Mobile Qualified Electronic Signatures and Certification on Demand

Heiko Rossnagel[1]

[1] Chair of Mobile Commerce and Multilateral Security,
Johann Wolfgang Goethe University Frankfurt, Gräfstr. 78,
60054 Frankfurt, Germany
heiko.rossnagel@m-lehrstuhl.de
http://www.m-lehrstuhl.de

Abstract. Despite a legal framework being in place for several years, the market share of qualified electronic signatures is disappointingly low. Mobile Signatures provide a new and promising opportunity for the deployment of an infrastructure for qualified electronic signatures. We analyzed two possible signing approaches (server based and client based signatures) and conclude that SIM-based signatures are the most secure and convenient solution. However, using the SIM-card as a secure signature creation device (SSCD) raises new challenges, because it would contain the user's private key as well as the subscriber identification. Combining both functions in one card raises the question who will have the control over the keys and certificates. We propose a protocol called Certification on Demand (COD) that separates certification services from subscriber identification information and allows consumers to choose their appropriate certification services and service providers based on their needs. We also present some of the constraints that still have to be addressed before qualified mobile signatures are possible.

1 Introduction

In the directive 1999/93/EC of the European Parliament and of the Council [ECDir1999] legal requirements for a common introduction of electronic signatures in Europe were enacted. The directive sets a framework of requirements for security of technology used for electronic signatures. Based on certificates issued by certification authorities, which certify public keys for a person registered by a registration authority, electronic signatures can be created with a so-called "secure signature creation device" (SSCD), carrying the private keys of a person.
The EC-directive distinguishes between "electronic signatures" and "advanced electronic signatures" [ECDir1999]. An advanced electronic signature is defined as an electronic signature that meets the following requirements:

"(a) it is uniquely linked to the signatory;

(b) it is capable of identifying the signatory;

(c) it is created using means that the signatory can maintain under his sole control; and

(d) it is linked to the data to which it relates in such a manner that any subsequent change of the data is detectable;" [ECDir1999]

Certification Service Providers can issue certificates for advanced signatures that will be qualified if they meet the requirements of Annex I of the directive. Those advanced signatures with qualified certificates will be refered to in this paper as qualified signatures.
In Germany and Austria, the local implementation of the EC directive requires evaluation of the SSCD to be done against ITSEC E4 or CC EAL 4+ levels [FuFr2000]. For directory services, stringent 24/7 availability and durability is required. Revocation lists and other feasible technology must be available to all accepting parties of signed documents. The EU suggests the implementation of a public evaluation infrastructure under control of a government authority. Germany has already implemented a system of evaluation service companies, evaluation consulting companies and the Regulatory Authority for Telecommunications [RegTP2004] as the responsible government authority.

The deployment of signature card products focused so far on smart cards with evaluation against the requirements for lawful electronic signatures. Based on these, personal computer based signature applications have entered the market. These applications require smart card readers attached to the workstation, thereby preventing user mobility.
The market share of EC-directive conforming smart cards is disappointingly low, failing to meet any involved party's expectations. This has partly been blamed on the incompatibility and missing standards of existing products. Also the lack of customers prevents companies from investing in signature products. As a result almost no commercial usage for qualified electronic signatures exists. Consequently no customers seek to obtain signature products.
There are numerous activities trying to enlarge the potential consumer base like putting key pairs on national identity cards [FSEID2004]. Lately there have been some efforts towards mobile signatures [ETSI] [Raddic2004] and this approach might have a chance to break up the deadlock of missing customers and missing applications.
However, there are numerous problems to be solved, before qualified signatures can be created with a mobile device.
The first part of this paper (Section 2) gives an overview on the possible approaches for mobile signatures especially focused on SIM[1]-based signatures and it's challenges

[1] Subscriber Identity Module

in detail (section 3). In section 4 we present a protocol for SIM-card deployment solving most of these challenges. Section 5 provides an outlook on possible usage scenarios of that protocol and section 6 focuses on the security of mobile devices. In section 7 we examine special constraints of mobile signatures and section 8 concludes our findings.

2 Mobile Signatures

Mobile signatures are electronic signatures which are created using a mobile device and rely on signature or certification services in a location independent telecommunication environment. They allow signatory mobility beyond fixed, secure desktop workstation with trusted, personal signing equipment [FrRaRo2003]. Although using mobile devices for signature creation has several shortcomings (e.g. display size, communication costs, limited computing power), the high market penetration of cell phones [GSM2004] and the mobility gained make this effort potentially successful and promising.
Two possible signing approaches in the mobile environment have been proposed in the past: signatures created in centralized signing server environments located at service providers like mobile network carriers; and electronic signatures created inside the signer's mobile device using a smart card.

2.1 Server Based Electronic Signatures

Server based electronic signatures are signatures created by a service provider for a specific customer. Figure 1 illustrates such a server infrastructure.

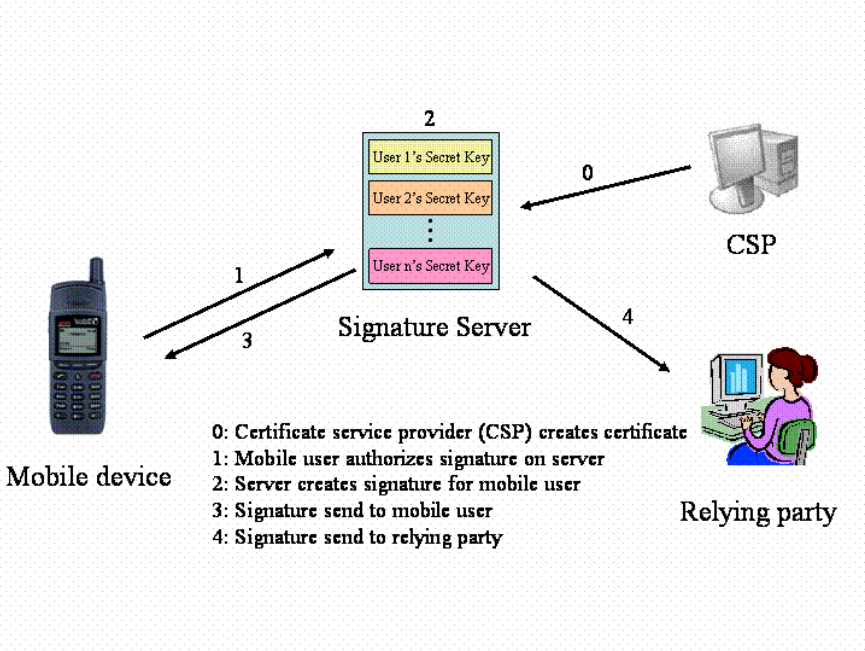

Fig. 1: Sever based electronic signature infrastructure

With server based signatures it is important to distinguish between signatures that have a corresponding certificate issued under the name of the customer and signatures with certificates issued under the name of the service provider or an employee of this provider.

In the first case it is necessary that the customer transfers his private key to the service provider. However, according to Art.2, 2(c) the signature has to be created by means that the signatory can maintain under his sole control to achieve the status of an advanced signature [ECDir1999]. By giving away his private key this premise can not be fulfilled [FrRaRo2003]. In the case of signatures whose certificates are issued under the name of the service provider you can not assume these signatures to be legal signatures of the customer. They are signatures of the signature service provider and only enable an identification of the provider. Those signatures can achieve the status of advanced signatures with qualified certificates as long as they fulfill the requirements of Annex I and are provided by certification service provider who fulfills the requirements of Annex II. Therefore, the signature service provider acts as a replacement for the customer. However, based on the signature of the provider it can not be verified that the customer really authorized the signature. Neither the integrity nor the fact that the user authorized it can be proven. There are possible technical solutions to accomplish the integrity and accountability of his authorization but they would require a security environment on mobile devices that would enable the device to create qualified signatures itself [RaFrRo2003].

2.2 Client Based Electronic Signatures

Signatures can be created inside the mobile device using a secure signature creation device which has to fulfill the requirements of Annex III. Using a multiple smart card solution, the signature smart card, certified by a certification provider, is inserted into the mobile device which already contains the usual SIM-card. Therefore, the signature process takes place on the mobile device and the user is able to use basically any signature card available on the market. This can be achieved by either exchanging the SIM-card with the signature card (Dual Chip) or by having an additional chip card reader within the mobile device (Dual Slot). The first solution is very inconvenient for the signatory since he has to switch of the phone to exchange the cards for the signature creation and again to use the phone functionality. In the latter case a specialized mobile phone is required that has multiple smart card slots which almost none of the current mobile phones do.
It would also be possible to use a single smart card that contains the SIM telephone functions, as well as the secure signature creation device. This can be achieved either by leaving some free space on the SIM-card, on which the components of the signature creation device can be installed later on, or by shipping SIM-cards with preinstalled signature functionality that has to be initialized and activated.
We propose the usage of evaluated smart cards suitable for qualified electronic signatures which are extended by the SIM functionality and usable through a unified interface, e.g. with the USIM[2] specification TS 21.111 [3GPPSpec]. Another approach might be the migration and evaluation of USIM with a full WAP[3]/WIM[4] implementation for the purpose of lawful mobile signing [WAPF2004]. Evaluation must be carried out with ITSEC or Common Criteria within an evaluation process similar to the evaluation summarized in [FuFr2000].

3 Challenges of SIM Based Signatures

Using a single smart card for both functionalities provides the most convenient solution for the signatory. He can sign documents and distribute them via communication services of his cell phone like GPRS[5] or UMTS[6]. To ensure that the requirements of Art.2 2(c) are met, it is necessary to provide some sort of reliable access control to the signature functions. The usual PIN used to control the access to the telephone functions is not sufficient, since users can keep their phones and SIMs unlocked for convenience. Like traditional signature cards, SIM-cards can be certified according to security evaluation criteria and are under control of the user.
However, using a single smart card for multiple purposes raises new questions and challenges. The SIM-card is issued by the telecommunication provider, while the

[2] Universal Subsriber Identity Module
[3] Wireless Application Protocol
[4] Wireless Identity Module
[5] General Packet Radio System
[6] Universal Mobile Telecommunication System

SSCD is issued by a certification service provider. Combining both functions in one card raises the question who will have the control over the keys and certificates.
The simple solution is that the deploying carrier also initializes the signature secrets to act as a trust provider for their customers. This seems to be reasonable at first glance, since some of the european carriers already own and maintain trust centers (i.e. Deutsche Telekom), but there are several shortcomings, which make this approach unpractical.
First of all the customer wants to leave the store with his SIM-card right away, so he can use his mobile phone instead of waiting several weeks for the certification process to be completed. Furthermore, binding the keys to a carrier creates a great hindrance for the customer to switch to a cheaper carrier in the future. From the carriers point of view this would of course be a positive effect. From the customer's perspective, however, it would be much better to be able to choose freely between different certification service providers.
Also due to the lack of success of the signature market so far most providers probably do not want to invest in building and maintaining their own trust center to provide certification services. In addition, they don't want to change their distribution channels unless they expect an increase in revenue.
Therefore, a different solution for mobile signing and certification is needed, that allows separation of subscriber information and certification services.

4 Certification on Demand

The mobile operator could sell SIM-cards equipped with a key generator for one or more key pair(s) which can be used for the signing functionality. After obtaining the SIM-card from the mobile operator, the customer can then generate the keys and activate the signature component and the public key(s) can be certified by any Certification Service Provider on demand.
Through the separation of the telephone functionality and the (possibly later) certification of the user's identity by a certification service provider, both functions can be sold separately and can be obtained from different providers.
The carrier will probably face increased costs for the signature capable SIM-card but can also expect increasing traffic caused by signature services. All distribution channels will remain unchanged.
Figure 2 illustrates the necessary steps for the distribution of the SIM-card and the certification process.

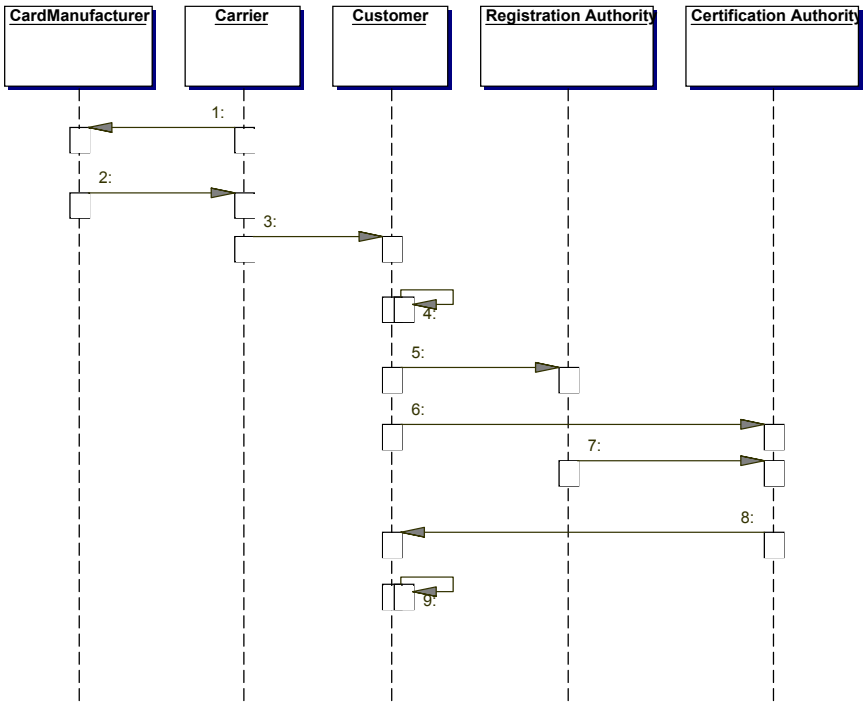

Fig. 2: Certification on Demand Protocol

1. The carrier gives his IMSI[7]/Ki[8] pairs to a card manufacturer.
2. The card manufacturer returns a SIM card containing an IMSI/Ki pair, a key generator for the signature application and the public key of the RootCA to the carrier.
3. The SIM card is sold to the customer and the carrier provides a nullpin that is used to generate the keys and activate the signing functionality.
4. The customer generates the keys and activates the signing functionality by entering the nullpin.
5. The customer registers at a Registration Authority of his choice, providing identification information and his public key.
6. The customer sends his identification information signed with his private key over the air to the Certification Authority.
7. The Registration Authority sends the public key and the identification information to the Certification Authority.

[7] International Mobile Subscriber Identity
[8] Individual subscriber authentication key

8. If the information provided by the customer and the Registration Authority match, the Certification Authority issues a certificate for the customer and sends it over the air to his mobile phone.
9. The user can verify the validity of his certificate by checking the certificate issued by the RootCA of the Certification Service Provider.

This protocol makes no changes to the existing distribution infrastructure of mobile operators. The steps 1 to 3 remain the same way they used to be before, apart from the fact that the card manufactures puts additional information and functionality (signature key generator, public key of RootCA) on the SIM card. In order to ensure that the card manufacturer does not know the private key of the user the key generation should be done by the card. The customer is not forced to certify his keys and can use the SIM for telephone functionality only. He could also activate the signing functionality without going through the certification process for example as a security token. If he wants to be able to make legal binding electronic signatures, he has to go through the complete process to obtain a qualified certificate. He can do this by freely choosing the CSP.
The nullpin to generate the keys and activate the signing functionality in step 4 is used to ensure that no signatures can be created before the customer has control over the SIM card. If the signature application has been activated before, the user will recognize this when entering the nullpin.
Step 6 could be omitted but serves as insurance for the customer to ensure him that the integrity of his identification information will be preserved.
If the customer wants to change his CSP, he only has to repeat steps 5 to 9 with his new CSP. If the customer wants to change his carrier, he has to go through the whole protocol again, but can register with his current Certification Service Provider.

5 Possible Applications

5.1 Enabling Security Infrastructures

There is a need of corporations to provide their mobile workforce with secure access to the corporate backend. So far security tokens have been used to allow this functionality. These tokens are expensive and stored on extra hardware that needs to be carried around and can easily be lost. Putting these credentials on a SIM, that will be placed in the mobile phone, reduces the risk of losing the credential as well as the costs. But some corporations greatly object to leave their private keys and certificates in the hands of their mobile operator.
With Certification on Demand the corporation's IT security department can obtain COD enabled SIMs from the corporation's cellular contractor and initializes them for the corporate mobile security infrastructure. The WiTness project [WiTness] sponsored by the European Union implements such an infrastructure.

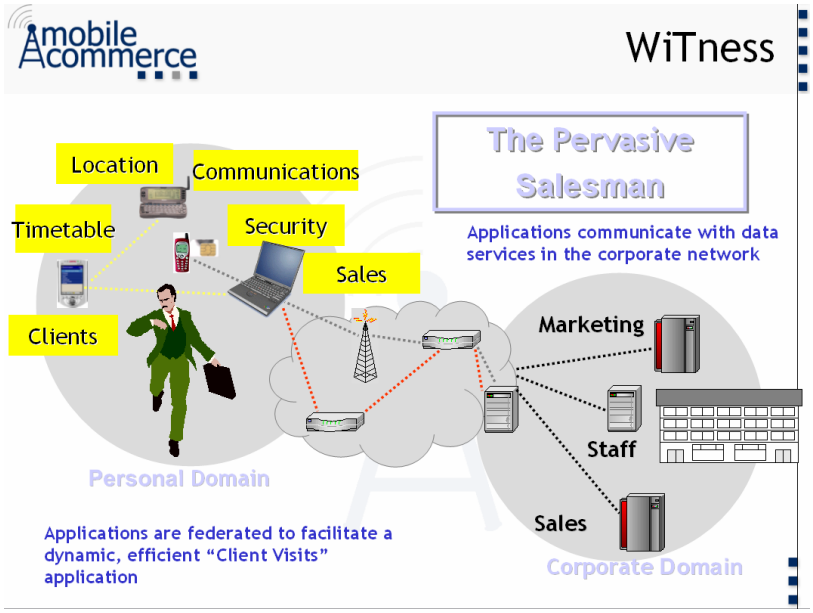

Fig. 3: WiTness Pervasive Salesman Scenario [WiTness]

Figure 3 shows an application scenario where a "pervasive salesman" has secure, corporate-controlled access to all data available to him in the corporate information system. Access is controlled by a security module based on a SIM with additional security functionality.

5.2 Multilateral secure financial transactions

Storing bank credentials on a SIM may help the migration from plastic to mobile phones. A COD infrastructure allows financial institutions to certify and enable mobile subscribers to use banking services online through their mobile terminal and SIM. Credentials could be certified by the bank itself, like the credentials used on bank cards. Therefore, the bank can still have the control over the credentials while the mobile operator still can issue the SIM cards without giving their IMSI/Ki pairs away to the bank. This would enable the bank to offer services like transactions, brokerage or checking the account balance based on the credentials stored in the SIM. This functionality can and has been realized without the Certification on Demand protocol but only if the banks and carriers are willing to cooperate. In the Czech Republic T-Mobile [TMO2004] and the Czech banks agreed to send their critical information to Giesecke&Devrient, a card manufacturer who started producing banking enabled SIMs [GuD2004]. However, the COD protocol would enable banks to use SIMs as credentials without having a contract with the mobile operator.

5.3 Enabling mobile electronic consent and identity management

Many mobility applications rely on a user's consent towards reducing his privacy for a particular service. Examples are location based services on cellular networks, situation based marketing scenarios and tracking technology following users to support them with information they need in-time and in-place. A secure provable electronic consent of users can be achieved using electronic signatures on SIM-created credentials that may contain information about time, intent and recipient of the electronic consent. Research has found SIMs to be on the edge of a global identity management infrastructure [Rann2003]. In the near future, personal or role attributes customized for particular application areas (e.g. online dating, identity management) could be managed on SIMs on demand from their owners.

5.4 Using COD in deployment of electronic identity cards

If the signature credentials are stored on an identity card issued by the government, the same problems as described in section 3 occur. The Government has to issue the identity card but does not want to act as a certification service provider. Using the COD or a similar protocol enables current CSPs to certify the keys of the recipients of the identity cards.

6 Trusted Mobile Devices

The mobile device serves as the card reader, storage device for the document to be signed and as a display for the signature application. Therefore, it must be ensured that the data shown on the display is identical with the data signed by the signature card. This is commonly known as "What You See Is What You Sign" (WYSIWYS). The operating system used on the mobile device has thus a pivotal importance to ensure the integrity and accountability of the electronic signature.
If the authorization mechanisms, memory protection, process generation and separation or protection of files in an operating system are flawed, an attacker may gain access to the different internal processes. He might take advantage of this situation to generate forged signatures.

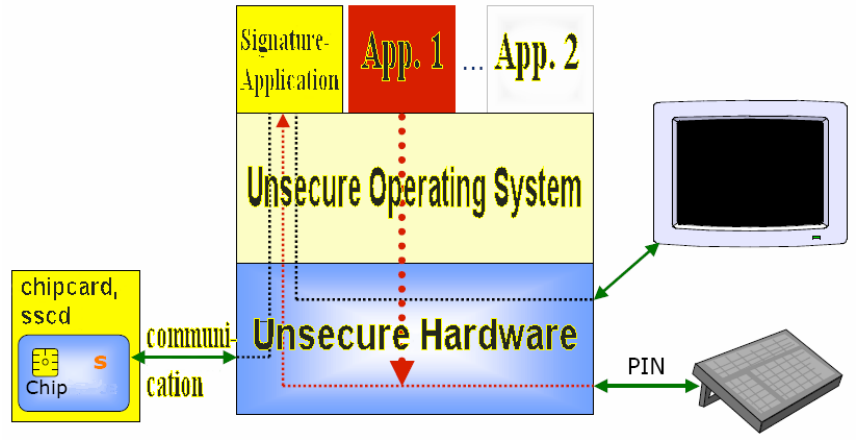

Fig. 4: Manipulated digital signature [Federr2003]

Figure 4 illustrates that application 2 as a malicious program can for example intercept the PIN. An even considerably higher risk exists, however, if the malicious application changes the data to be signed after they are displayed to the user. Due to the virtual unrestricted hardware access, a malicious program is able to manipulate all data transmitted to the signature application before the actual signature takes place.

In the past mobile phones were "closed" devices that gave the user no possibility of installing programs or storing data apart from address book entries. But with increasing computing power and storage capabilities, new and open operating systems like PocketPC [Pocket2004] and Symbian [Symbian2004] were developed, which allow users to install any program they like. This of course raises the possibility that malware or Trojan horses are installed by the user or a third person.

Although a tamper resistant mobile phone could be build and certified most of the features of present phones would probably not be available for this phone. Therefore, it will probably fail to get a high market penetration. The only solution that seems to be promising is to have a small microkernel as a secure platform which runs the signature application and an additional common operating system running on top of the security kernel.

The "Open Source" project at Saarland University Perseus develops such a system [Perseu2004]. It provides a small microkernel as a secure platform. The microkernel is responsible for the administration of the device, files, memory and processes and is loaded directly after booting. It is aimed at protecting security-critical applications by isolating the individual processes from each other. Perseus is based on the approach that a normal operating system runs like an application, and therefore the Perseus kernel lies below the operating system in the layer architecture. Only by being embedded below the operating system, which is still needed for ordinary applications, Perseus can permit isolated processes to take place system-wide between the applications. Isolated processes are not possible for applications within the standard operating system, however, but only between the individual "secure applications" and the Perseus operating system.

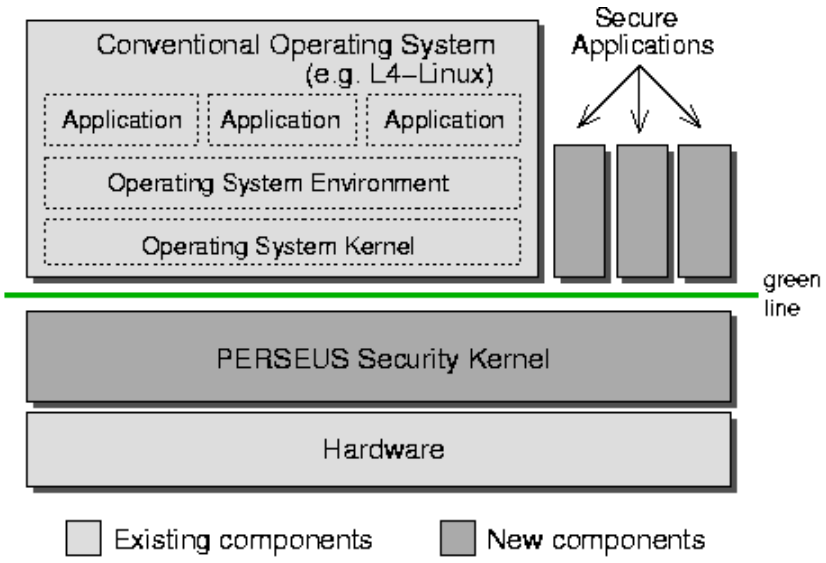

Fig. 5: System Architecture Perseus [Perseu2004]

In the Perseus prototype, the trustworthy user interface reserves a line in the upper section of the screen that is permanently under the control of the security kernel.
As the line or LED is under the sole control of Perseus, it cannot be misused by a compromised operating system. If the display indicates that the user is communicating with the Perseus kernel, the control of the full display and keyboard solely lies with the security kernel.

7 Mobility and Signing

Mobile signatures are made with mobile devices and therefore constraints have to be addressed that are not present in traditional signing infrastructures.

7.1 Data Transfer

First of all any traffic that is necessary will be accounted to the customer's bill. Therefore, it is essential to create as little data traffic as possible in order to get the customer to accept the additional costs. In the case of the signature creation, traffic is only necessary for the download of the document to be signed, if at all. In the process of signature verification, several documents, especially the keys of all CA's involved have to be downloaded in order to ensure the integrity of the verification process.

Revocation lists are a particular concern that has to be met. In order to be up to date with actual revocation lists the customer has to be "online" to be able to get access to the actual status of all the involved signatures and certificates. This could lead to lots of data being transferred and a lot of additional costs. Standards like ISIS-MailTrusT [ISISMTT] can be useful as well as concepts of server centric support in document verification [Fritsch2002].
It would also be possible for the CSP to sponsor the additional data traffic in order to get customers to accept and use mobile signatures [FSMR 2003].

7.2 Storage

Mobile devices usually have a rather fixed amount of storage space. This is even more relevant, if one has to store the data on the SIM-card itself, for whatever reason possible. Therefore, the mobile signature application should whenever possible try to store the necessary information on a server of the service provider. This of course is in contrast to the goal of minimizing the necessary traffic for signature applications. Therefore, a trade off between cached information and information to be transferred has to be found. This is particularly important for the storage of root certificates, certification chains and certificate revocation lists for offline-verification. This problem might be solved by increasing storage space on mobile devices and the ability of modern devices to use external storage like SDCards.

8 Conclusion

Mobile Signatures are a promising approach to break the deadlock between missing customers and missing applications. The high market penetration of mobile phones enables certificication service providers to target millions of potential customers. We analyzed two possible signing approaches (server based and client based signatures) and conclude that SIM-based signatures are the most secure and convenient solution. However, using the SIM as an SSCD seems to force the mobile operator to act as a trust provider and therefore to challenge the existing CSPs in a market that hasn't been successful so far. We proposed a protocol called Certification on Demand that seperates subscriber information from certification services and therefore enables both industries to cooperate instead of compete with each other.
We also provided possible application scenarios that can be realized with Certification on Demand. But even with the certification problem solved, there are still a lot of open issues that have to be addressed before mobile qualified electronic signatures will be able to get market acceptance.

References

[3GPPSpec] Specification of GSM, http://www.3gpp.org/ftp/Specs/archive/

[ECDir1999] European Union: DIRECTIVE 1999/93/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 13 December 1999 on a Community framework for electronic signatures

[ETSI] ETSI MCOMM Specialist Task Force 221

[Federr2003] H. Fedderath: Digitale Signatur und Public Key Infrastruktur, http://www-sec.uni-regensburg.de/security/5PKI.pdf

[FSEID2004] Project "Feasibility Study Electronic Identity Card", http:// www.uni-kassel.de/fb10/oeff_recht/english/projekte/projekteDigiPerso_eng.ghk

[Fritsch2002] L. Fritsch.: A secure, economic infrastructure for signing of web based documents and financial affairs; CBL – Cyberbanking & Law, issue 2/2002;

[FrRaRo2003] L. Fritsch, J. Ranke, and H. Rossnagel: Qualified Mobile Electronic Signatures: Possible, but worth a try? In: Information Security Solutions Europe (ISSE) 2003 Conference, Vienna Austria

[FSMR2003] S. Figge, G. Schrott, J. Muntermann, and K. Rannenberg: EARNING M-ONEY – A Situation based Approach for Mobile Business Models; In: Proceedings of the 11th European Conference on Information Systems (ECIS) 2003

[FuFr2000] T. Fuchß, L. Fritsch: Security Certificates as a tool for reliably software engineering; Datenschutz und Datensicherheit 9/2000, pp.514ff.

[GuD2004] Giesecke & Devrient: STARSIM® Applications, STARSIM®banking; www.gdm.de/eng/products/04/index.php4 ?product_id=386

[GSM2004] GSM Association: GSM Statistics www.gsmworld.com/news/statistics/index.shtml

[Perseu2004] B.Pfitzmann, C. Stüble: PERSEUS: A Quick Open-Source Path to Secure Electronic Signatures, http://www.perseus-os.org/

[Pocket2004] Windows Mobile – based Pocket PCs, http://www.microsoft.com/windowsmobile/products/pocketpc/default.mspx

[Raddic2004] Radicchio, http://www.radicchio.org

[RaFrRo2003] J. Ranke, L. Fritsch, H. Rossnagel: M-Signaturen aus rechtlicher Sicht. In: Datenschutz und Datensicherheit 27 (2003) 2, p.95-100, Vieweg & Sohn

[Rann2003] K. Rannenberg: Identity Management in Mobile Applications In: Datenschutz und Datensicherheit 27 (2003) 9 (DuD), pp.546-550, Vieweg & Sohn

[RegTP2004] Regulierungsbehörde für Telekommunikation und Post (RegTP) der Bundesrepublik Deutschland; http://www.regtp.de/

[Symbian2004] Symbian OS – the mobile operating system, http://www.symbian.com

[TMO2004] T-Mobile: Czech Republic: m-payment becomes a universal payment tool for customers; www.t-mobile.net/ CDA/news_details,20,0,newsid-1799,en.html?w=925&h=588

[WAPF2004] WAP Forum: Specifications of WAP, WIM; http://www.wapforum.org/

[WiTness] European IST Project „Wireless Trust for Europe“(WiTness), www.wireless-trust.org