E-authentication using chosen biometric characteristics

Jurica Ševa, Miroslav Bača, Ivan Kljaić
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
{jurica.seva, miroslav.baca, ivan.kljaic}@foi.hr

Abstract

With the development of the Web and the services it provides to the end user one of the aspects one has to take in mind is the need for an account for each of its services. Until now the accounts used a username and password model for the user authentication but with the development of the biometric science and high-speed Internet there are other possibilities that provide a more secure solution. Instead of using username and password we can use behavioral or psychophysical biometrical characteristics to prove if the person is a legal or illegal user. Many today's computers or notebooks have fingerprint readers (usually on mouse) or have some cameras as an addition to standard IO (input/output) devices, keyboard and mouse. Using only these two characteristics a model for so called e-Authentication, based on multimodal biometrics systems, can be developed. This multimodal biometrics system can be supported by soft biometrical characteristics (like eye color or hair color) or some other biometrics system like keystroke dynamics. Possible applications include securing password managers, points of single logon etc. In this paper we described the development of multimodal biometrics system, called eBAWS (e-Biometric Authentication for Web Systems), based on standard IO devices based on keystroke dynamics, shape drawing, shape order and handwriting verification[1].

Keywords. Multimodal biometric system, web site authentication, handwriting, signature, keystroke dynamics, shape drawing, shape order, CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart)

1 Introduction

According to [1] approximately 1.5 billion people are using Internet today with a yearly usage growth of 290 %. Each of the existing as well as new users use one or the other service available on the Web, from mail services, forums and chat rooms to information sharing sites where each of these services needs authentication mechanism that will allow the user to enter and use the service. The most popular and commonly used authentication method is based on the username/password information pair which is generated upon user's registration (and is in general generated based on users choice). But there are some disadvantages to this authentication method. One important disadvantage is that over time a single user has to handle such information for numerous accounts (e.g. I have more than 30 active accounts) and has to find a way to either memorize or (securely!) write down login information for each of the accounts. A common way to avoid this problem is to use the same or similar login information for each newly created account, but this can lead to security level reduction not for one but to more accounts. Other problem regarding username and password pair is the possibility of using an algorithm that, in reasonable time, can break the code and access the account.

Besides security issues with such authentication methods, as the Web evolves and becomes "bigger and better", more and more of personal information accounts will exist (e.g. Internet banking, e-Commerce services, e-Government services) that will need to be sure that the user is who he claims to be so he can access certain web service. One can improve username and password method by building in additional authentication steps and related methods

[1] Shown results came out from scientific project (Methodology of biometrics characteristics evaluation 016-0161199-1721), supported by Ministry of science, education and sport Republic of Croatia

(e.g. CAPTCHA systems) but the overall effect will be two sided: on one hand the security levels will increase but on the other hand the user experience will decrease which can present a problem in the overall acceptance of Internet as a tool in our everyday life.

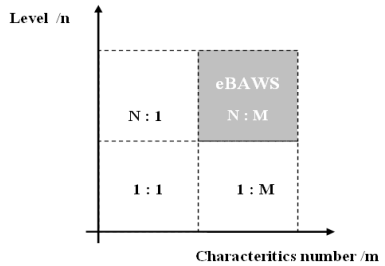
In this point we have to explain the difference between user identification and user authentication. Identification can be described as the process of determining if the person is who he or she says he or she is. In simpler words, it is the process of asserting ones identity. In a computer system this is defined as a 1: N system where user entered information (1) is compared to N database records.

Authentication is the process of confirming someone's identity and is based on some evidence. In computer systems this is defined as a 1:1 system where user entered information is compared to 1 record in the database.

With the use of biometrics, which can be defined as the automatic recognition of a claimed identity using certain physiological or behavioral traits associated with the person [2], there is a way to improve the overall security and also improve the user experience by replacing numerous security check stages with the use of different biometric systems. In these paper we will present a model for a web based biometric authentication system that implements a two level security check based on user input over the standard IO (keyboard and mouse actions). The details will be discussed in the following chapters.

2 E-authentication model

Based on the number of characteristics to be considered we distinguish unimodal and multimodal biometric systems. Unimodal systems take in consideration only one biometric trait from which we extract only one or more biometric structures where as with the multimodal systems we consider multiple biometric characteristics for user verification, as shown in picture 1. The same way it is possible to use multi- level security verification.



Picture 1: Security level and characteristics number relation

For the development of this model numerous biometric traits have been taken into consideration and the most suitable ones have been chosen for authentication verification. The main criteria for accepting or denying a biometric trait were the available computer periphery devices. Since no one can ensure that every user has the same devices, e.g. web cam, microphone, fingerprint reader (although new laptops do come with such devices already intergraded), the focus was put on the standard IO devices which every user has – keyboard and mouse devices. Other methods, although possibly efficient in web environment, have been rejected because of these technology availability constraints. In the following paragraphs is the list of all biometric traits and the reason for their acceptance or rejection for biometric trait capturing.

2.1.Available methods

Currently there are more than 50 biometrical traits taken into consideration for identification and/or authentication a user, a detailed list of which you can find in [6]. Our approach is focused on biometric traits that can be gathered, and afterwards processed, using the before mentioned standard IO devices, as shown in table 1.

The rejected traits have been rejected mostly because the dislocation of the user from the server that runs the needed services, explained in later chapters. Because of that one can't guarantee a live sample (e.g. face picture). Since no one can guarantee that all users have non standard IO devices listed before, verification using such biometric traits can't be implemented. In later stages of development one of the goals is to actively include face and voice recognition systems into online environment.

With the accepted methods as well, once again because of dislocation of the client – server computers, one can't guarantee a live sample (e.g. Internet bot trying to login) but the possibilities can be increased to accepted levels with Internet bot prevention systems (e.g. CAPTCHA system). The other important reason is that everyone can be authorized based on them. Since there are some exceptional situations where keyboard dynamics can't be used (e.g. users hand injuries) the system will offer a different approach in such situations.

2.2.Chosen methods

All the requirements are fulfilled for the method chosen to be integrated in to the model. Since all computers (either desktop or laptop) have a mouse and a keyboard for communication one can use keystroke dynamics and mouse gesture movements for user action tracking and data collection that provide enough information to authorize a user. A more detailed description of the processes will be given in the following subchapters.

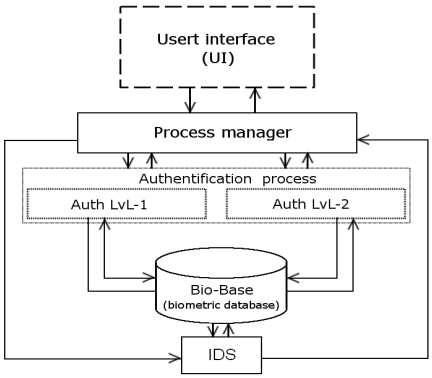
Table 1: Selected biometric traits [6]

Trait	Structure	Structure extraction	Recognition
Keystroke dynamics	Rhythm	Neural networks	Template matching
	Structure		Neural networks
	Timing interval		
	Key press time		
	Data entry method		
Mouse gesture	Mouse key usage and movement specificity	Neural networks	Template matching
	Mouse key usage specificity		Neural networks

In our research we have discovered previous work dealing with specific matters of each of the methods and this model is based on those results.

2.3.System model

The proposed eBAWS system consists of a "Process manager" (PM) module that is monitoring user's actions on the system and accordingly calls specific methods. This system proposes a two level security system based on multimodal biometric characteristics, as shown on picture 2.



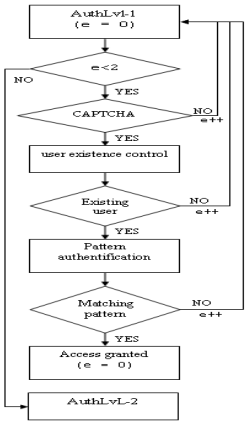
Picture 2: eBAWS system model

If the user wants to login, PM calls AuthLvL-1 method responsible for the user verification. In the case of two unsuccessful login attempts a backup authentication system, AuthLvL-2, is called and the user has a third and last login try. Otherwise, the PM banns the given username for the according IP address. This way we can minimize the effects and efficiency of a possible Botnet attack. A successful login attempt will trigger the underlying IDS to monitor his actions and gather statistical information that will be used for future login attempts. If the IDS suspects a security breach, it generates a call to PM which calls AuthLvL-2 method. AuthLvL-1 and AuthLvL-2 methods are described in following subchapters.

2.4.AuthLvL-1

The first authentication level is the basis for a quick but secure way for a user to prove his identity. The security of each user depends on the registration data the user has entered during the process of registration that is not shown on the model above.

During the first authentication level biometric methods used are keyboard dynamics and mouse gesture movements. For the purpose of minimizing the vulnerability of a dislocated attacker the system uses a modification of the regular CAPTCHA system which asks the user to recognize some kind of a pattern and to describe it. During the authorization process the user introduces him to the system using a simple username. Instead of using a classical approach with a password which easily can be tracked this system uses the keystroke dynamics while typing the user name to increase the security level. That way it is demanded that the user knows his own username and also knows how to type it by his regular typing style. Alduk and Bakliža showed in [3] that although the username/password combination is publicly known, systems based on keystroke dynamics have a significantly smaller probability of unauthorized access.



Picture 3: AuthLvL-1 method

To improve the security level of a username based system we suggest combining the above mentioned biometrical method with a mouse input-based method. For this method the user decides, during the registration, whether to use virtual handwritten signature, as described in [4], either by using a mouse, graphical tablet or sensitive touch screen, or to draw a simple geometrical picture, introduced in [5], for the applied use of the touchpad or simple mouse. The first step of AuthLvL-1 algorithm is ensuring that the

subject filling the form is human by means of a modified CAPTCHA system that not only asks the user to repeat a patter generated by the system but also to do it in an intelligent way (e.g. recognize a given text from a picture).

We can describe the login algorithm as follows:

$$l_1 = x \cap (a \cap b), \qquad (1)$$

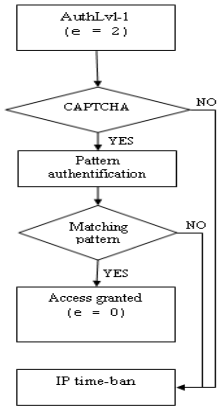
where:

- x describes the Boolean value of the CAPTCHA function,
- a describes the Boolean value of the keystroke dynamics
- b describes the Boolean value of the mouse gesture recognition
- l_1 describes the Boolean value of the level 1 authentication procedure

2.5.AuthLvL-2

Because of several situations that could affect the user (e.g. hand injuries, long user inactivity etc.) it is necessary to develop a backup authentication algorithm so the user would not be cut off of the service. On this level as well all the user input is received over the keyboard and mouse.

As the number of signs in the pattern is proportional to the security level (more signs = better keystroke dynamics result) it is necessary to use a larger text pattern than the username length. We suggest that the pattern length should be between 30 and 50 signs because a shorter pattern would cause a less secure environment and a longer pattern would decrease. To disable the possibility of cashing entry field values is the common browser we suggest that a random text pattern should be generated from a random source (e.g. Google search etc.) depending on the language(s) the user selected during the registration process.



Picture 4: AuthLvL-2 method

As a final security method we combine the CAPTCHA system with mouse gesture movements, discussed in [4] and [5]. This security authentication method increases the security level by randomly generating a series of animal pictures where each of the pictures belongs to one of 4 groups of animals. The user is asked to draw a simple geometrical (e.g. circle) or biological (e.g. flower) shape on the pictures belonging to a specific animal group. This way the sign scope significantly increases and raises the security level not only by watching it being entered but also by the drawing position, size and drawing attempts (how many strokes etc.). By using random tagged pictures of animals from the Internet, and not from a database, it is possible to eliminate the security breach of database estrangement with limited and predefined pictures.

We can describe the login algorithm as follows:

$$l_2 = c \cap (d \cap y), \qquad (2)$$

where:

- c describes the Boolean value of the random text keystroke dynamics
- d describes the Boolean value of the shape order and drawing recognition
- y describes the Boolean value of the random image CAPTCHA function
- l_2 describes the Boolean value of the level 2 authentication procedure.

3 Conclusion

Using biometric traits for authentication is a prospective alternative to standard login methods (login and password combination). In a long term it could contribute to the trustworthiness in Web and its services.

Unfortunately there are still some obstacles that we have to cross over. In the first place biometric security demands a significant financial investment in hardware and software development or/and implementation. Secondly there are still some uncertainties as to the level of security it provides since a large scale testing on a large public sample have not been done. Finally there is a small amount of doubt in general public regarding the benefits and risks that someone enters with providing biometrical data. Issues as personal privacy and the general “Big brother is watching us” fears are effecting the overall popularity of biometric methods.

By developing and implementing eBAWS system we hope to cross over some of these obstacles and contribute to the general use of biometric methods in

a networked society, a society that is already under construction.

4 Future research

The development of such a system would enable other services to encapsulate eBAWS into its own authentication process while considering each of these processes as a semantic web object.

The first step is the implementation and embedding for overall use to according W3C standards and adopts exiting state of art technologies. To test the concept of biometrical authentication and behavioral IDS an existing project TaOPis [7], already underway at the Faculty of organization and informatics (FOI), will be used as the implementation foundation.

During the testing phase each of the above mentioned methods included in eBAWS will be subjected to testing and benchmarking separately according to their performances.

5 References

[1] World Internet Usage Statistics News and World Population, on-line: http://www.internetworldstats.com/stats.htm, loaded: June 10th 2008.

[2] González-Agulla, E., Otero-Muras, E., García-Mateo, C., Alba-Castro, J.L., A multiplatform Java wrapper for the BioAPI framework, Computer Standards & Interfaces, 2007.

[3] Alduk,Ž., Bakliža, O., DINAMIKA TIPKANJA, Menadžment i sigurnost M&S 2006. Zbornik radova 1. znanstveno-stručne konferencije s međunarodnim sudjelovanjem Čakovec, Čakovec, Hrvatska, 18.5.-18.5.2006.

[4] Scheidat, T., Vielhauer, C., Dittman, J., Handwriting verification – Comparison of a multi-algorithmic and a multi-semantic approach, Image and Vision Computing, 2007.

[5] Shirali-Shahreza, M.; Shirali-Shahreza, S., Drawing CAPTCHA, Information Technology Interfaces, 2006. 28th International Conference on Volume , Issue , 2006 Page(s):475 – 480

[6] Schatten, M.: Zasnivanje otvorene ontologije odabranih segmenata biometrijske znanosti, Master's thesis, Fakultet organizacije i informatike, Varaždin, 2007.

[7] TaOpis, on-line: http://autopoiesis.foi.hr/, loaded: June 10th 2008.