User Authentication Architecture for Mobile Devices: A Composite Approach

N.L.Clarke[1] & S.M.Furnell[1,2]

[1] *Network Research Group, School of Computing, Communication and Electronic, University of Plymouth, Plymouth, United Kingdom*

[2] *School of Computer and Information Science, Edith Cowan University, Perth, Western Australia*

info@network-research-group.org

Abstract

With the ever-increasing functionality and services accessible via mobile telephones, there is a strong argument that the level of user authentication implemented on the devices should be extended beyond the Personal Identification Number (PIN) that has traditionally been used. This paper proposes the use of more advanced biometric methods as an alternative, and proceeds to explain how, through the use of a portfolio of authentication techniques it is possible to provide a robust, accurate and transparent authentication mechanism for mobile devices. An Intelligent Authentication Management System (IAMS) is proposed that provides a continuous confidence level in the identity of the user, removing access to sensitivity services and information with low confidence levels and providing automatic access with higher confidence levels.

Keywords: User Authentication, Biometrics, Security

1. INTRODUCTION

Mobile devices such as cellular phones and Personal Digital Assistants (PDAs) are now allowing access to an increasing range of data-centric services. Users of such devices can now pay for products using micro-payments, surf the Internet, buy and sell stocks, transfer money and manage bank accounts. In order to enable delivery of such services, mobile devices have become increasingly powerful: phone handsets in particular have evolved from relatively basic terminals, that would handle analogue telephony communications, to digital handsets capable of providing a host of data-centric services, turning the handset into a multimedia, multi-purpose, mobile communications tool, providing much of the functionality of today's PDAs.

With more applications being accessible, and more data being stored, it can be argued that users are now carrying devices that require correspondingly greater levels of protection. Specifically, the reasons for this will include:

1. More technologically advanced mobile handsets – future handsets will be far more advanced than current mobile phones, increasingly incorporating much of the functionality of PDAs, MP3 players, and other portable devices. As such, they will be more expensive and attractive to thieves, resulting in a financial loss to the subscriber.

2. Availability of data services – cellular and wireless networks will provide the user with the ability to download and purchase a whole range of data services and products that would be charged to the subscriber's account. Theft and misuse of the handset would result in financial loss for the subscriber.
3. Sensitive Information – Devices will store much more information than current handsets. Proposed applications could result in a whole range of personal, financial and medical information being held, alongside records of business and personal communications conducted by the user (e.g. via emails and multimedia messages). As a simple example of how such evolution has already occurred we need only consider the contact list on a typical handset. Whereas devices a few years ago would simply hold names and phone numbers, current devices can store full home and business address details for each contact, as well as multiple phone numbers, date of birth and other family information (e.g. names of spouses and children). As such, the compromise of the device would reveal a far greater degree of personal data.

The increasing requirement for protection is evidenced by a survey of 230 business professionals, which found that 81% considered the information on their PDA was either somewhat or extremely valuable. As a result, 70% were interested in having a security system for their PDA, with 69% willing to pay more for a PDA with security than one without [1]. With this in mind, it is relevant to consider the degree to which related security measures are already provided and utilised. Currently, the most widely deployed authentication methods are passwords and PINs (Personal Identification Numbers) - secret knowledge approaches that relies heavily upon the user to ensure continued validity. For example, the user should not use the default factory settings, share their details with others, or write the information down. However, the poor use of passwords and PINs has been widely documented [2], and many mobile users do not even use the security which is available. For example, a survey assessing authentication and security practices on mobile handsets found that 34% of the 297 respondents did not use any PIN security [3]. In addition, even for those respondents who did use the PIN at switch-on only, 85% would leave their handset on for more than 10 hours a day, thereby mitigating any security the PIN might provide. Interestingly however, it would appear users do have an appreciation of security, with 85% of respondents in favour of additional security for their mobile device. These findings introduce an interesting and somewhat contradictory view of security, with users willing to adopt new security but not willing to utilise current functionality.

It is widely recognised that authentication can be achieved by utilising one or more of three fundamental approaches: something the user *knows* (password); something the user *has* (token) and something the user *is* (biometric) [4]. The downside of the first approach has already been indicated, with the use of PINs found to be somewhat lacking in practice. Similarly to secret knowledge techniques, token based approaches fundamentally rely upon the user to remember something to ensure security, with the token needing to be physically present in order to access the device. However, it is considered that this does not lend itself particularly well to the mobile device context either. The most likely scenario is that users would simply leave the token within the mobile handset for convenience. Indeed, this is the case with the Subscriber Identity Module (SIM) in mobile handsets, which already exists as a token and could be physically removed from a phone when not in use. Users typically do not do this because it is inconvenient, and increases the risk of losing or damaging the SIM card. In contrast to the other methods, the third approach to authentication does not rely upon the user to remember anything – it just requires them to be themselves. Such techniques are collectively known as biometrics, and it is here that the most suitable alternatives for going beyond the PIN can be found.

This paper introduces the concept of advanced user authentication for mobile devices through the application of biometrics in a composite, transparent and continuous fashion. Given the wide variety of mobile devices that exist, with different hardware configurations and processing capabilities, it is clear that no single authentication technique would be suitable for all situations. Rather it would be far more appropriate to provide a suite of authentication techniques that could provide an overall authentication approach for mobile devices. This paper describes how such an approach can be achieved, fulfilling the objectives of a more secure, transparent and continuous authentication mechanism. The paper begins by introducing an architectural overview of a composite authentication mechanism, describing the key functionality before proceeding to present the two security processes that operate at the core of the system. The function of the security processes is to maintain the trade-off between the security required by the system and the level of user convenience - ensuring the users interaction with the mobile device is not adversely affected. The paper proceeds to illustrate the performance expectations of the system versus traditional standalone authentication mechanisms and concludes with an evaluation of the approach.

2. ACHIEVING A COMPOSITE AUTHENTICATION MECHANISM

It is envisaged that a successful authentication mechanism for mobile devices must satisfy a number of objectives:

- to increase the authentication security beyond secret-knowledge based approaches;
- to provide transparent authentication of the user (within limits) to remove the inconvenience factor from authentication;
- to provide continuous or periodic authentication of the user, so that confidence in the identity of the user can be maintained during usage of the device rather than simply at switch on;
- to provide an architecture that would function (to one extent or another) across the complete range of mobile devices, taking into account the differing hardware configurations, processing capabilities, and varying levels of network connectivity.

The underlying mechanism utilises a combination of secret knowledge and biometric techniques within an appropriately flexible framework. The framework operates by initially providing a baseline level of confidence in the user, using secret knowledge approaches, which progressively increases as the user interacts with their device and biometric samples are captured. Although user authentication will still begin rather intrusively (e.g. when the device is switched on for the first time), with the user having to re-authenticate periodically, the system will however quickly adapt, and as it does so the reliance upon secret knowledge techniques is replaced by a reliance upon biometrics – where the user will be continuously and non-intrusively authenticated. The result is a highly modular framework that can utilise a wide-range of standardised biometrics, and which is able to take advantage of the different hardware configurations of mobile devices – where a combination of cameras, microphones, keypads etc can be found. Therefore any given device will have a range of authentication techniques that the system can utilise to maintain security. It is important to note, however, that due to the different performance rates biometrics achieve, each of the techniques is assigned a confidence level tied to the performance rate. This will give rise to mobile devices being able to provide different levels of security based upon the hardware available to capture the biometric samples.

Architecturally this system could take many forms, but it is envisaged a number of key components would be required, such as an ability to capture and authenticate biometric samples, an intelligent controller, administrative capabilities and storage of the biometric profiles and authentication algorithms. To satisfy these requirements the authors propose the Intelligent Authentication Management System (IAMS). Built around a server-client topology, the system also has the flexibility of operating in an autonomous mode to ensure security is maintained even during periods with limited or no network connectivity. Figure 1 outlines the functional components of the server topology. The Authentication Manager has overall control of the authentication system, determining both when authentication should take place and what the current state of security is. The process engines provide the computational power of the system, with an Authentication Engine to authenticate users, a Biometric Profile Engine to generate and train the relevant biometric templates required for subsequent classification, and a Communications Engine to communicate and synchronise data with the client device. To supplement these process engines, a number of storage elements are utilised. As the Operating System (OS) and hardware on mobile devices tend to vary considerably, devices will not automatically be supported. The Hardware Compatibility database contains information about which mobile devices are configured to work with the architecture, along with a list of supported biometrics. The system administrator will utilise this information, in addition to a number of system parameters to generate a client profile, which is stored in the Client database. This database holds a master list of clients enabled, along with individual user information such as performance rates, confidence levels and history of the relevant authentication techniques

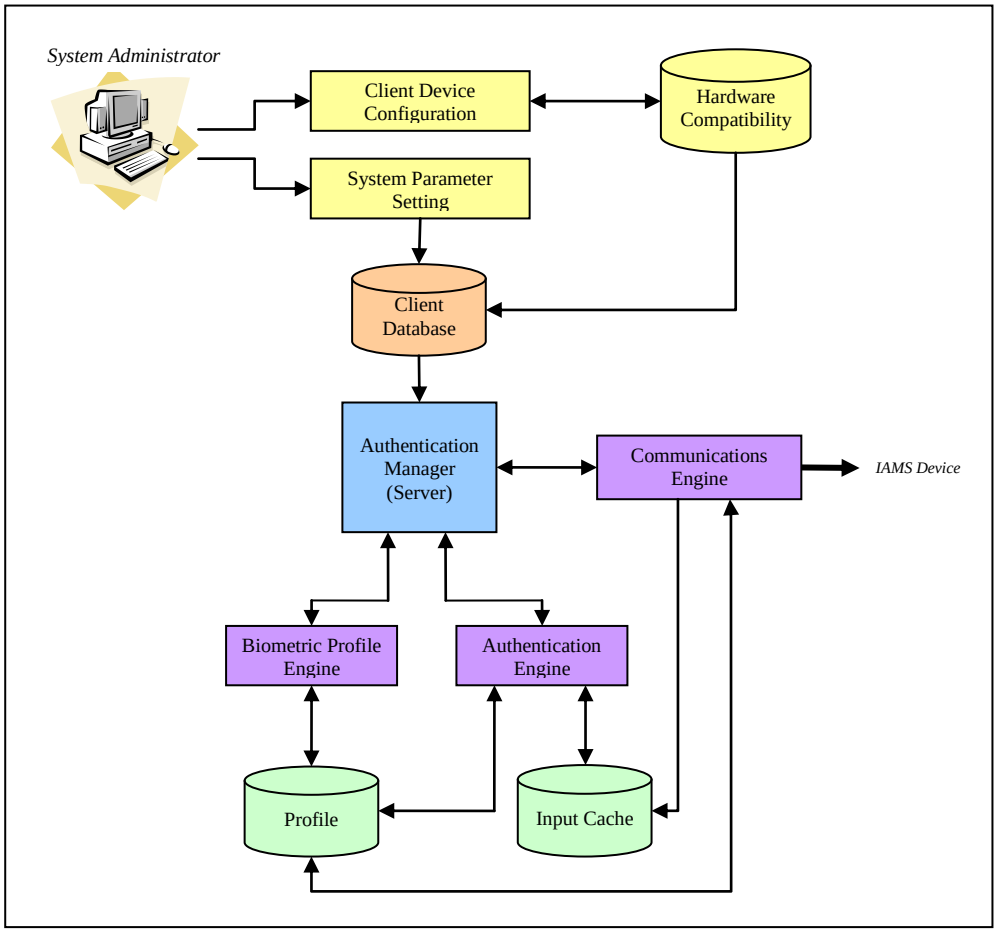

Figure 1: IAMS Server Architecture

The majority of the device topology, as illustrated in Figure 2, is identical to the server architecture, with the operation of the process engines, storage elements and Authentication Manager remaining (in principle) the same. The device topology does however introduce a number of additional components that provide the input and output functions of the system. The fourth process engine in the form of the Data Collection Engine is included on the device topology and provides the input mechanism, which collects and processes users' device interactions. The output components consist of an Intrusion Interface and Security Status. The former provides the IAMS to OS connection for restricting user access and provides user information as and when required, and the latter provides an overview to the system integrity and security of the device.

The implementation of the architecture will differ depending upon the context that a device is being used within. For instance, in a standalone implementation the device has no use for the Communications Engine – as no network exists to which it can connect. Meanwhile, in a client-server topology the components required will vary depending upon the processing split between the server and client. There are numerous reasons why a network administrator may wish to split the processing and control of IAMS differently, such as network bandwidth and availability, centralised storage and processing of the biometric templates, and memory requirements of the mobile device. For example, in order to minimise network traffic, the network administrator may require the host device to authenticate user samples locally, or conversely, the administrator may wish the device to only perform pre-processing of input samples and allow the server to perform the authentication, thus removing the majority of the computational overhead from the device, but still reducing the sample size before transmitting across the network. More detailed information on this architecture can be found in [3].

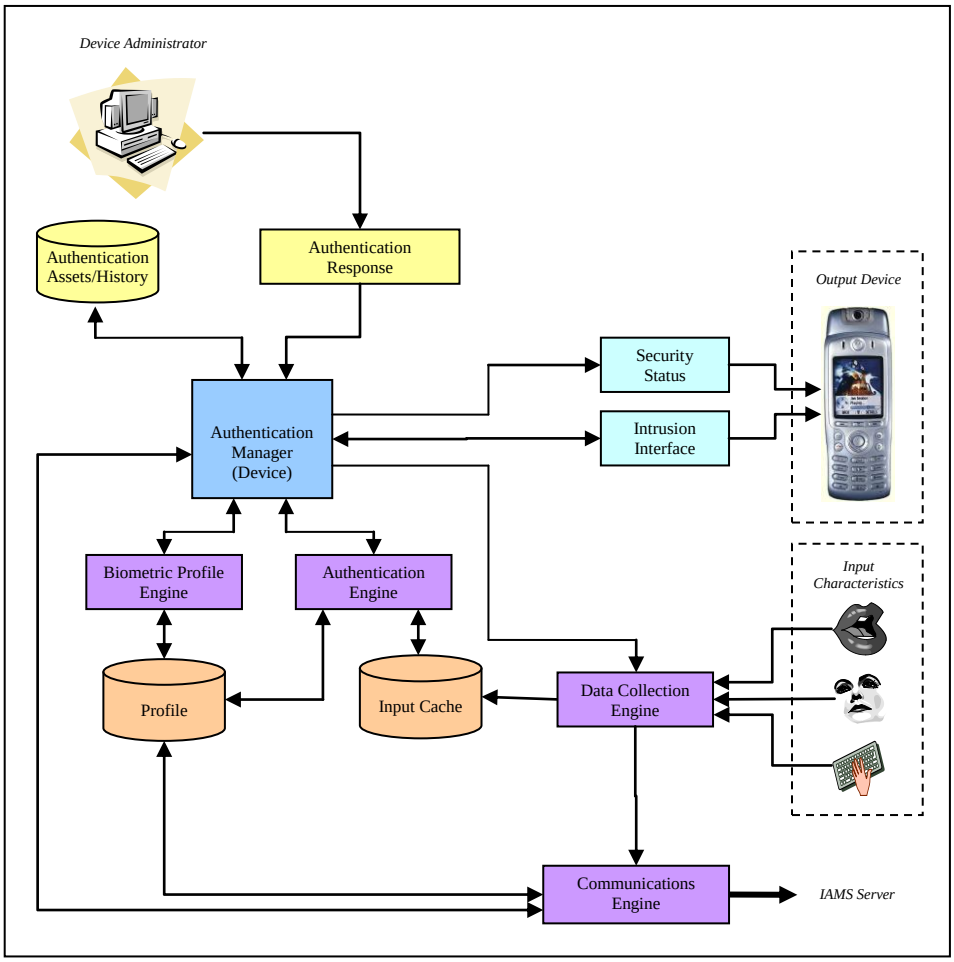

Figure 2: IAMS Client Architecture

3. SECURITY PROCESSES WITHIN IAMS

To maintain security within this system, two security mechanisms are considered imperative:

1. Alert Level
2. System Integrity

The Alert Level is controlled via the process illustrated in Figure 3, and has four possible states: normal, authenticate on next input, authenticate with strong biometric, and lock device from use. The level of authentication required is increased when previous requests are failed, until the point at which the device is locked (requiring an administrative password or PUK (Personal Unblocking Key) code from a cellular network provider before the user can regain access). The general operation of the system is to periodically poll the device with an authentication request. The system will subsequently retrieve the last and highest (in terms of confidence value – different authentication techniques will be more reliable than others) set of user's inputs (i.e. a camera image from a video conference call, or a sound file from voice dialling). If the request is passed, the system goes back into a monitoring mode. If not, then the system makes another authentication request, but using the remaining data that has been stored within a specified time using the highest confidence level technique. If no additional data is present or the response is a fail, the system increases the Alert Level and will request

authentication on the next input sample to the device – the user would now not be able to use any of the more sensitive and protected services a mobile device might have until this stage had been completed. If the user passes this, or any of the previous stages, then the system goes back into a monitoring/collection mode. If the request is a fail, however, the system will issue an explicit authentication request to the user. The system will use a biometric technique with the highest confidence value in order to minimise the risk of a false acceptance. If, and only if, no biometric techniques are supported by the device, or no templates exist with a high confidence value, then the user will be requested to enter their PIN, password or answer a cognitive question. If they pass this, and the PIN or password has a corresponding keystroke analysis template, then this will also be utilised in order to provide a stronger two-factor authentication mechanism [5]. If the keystroke analysis template exists, and the user passes the biometric authentication, then the system will revert back to a monitoring mode. If the biometric fails, or the template does not exist, then the Alert Level will remain at a heightened status of "authenticate on next input". If an intrusive authentication request is passed, the previous biometric samples that were failed are deemed to be in fact from the authorised user and incorrectly failed. As such, these samples are added to a Profile database for subsequent re-training and are not deleted.

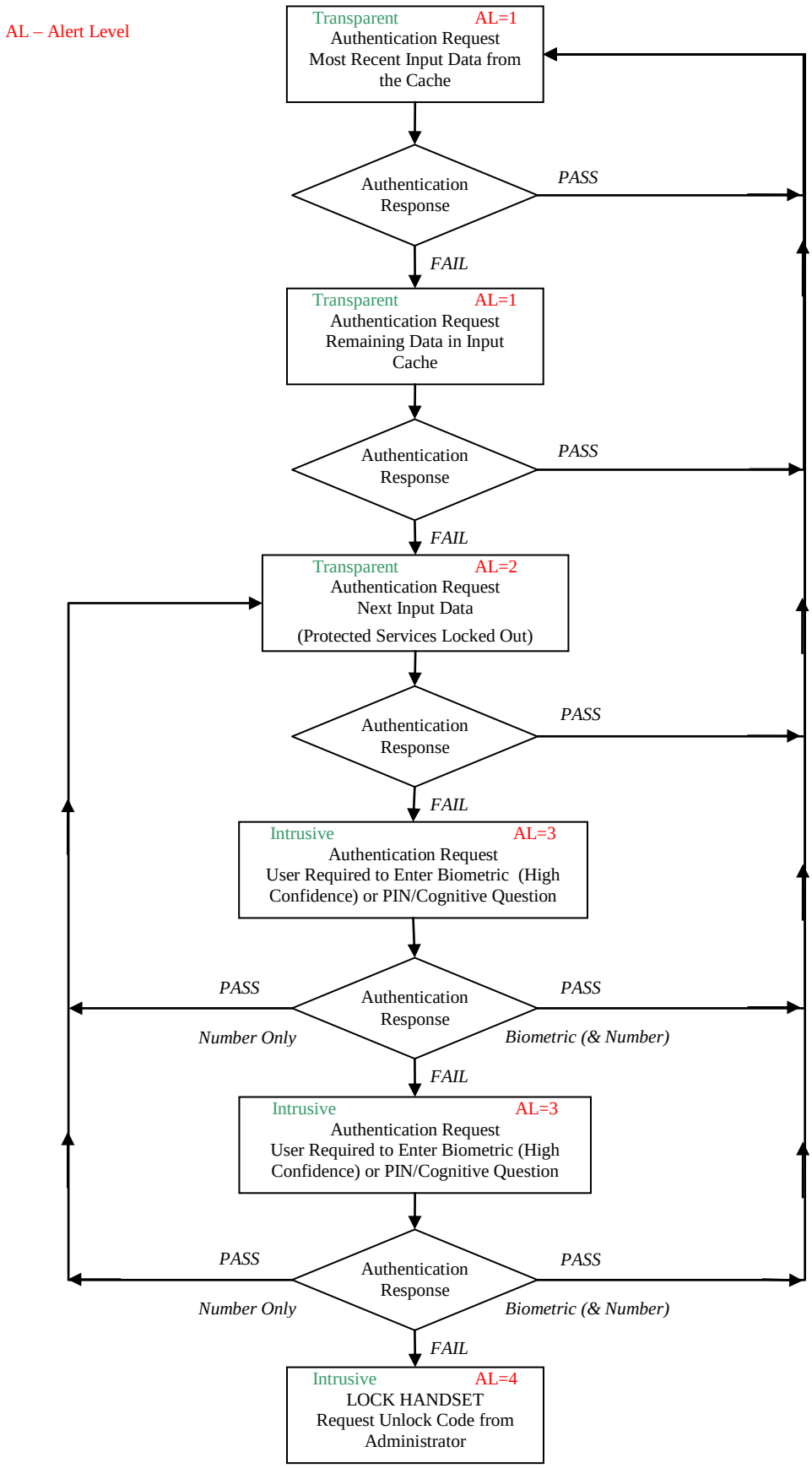

Figure 3: The authentication process, showing the increase in the Alert Level in response to failed authentication attempts

The Alert Level is inherently biased toward the authorised user, as they are given three non-intrusive chances to authenticate themselves correctly, with two subsequent additional intrusive chances. This enables the system to minimise inconvenience from the legitimate user's perspective. However, due to the trade-off between the error rates, this has a detrimental effect on the false acceptance rate, increasing the probability of wrongfully accepting an impostor every time an authentication request is sent. For an impostor to be locked out of the device they must have their authentication attempt rejected 5 consecutive times. However, this is where the second security mechanism operates. The probability of an impostor continually being accepted by the system becomes very small as the number of authentication requests increases. This indicates that the impostor would be identified as such more often than not (even if not consecutively as required by the Process Algorithm). The System Integrity is a sliding numerical value between -5 and +5[1], with -5 indicating a low security, 0 a normal 'device switch-on' level, and +5 indicating a high security level. The System Integrity changes depending upon the result of the authentication requests and the time that has elapsed between them. Each of the biometric techniques confidence levels are given a number which is added or subtracted from the System Integrity dependent upon whether the technique has passed or failed the input sample, up to a defined maximum level (to ensure weak authentication techniques do not provide a mechanism for obtaining high System Integrity values). This ensures a user with a System Integrity Level of 5 has not only had consistent successful authentication requests during their session, but has also recently been authenticated by a biometric technique with a high confidence value. Access to the applications and services found on a mobile device can then be tied to the System Integrity level, such that immediate access is only given to a user if they have the required level or greater. For instance, when a user attempts to access a protected service or file location, if they do not have the required integrity level, the system will intrusively request them to authenticate using a technique with the required confidence value to permit access to the file or service. In this case, should the Alert Level reside at "normal" or "authenticate on next input", the authentication request can be used as the next authentication request in the Process algorithm. Should the request succeed then the user is given access to the information or service they require. However, should the request fail, the user will be blocked from using the file or service and the Alert Level will proceed to the next stage. The trade-off existing within these processes is between user convenience and device misuse. Although an impostor will not be rejected from the system immediately under this process, the degree of misuse has been limited by the presence of the System Integrity. In a practical situation, it is likely an impostor will be able to make a telephone call or send a text message before the system locks down (the actual range of services available to the impostor will largely depend upon the authentication techniques available). However, all of the key sensitive and expensive services will be locked out of use. By permitting this limited misuse of the device, it is possible to achieve a much higher level of user convenience at minimal expense to the security.

4. PERFORMANCE CHARACTERISTICS

The performance of such a composite authentication mechanism will be largely dependent upon the authentication techniques available to a particular mobile device. Those with stronger techniques will be more capable of successfully detecting an authorised and unauthorised user than their counterparts. Table 1 illustrates the performance achieved for a

[1] The boundaries defined on the numerical scale are only provided as a suggestion. Practical evaluation might result in a redefinition of these limits.

number of test cases based upon the authentication techniques that could potentially be available given their specific hardware configuration. The devices themselves are illustrated for reference in Figure 4. As this composite mechanism involves multiple authentication requests and multiple authentication techniques it is difficult to obtain a single FAR and FRR. Table 1 presents the FRR at the point where the authorised user is essentially locked-out from using the device, and the FAR of an unauthorised user achieving a System Integrity level of +5, which would permit the user to access the most sensitive services and information. The FAR and FRR for the authentication techniques which the subsequent system level performances were calculated were derived from studies performed on keystroke analysis [6] and the National Physical Laboratory [7].

(a) Sony Ericsson T68i

(b) HP IPAQ H5550

(c) Sony Clie PEG NZ90

Figure 4: Mobile Devices

Worked Example:

FRR at Stage 4 of the Process Algorithm:

Best Case Probability = Voice FRR x Voice FRR x Voice FRR x PIN FRR x PIN FRR
= 0.04 x 0.04 x 0.04 x 0.4 x 0.4
= 0.0000102
*= **0.001%***

Worst Case Probability = Tele FRR x Tele FRR x Tele FRR x PIN FRR x PIN FRR
= 0.29 x 0.29 x 0.29 x 0.4 x 0.4
= 0.00039
*= **0.04%***

Mobile Device	Authentication Techniques	FRR at Stage 4 of the Process Algorithm (%)	FAR at a System Integrity level of +5 (%)
Sony Ericsson T68	Keystroke Analysis Voice Verification	0.001-0.4	0.000001-0.00002
HP IPAQ H5550	Facial Recognition Fingerprint	0.00003-0.0001	0.00000007-0.0000008

	Scanning Voice Verification		
Sony Clie PEG NZ90	Facial Recognition Keystroke Analysis Voice Verification	0.0002-0.4	0.0000008- 0.00002

Table 1: Composite Authentication Performance

Even with devices such as the cellular handset, with limited authentication techniques, the levels of FAR and FRR achieved are sill stronger than many individual authentication techniques alone, with a (worst case) probability of an authorised user incorrectly being rejected of 0.4% (equivalent FRR) and a (worst case) probability of an unauthorised user gaining entry to the most sensitive services of 0.00002% (equivalent FAR).

The results from the theoretical system performance illustrate how difficult it is for an impostor to obtain access to sensitive services, with a FAR in the range of 0.00000007-0.000001% compared with the best FAR of 0.1% using a fingerprint technique. The false rejection probability has also improved, with a worst case of 0.4% and a best case of 0.00003%. Although it is difficult to directly compare the performance of this composite system against individual techniques (as the probability of successfully authenticating a person depends on various stages of the security algorithms), a comparison of these results against individual results, as presented in Table 2, illustrates the improvement in performance this mechanism experiences.

5. CONCLUSIONS

With mobile device functionality increasing, the ability to perform suitable user authentication becomes evermore important. Existing PIN-based techniques are under-utilised, and in any case provide an inadequate level of protection when compared to the sensitivity of data and services accessible through the devices. Individual techniques such as keystroke analysis can provide valuable enhancements in certain contexts, but are not suited to all users and scenarios. However, the use of multiple authentication techniques, bound within a wider framework, enables the system to compensate for potential weaknesses of one technique by using the strengths of others.

In a worst case, the proposed mechanism enhances PIN/password-based authentication with keystroke analysis that periodically asks the user to re-verify their identity. At best, this mechanism can provide completely transparent authentication of the authorised user throughout the duration of the day protecting key services and information from misuse. Both scenarios increase the level of authentication beyond that currently available from the standard point-of-entry PIN/password technique.

Through adding a level of intelligence to the authentication process, it is no longer a matter of providing a pass or fail response, but a probability level indicating the confidence the system has in the identity of the user, with the system's behaviour becoming dependent upon the

result. With a low confidence, the system removes automatic access to key services and information and increases the level of monitoring of the user. With a high confidence level, the user has the ability to interact and access the complete range of services and applications provided by the mobile device without hindrance.

6. REFERENCES

[1] Shaw, K. 2004. "Data on PDAs mostly Unprotected". Network World Fusion. http://www.nwfusion.com/

[2] Denning, D. 1999. *Information Warfare & Security*. ACM Press, US.

[3] Clarke, N. 2004. *Advanced User Authentication for Mobile Devices*. PhD Thesis. University of Plymouth, UK.

[4] Nanavati, S., Thieme, M., Nanavati, R. 2002. *Biometrics. Identity Verification in a Networked World*. John Wiley & Sons.

[5] Monrose, R., Reiter, M., Wetzel, S. 1999. "Password Hardening Based on Keystroke Dynamics". Proceedings of the 6th ACM Conference on Computer and Communication Security, pp. 73-82 Singapore, November 1999.

[6] Clarke, N., Furnell, S., Lines, B., Reynolds, P. 2003. "Using Keystroke Analysis as a mechanism for Subscriber Authentication on Mobile Handsets*". Security and Privacy in the Age of Uncertainty*, International Federation of Information Processing, pp. 97-108.

[7] Mansfield, T., Kelly, G., Chandler, D., Kane, J. 2001. "Biometric Product Testing: Final Report". Crown Copyright.

COPYRIGHT