ADVANCED SUBSCRIBER AUTHENTICATION APPROACHES FOR THIRD GENERATION MOBILE SYSTEMS

N.L.Clarke[†], S.M.Furnell[†], P.L.Reynolds[‡] and P.M.Rodwell[†]

[†] University of Plymouth, United Kingdom
[‡] Orange Personal Communications Services Ltd, United Kingdom

INTRODUCTION

Recent years have witnessed substantial and well-documented growth within the mobile communications sector, with global mobile subscribers forecast to rise from 768m in 2001 to 1,848m by 2004 [1]. However, in parallel with this rise in ownership there has been a rise in mobile related abuse, suggesting the need for greater security measures to prevent unauthorised use. Mobile handsets are already recognised as being prime targets for theft and research findings published by the UK Home Office has estimated that over 700,000 handsets were stolen from subscribers during 2001 [2]. It can be conjectured that the more advanced capabilities of third generation (3G) handsets will make them even more desirable targets in this respect. Additionally, the increased bandwidth available in 3G will enable service providers to support significantly wider application scenarios than the rudimentary voice and data services of current cellular networks. This expansion of services and subsequent increase in private data, will demand a corresponding increase in the level of protection provided by the terminal and network hardware.

This paper presents consideration of the security requirements for 3G, considering specifically the issue of subscriber authentication. The discussion includes the results of a survey of current mobile subscribers, to determine their attitudes towards existing security measures, leading into architectural considerations and initial experimental results in relation to future techniques.

SUBSCRIBER SECURITY REQUIREMENTS FOR 3G SYSTEMS

Security provisions within 2G networks such as GSM are largely geared towards secure communication (i.e. radio interface encryption) and terminal-based authentication. The latter is achieved via the combination of the SIM (Subscriber Identity Module) and the IMEI (International Mobile Equipment Identifier), and enable the legitimacy of the terminal to be determined before allowing it to utilise the network. By contrast, relatively little attention is devoted to ensuring the legitimacy of the current user, and subscriber authentication provisions in the vast majority of devices rely upon Personal Identification Number (PIN) based methods. This facility must be enabled by the subscriber before any level of protection is provided. Even assuming this is done, the level of protection can still vary between devices depending on the handset manufacturer's implementation. Nonetheless, it can be argued that the level of security delivered by the PIN is commensurate with the requirements of the devices, as the potential consequences from theft or impostor access can be broadly categorised as financial loss (which the legitimate user can limit by reporting the theft of the phone and getting it blocked by the operator) and breach of personal privacy, due to the impostor gaining access to contact details and text messages held on the device. However, it is acknowledged that this is a fairly limited amount of information, the disclosure of which would not normally be considered highly sensitive. Stored text messages may potentially have more significance, but would not generally represent a significant body of information. By contrast, the proposed services 3G networks demand a more secure subscriber-based authentication system in order to protect personal information in the event of masquerade attacks. The primary reason for this is the hastening convergence of mobile devices with Personal Digital Assistant (PDA) devices, and the subsequent expansion in the range of possible services enabled as a result. The potential consequences of a masquerade will, therefore, become far more severe owing to the additional and more private information that these hybrid devices will store:

- financial details enabling mobile electronic commerce transactions
- electronic certificates for digital signatures
- full contact details of family and associates
- commercially sensitive miscellaneous information (e.g. scheduler/notepad files)
- medical records as a result of telemedicine or teleconsultations.

ATTITUDES TO SECURITY PROVISION WITHIN MOBILE NETWORKS

A survey was conducted to determine current mobile subscribers' attitudes towards the security provisions within their devices, and possible areas of improvement. The survey was distributed to a broad range of mobile phone users, in both printed and online formats, yielding a total of 161 responses. Full details of the results can be found in [3], but relevant summary information is presented below.

As discussed previously, the primary method of personal security within a mobile phone is the PIN. Although 89% of respondents knew about this facility, only 56% actually used it. The survey showed that 76% of respondents had phones with only a single level of security (at power on). Of those users that had the facility to PIN protect the phone in standby mode, only 36% used it. Other key findings included:

- 11% of respondents did not even know about the PIN facility. Scaled up this could represent up to 84.5 million subscribers worldwide.
- Of the 44% of respondents who did not use the PIN facility, 65% gave the reason as being its inconvenience.
- Providing additional levels of security does not necessarily mean that a subscriber will actually use them, as evidenced by those users who did not use the PIN to lock phones in standby.
- A large number of respondents, 41%, have little confidence in the protection offered by the PIN facility, believing their phone is still at risk even with the facility active.

Given these results, even in a 2G context, the prognosis for the successful application of the same methods in 3G is not encouraging. At the same time, the survey also revealed that 88% of users wanted to be able to access additional data services, such as m-commerce, video conferencing and web browsing, from their devices – highlighting the need for better authentication in future devices.

Despite their reluctance to use the existing PIN-based methods, the survey results revealed that respondents recognised the need for security, with 81% believing it would be either *good* or *very good* to have increased protection. Only 2 respondents thought it a bad idea.

Responses to the implementation of additional security showed a strong preference towards fingerprint analysis; over 70%. Voiceprint and iris scanning also achieved good responses. Analysing the results does however indicate that respondents have possibly reacted more positively to those authentication mechanisms that they are already aware of. Fingerprints have, for a long time, been known to provide a reliable means of identification. In fact fingerprint recognition has already being demonstrated by Sagem for advanced e-commerce authentication purposes in mobile phones [4]. Voice print analysis has also attracted much attention recently through computer software and also in the phone industry as a means of dialling numbers. Techniques such as ear geometry and typing style (keystroke analysis) are more recent and, as such, less research has been done on them. Preliminary results in relation to keystroke analysis on mobile phones will be considered later in the paper.

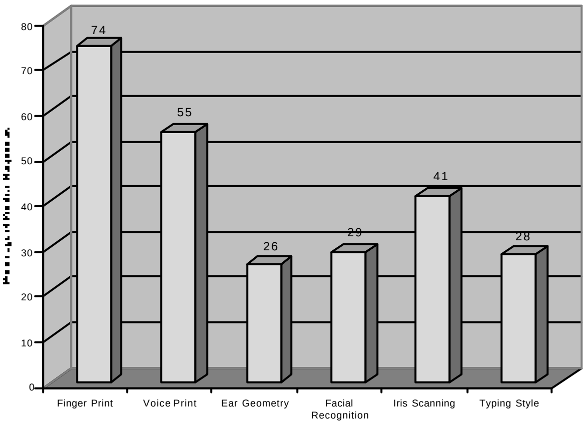

Figure 1: Positive responses to six main authentication techniques

Of the respondents who indicated that they would like more security, 63% also felt that a continuous technique during normal phone use would be a good idea. This apparent acceptance of continuous authentication is compatible with one of the stated requirements for secure 3G service provision - namely that it should be possible for service providers to "authenticate users at the start of, and during, service delivery" [5]. Authentication *during* service delivery represents a departure from the approach in 2G systems, and again implies the need for some form of transparent measure to avoid disrupting a subscriber's legitimate activity. Options for achieving this may be related to periodic or continuous supervision, utilising profiling techniques or biometric monitoring. Some authentication methods clearly lend themselves to this much better than others, and it would be important from the user acceptance perspective to ensure that chosen method(s) could be applied in a non-intrusive manner.

AN ARCHITECTURAL FRAMEWORK FOR AUTHENTICATION

Authentication could most usefully be handled within a flexible security framework, which is able to intelligently monitor the available characteristics based upon the current activity of the subscriber. For example, voice verification could be utilised during a voice call, but during an e-commerce transaction it could be replaced by other characteristics that are more appropriate to the context, such as keystroke analysis (see later discussion). The monitoring system would determine which characteristics, from those available on the terminal, should be assessed at any given time and then pass on the relevant data for analysis. The concept of such an arrangement is illustrated in Figure 2. The approach would be non-intrusive in the sense that the terminal user would be unaware of the security system unless compromise was suspected.

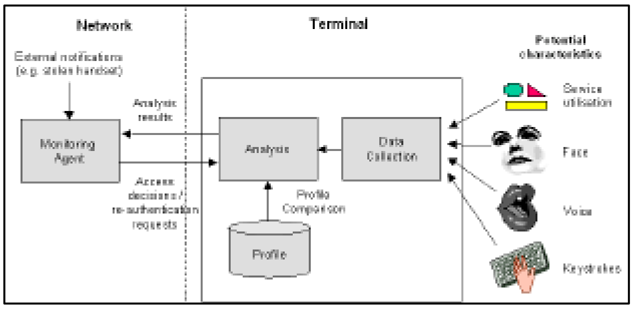

Figure 2: Potential Subscriber Monitoring Scenario

It can be seen from Figure 2 that elements of the functionality are split between the network and the terminal. However, the approach depicted in the diagram is by no means the definitive solution, and a fundamental issue is whether the security monitoring should be decentralised within the subscriber handset (and SIM), or centralised within the mobile network. Compared to GSM, UMTS does not share the concept of a home network – the 'universal' aspect suggested in the name is based upon roaming between operators to suit the service required. This raises a number of important issues, not least of which being that any security system needs to transcend the technology infrastructures of both software and hardware, raised by the different operator networks.

Terminal Issues

A terminal-based approach, where the subscriber's biometric profile is held within the handset, or more likely within the SIM card, places responsibility for the security of the profile, and consequently security of the network portal, in the hands of the subscriber. From an operator's perspective, this negates the need for additional government confidentiality legislation or network server security; there is also less need for a legal pathway of non-repudiation. As biometric authentication and supervision would be performed within the handset, it can be achieved without imposing an additional network traffic overhead and enabling authentication to be performed independent of link availability.

There are also hardware issues relevant to the terminal-centric solution. For example, any additional processing within the handset would consume valuable CPU cycles and potentially reduce the performance for other tasks. It would also have an associated impact upon battery life and subsequent recharge interval, especially if the technique were to be applied continuously.

Network issues

There are some very strong arguments for introducing a centralised network based security system. It can be argued that placing security into the handset, and effectively into the hands of the subscriber, is inherently insecure to begin with. It could, for example render the profile more vulnerable to misuse or compromise if the terminal is stolen. From the operator's perspective, holding the profile, and performing any analysis, within the network may represent a more trusted solution.

Another potential advantage offered by the network centric system is that of increased personal mobility, where a subscriber may register with any network terminal in order to access their network operators services under their personal subscription. Although GSM was originally designed to offer personal mobility via the SIM, the reality is that hardware, software, and operator network incompatibilities, as well as network locking agreements, have often restricted mobility to within a subscriber's own network. Additionally, the modern plug-in SIM generally resides behind the battery and is inherently too small and inconvenient to be of practical use as a token of true personal mobility. By replacing the token with a biometric, and centralising the authentication system, true personal mobility can potentially become a reality.

There are however some consequences to the personal mobility scenario, two primary drawbacks being increased data traffic over the wireless link and subscriber signature confidentiality. Taking the first point, the increased bandwidth of proposed next generation networks should have little trouble handling the extra handshaking required by any biometric and potentially continuously monitoring security system, especially compared to the bandwidth of a video signal. The second point of subscriber signature confidentiality reverses the trust issue mentioned earlier. Now the onus on data confidentiality is in the hands of the network operators, with subscribers trusting their sensitive and personal biometric data to not necessarily their network operator but perhaps even a third party associate.

In reality, the network-centric solution is more elegant, enabling network operators to better protect themselves against rogue users, and ultimately offering the more secure option for both of the legitimate parties.

In either scenario, the approach requires suitable biometrics to be available to monitor from the handset. Some experimental findings in relation to one such technique are presented in the next section.

MOBILE SUBSCRIBER AUTHENTICATION VIA KEYSTROKE DYNAMICS

Keystroke Dynamics is the term given to a biometric authentication technique that is able to classify, authenticate or identify a person according to their typing pattern. The use of keystroke dynamics as an authentication technique for mobile phones has two distinct advantages over other biometric techniques, in that it requires no additional hardware and can be implemented in a completely transparent environment.

The principal typing feature that is used to characterise behaviour in keystroke dynamics is the inter-keystroke latency. Many studies have taken place over the years, dating back to the 1980's, such as Joyce and Gupta [6], Leggett and Williams [7], Napier et al [8], which have all demonstrated that characteristic patterns can be discerned from an individual's typing style, which in many cases can be used to distinguish that individual from a would-be impostor. However, these studies have all centred upon the verification or identification of a user typing on a full computer keyboard. One previous study investigated authenticating users from a numerical input entered on a standard numerical keypad [9]. Although not identical to a mobile phone keypad, due to tactile qualities and typing context, the study concluded successfully, suggesting the potential for further experimental evaluation in a telephony context.

Experimental investigations and results

From the foundation provided by previous studies, a series of investigations were designed to examine the feasibility of using keystroke dynamics on a mobile handset. Three experiments were conducted, each involving a total of 16 participants:

1. the entry of a four digit number, analogous to the PINs used on current devices;
2. the entry of a series of varying telephone numbers;
3. the entry of a fixed telephone number.

The first and third investigations required the participants to enter the numeric keystroke sample thirty times, with twenty samples then being used to create a reference profile, and the remaining ten for subsequent testing. The second investigation required a larger number of samples due to the changing nature of the input string, and thus the need to train the authentication system more accurately. Fifty samples were taken, with thirty for training and twenty for testing.

Previous studies have shown neural networks to provide an effective foundation for keystroke analysis [9,10], and they have consequently been used in these investigations. The neural network structure is constructed on the feed-forward back-propagation network [11]; best exemplified for pattern recognition techniques.

Results. Brief analysis of the input data has identified two types of variance that enable or inhibit the classification process. The inter-user variance, which is essentially a measure of similarity between users and ideally would be as large as possible, and the inter-sample variance, which is a measure of similarity between individual samples of a particular user, and would ideally be zero. Neither of these variances are near their respective ideals, giving rise to the following results, as indicated in Table 1 and illustrated in Figure 3. Each investigation gives rise to a characteristic curve with two competing error rates. The False Acceptance Rate (FAR), the rate at which impostors are accepted by the system, and the False Rejection Rate (FRR), the rate at which the authorised user is rejected by the system. As can be seen from the figure it is possible to reduce one of the error rates only at the expensive of increasing the other. Therefore a decision has to be made between high security and low user acceptance (due to inconvenience), or low security and high user acceptance. Table 1 illustrates a threshold level chosen to have a compromise between error rates. The Equal Error Rate (EER) is also given as this can often be used as a performance measure when comparing biometric systems [12]. The figures in the table are, of course, averages across all of the test subjects involved. It is relevant to note, however, that some individual networks performed as well as 0% FRR and 1.3% FAR – showing that in some cases the technique has a much more significant potential than the average results would seem to imply.

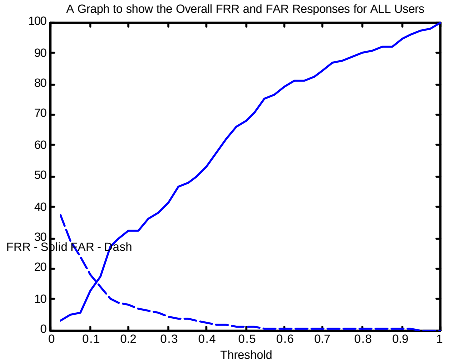

Figure 3: Performance Curves for the PIN Investigation

Investigation	*FAR (%)*	*FRR (%)*	*EER (%)*
PIN Code	18.1	12.5	15
Varying Telephone	36.3	24.3	32
Fixed Telephone	16	15	15

Table 1: Keystroke Dynamics Investigation Results

The results demonstrate the potential to distinguish authorised users from impostors, although arguably not to any great accuracy. However, the experimental procedure used in this study was performed under controlled conditions, with users all entering the same input data - a condition that is unlikely in the real world. Additionally, the design, and implementation of the neural network used for classification was primitive and un-optimised. Continuation of the study beyond this feasibility stage requires variables such as pre-processing, generalisation, network sensitivity and network configuration to be considered and analysed.

Further development of the technique will also consider other forms of user interaction with mobile handsets, in order to attempt to profile behaviour in different contexts. For instance, the way in which someone types when entering an SMS message is likely to be different to the way in which they enter a telephone number. Some users will use certain applications or functionality on the phone more often than others; will dial certain number more than others; and equally as important will not use or dial certain people or services. All of these factors could potentially be used as discriminating characteristics, leading to a stronger overall verification technique.

CONCLUSION

The capabilities of 3G mobile systems will open up a range of new service opportunities and, as a consequence, will impose new requirements for security. The survey findings indicated a weakness of the current provisions, in that the authentication technology is optional and, therefore, often unused. However, subscribers have shown the desire for additional security, and have responded positively towards a number of alternative techniques. Given that many respondents do not use the current security techniques that are available to them, it can be assumed that a non-intrusive method of authentication may prove to be most acceptable and widely utilised by end users. Viable architectural frameworks can be specified to support this, and appropriate biometric measures can be identified to provide the underlying authentication methods.

Keystroke dynamics can only be considered to be a transparent technique when the user is interacting with the keypad. If the user begins a voice or video conference call, the approach becomes an intrusive and impractical method for continuous authentication. In order to overcome this, the use of two or more biometric techniques could be used in a hybrid non-intrusive manner, i.e. keystroke dynamics for typing authentication, voice recognition whilst speaking, and facial recognition for video conferencing. The effective and intelligent management of these biometrics would provide the necessary security required in a 3G environment.

REFERENCES

[1] Giussani, B. 2001. *Roam – Making Sense of the Wireless Internet*. Random House Business Books, London, UK.

[2] BBC. 2002. “Huge surge in mobile phone thefts”, BBC News report, 8 January 2002. http://news.bbc.co.uk/hi/english/uk/newsid_1748000/1748258.stm

[3] Clarke, N.L., Furnell, S.M., Rodwell, P.M. and Reynolds, P.L. 2002. “Acceptance of subscriber authentication methods for mobile telephony devices”. *Computers & Security*.

[4] Sagem. 2000. *Sagem MC-959-ID*. Smart Card 2000 Show, London. http://www.sagem.com

[5] 3GPP. 1999. *3G Security: Security Threats and Requirements*. 3G Partnership Project. Technical Specification Group Services and System Aspects. Document 3G TS 21.133 version 3.1.0.

[6] Joyce, G., Gupta, G. 1990. "Identity Authentication Based on Keystroke Latencies". Communications of the ACM, Vol 39. pp168-176.

[7] Leggett, J., Williams, G. 1988. "Verifying Identity via Keystroke Characteristics". International Journal of Man-Machine Studies, Vol. 28.

[8] Napier, R., Laverty, W., Mahar, D., Henderson, R., Hiron, M., Wagner, M. 1995. "Optimising digraph-latency based biometric typist verification systems: Inter and intra typist differences in digraph latency distributions". Int. Journal of Human-Computer Studies, Vol. 43, pp579-592.

[9] Ord, T. and Furnell, S.M. 2000. “User authentication for keypad-based devices using keystroke analysis”, Proceedings of the Second International Network Conference (INC 2000), Plymouth, UK, 3-6 July 2000: 263-272.

[10] Cho, S., Han, C., Hee Han, D., Kim, H. 2000. "Web based keystroke dynamics identity verification using neural networks". Journal of Organisational Computing & Electronic Commerce, Vol. 10, No. 4, pp. 295-307.

[11] Bishop, C. 1995. *Neural Networks for Pattern Recognition*. Oxford University Press.

[12] Ashbourn, J. 2000. *Biometric. Advanced Identity Verification. The Complete Guide*. Springer.