Chapter 9
Keystroke Analysis as a Tool for Intrusion Detection

Daniele Gunetti
University of Torino, Italy

Claudia Picardi
University of Torino, Italy

ABSTRACT

The original dream of Keystroke Analysis was the same as that of other biometric techniques: replacing traditional authentication methods with techniques based on the analysis of the typing dynamics of users. Unfortunately, until now a sufficient level of accuracy and user friendliness for practical applications has not been achieved.

However, typing rhythms are available throughout an entire login session and, if we can perform Keystroke Analysis of free text, we can implement one of the best applications of continuous authentication: Intrusion Detection. In this case, an accuracy that is still not acceptable for access control can be more than sufficient as part of an Intrusion Detection policy, where (1) alarms that have been raised must in any case be validated by a human (e.g., a system administrator); (2) intrusions are normally detected joining together different techniques, in order to improve the resulting accuracy.

In this chapter we discuss the potentialities of Keystroke Analysis as a tool for Intrusion Detection and other security applications, and investigate experimentally how the accuracy of the analysis scales with the increase of the number of individuals involved, a fundamental issue if we want to add Keystroke Analysis to the set of tools that can be used to improve the security of our computers and networks.

DOI: 10.4018/978-1-61350-129-0.ch009

INTRODUCTION

There is an evident and growing interest for security applications based on biometric features, such as fingerprints, retina, iris, signature, palm and voiceprint (Jain et al. (eds.) 2008). The usual argument brought in favor of the use of biometric information for identity verification against more traditional techniques is that passwords (security based on *what you know*) can be forgotten or forged, and ID cards (*what you have*) can be lost or replicated, whereas biometric features (*what you are*) are unique and cannot be (easily) replicated, shared or stolen.

Within computer security, a natural biometric is represented by the dynamics shown by individuals when typing at the keyboard, and some peculiarities make keystroke analysis somewhat unique w.r.t. other biometrics. First, keystroke analysis is not intrusive, since users will be typing at the computer keyboard anyway. Second, it is inexpensive to implement, since the only sensor required is the keyboard itself. Third, keystroke patterns of an individual cannot be easily replicated or stolen by an impostor, even if such patterns are known. Finally, unlike other biometric features, typing rhythms are available throughout an entire login session (i.e., after an authentication phase has been passed, or fooled), since keystrokes exist as a mere consequence of users using computers.

The original dream of keystroke analysis was the same as that of other biometric techniques: replacing traditional authentication methods with the analysis of the typing dynamics of users. Unfortunately, such a dream has never become reality: although there have been a number of studies on the possibility of verifying personal identity through the analysis of the typing dynamics of short, pre-defined passphrases (e.g., Joyce and Gupta. (1990), Bleha et al. (1990), Brown and Rogers (1993), Obaidat and Sadoun (1997), Clarke et al. (2003), Clarke and Furnell (2007)), a sufficient level of accuracy and user friendliness for practical and commercial applications has never been really achieved. Even when keystroke analysis is used not to replace, but to strengthen passwords (e.g., Reiter et al. (1999), Ong and Lai (2000), Revett et al. (2005), Chang (2005), Rodrigues et al. (2005), Hocquet et al. (2006), Mészáros et al. (2007)) one wonders whether a well-chosen password would be sufficient to reach the same level of security provided by a poor chosen password whose performance we try to improve by the analysis of typing rhythms of that password. All the above methods and systems show an error rate of a few percentage points. As a consequence, a well chosen password will always outperform any combinated technique, in terms of simplicity of implementation and false alarms.

However, the aforementioned peculiarities of typing rhythms, together with recent advances in the analysis of typing dynamics of free text, may suggest a change of perspective. Passwords will probably never be replaced by keystroke analysis, but the growth in the number of crimes perpetrated through and on computer and network resources, and on the Internet, call for new and improved security techniques to investigate and fight such illegal activities.

In this chapter we argue that computer and network intrusion detection, as well as identity tracing over anonymous Web-based resources, and password recovery, are only some of the possible situations where the analysis of typing rhythms may deserve a second chance.

BACKGROUND

Identifying users through the way they type on a keyboard is difficult, because keystrokes convey little information: just the time when keys are depressed and released. It is true that special keyboards may allow to measure additional features such as the acceleration or the energy impressed to the keystrokes, but such keyboards are normally not available, and are expensive. Moreover, recent sophisticated devices may use virtual keyboards