Global Information Assurance Certification Paper

Authentication by Typing Rhythm

Philip Marriott (philip.005)
SANS GIAC GSEC Practical 1.4c
29 September 2004

Abstract

The application of *Typing Rhythm*, also known as Keystroke Dynamics / Timing / Latencies, to user identification and authentication holds great promise. There is sufficient evidence to support its effectiveness, and a number of implementations have emerged in recent years including the commercial BIOPASSWORD software by BioNet Systems Inc. Despite all of this, Biometric Typing Rhythm as means of authentication, has had seemingly little impact on users, government, and corporations worldwide. This paper investigates the nature and appropriateness of *Typing Rhythm* to PC-based authentication through a review of relevant literature and recent published evaluations of the technology. Explanations and recommendations are provided which help situate this technology in its present context.

Introduction

User authentication typically involves up to three fundamental pieces of information: who you are, what you have and what you know (Patrick, 2002). If some or all of these are satisfied then the user is admitted to the system. What you have is proven by the possession of some artefact such as a smart card, key, or digital certificate. What you know is usually achieved by possession of a shared secret in the form of a password. Who you are is usually determined by a userid (name, email address etc) however this is no real assurance of identity and the field of Biometrics can provide assistance with retina scans, finger print analysis, voice recognition, and typing rhythm as a means of increasing the probability of correct user identification. (Patrick, 2002)

Biometrics, for convenience, can divided into two main categories: physiological (face, eye, fingerprints, hand geometry, thermal images) and behavioural (voice,

written signatures, and typing rhythm) (Bergadano, Gunetti, and Picardi, 2002). Typically Biometric authentication schemes using physiological features have been the most successful, this is due to the stability of physiological features over time, whereas behavioural features such as typing rhythm and to a lesser extent voice may change substantially between consecutive samplings.

Many biometric techniques, usually the physiological ones, require additional hardware that is inconvenient and expensive to implement and deploy. The biometric techniques that are the least intrusive, require no special hardware, and have the potential for wide-scale deployment are ones based on typing rhythm eg. (Bleha, Slivinsky, and Hussien, 1990; Obaidat, and Sadoun, 1997).

The most commonly implemented means of gaining access to a computer system is by a username and password (Joyce, and Gupta, 1990). There are inherent problems with this and most can be attributed to human factors (Adams, 1999).

> 'Human factors are perhaps the greatest current barrier to effective computer security. Most security mechanisms are simply too difficult and confusing for the average computer user to manage correctly' (Whitten, and Tygar, 1998 :1)

Users have difficulty with passwords, often choosing easy to guess or crack passwords. They have difficulty remembering passwords and often write them down in easy to access locations. The end result of this is that it can be possible for a third party to gain possession of a username and password and thus gain unauthorised entry to a computer system.

Typing Rhythm (also known as Keystroke Dynamics (Bergadano *et al.*, 2002) and Keystroke Latencies (Joyce *et al.*, 1990)) is one biometric means of addressing the problem of unauthorised access to computer systems. In the commercial implementation BIOPASSWORD by BioNet Systems/Net Nanny (Konicki, 2000); even if the third party is in possession of a valid username and password, a password must be typed in using a rhythm that is characteristic of the person associated with the username / password before access is granted. Typing Rhythm as a means of Biometric identification and authentication is claimed to be effective, cost effective, and deployable (Altman, 2002; Bragg, 2002; Yu, and Cho, 2004). Interestingly, it has not received widespread acceptance as an authentication measure. Bergadano (2002 :396) succinctly summarises the current state of play: 'Keystroke dynamics is the most obvious kind of biometrics available on computers, but it has not yet led to real security applications, if compared to other biometric measures.'

Evaluation Results

Evaluation of Typing Rhythm technology in the literature falls into three main areas:

1. Researchers reporting on trials of their particular implementations of typing rhythm for the purposes of improving the methods of matching and identification. This is general research and is not linked to any existing security measures and should not be confused with the password hardening research in 2.

2. Researchers reporting on their implementations of password hardening using typing rhythm. This is specific research aimed at improving user authentication using the results from 1 above.

3. Reviews by Press and Users of the commercial BIOPASSWORD software, the only commercial implementation of typing rhythm discussed in the literature.

1. General Research

Reporting of research into typing rhythm began in the early 1980's with the often cited pioneering paper by Gaines et al (Gaines, Lisowski, Press, and Shapiro) from the RAND Corporation. They measured the effectiveness of their typing rhythm system by two parameters (still in use today) : False Alarm Rate (FAR), the rate that a keyboard rhythm is falsely identified as belonging to an imposter; and Imposter Pass Rate (IPR), the rate that an imposter's keyboard rhythm is incorrectly identified as belonging to a legitimate user (Bergadano *et al.*, 2002). The ideal situation is for both these parameters to be as close to zero as possible – usually it is more acceptable to have a higher FAR than IPR if a secure environment is the goal.

In Gaines et al's (1980) experiments, seven secretaries were asked to retype the same three paragraphs at two different times over four months and keystroke timings were compared. Their results showed a FAR of 4% of an IPR of 0%. While they proved the concept of user identification by keyboard timings as viable it is difficult to evaluate the effectiveness of their methods due to the limited scale of their experiments.

Joyce and Gupta (1990) describe their identity 'Verifier' which is based on keystroke timings. In their experiments, 33 users each provided a reference signature by typing in their login name, password, first name and last name eight

times. The user then tried to login to their account five times and the data collected. Six of the users acted as imposters and tried to log into the remaining 27 accounts. They achieved a False Alarm Rate (FAR) of 16.7% and an Imposter Pass Rate (IPR) of 0.25%. The high FAR of 16.7% is equal to a rejection of 1 in 6 login attempts requiring another attempt. Joyce and Gupta note that the FAR could be reduced if a higher IPR was considered acceptable. They also note that significant reductions in FAR can be achieved with only slight increases in IPR if thresholds are manipulated in certain ways. Interestingly the samples for their experiments were taken using the same computer system and therefore the same keyboard. Joyce and Gupta recommend that further research is done to see the effect of other systems on their results - particularly seeing that the ability to record accurate timings is an essential part of their algorithm; and this may not be available in a distributed or online environment.

Monrose and Rubin (1997) acknowledge the work of Joyce and Gupta and extend their research by: 1. examining the use of keystroke durations (length of time keys are depressed) in addition to keystroke latencies (time between successive keystrokes); 2. exploring the long term measurement of keystroke dynamics over weeks; 3. and measurement of keystroke dynamics using the user's own computer. Their results showed that all three aspects could be achieved within a workable framework. Of particular interest is their foundation work on the design of a dynamic authentication system (see also (Leggett, Williams, Usnick, and Longnecker, 1991)) that authenticates a user over time using the unstructured text typed by a user in their normal work practices.

Obaidat and Sadoun (1997) report on their work using keystroke durations and latencies and neural nets to determine a user's identity based solely on their userID. They claim very low False Alarm and Imposter Pass rates and make the observation that the keystroke durations (hold times) are more significant than the keystroke latencies (time between key presses). Significantly, they have achieved good recognition using very short strings (10 characters). What is not clear from their paper is amount of training required before their system is able to perform the verification recognition that they claim. Additionally it is of concern that both the imposter's and owner's typing patterns were used for learning which is not applicable to most network situations.
Robinson et al (1998) also conducted work on verification of userIDs, with reference to Obaidat and Sadoun (1997), that achieved a False Alarm Rate (FAR) of 10% and a Imposter Pass Rate (IPR) of 9%. They used both keystroke durations and latencies and the mean userID length was 6.4 characters. While impressive, they caution that a FAR of less than 1% is required before this type of security measure could be considered non-invasive.

Bergadano et al (2002) report on their keystroke analysis technique, which takes into account problems associated with variability of typing and typing errors, and produces a False Alarm Rate (FAR) of 4% and an Imposter Pass Rate (IPR) of less than 0.01%. This was achieved by using the same sampling text of 683

characters per user (entered 5 times), allowing typing errors, and in a simulated online environment. Interestingly, once again, the samples were all collected in the same room on the same computer and therefore the same keyboard. The authors state that they are unsure of the effect of variability on keyboard type and condition, and this may be a weakness in their method.

There would seem to be substantial evidence that Typing Rhythm as a method of authentication is proven to be viable. Ongoing research is clearly needed to reduce both False Alarm Rates (FAR) and Imposter Pass Rates (IPR) to levels that become transparent to the user.

2. Password Hardening

Research oriented towards improving the security of passwords is presented in this section. These systems integrate and augment the security provided by conventional username/password systems.

Cho et al (2000) propose a web-based java applet system for verifying authenticity of passwords using keystroke dynamics and neural nets for analysis. The system is described as follows:

> 'When a client tries toaccess a home page, for example, say a firm's online shop, located in a server, the user types the already registered user ID. Then the server sends the client a Java applet code that can measure the user's password keystroke timing vector. Once the Java applet running in the client system gathers the user's keystroke timing vector, it sends it back to the server. Then the autoassociative neural network located in the server can verify whether the user is the person he/she claims to be. Because thecode is programmed in Java, any client system that has a Java browser can be connected to the server.' (:305)

Chow et al report a FAR of 1% which is within the specification for acceptance by users suggested by Robinson et al (Robinson *et al.*, 1998). However they rejected the results from some inexperienced typists which they claim improves the FAR results of their experiments. They recommend that further investigation on role of typing experience is conducted. Yu and Cho (Yu *et al.*, 2004) refect on this work in their later paper and identify two significant problems: 1. Training time was excessive; 2. The data set required was too large. They propose a solution that addresses these problems while still retaining similar FAR and IPR results.

Monrose et al (2002) present a system where a user's keystroke latencies and durations are combined with the user's password to form a hardened password that is more secure that a conventional password. Their scheme automatically adapts to gradual changes in a user's typing patterns while still maintaining the same hardened password across multiple logins. Initially the password is as secure as a conventional password and is gradually hardened as biometric information becomes available. They identify the main limitation of their system is

the situation where a user, whose typing patterns change substantially between successive logins, possibly due to an unfamiliar keyboard, fails to generate the correct hardened password and is locked out of the system. They recommend that hardened passwords are therefore most suited to users operating consistent hardware.

Monrose et al also claim that their system improves on other existing password hardening systems, in particular the commercial BIOPASSWORD system, by generating a repeatable key from the biometric component of the hardened password that is stronger than the password itself. Other systems, they argue, are able to be compromised if the hardened password is captured and attacked; although one would expect this to take significantly longer than with a conventional password. While their results are very encouraging they provide a cautionary note that the trial was limited to 20 users and 1 password – they strongly recommend that further research is conducted in this area.

While the research on password hardening using keystroke dynamics is limited, it is clear that as a means of improving the security of username/password authentication while still working within existing frameworks, the method is viable in a networked environment.

3. Reviews of BIOPASSWORD

BIOPASSWORD, www.biopassword.com , is a commercial implementation of typing rhythm for securing networks, and standalone PCs using a standard username / password logon. It is middleware that replaces the normal logon screen of a PC. It is sold by BioNet Systems who recently purchased the rights and technology from Net Nanny Inc. BIOPASSWORD is a deriviative of the pioneering work conducted by the RAND Corporation (Gaines *et al.*, 1980) and is protected by a number of patents.

In the network version of the software, special server software is installed on a windows NT/2000/2003 domain controller, which then controls the logon of domain member computers. New users are required to enter their username and password 15 times (default value) to enable keystroke dynamics to be recorded, this is called the training cycle. A security level can be set for each user; this appears to be a threshold for balancing False Alarm Rates (FAR) and Imposter Pass Rates (Patrick, and Mu, 2004).

BIOPASSWORD has received a number of favourable reviews from the IT press (Altman, 2002; Bragg, 2002; Distance-Educator.com, 2001). It would seem that reviewers on the whole found the security offered by the system to be reliable and effective with none of the reviewers able to generate Imposter Pass errors. Reviewers also found the learning phase to be acceptable. There was mixed opinion on the ease of installation with one reviewer lamenting on the high

knowledge of Windows Domain structures required. Because of the middleware nature of the software, one reviewer was able to bypass security by using 'run as' privileges – however it was suggested that would be fixed in a later version of the software.

Of particular interest is the timing of the review articles which all occurred around the launch of the software in 2001/2002. There were also a number of announcement type articles written around this time also. Since this time the BIOPASSWORD software seems to have been largely forgotten by the IT and popular press. This could be interpreted to mean that it has not yet made the market penetration that was heralded in its initial release, however it could also indicate that the media has simply turned its attention to more newsworthy items – time of course will tell.

BIOPASSWORD comes in a Software Development Kit (SDK) version and two commercial products claim to incorporate it into their products. ContinuedEd.com claimed (Distance-Educator.com, 2001) in 2001 to have entered into a licensing agreement with NetNanny Inc (the then owners of BIOPASSWORD) to incorporate BIOPASSWORD technology into their online verification system. Whether this actually occurred (and is still in use) cannot be determined from available online information.

Symetric (Symetric Sciences, 2002) is software that manages clinical trial data. It has incorporated biometric user authentication features developed using the BIOPASSWORD SDK since 2001. According to their website, www.symetric.ca , the latest version of the software still has this feature.

The Credit Union Times (Gentile, 2004) report that San Antonio City Employees Credit Union has recently introduced BIOPASSWORD security to their laptops. BioNet Systems themselves claim on their website, www.biopassword.com , to be in partnership with large corporations such as Novell and Citrix and are actively developing products incorporating their BIOPASSWORD technology.

It would be reasonable to conclude from the net presence and reviews of BIOPASSWORD that the product effectively performs its authentication role, albeit with some usability issues. It can be deployed with relative ease, yet requires ongoing system administration. However it does not seem to have been widely adopted by users, government and corporations. Despite its great promise, BIOPASSWORD, seems to be in a holding pattern at present.

Lessons learned

A number of lessons can learned from the findings from the literature review conducted in this paper – these are presented here.

Complexity of passwords is a factor in the ability of imposters to fake a typing rhythm, this is analogous to trying to forge a simple written signature. ‘Easy-to-type’ strings which are usually short in length are the most likely to be faked (Joyce *et al.*, 1990). The recommendation here is for the enforcement of a minimum password length.

It has also been noted that there is a correlation between a person’s typing skills and the ease of which an imposter can learn to mimic their typing rhythm (Bleha *et al.*, 1990). It would seem that the greater the speed at which a phrase is typed tends to make it more difficult to mimic the keystrokes of a user. There is also an indication that known phrases are easy to type and can be typed quicker making them more difficult to learn by an imposter. This tends to contradict the findings of Joyce (Joyce *et al.*, 1990) above and goes against good password policy. Robinson et al (Robinson *et al.*, 1998) in their experiments found that the rate of typographical errors made by users when typing in userIDs was high – around 24%. In their implementation, even if corrected, these errors would lead to a rejection or false alarm. This high rate of rejection would be unacceptable for users and Robinson et al suggest that users would resort to slow, deliberate, typing that would be easy to mimic by imposters. Some systems claim to accept errors (Bergadano *et al.*, 2002); however it is not clear whether they accept them in the testing as well as the learning phase.

Probably the recommendation should be for a minimum password length with ‘not too difficult’ to type phrases for that particular user. Interestingly, over time, it would be expected that ‘difficult to type’ phrases will become ‘easy to type’ phrases and the associated keystroke timings will change. This points to the need for regular sampling as a means of maintaining a record of the reference typing rhythm.

There is a ‘learning’ phase of any typing rhythm system and this can vary in complexity and time. Some systems described in the literature require an extensive learning phase requiring up to a few thousand keystrokes collected over multiple sessions(Gaines *et al.*, 1980), other systems require only a couple of strings collected on 5-10 occasions (Joyce *et al.*, 1990). Clearly high complexity of the learning phase is detrimental to usability and systems with short learning phases are more acceptable to users. Interestingly there seems to be ‘diminishing returns’ when comparing learning complexity to False Alarm Rates (FAR) and Imposter Pass Rates (IPR), with acceptable results available with modest learning phases in some systems eg. (Bergadano *et al.*, 2002; Joyce *et al.*, 1990). Ideally security systems that rely on typing rhythm should have a learning phase that is ongoing and transparent to the user.

Future and Recommendations

Despite the great promise of biometric techniques using Typing Rhythm as a means of improving authentication, there seems to have been a disproportionately low penetration of the method into mainstream authentication. Nearly all the papers, reviewed in this report, lament on this situation. A number of plausible explanations can be offered that centre of the theme that once a Typing Rhythm system of authentication is deployed, life is made more difficult for everyone involved including users, administrators, and support staff. For example:

- The technology usually requires the installation of middleware which is an additional expense and additional drain on IT administrative and support resources. Middleware introduces more complexity into the logon procedure and creates a greater opportunity for failure and attack vectors (Bragg, 2002).

- The technology makes the logon procedure more difficult for users, particularly when False Alarm Rates are high. This will impact on Help Desks, already receiving half their workload as password related issues (Patrick, 2002); who could risk having even more password-related support calls.

- The use of such biometric techniques needs to be coordinated across a user group and this requires setup and maintenance resources. The use of such techniques, with its greater reliance on acceptable password construction may expose existing weaknesses in IT policy and implementation in a workplace.

- The technology is new and there may be a resistance and lack of trust towards such an innovation. Conservative organizations may be waiting until other organizations adopt such procedures.

- There is probably a lack of government/legislative requirements/incentives to improve authentication to the level afforded by Biomentric Typing Rhythm.

Of course these objections will be tolerated if the value of the improved security is a high enough priority for the organization. It would be reasonable to say that at present most users and organizations are not motivated enough to adopt such measures as Biometric Typing Rhythm authentication.

It may be that Typing Rhythm will make inroads into niche applications such as determining whether a user is exhibiting uncharacteristic keystroke patterns

indicative of fatigue, drug and alcohol use eg. (Monrose *et al.*, 1997); or authentication for high security environments. However, whether this technology will be part of mainstream authentication is indeterminate at present.

References

Adams, A., and Adams, A. (1999). Users are not the enemy. *Communications of the ACM* 42, 40-46.

Altman, A. (2002). *BioPassword 4.5*. [accessed online 21/9/04: http://www.biometritech.com/features/022502review.htm]

Bergadano, F., Gunetti, D., and Picardi, C. (2002). User Authentication through Keystroke Dynamics. *ACM Transactions on Information and System Security,* 5, 367-397.

Bleha, S., Slivinsky, C., and Hussien, B. (1990). Computer-Access Security Systems Using Keystroke Dynamics. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12, 1217-1222.

Bragg, R. (2002). *Biometric Security Products: BioPassword 4.5*. [accessed online 19/9/04: http://mcpmag.com/reviews/products/article.asp?EditorialsID=253]

Cho, S., Han, C., Han, D. H., and Kim, H.-I. (2000). Web-Based Keystroke Dynamics Identity Verification Using Neural Network. *Journal of Organizational Computing & Electronic Commerce* 10, 295-308.

Distance-Educator.com. (2001). ContinuedEd.Com Licenses Net Nanny's BioPassword Technology To Provide Strong User Authentication For Online Driving Safety Courses.

Gaines, R., Lisowski, W., Press, S., and Shapiro, N. (1980). Authentication by keystroke timing: Some preliminary results (Report Number R-256-NSF). Rand Corporation.

Gentile, P. (2004). San Antonio City Employees FCU Takes Security to a Higher Level By Analyzing Unique User Keystrokes. [accessed online 19/9/04: http://www.biopassword.com/assets/cu-reprint.pdf]

Joyce, R., and Gupta, G. (1990). Identity authentication based on keystroke latencies. Communications of the ACM 33, 2

Konicki, S. (2000). Typing Rhythm Can Protect Computers From Intruders. [accessed online 19/9/04: http://www.informationweek.com/showArticle.jhtml?articleID=6509734]

Leggett, J., Williams, G., Usnick, M., and Longnecker, M. (1991). Dynamic identity verification via keystroke characteristics. International Journal of Man-Machine Studies 35, 859-870.

Monrose, F., and Rubin, A. (Year). "Authentication via keystroke dynamics." Paper presented at the Proceedings of the 4th ACM conference on Computer and communications security, Zurich, Switzerland, 1997.

Monrose, F., Reiter, M. K., and Wetzel, S. (2002). Password hardening based on keystroke dynamics. International Journal of Information Security 1, 69-84.
Obaidat, M. S., and Sadoun, B. (1997). Verification of computer users using keystroke dynamics. IEEE Transactions on Systems, Man & Cybernetics: Part B 27, 261-270.
Patrick, A. (2002). Human Factors of Security Systems: A Brief Review. [accessed online 16/9/04: http://www.andrewpatrick.ca/passwords/passwords.pdf]
Patrick, A., and Mu, S. (2004). Usability and Acceptability of Biometric Security Devices. [accessed online 16/9/04: http://www.andrewpatrick.ca/biometrics/index.shtml]
Robinson, J. A., Liang, V. M., Chambers, J. A. M., and MacKenzie, C. L. (1998). Computer user verification using login string keystroke dynamics. IEEE Transactions on Systems, Man & Cybernetics: Part A 28, 236-242.
Symetric Sciences, I. (2002). Symetric. [accessed online 21/9/04: http://www.symetric.ca/Software.htm]
Whitten, A., and Tygar, J. D. (1998). Usability of Security: A Case Study (Report Number CMU-CS-98-155). Carnegie Mellon University.
Yu, E., and Cho, S. (2004). Keystroke dynamics identity verification—its problems and practical solutions. Computers and Security 23, 428-440.