Display and Monitoring. In *Proceedings of the* 5^{th} *International Conference on Enterprise Information Systems*, volume 4, pages 21–27, 2003.
5. Hugo Gamboa and Ana Fred. An Identity Authentication System Based On Human Computer Interaction Behaviour. In *Proceedings of the* 3^{rd} *International Workshop on Pattern Recognition in Information Systems*, pages 46–55, 2003.
6. J. Gupta and A. McCabe. A review of dynamic handwritten signature verification. Technical report, James Cook University, Australia, 1997.
7. Arnaud Le Hors, Philippe Le Hgaret, and Lauren Wood. Document object model level 2 core. Technical report, W3C, 2000.
8. Anil K. Jain, Robert P. W. Duin, and Jianchang Mao. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1), 2000.
9. Vaclav Matyas Jr and Zdenek Riha. Biometric authentication systems. Technical report, ECOM-MONITOR, 2000.
10. Tony Mansfield and Gary Roethenbaugh. 1999 glossary of biometric terms. Technical report, Association for Biometrics, 1999.
11. Tom Pixley. Document object model (dom) level 2 events specification. Technical report, W3C, 2000.
12. Thomas Ruggles. Comparison of biometric techniques. Technical report, California Welfare Fraud Prevention System, 2002.