Flexible and Transparent User Authentication for Mobile Devices

Nathan Clarke, Sevasti Karatzouni, and Steven Furnell

Centre for Information Security & Network Research, University of Plymouth, Plymouth, PL4 8AA, United Kingdom
info@cisnr.org
http://www.cisnr.org

Abstract. The mobile device has become a ubiquitous technology that is capable of supporting an increasingly large array of services, applications and information. Given their increasing importance, it is imperative to ensure that such devices are not misused or abused. Unfortunately, a key enabling control to prevent this, user authentication, has not kept up with the advances in device technology. This paper presents the outcomes of a 2 year study that proposes the use of transparent and continuous biometric authentication of the user: providing more comprehensive identity verification; minimizing user inconvenience; and providing security throughout the period of use. A Non-Intrusive and Continuous Authentication (NICA) system is described that maintains a continuous measure of confidence in the identity of the user, removing access to sensitive services and information with low confidence levels and providing automatic access with higher confidence levels. An evaluation of the framework is undertaken from an end-user perspective via a trial involving 27 participants. Whilst the findings raise concerns over education, privacy and intrusiveness, overall 92% of users felt the system offered a more secure environment when compared to existing forms of authentication.

1 Introduction

Recent years have witnessed a considerable increase in the power and capabilities of mobile devices, with the users of today's smartphones and PDAs having access to a far richer range of features and functionality than they enjoyed a few years ago. Although offering a number of clear benefits, this transition poses serious security considerations for mobile users. With the ability to access and store a wide variety of more sensitive information, the need to ensure this information is not misused or abused is imperative. Whereas the replacement cost arising from loss or theft might previously have been the principal risk associated with mobile devices, unauthorized access to its data could now be a far more significant problem (introducing threats ranging from personal identity theft through to serious corporate loss and increasingly liability).

Given the changing nature of the mobile device and network, it is necessary to consider whether the current authentication on mobile handsets is capable of providing the level of security that is necessary to meet the changing requirements. Even with increasingly large amounts of literature suggesting that secret-knowledge techniques are

D. Gritzalis and J. Lopez (Eds.): SEC 2009, IFIP AICT 297, pp. 1–12, 2009.

ineffective [5,8], the Personal Identification Number (PIN) is still the most widely used approach on mobile devices. 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 [10].

Looking beyond secret-knowledge, two other forms of authentication are available, namely tokens and biometrics. However, only the latter are able to realistically provide more secure mechanisms for user authentication. Tokens rarely authenticate the user, but rather authenticate the presence of the token; with the assumption being the legitimate user is in possession of the token. Moreover, its application within a mobile device context would require a user to remember both the device and token or more commonly simply leave the token in situ within the device (e.g. the use of the SIM card). However, given the evolving nature of mobile devices, simply replacing one authentication mechanism with another is arguably not sufficient. Rather, only through an analysis of the requirements can an effective solution be proposed.

This paper presents the results from a two-year study investigating and proposing a new user authentication approach for mobile devices. The paper begins by presenting the research undertaken to develop and understand the requirements in order to derive the objectives of the system. Section 3 then broadly describes the proposed framework; in particular, focusing upon the key processes that enable security and usability. Section 4 presents the end-user trial of the system, with the final section describing the conclusions and future work.

2 Analysis of Stakeholder Requirements

In order to establish an understanding of stakeholder requirements, a qualitative and quantitative research methodology was undertaken. Stakeholders were largely divided into two groups: end-users of mobile devices and managers of mobile devices/networks (e.g. network operators, system administrators). It was determined that the end-user group, representing the principle stakeholder group, it would be assessed both quantitatively through a survey and qualitatively through a focus-group. It was felt, due to the specialist nature of the other group of stakeholders and getting sufficient access to them, a qualitative focus-group based methodology would be most appropriate. To this end, two activities were undertaken:

1. A survey of end-user attitudes and opinions towards current and future forms of user authentication technologies. A total of 297 participants took part in the survey and complete published results can be found in [1].
2. A focus group activity involving all stakeholders. A total of 12 participants took part and a series of questions were put forward regarding current authentication and the security requirements of current and future services. In order to maximise the usefulness of the focus group, this activity was devised based upon the analysis and findings of the survey. Detailed information on the focus group and its outcomes can be found in [6].

In summary, the survey found that 34% of the 297 respondents did not use any PIN security. 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 undermining any security the PIN might provide. Interestingly, however, it would appear that users do have an appreciation of security, with 85% of respondents in favour of additional security for their device.

Within the focus group these findings were not so evident, with the end-user group finding it difficult to understand why such protection was required. Whilst this was somewhat expected given current usage (with most end-users simply using their device for telephony or texting); the few enterprise-level users of devices (using advanced features such as email and corporate network access) that participated in the focus group understood and agreed with the need for better protection. Moreover, once the possible future uses of the mobile devices were explained to end-users (for instance micro-payments and accessing back accounts), they also understood the need for better security. From the other stakeholder groups, it became evident that existing controls were not sufficient, with system administrators particularly concerned regarding the increasing integration of mobile devices within their organisations network and the effective control and management of them.

When taking the feedback into consideration and reflecting upon all the other requirements, such as: varying hardware configurations and processing capabilities of mobile devices; network versus device centric operation; an enormous end-user population of approximately 2.7 billion [7]; privacy of end-user data (particular biometric based); it became evident that a flexible authentication scheme would be preferable. As no single authentication technique would be suitable for all situations it would be far more appropriate to provide a suite of authentication techniques within an appropriate framework that could provide an overall authentication approach for mobile devices.

From the analysis of stakeholder requirements, it is envisaged that a successful authentication mechanism for mobile devices must address a number of requirements:

- 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 the confidence in the identity of the user can be maintained throughout the life of the device;
- to link security to service provision, so that for instance the risk associated with sending a text message and accessing a bank account can be understood and be incorporated with the decision making process;
- 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 network connectivity.

From these requirements a Non-Intrusive and Continuous Authentication (NICA) system was devised.

3 Non-Intrusive and Continuous Authentication (NICA) for Mobile Devices

NICA operates by utilising a combination of secret knowledge and biometric techniques within a flexible framework. The framework operates by initially establishing a baseline level of security, using secret knowledge approaches, which progressively increases as the user interacts with their device and biometric samples are captured. Although user authentication will 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 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.

3.1 Proposed Framework

Architecturally this system could take many forms, but it is proposed that 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. Although principally conceived around a client-server 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 architecture.

The client-side includes all of the components illustrated in figure 1 and the server-side architecture includes all but the input and output components (the Data Collection engine, Security Status and Intrusion Interface). 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 NICA 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.

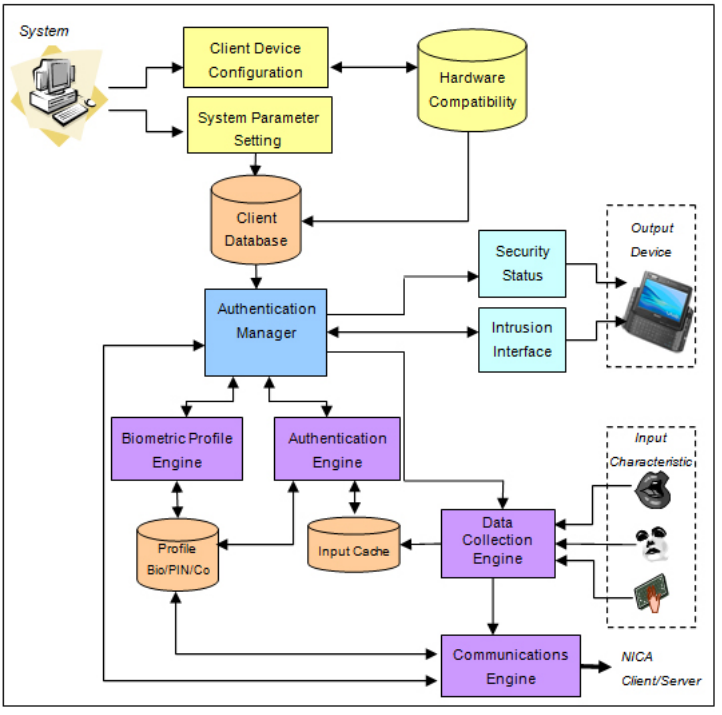

Fig. 1. NICA Architecture

3.2 Security and Usability Processes

The principal objective of the system is to maintain the level of security required commensurate with the services being provided by the device and to achieve this in a user friendly and convenient fashion. To this end, two key processes operate to ensure this:

•Authentication Confidence Level •Alert Level

The Authentication Confidence Level (AuCL) process assists in ensuring security through maintaining a continuous level of confidence in the identity of the user. It is a sliding numerical value between -5 and +5 (these values are merely suggestions rather than definitive values), with -5 indicating low security, 0 a normal 'device switch-on' level, and +5 indicating a high security level. The confidence level is modified depending upon the result of authentication requests and the time that has elapsed between them. The magnitude to which the AuCL is modified is dependent upon the authentication technique – recognising that a difference exists between strong biometrics such as face and fingerprints and weaker biometrics such as keystroke analysis. A protection mechanism also exists to ensure a user utilising a weaker biometric is unable to achieve

high levels of confidence. This confidence level is then associated with the services and information the device is capable of providing, so that a user who already has sufficient confidence to access a service is automatically provided access. However, should a user request access to a service for which they currently do not have sufficient confidence for, a subsequent intrusive authentication request will be made.

The Alert Level is the second of the key security processes working at the core of this framework. Its purpose is to ensure continuous identity verification of the user in a transparent and therefore convenient fashion. There are six levels (depicted in table 1) with the level of authentication security being increased until the device is locked (requiring an administrative password or PUK code from a cellular network provider). The number of stages was determined by a compromise between requiring a good level of user convenience and better security. Through mixing transparent and intrusive authentication requests into a single algorithm it is intended that the majority of authorised users will only experience the transparent stages of the algorithm. The intrusive stages of the algorithm are required to ensure the validity of the user by utilising the stronger authentication tools before finally locking the device from use.

The Alert Level algorithm is inherently biased toward the authorised user, as they are given three non-intrusive chances to authenticate correctly, with two subsequent additional intrusive chances. This enables the system to minimise inconvenience from the authorised user 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. With this in mind, for an impostor to be locked out of the device they must have their authentication request rejected a maximum of 5 consecutive times. However, this is where the companion process, the AuCL, has a significant role. The probability of an impostor continually being accepted by the framework becomes very small as the number of authentication requests increase. This would indicate that the impostor will be identified correctly more often than not (even if not consecutively as required by the Alert Level), reducing the AuCL value to a level where the majority if not all of the services and file access permissions have been removed – essentially locking the device from any practical use. In a practical situation, it is likely an impostor will be able to undertake tasks with a low risk, such as, a telephone call or sending a text message, for a short period of time before the system locks down. 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.

3.3 NICA Prototype

A proof-of-concept prototype was developed in order to assess the effectiveness of the proposed framework. The prototype, based upon the client-server model, comprised of four software systems:

1. Authentication Manager – providing the entire server-side operational functionality, including, biometric profiling, authentication and data synchronization.
2. Administrative Console – containing all the administrative and system settings, and providing a visualisation of active devices and their operational status.

Table 1. Escalation of the alert level

Alert Level	NICA Authentication action
1	Perform transparent authentication using the most recent data in input cache.
2	Perform transparent authentication using remaining data in input cache.
3	Perform transparent authentication using the next available user input.
4	Issue an intrusive authentication request using a high-confidence method.
5	Issue a further intrusive authentication request using a high-confidence method.
6	Successive authentication failure invokes a system lock.

3. Client-Side Interface – providing the simulated mobile handset functionality, data capture and intrusion control.
4. Databases – an SQL server containing all the server-side databases.

The hardware utilised for the prototype included a Samsung Q45 that acted as the Authentication Manager, Console Manager and contained the databases. The nature of these components meant they could be deployed in separate systems. The clients were deployed on a Sony Vaio UX1 and HP Mini-Note 2133 running Microsoft Vista and XP platforms respectively. Whilst these client devices are classed as mobile devices, they do not represent the traditional mobile handset that the framework was devised for. The decision to utilise these platforms over mobile handsets was largely due to development constraints within the timeframe of the funded project, as mobile platform development would have had to been undertaken using unmanaged code in C++, rather than rapid prototyping languages such as Visual Basic.

Having undertaken a thorough examination of biometric technologies and the commercial products that were available, it was determined that few suitable commercial biometric solutions existed for integration within NICA. The principal reason for this was the lack of available Software Development Kits (SDKs), with vendors preferring to design bespoke solutions for customers rather than license their biometric solutions for development. The project therefore identified some facial and voice verification algorithms developed in MatLab and sought to modify these for use within NICA [9]. These were accompanied by keystroke analysis algorithms previously created by the authors [2]. It was considered that these biometric approaches would provide the appropriate variety of transparent and intrusive authentication required for the proof-of-concept.

4 End-User Trial of NICA

In order to evaluate the approach, a user trial was conducted that ultimately involved 27 participants. The trial activity was split to two phases:

- **Enrolment Phase:** The participants used the prototype to provide face, voice and keystroke biometric samples that would be subsequently used to create their biometric profiles and also define two cognitive questions. A simple to use and intuitive interface was used to capture the samples. 8 samples for face, 9 for voice and 15

for each cognitive response they gave (which they were asked to provide 2) from which keystroke information was extracted. The enrolment process took no more that 15 minutes per person and at the end the participants were asked to complete the first questionnaire that looked to assess their experience.

- **Usability Phase:** Each participant was asked to follow a series of steps that would force an interaction with the device while the authentication prototype was running on the background. This would enable for biometric samples to be captured transparently as well as force access to services set to be of high security in order to test the operation of the alert level algorithm and the authentication mechanism in general. In order to ensure that the participants would have something to do during the 'usability' phase of the trial, and to ensure that contexts would occur in which different aspects of the prototype could be utilised, each user was asked to work through a given set of tasks such as using Instant Messenger, Microsoft Word, Microsoft Excel and an Internet Browser. The length of this phase varies as each user took different periods of time to interact with the device and complete the tasks. The average time of this phase was 45 minutes and on average over 60 biometric samples were captured from each participant during the usability phase of the trial. After completion of the scenario, the user was asked to fill in a questionnaire assessing their experience and the system. After that, the participants were asked to play the role of an impostor on the same device using the profile of another person and through using the same steps see how quickly the system would recognise that they were not the legitimate users.

The results from the evaluation overall demonstrated a positive opinion of the authentication system, with 92% of the users considering that it offered a more secure environment in comparison to traditional forms of authentication. The participants were also asked to evaluate how convenient the system was in a scale of 1 to 5, the results of which appear in figure 2. Although the responses were mixed, a slight skew towards the system being convenient exists on average. It is worth noting that through observation of the evaluation, participants' opinions were affected by the delays that occurred on the system while trying to manage all the processing. These occurred in some cases where applications might have been initialising concurrently and thus giving extra overhead to the system with NICA running in the background. This was a function of the prototype and a real system would not have such significant delays.

Furthermore the above views were also affected by the transparency of the system which was not always ideal. The lack of robust biometric algorithms caused a lot of transparent authentication requests to fail, prompting some of the users to experience more intrusive requests that they would normally get. Unfortunately the biometric techniques being utilised were largely developed in-house due to a lack of availability of commercial algorithms. In order to mitigate the errors a manual trimming of the threshold was taking place during the experiment in order not to allow the lack of accuracy from the biometric algorithms to affect the performance of the actual system. Nevertheless, what also happened in the experiment was that the scenario included access to a number of protected services in a small amount of time causing even more intrusive requests to occur but not necessarily having the chance to build the required confidence in the user while authenticating them transparently. Unfortunately, it was not possible

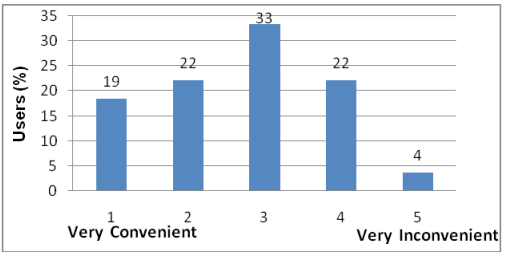

Fig. 2. Perceived convenience of the NICA prototype

to have the participants use the system for a prolonged period of days, so therefore the experimental study had to artificially include a number of steps to fully evaluate the prototype. It is likely this artificial environment resulted in a more negative attitude towards the system than what would have occurred in practice. The responses of the participants with regards to the transparency of the system are illustrated in figure 3.

With regard to the individual techniques that were utilised, there was a slight preference towards voice verification and keystroke analysis. From verbal feedback from participants there was a strong preference to techniques that did not require much explicit user interaction and were not very time consuming. As such, cognitive responses as an intrusive means of authentication were not very popular. The same occurred with face recognition as the algorithm utilised in the prototype required more time than other techniques to perform the authentication and the user also had to keep facing the camera until a sample was captured. At the same time voice verification (in its intrusive form) appeared to be more preferable as the user only had to repeat a small phrase with a subsequent quick response from the NICA server. Although many of the above were affected by the robustness of the algorithms utilised it still provides an insight that users prefer to have a higher level of security with the least overhead in their interaction. Usability and convenience were stronger preferences than security.

Regardless of the aforementioned problems regarding the convenience of the system, the majority of the users (70%) registered a preference to the use of transparent and continuous authentication as a protection mechanism. Although many of the participants suggested that the requests were too frequent the idea of being constantly protected and

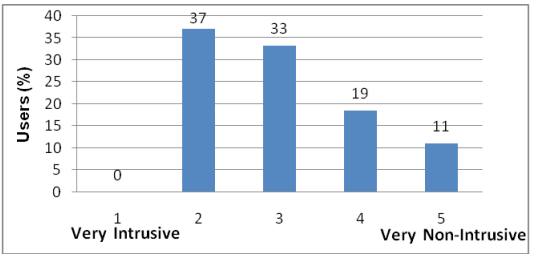

Fig. 3. Perceived intrusiveness of the new authentication system

specifically having extra security for highly sensitive information was very appealing to them. As such, 81% of the users said that they would use such system in practice as they would feel more protected than using traditional means of authentication. Although the remaining 19% stated they would not use it, their justification was that although they believed the system would offer higher security they do not perceive that their current use of their mobile device actually required a higher level of protection as they do not store or access personal information. This was actually an opinion that had arisen on a number of occasions during discussions with stakeholders. A body of users exist for which the mobile device is only (and will remain only) a telephony-based device. They have no desire to use it for any other purpose and as such do not perceive the need for additional security.

When the evaluation came to the participants acting as impostors it must be noted that although a number of users were not very positive when acting as the authorised user, their opinion became more positive when they saw the performance of the system reacting to an impostor. When the participants were asked whether the system managed to detect them and locked them out in a timely manner, 81% said yes. When the users where asked on how secure the system was their answers were very positive with 86% leaning to being secure or very secure.

5 Conclusions and Further Work

The research has resulted in the development of an operational proof-of-concept prototype, which is not dependent upon specific hardware and is functional across Windows XP and Vista platforms. It is able to operate in both client-server and standalone modes, and has successfully integrated three biometric techniques.

The evaluation of NICA clearly demonstrates the strengths and weaknesses of the proposed system. It is evident from the findings that such a transparent and continuous system has real merit and a large proportion of the participants felt it would provide the additional security they desire for their mobile devices. Unfortunately, with almost half of the world's population having a mobile device, it is difficult to establish an approach that satisfies all users. NICA has specifically considered this and developed a flexible approach that can utilise a variety of biometric and other authentication techniques and through a series of operational settings that can vary the level of security both transparent and intrusive being provided. Through this flexibility it is hoped the majority of users will be able to find a suitable mixture of settings and techniques they prefer and desire.

Whilst the prototype and subsequent evaluation has illustrated a number of key findings, it is important to highlight that if the system was operating within specification (i.e. the performance of the biometric techniques was good and the operational performance of the server was managed rather than everything operating for a single server) the nature of the transparency would mean few users would ever experience intrusive authentication. During the evaluation, however, the framework was configured to perform authentication on a more frequent basis than normal in order to ensure that sufficient judgments were made during the trial session. This was done in order to ensure that participants would see the full extent of the system in operation, but the consequence

was that they also encountered more intrusive authentication requests than would normally be expected. In some trial sessions, these requests were too frequent and time consuming, and participants therefore formed a more negative impression of the prototype.

The study has accomplished its overall aims of developing a next generation user authentication system. It has taken into account stakeholder considerations of usability, flexibility and convenience and provided a system that can improve the level of security in a continuous and transparent fashion – moving beyond traditional point-of-entry authentication. Whilst the prototype has a number of operational shortcomings, it is not anticipated that any of these would actually prevent a NICA-type approach from being operationally viable in the future. The project has also identified a host of additional avenues that require further consideration and research. In particular future work will focus upon three aspects:

1. Transparency of biometric techniques – Developing biometric approaches that will not only operate in point-of-entry mode but in a transparent fashion with varying environmental factors.
2. Privacy of biometric samples – the importance of this data is paramount and large adoption of any biometric system will only occur when such issues can be resolved to the satisfaction of all stakeholders.
3. Developing a risk assessment and management strategy for mobile devices. Given the wide-stakeholder group, varying responsibilities from general users to network operators and system administrators, it is imperative that an approach is designed so that the level of risk associated with a particular service request can be better understood and therefore protected.

The authors have already begun to consider the issue of transparency with respect to facial recognition, signature recognition and keystroke analysis [2,3,4] and will continue to address other key biometric approaches.

Acknowledgement. This research was support by a two year grant from the Eduserv Foundation.

References

1. Clarke, N.L., Furnell, S.M.: Authentication of Users on Mobile Telephones - A Survey of Attitudes and Practices. Computers & Security 24(7), 519–527 (2005)
2. Clarke, N.L., Furnell, S.M.: Authenticating Mobile Phone Users Using Keystroke Analysis. International Journal of Information Security, 1–14 (2006)
3. Clarke, N.L., Mekala, A.R.: Transparent Handwriting Verification for Mobile Devices. In: Proceedings of the Sixth International Network Conference (INC 2006), Plymouth, UK, 11-14 July, pp. 277–288 (2006)
4. Clarke, N.L., Karatzouni, S., Furnell, S.M.: Transparent Facial Recognition for Mobile Devices. In: Proceedings of the 7th Security Conference, Las Vegas, June 2-3 (2008)
5. Denning, D.: Information Warfare & Security. ACM Press, US (1999)

6. Karatzouni, S., Furnell, S.M., Clarke, N.L., Botha, R.A.: Perceptions of User Authentication on Mobile Devices. In: Proceedings of the ISOneWorld Conference, Las Vegas, CD-Proceedings (0-9772107-6-6) (2007)
7. GSM World GSM Subscriber Statistics. GSMWorld.Com (2002), http://www.gsmworld.com
8. Lemos, R.: Passwords: The Weakest Link? Hackers can crack most in less than a minute. CNET News.Com (2002), http://news.com.com/2009-1001-916719.html
9. Rosa, L.: Biometric Source Code. Advanced Source Code (2008), http://www.advancedsourcecode.com
10. Shaw, K.: Data on PDAs mostly unprotected. Network World Fusion (2004), http://www.nwfusion.com