K

# Keystroke Dynamics and Graphical Authentication Systems

**Sérgio Tenreiro de Magalhães**
*University of Minho, Portugal*

**Henrique M. D. Santos**
*University of Minho, Portugal*

**Leonel Duarte dos Santos**
*University of Minho, Portugal*

**Kenneth Revett**
*University of Westminster, UK*

## INTRODUCTION

In information systems, authentication involves, traditionally, sharing a secret with the authenticating entity and presenting it whenever a confirmation of the user's identity is needed. In the digital era, that secret is commonly a user name and password pair and/or, sometimes, a biometric feature. Both present difficulties of different kinds once the traditional user name and password are no longer enough to protect these infrastructures, the privacy of those who use it, and the confidentiality of the information, having known vulnerabilities, and the second has many issues related to ethical and social implications of its use (Magalhães & Santos, 2005).

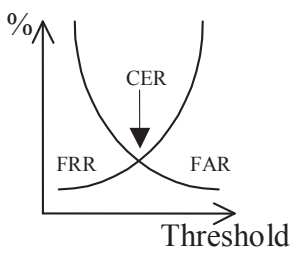
Password vulnerabilities come from their misuse that, in turn, results from the fact that they need to be both easy to remember, therefore simple, and secure, therefore complex. Consequently, it is virtually impossible to come up with a good password (Wiedenbeck, Waters, Birget, Brodskiy, & Memon, 2005). On the other hand, once users realize the need for securing their authentication secrets, even fairly good passwords become a threat when the security policies (if at all existing) fail to be implemented. The results of an inquiry made by the authors in 2004 to 60 IT professionals show that, even among those that have technical knowledge, the need for password security is underestimated (Magalhães, Revett, & Santos, 2006). This is probably one of the reasons why the governments increased their investment in biometric technologies after the terrorist attack of 9/11 (International Biometric Group [IBG], 2003).

The use of biometric technologies to increase the security of a system has become a widely discussed subject, but while governments and corporations are pressing for a wider integration of these technologies with common security systems (like passports or identity cards), human rights associations are concerned with the ethical and social implications of their use. This situation creates a challenge to find biometric algorithms that are less intrusive, easier to use, and more accurate.

The precision of a biometric technology is measured by its false-acceptance rate (FAR), which measures the permeability of the algorithm to attacks; its false-rejection rate (FRR), which measures the resistance of the algorithm to accept a legitimate user; and its crossover error rate (CER), the point of intersection of the FAR curve with the FRR curve that indicates the level of usability of the technology (Figure 1). For a biometric technology to be usable on a stand-alone base, its CER must be under 1%. As an algorithm becomes more demanding, its FAR is lower and its FRR is higher. Usually the administrator of the system can define a threshold and decide what the average FAR and FRR of the applied algorithm will be according to the need for security, which depends on the risk evaluation and the value of what is protected; also, the threshold can be, in theory, defined by an intrusion detection system (software designed to identify situations of attack to the system).

Establishing the error rates of a biometric technology is a complex problem. Studies have been made to normalize

*Figure 1. Crossover error rate*



their evaluation, but the fact is that the results are strongly dependent on the number of individuals involved in the process and, what is worst, on who is chosen. This means that, even with a large amount of data collected, the results can be very different if we change the evaluated group. The lack of trust in the precision evaluation methodologies and values is one of the reasons why the human rights associations are opposing the generalization of use of biometric technologies and their acceptance as standards for authentication procedures (Privacy International, Statewatch, & European Digital Rights, 2004). Even so, in an inquiry made by Epaynews (http://www.epaynews.com), 36% of users stated that they would prefer to use biometric authentication when using credit cards, a value only comparable to the use of personal identification numbers (PINs) and much higher than the 9% of authentication obtained by signature.

Considering all the advantages and disadvantages of biometric procedures, it seems that the only way is to allow the user a choice. Being so, the traditional password systems must be enhanced both in the biometrical way and in another completely different way. On the biometric component we propose keystroke dynamics, a biometrical authentication algorithm that tries do define a user's typing pattern and then verifies in each log-in attempt if the pattern existing in the way the password was typed matches the user's known pattern; it is the only biometric technology that can be used with the existing log-in and password systems without requiring any extra hardware. On the nonbiometric component, we propose the use of a graphical authentication system, a log-in system that verifies the user's knowledge of specific images or parts of images to grant or deny successful log-in, because it has been proven that it provides a wider key space and because it can be used to generate complex secret strings from simple passgraphs (the user's secret code to access a system protected by a graphical authentication system, constituted by a sequence of points where the user must click in order to obtain a successful log-in).

# BACKGROUND

## Keystroke Dynamics

As in many other problems, there have been two different approaches to the challenge of finding an algorithm for keystroke dynamics that minimizes the CER: machine learning and deterministic algorithms.

Among the solutions based on machine learning, we can find the work presented by Ord and Furnell (2000) that tested this technology with a 14-person group to study the viability of applying it to the simple use of PINs typed on a numeric pad. Unfortunately, the results suggest that, for large-scale use, the technology is not feasible. Deterministic algorithms have been applied to keystroke dynamics since the late '70s. In 1980, Gaines et al. (1980) presented a report on the study of the typing patterns of seven professional typists. The small number of volunteers and the fact that the algorithm is deducted from their data and not tested for other people later results in lower confidence in the FAR and FRR values presented. However, the method used to establish a pattern was a breakthrough: the study of the time spent to type the same two letters (digraph) when together in the text. Since then, many algorithms based on algebra and on probability and statistics have been presented. Joyce and Gupta presented in 1990 an algorithm to calculate a value that represents the distance between acquired keystroke latency times and correspondent times previously stored. In 1997, Monrose and Rubin used the Euclidean distance and probabilistic calculations based on the assumption that the latency times for one digraph exhibits a normal distribution. Later in 2000, they also presented an algorithm for identification based on the similarity models of Bayes, and in 2001 they presented an algorithm that uses polynomials and vector spaces to generate complex passwords from a simple one using the keystroke pattern (Monrose et al., 2001).

In 2005, Magalhães, Revett, and Santos presented an improvement of the Joyce and Gupta algorithm and tested it with 170.391 attacks to 143 patterns, obtaining a 0% FAR with an FRR of 26%, and an estimated CER below 5%.

## Graphical Authentication Systems

A graphical authentication system is a log-in system that verifies the user's knowledge of specific images or parts of images to grant or deny successful log-in. Greg Blonder (1996) was the first to describe graphical passwords, presenting in a United States patent a system that would allow users to choose a picture, the number of regions to be clicked, and their sizes and positions. Since then, many variations of this system were presented and images have gained their way into the authentication processes.

Among the most popular graphical authentication systems, we find Passfaces™ from the Passfaces Corporation (2005), a commercial system where the user chooses a previously selected face from a set of faces and repeats this process for different faces in different sets for a defined number of times. However, being popular does not imply being secure, and a study of the users' choices demonstrated that they are, in some cases, similar for all users. For instance, 10% of the passwords of males could have been guessed with only two attempts (Davies, Monrose, & Reiter, 2004).

The déjà vu scheme involves a matrix of $m$ images in a set, where $n$ images are part of the user's portfolio, previously chosen from a set of proposed images. The user must identify those $n$ images to log in.

The draw-a-secret (DAS) scheme is a graphical authentication system with an approach completely different. In DAS, the user draws something over a grid that becomes the