THE USE OF KEYSTROKE ANALYSIS FOR CONTINUOUS USER IDENTITY VERIFICATION AND SUPERVISION

Steven M Furnell, Peter W Sanders and Colin T Stockel
Network Research Group
Faculty of Technology
University of Plymouth
Plymouth, United Kingdom
E-mail : stevef@soc.plym.ac.uk

ABSTRACT

This paper proposes the concept of dynamic keystroke analysis as a means of enhancing user authentication in modern information systems. Whilst existing password-based schemes normally rely upon a single authentication judgement, the use of keystroke analysis would allow supervision to occur continuously throughout user sessions. In addition, the concept may be implemented transparently so as not to unnecessarily disrupt user activity. These points make it suitable for application in modern, user-friendly contexts such as multimedia.

The theoretical discussion is supported by the findings of an experimental study mounted within our group using 26 typists and a prototype authentication system. The results demonstrate considerable success, with an impostor detection rate of 85%. However, a number of potential problems identified in the discussion suggest that keystroke analysis would be best implemented alongside other supervision techniques rather than as a standalone system.

INTRODUCTION

A key issue in the implementation of secure information systems is user authentication. The password remains the popular and widespread technique (National Computing Centre 1994), having the advantage of simplicity for both systems designers and end users. However, a disadvantage is the ease with which its protection is often compromised, either deliberately, by accident or by guesswork. In recent years the reliability of passwords has been repeatedly questioned (Jobusch and Oldehoeft 1989) and it is now widely accepted that stronger means of authentication may be necessary, using techniques that are more difficult to forge. In addition, password techniques can only verify user identity at discrete points within a session (and are normally only incorporated at the beginning). With the increasing advancement of information systems, as witnessed by the progression to multimedia, it is both desirable and appropriate to have a means of identity verification that can deliver a continuous assessment of user legitimacy (and thereby provide greater protection against compromise).

This paper proposes a behavioural biometric measurement based upon the analysis of users typing characteristics. It has been established that users may exhibit significant differences in terms of typing styles and abilities (Card et al. 1980), which may consequently be used to determine reasonably unique typing "signatures" (analogous to those which can be identified with normal handwriting (Fairhurst et al. 1994)). These signatures may then be used as the basis for real-time user supervision, providing a continuous and transparent (i.e. non-intrusive) means of verifying user identities in conjunction with their normal working activities.

CONCEPTUAL SUMMARY

Several typing characteristics may be considered as the basis for determining keystroke signatures, including the intervals between keystrokes, the duration of keystrokes and the frequency of typing errors. The chosen factors must be assessed to create a typing *profile* for each legitimate user. Subsequent authentication / supervision is then based upon a comparison of the current users typing characteristics against the profile associated with his / her claimed identity (with any significant departures triggering impostor alerts).

Keystroke analysis may be implemented in two ways (referred to as *static* and *dynamic* verification approaches), which differ in how they attempt to use the technique. In the *static* scenario, analysis is based upon a constant text string and is normally used for a single authentication judgement (e.g. in conjunction with the entry of a normal user id and password). By contrast, the *dynamic* approach attempts to analyse any arbitrary text input, allowing much greater scope for user supervision as the authentication period may become continuous. The majority of previous studies have concentrated upon the *static* verification approach (Bleha et al. 1990; Joyce and Gupta 1990).

As with other biometric systems, the effectiveness of keystroke analysis is judged on the basis of two factors :

- *False Acceptance Rate (FAR)*
  The proportion of cases in which impostors are falsely authenticated by the system (also referred to as Impostor Pass Rate).

- *False Rejection Rate (FRR)*
  The proportion of cases in which legitimate users are rejected by the system (also referred to as False Alarm Rate).

Acceptable figures for these measures are heavily dependant upon whether a static or dynamic verification strategy is employed. In the static scenario, minimising the FAR should be the most important consideration, as any successful impostor could potentially go unchecked for a whole session. However, in the dynamic scenario, with continuous assessment, a greater window for impostor detection is available and so the prime concern becomes to minimise the FRR (as rejections *during* a session could irritate and disrupt a legitimate user more significantly than occasional false login failures).

PRACTICAL STUDY

This section details the research teams implementation of a prototype keystroke authentication system based on the dynamic verification approach.

Experimental System

An experimental system has been developed for the PC environment to allow an evaluation of the concept in practice. It is comprised of three modules, as follows :

- *Profiler*
  Accepts the initial typing samples used to create profiles for legitimate users. PC hardware interrupts are used to detect key depression and release with one millisecond accuracy.
- *Sampler*
  Accepts user test samples and stores all keystrokes and associated timings for later use.
- *Monitor*
  Compares the test samples against typing profiles to determine the effectiveness of the system.

Typing profiles were based upon inter-keystroke times for specific character pairs (digraphs), storing the mean time and standard deviation for each profiled digraph (note : inter-keystroke time was found to be the most distinctive typing characteristic in a provisional study, with the keystroke duration and typing error frequency measures exhibiting FARs significantly high enough to warrant exclusion from further investigation). Analysis was restricted to digraphs involving alphabetic and "space" characters, as these were considered the most likely to reveal any characteristic keystroke rhythm and were found to produce the best results in a previous study which conducted a comprehensive investigation of this aspect (Leggett and Williams 1988).

The profiling procedure demanded that users enter two samples of a 2200 character *reference text*. A significant length was necessary to ensure that each users "natural" typing style emerged and that sufficient samples of each digraph were obtained to enable appropriate mean and standard deviation values to be established (note that at least five samples were required for profile entries to be usable in monitoring, as any less could result in them being unrepresentative of the users normal style). Another property of the reference text was that the relative frequencies of character digraphs within it corresponded closely to those of normal English (Beker and Piper 1982), with the 30 most common digraphs all significantly represented (ensuring strong profile entries for the digraphs most likely to be encountered).

The profiler attempted to further ensure representative profiles by filtering out potentially uncharacteristic typing. This was achieved in two ways : firstly, deleted keystrokes were ignored, as any entries resulting from mis-strokes could be unrepresentative. Secondly, inter-keystroke times exceeding 750ms (i.e. Card et al's speed classification for a user unfamiliar with the keyboard) were disregarded, being considered more likely to represent unnatural pauses than part of the users typing rhythm.

The monitoring / supervision system compared incoming inter-keystroke times (from the test samples) against user profiles, with times being judged invalid if they fell outside 1.5 standard deviations of the relevant profiled value. Invalid keystrokes were then analysed in two ways to detect intrusions :

1. monitoring the percentage of invalid keystrokes during the 100 most recently typed;

2. monitoring the number of consecutive invalid keystrokes.

However, even legitimate users will generate some invalid keystrokes and, as a result, the monitor incorporates user-specific *authentication thresholds*

which specify the maximum levels for percentage invalid keystrokes and consecutive invalid keystrokes that are tolerated against each profile (note that the use of common threshold levels for all users was found to be less effective). The appropriate levels were determined using the two further text samples (of 574 and 389 characters) entered by each user, which were run against their initial profile. The peak values observed for percentage invalid keystrokes and consecutive invalid keystrokes across the two tests were then used as the basis for establishing the thresholds. If either threshold was exceeded during monitoring, an intrusion alert was generated.

Given that the *dynamic* verification approach was being tested, minimising the FRR was considered important. The user-specific thresholds were, therefore, set to ensure that no false rejections would arise from the test samples. The advantage of this was that the resulting FAR would then effectively represent a "worst case" figure.

Test Subjects

A total of 26 subjects were involved in the tests, with abilities ranging from experienced typists to comparative novices.

The two additional text samples that had been used to determine the authentication thresholds for legitimate users were also used to represent impostor attempts (by running them against all other profiles). The final results were, therefore, derived from approximately 1300 impostor attempts.

Results and Analysis

With the FRR having been eliminated, the aims of the study were to determine the FAR and the speed of successful impostor detection.

In terms of overall impostor detection effectiveness, the experimental system exhibited a FAR of 15%, as shown in figure 1. However, given that each subject provided two test samples, it was also possible to investigate impostor consistency. This was established by sub-dividing the test samples into the pairs that were generated by the same subjects and then determining the proportion of cases where both samples were able to pass as another user against those where only one attempt was successful. This information is also illustrated in figure 1. It can be conjectured that, given longer test samples, the impostors who were successful in only one attempt would eventually be detected at some point (albeit after a significant number of keystrokes) and that the FAR would, therefore, be somewhat less.

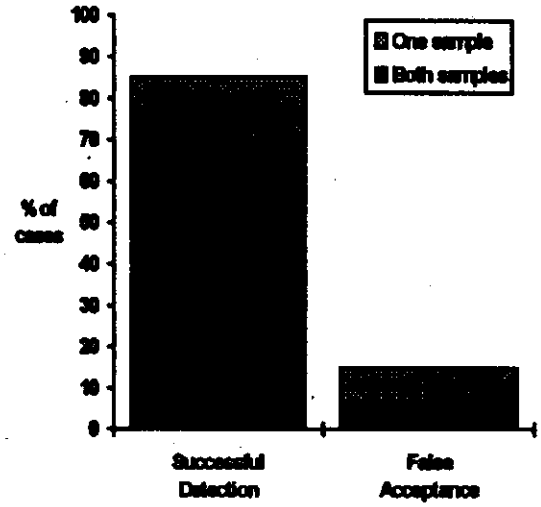

Fig. 1 : Impostor detection performance

The performance of the two detection methods employed was found to be very similar, with 49% of impostors being detected as a result of their percentage of invalid keystrokes, against 51% due to consecutive invalid keystrokes. As such, both methods can be considered to be useful authentication measures.

Given that impostor detection is actually possible, the next most important consideration is the speed with which it can be achieved (i.e. how many keystrokes would an impostor be able to enter before being detected - a factor which does not appear to have been addressed in previous studies). The experimental findings on this aspect are shown in figure 2 below. This shows the percentage of impostors detected within five distinct keystroke ranges, with cumulative values also indicated.

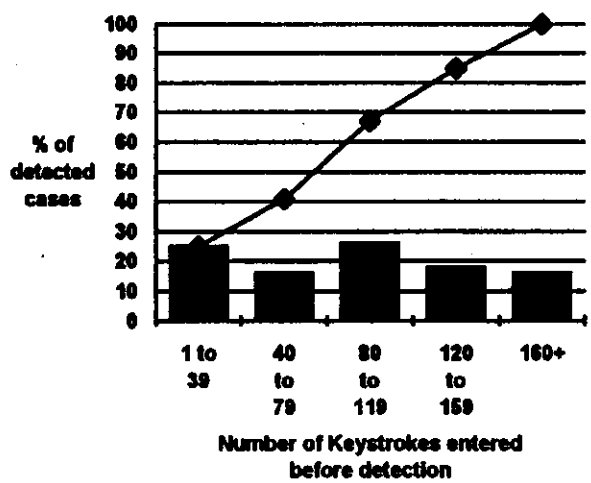

Fig. 2 : Keystrokes before impostor detection

These results indicate that the vast majority of impostors would be detected within 160 keystrokes (the equivalent of 2 standard lines of text), with detection in under 40 keystrokes in 25% of cases. Whilst this may not combat the most destructive scenarios (e.g. the immediate entry of "delete *.*" would very likely be unchallenged), it should be sufficient to identify the more common types of intruder who generally require sustained access in order to effect a serious breach.

It should also be noted that these figures essentially characterise the impostor detection performance that would be observed from the point of initial login (i.e. beginning with 0% invalid keystrokes). However, in scenarios where an impostor takes the place of a legitimate user it is likely that detection would be quicker and more frequent, as a certain percentage of invalid keystrokes would already have been registered (by the legitimate user) and, therefore, the rejection threshold would be reached more easily.

A FAR of 15% would be of less significance if the preliminary user identification phase was still to include some other form of authentication as well (e.g. a standard password system) as the combination would almost certainly serve to foil the majority of intrusion attempts.

A FRR of 0% is of course somewhat artificial, as some false rejections would be almost bound to occur in practice. However, with authentication thresholds set correctly, it is envisaged that these cases would not be frequent enough to significantly trouble legitimate users.

ADVANTAGES OF KEYSTROKE ANALYSIS

The principal advantages of the approach are improved security, reduced cost and user convenience - some of which cannot be claimed for many alternative authentication methods.

Improved security is advantageous in any information system, and is achieved here as authentication is no longer restricted to a single judgement, but may become continuous throughout the session. In addition, the biometric nature of the approach makes it more difficult for users themselves to undermine security (e.g. by allowing colleagues unauthorised access to their accounts) as typing abilities cannot be passed on to someone else in the same way as a password.

Cost advantages result from the fact that it is possible to implement the concept entirely in software (with the necessary recognition hardware already present in the form of existing PCs), whereas many frequently suggested authentication enhancement schemes (e.g. Smart cards, other biometric methods) are reliant upon specialised equipment. This makes the technique particularly suited to financially constrained environments. Cost may also be an important consideration in multimedia systems, as these require expensive base technologies which may leave little scope for additional expenditure on security.

Finally, user convenience comes from the fact that identity verification can be performed transparently, in a non-intrusive manner. This is an important consideration, particularly in a multimedia context, and is illustrated in figure 3. This shows a potential means of implementing keystroke analysis, with the existence of the monitor remaining transparent to the user unless an intrusion is suspected.

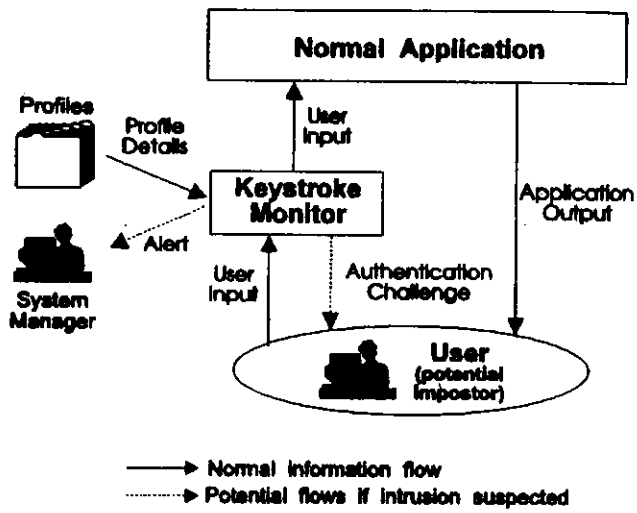

Fig. 3 : Implementation of Keystroke Analysis

This approach may again be contrasted with other authentication methods, which often place an increased burden upon the user (e.g. requiring that additional tasks be performed in order to be authenticated), which may be both time consuming and generally inconvenient (Sherman 1992). However, modern multimedia-based information systems demand security mechanisms that are as transparent as possible in order to complement the otherwise user-friendly nature of the environments.

POTENTIAL PROBLEMS

In addition to the false acceptance window, a number of further potential problems can be identified with keystroke analysis. These were outside the scope of this investigation, but will need to be addressed in future work.

- *Consistency of users*
  Users typing performance may be adversely affected by many factors (e.g. fatigue, injury, keyboard variations, interruptions), leading to departures from their profiled level.
- *Mimicry*
  It may be possible for impostors to deliberately imitate the keystroke "signatures" of legitimate users (particularly poor typists).
- *Timing accuracy*
  The concept can only be implemented in networked environments if accurate inter-keystroke timings can be obtained at the local terminals.
- *User acceptance*
  Some users may object to the idea of their activities being continuously monitored, leading to potential resistance.
- *General applicability*
  A keyboard-intensive context is required if monitoring is to be effective. However, some applications (particularly in multimedia) significantly reduce the role of the keyboard.

It would be possible to compensate for this last point by specifically profiling and monitoring words or key combinations that are still known to be frequently typed (thereby applying a static verifier approach in a dynamic context).

CONCLUSIONS

We believe that the experimental study has served to confirm the significant potential of keystroke analysis as a means of user supervision. Whilst it cannot be regarded as a panacea to the authentication issue, it should, at the very least, provide considerable protection over the use of a simple password alone.

The experimental system is currently being enhanced to enable more extensive investigation. Firstly, a full implementation of the system has been developed that runs transparently on a user workstation. In this scenario keystroke data is collected locally and then analysed by a monitoring system operating on another machine. Secondly, neural network techniques are being incorporated to allow the system to *learn* how best to conduct its analysis (for example, to determine which character digraphs are the most distinctive for a particular user). Once these enhancements have been completed, the resulting system will provide a much better indication of the concepts real-world potential.

It is considered that the FAR could be reduced by generating more representative profiles of legitimate users. Whilst this would require larger text samples (which could be collected via a background process to reduce the user burden), it would potentially allow more accurate authentication thresholds to be set and reduce the number of unrepresented digraphs in the profiles (therefore allowing more keystrokes to be analysed).

Despite this, it is felt that keystroke analysis would be most effectively used in conjunction with other forms of supervision, as a supplementary means of authentication (with passwords, or some other appropriate technique, still being employed as the primary mechanism). This would provide an opportunity to combat the FAR and could also reduce the significance of the potential problems identified above. As such, the eventual aim of the research is to incorporate the concepts into a more comprehensive intrusion monitoring framework, using a number of additional behaviour parameters to identify departures from normal system usage.

REFERENCES

Baker, H. and F.Piper. 1982. *Cipher Systems: The protection of communications.* Northwood Books, London, UK.

Bisha, S.; C.Slivinsky; and B.Hussien. 1990. "Computer-Access Security Systems Using Keystroke Dynamics", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12, no. 12 (Dec.): 1217-1222.

Card, S.K.; T.P.Moran; and & A.Newell. 1980. "The keystroke level model for user performance time with interactive systems", *Communications of the ACM* 23, no.7 (Jul.): 396-410.

Fairhurst, M.; K.Cowley; and E.Sweeney. 1994. "KAPPA Automatic Signature Verification: Summary", British Technology Group Ltd, London, UK (May).

Jobusch, D.L. and A.E.Oldehoeft A E. 1989. "A Survey of Password Mechanisms : Weaknesses and Potential Improvements. Part 1", *Computers & Security* 8, no.7: 587-604.

Joyce, R. and G.Gupta. 1990. "Identity Authentication Based on Keystroke Latencies", *Communications of the ACM* 33, no.2 (Feb.): 168-176.

Legget, J. and G.Williams. 1988. "Verifying identity via keystroke characteristics", *International Journal of Man-Machine Studies* 28: 67-76.

National Computing Centre. 1994. *IT Security Breaches Survey Summary.* National Computing Centre, United Kingdom.

Sherman, R L. 1992. "Biometric Futures", *Computers and Security* 11, no. 2: 122-133.