



Mobile PKI

A technology scouting for security and use of mobile authentication technologies

Authors: Martijn Oostdijk, Maarten Wegdam (Novay) for SURFnet

December 2009



Mobile PKI for SURFnet

*A technology scouting for security and use of
mobile authentication technologies for
SURFnet*



Colofon

Date: December 8th 2009
Version: 2.0
Changes: -
Project reference: TS Mobile PKI
Novay reference: -
Company reference: -
URL: -
Authorisation: -
Status: Final
Editor: -
Company: SURFnet
Author(s): Martijn Oostdijk, Maarten Wegdam



THIS REPORT IS PUBLISHED UNDER THE CREATIVE COMMONS
"ATTRIBUTION-NONCOMMERCIAL-SHARE ALIKE 3.0 NETHERLANDS"
LICENSE. FURTHER DETAILS ON THIS LICENCE ARE AVAILABLE AT
[HTTP://CREATIVECOMMONS.ORG/LICENSES/BY-NC-SA/3.0/NL/](http://creativecommons.org/licenses/by-nc-sa/3.0/nl/)

Management Summary

A GSM/UMTS telephone has a SIM card. This is a standardised smartcard that is issued to the user by the telecom operator and is primarily used to authenticate the user on the mobile network. However, the SIM card has more potential uses. For instance, it allows for secure storage of digital keys that can be used for online authentication and digital signatures. This is referred to as Wireless PKI and Mobile PKI.

This report is an assessment of Mobile PKI technology and its potential application for authentication in education. This assessment focuses on its security and its application within the educational domain, with a specific emphasis on applications for SURFfederatie.

Mobile PKI employs encrypted SMS text messages that are used to represent authentication or a digital signature. The user has to express consent by entering a PIN code that secures the private key and which typically needs to be entered for each transaction separately. The relevant standards for this are well established and are supported on all mobile phones. This has advantages compared to other secure means of authentication. For instance, no additional authentication device is required, which also means that no software needs to be installed by the user on either the phone or on other client devices such as a PC. Neither is there a need to manually enter codes, as in the case of one-time passwords via SMS text messages. This improves user-friendliness. Malware such as viruses and key loggers that may have been installed on a PC cannot interfere with Mobile PKI.

This report considers the issue whether Mobile PKI is a secure means of authentication. The analysis identifies a “man in the middle”¹ attack as the most important threat for the internet channel. However, the authors of this report deem Mobile PKI to be more than sufficiently secure compared to other means of authentication and considering the kind of applications in (higher) education.

In our view the most important issues regarding Mobile PKI technology are not related to security or technology but have to do with the costs and the business model. In the Netherlands, Mobile PKI technology has only been deployed for limited pilots and it is therefore difficult to estimate the costs. These could turn out to be too high for many applications in the educational domain if there are no other large-scale deployments of Mobile PKI. A related aspect is the business model. Use of this technology requires the cooperation of the mobile operator, who is the owner of the SIM card. This means that the cooperation of all mobile operators is required for a large-scale deployment.

The final conclusion of this report is that Mobile PKI provides a secure means of authentication that in time will find wide application within the educational domain in the Netherlands. For the near future Mobile PKI will only be employed for services that require a high standard of security and that are used by a limited group of employees due to a) the expected costs, b) insufficient insight into the business model, and c) limited support from the mobile operators. It seems too early for a deployment for students or for general authentication for SURFfederatie or any other large-scale application for SURFnet, Kennisnet or other service. In the meantime it may be useful to consider one-time passwords via SMS text messages as step-up authentication or for password reset because this is cheaper and prepares users for Mobile PKI.

¹ A “man in the middle” attack is an attack in which a malevolent party intercepts and/or modifies the internet traffic between two parties. This allows the malevolent party to eavesdrop on the communications and to modify the content of messages to suit his own purposes.

Table of contents

Management Summary	v
1 Introduction	1
1.1 Background	1
1.2 Mobile PKI in de praktijk	1
1.3 Aims, focus and terminology	3
1.4 Overview	3
2 State-of-the-art	5
2.1 Current use	5
2.2 Standards	6
2.2.1 Smart card standards	6
2.2.2 SIM card standards	7
2.2.3 PKI standards	8
2.2.4 Standards for electronic signatures	9
2.2.5 Mobile Signature Service (MSS)	9
2.3 Conclusion	10
3 Architecture	12
3.1 Processes	12
3.2 Roles	13
3.3 Registration process	13
3.4 Authentication process	14
3.5 Technical analysis of various interfaces	15
3.5.1 Interface between AP and MSSP on the internet	16
3.5.2 Interface between handset and SIM	17
3.6 Conclusion	17
4 Security analysis	18
4.1 An ad-hoc threat analysis	18
4.2 ETSI security requirements	21
4.3 Security requirements from the Common Criteria protection profile	22
4.4 Conclusion	23
5 Application for higher education	24
5.1 Present situation: username/password	24
5.2 Criteria for the use of Mobile PKI	24
5.3 Potential services for using Mobile PKI	24
6 Conclusions and recommendations	26
References	30

1 Introduction

1.1 Background

The use of digital services requires trust on the part of both providers and users. In daily practice this trust is supported by security, access and identity management technologies such as the authentication of parties and the signing of documents and transactions by parties using a digital signature. Several solutions are available for implementing authentication and signing, including the use of username/password combinations, a smartcard or a smart USB token. The quality of a solution is determined mainly by its degree of security (resistance to fraud) and user-friendliness.

SURFnet offers several services that require authentication and, in some cases, signing. SURFfederatie is the most prominent example. This services allows students and staff in higher education to access a broad range of online services using their own account issued by their home institution. The vast majority of these cases are based on a username and password combination. This means of authentication is not very secure for several reasons, including vulnerability to phishing attacks.

Mobile PKI² is a technology that has the potential to authenticate people in a secure and relatively user-friendly manner. Mobile PKI employs the SIM/USIM card (hereafter referred to as SIM card) that is present in a GSM/UMTS mobile phone. The SIM card is the property of the mobile operator and is used primarily to authenticate a mobile phone on the mobile network. However, the mobile operator can install additional applications on this SIM card. These can include third-party applications. Mobile PKI makes use of this option by installing an application that allows users to authenticate themselves when accessing internet services, and to digitally sign electronic information.

A great advantage of Mobile PKI is that almost everyone in the SURFnet target group owns a mobile phone and that almost all of these phones are suited for Mobile PKI. A second advantage is that Mobile PKI is a standardised technology, which benefits both the security (as it is open) and the interoperability.

1.2 Mobile PKI in de praktijk

The pilot to test Mobile PKI is commissioned by SURFnet. Based on this pilot we describe the workings of Mobile PKI in actual practice.

Mobile PKI requires a suitable SIM card. Specifically, this means that the SIM card must have sufficient cryptographic processing power and sufficient memory, and the right software must be installed. This software is a so-called SIM Toolkit application. In the pilot we used the VMAC application by Valimo³. Figure 1 shows a photograph of the SIM card that was used. It shows not only the small size of the SIM card, but also the IMSI number that uniquely identifies a SIM card.



Figure 1: Photograph of a SIM card.

² Also referred to as Wireless PKI or SIM-based authentication.

³ Valimo Wireless Ltd. can be found on the web at <http://www.valimo.com/>.

A typical scenario for use is that someone uses a PC to access an online service that requires authentication or a digital signature. This proceeds as follows from the perspective of the user:

- The authentication service sends 'data to display' and 'data to sign'⁴. These are displayed to the user, who then has to confirm them by pressing 'ok'. The data are sent to the mobile phone as SMS text messages, but this is invisible to the user. From the user's perspective the messages appear automatically and the user only has to press 'ok', without going through any menus.
- The user then has to enter his/her PIN for authentication. The SIM Toolkit subsequently transmits a number of text messages. Depending on the settings of the mobile phone this may require explicit permission from the user, but generally this happens without involving the user.
- Finally, after successful authentication (or signing), the user receives a notification in the form of a text message.

The screenshots in Figure 2 present an overview of the authentication process.

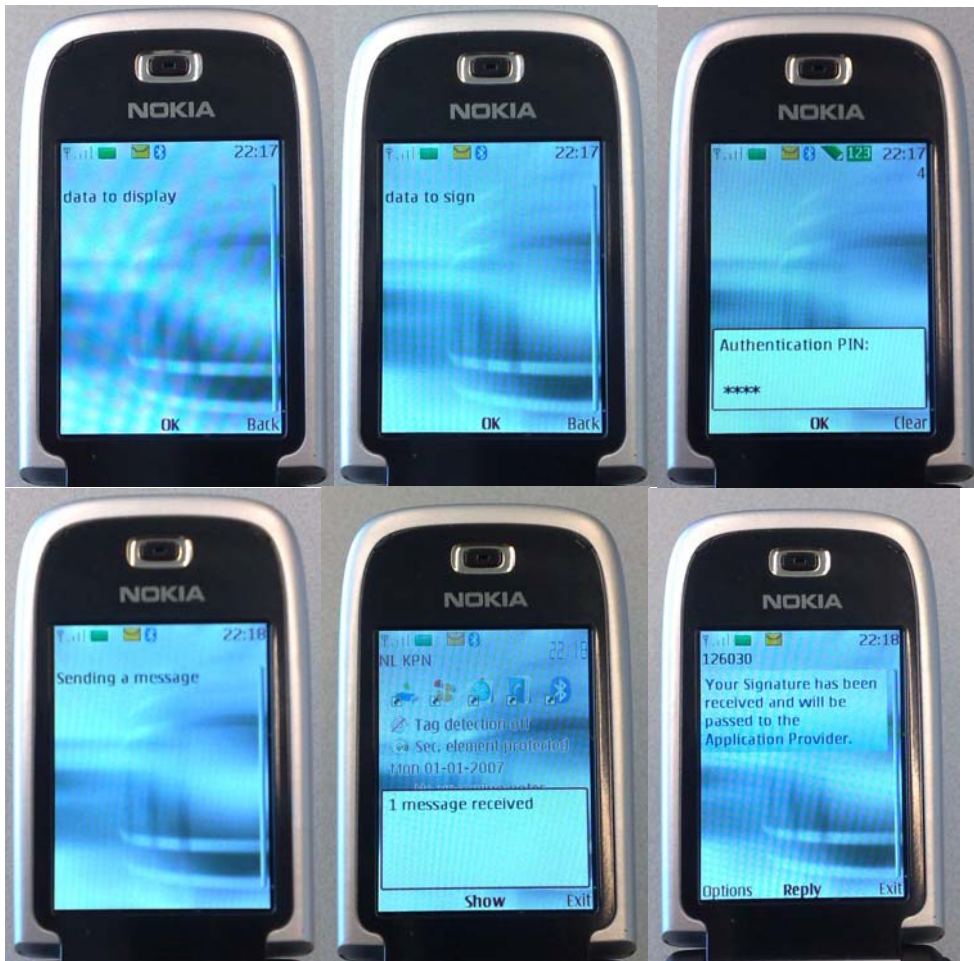


Figure 2: Screenshots of the steps in a transaction

⁴ The examples in this report will use the ASCII strings 'data to display' and 'data to sign' to represent the messages. In the case of authentication the second message consists of an unpredictable sequence of characters. In the case of signing the second message will consist of a reference to or a fragment of the document to be signed or of a transaction reference if a transaction is to be authorised.

Mobile PKI is relatively user-friendly in comparison to other secure means of authentication. For instance, it does not require the user to type in a code on his/her computer, as in the case of smartcard-based hardware tokens used by Dutch banks for online banking (the Rabobank “Random Reader” and the ABN-AMRO “E.dentifier”) and in the case of one-time password solutions via SMS text messages (SMS-OTP) that also use a mobile phone. Nor does Mobile PKI require any software to be installed on the PC.

1.3 Aims, focus and terminology

The aim of this report is *to present an analysis of the feasibility/applicability of Mobile PKI technology, specifically for SURFfederatie*. The emphasis of the analysis lies on the security and architectural aspects of this technology. In addition we also consider what kind of applications could benefit from this technology in the near future. This report is based mostly on existing materials and employs the quickscan methodology due to the limited resources available for compiling this report.

The following should be noted regarding the terminology used in this report. Many of the consulted sources are extremely technical in nature. This holds specifically for several standardisation documents. It is standard practice in such documents to introduce long lists of abbreviations and to use these in the rest of the document, creating a certain level of abstraction that allows the standard to be applied to multiple implementations. In this document, however, we use the concrete terms used in day-to-day communication. For instance, in this document we use the common terminology such as SIM, mobile phone, etc. instead of MSCA and MSCD. In spite of this, some parts of this report will be hard to read without the proper technical background knowledge.

1.4 Overview

Chapter 2 describes the state of the art in the field of Mobile PKI, including an overview of the relevant standards and examples from other countries where this technology is being used. Chapter 3 presents the architecture and an analysis of the most relevant interfaces. Chapter 4 undertakes a security analysis. Chapter 5 analyses the potential applications for education, based on the findings of the previous chapters. Chapter 6 includes a short comparison of Mobile PKI to other widely used means of authentication followed by the conclusions and recommendations.

The authors of this report express their thanks to Roland van Rijswijk (SURFnet), Joost van Dijk (SURFnet), Kick Willemse (Evidos / Diginotar) and Bram Sniekers (Diginotar) for their input and feedback.

2 State-of-the-art

The concept of Mobile PKI, with a main part to play for the SIM card, has been around since at least 2001. However, wide scale application of Mobile PKI appears to be taking place only recently (mainly in 2008, 2009). Valimo's Mobile PKI solution, used in the SURFnet pilot, surfaces more and more frequently in these efforts. A quick scan provides an overview of where Mobile PKI is already being deployed. The outcome of this scan is presented in the first part of this chapter.

Many of the technologies and protocols that are being used in Mobile PKI are documented in publicly available open standards. The second part of this chapter aims to provide an overview of these standards and how they are being employed to achieve a Mobile PKI solution. A detailed architecture, and the way in which the standards are incorporated into this architecture, is presented in chapter 3.

2.1 Current use

Table 1 provides a chronological overview of a quick scan of press releases and other (commercial) announcements and communications with respect to Mobile PKI products.

2001	Vodafone undertakes initial trials in the United Kingdom with mobile signing. [Link]
2001	Publication of the Common Criteria protection profile [3].
2002	Schlumberger (currently a subsidiary of Gemalto) publishes a Common Criteria security target for a Secure Signature Creation Device in the form of a Java Card applet. [Link]
2003	Publication of the ETSI standards for Mobile PKI [8][9][10][11].
2006	Giesecke & Devrient undertakes a Mobile PKI pilot together with Vodafone using the G+D's StarSIM product. [Link]
2006	TeliaSonera and Valimo announce the Tunnistuspalvelu Pro/Plus product. [Link]
2006	Elisa offers users the option to link their Finnish eID to a SIM card ⁵ . The technology is provided by SmartTrust and Valimo. [Link]
2007	Turkcell uses the Valimo solution for internet banking (at least 9 banks at launch, but 19 banks already by October 2008). [Link]
2007	Telefonica concludes a contract with Ericsson for Mobile Signatures. The technology is provided by Valimo. [Link]
2007	Valimo enters into a partnership with Gemalto for the VMAC product. [Link]
2007	Technology provider EMT from Estland introduces Mobiil-ID. [Link]
2008	Experian and Keynetics undertake a pilot with Valimo technology in

⁵ The SmartTrust solution is also mentioned in [15], which proposes an alternative method for using the national identity card.

	France. [Link]
2008	Pilot by Vodafone and Banco Sabadell. [Link]
2009	Oberthur includes the VMAC applet in Oberthur's SIM portfolio. [Link]
2009	Collaboration between Telenor and Valimo. [Link]
2009	Collaboration between LatTelecom and Valimo. [Link]
2009	Government initiative in Austria for mobile signatures in Q4. [Link]

Table 1: Announcements of Mobile PKI pilots etc.

An extensive market analysis⁶ is beyond the scope of this report, but some notable points present themselves when considering this table. The first pilots took place in 2001, even before the draft ETSI standards had been compiled, although Common Criteria protection profiles existed already for several types of signature technology at that time. A number of pilots were subsequently undertaken by SIM card providers and mobile operators (in collaboration with technical parties such as CAs and integrators). In the meantime, small-scale pilots in several countries had been succeeded by large pilots aimed at the consumer market. In recent years full-scale Mobile PKI implementations are being rolled out, specifically in Scandinavia⁷, the Baltic states⁸ and Turkey. Occasionally the government takes part in cases that involve so-called citizen-to-government (C2G) transactions that are facilitated by Mobile PKI. The large SIM card manufacturers (Gemalto, Oberthur, Giesecke & Devrient) all collaborate with Valimo in supporting Mobile PKI.

2.2 Standards

2.2.1 Smart card standards

The SIM smartcard plays a key part in Mobile PKI. All SIM cards comply with the ISO 7816-4 standard [13]. Some SIM cards are Java Card [14] smartcards, which mainly facilitates personalisation for the mobile operator. Managing applications on such Java Card smartcards is described in Global Platform [12].

- ISO 7816-1, 2, 3, 4, ...: These ISO standards describe smartcards, beginning at a very low level (the dimensions of a smartcard, the location of the contact chip, the meaning of the contacts, how many times a plastic card should bend before it is allowed to break), then moving to the Transport Protocol (so-called TPDU: Transport Protocol Data Units) and finally describing the Application Protocol Data Units in Part 4. A smartcard that complies with ISO 7816-4 is able to process standardised commands (for instance for navigating through the file system that is present on the card).

Higher layers in the ISO 7816 stack relate to inter-industry application functionality. For instance, ISO 7816-8 and 7816-15 describe how a PKI token may be implemented - standard commands for authentication, signatures, reading certificates stored in the file system, etc. and overlap with PKCS standards as described further on. The ETSI standards for Mobile PKI (also described below) do not necessarily employ these higher layer standards.

⁶ The presentation by Bill Nagel from Forrester Research at RSAcon'08 indicates that this involves a non-trivial analysis. See <https://365.rsaconference.com/docs/DOC-1172>.

⁷ Valimo Wireless Ltd is a Finnish company.

⁸ See the website of the Baltic WPKI forum: <http://wpki.eu>.

- Java Card: Java Card (see [14]) is a Java technology (and therefore a de facto Sun “standard”) for programmable smartcards. The language is much more limited compared to standard Java, both in terms of syntax and in API. Java Card provides a multi-application platform (several so-called applets can reside on a card, but only one will be active at any one time). A number of features of Java Card are not present in standard Java, mainly in the area of security and the memory model. Java Card allows the applet programmer to use the cryptographic API, but whether this is supported in the hardware depends on the specific smartcard.
- Global Platform: The Java Card standard does not address application management (such as installing new applications). This is fixed by Global Platform (GP, previously Visa Open Platform, see [12]). In principle GP is independent of Java Card (the standard discusses applications, not Java Card applets) but in practice one can regard GP as Java Card application management. In addition GP also offers functionality to applets that allows them to delegate certain (security related) tasks to the platform. Examples include establishing a secured end-to-end connection with a terminal (or underlying systems). A Java Card SIM in general can be expected to adhere to GP.

2.2.2 SIM card standards

The SIM card is the property of the mobile operator and is located in the user’s mobile phone. The SIM is secure in the sense that the mobile operator can restrict access to the data that is stored on the SIM. Neither the user nor potentially malevolent software installed on the mobile phone can access this data.

GSM SIM cards adhere necessarily to standards. After all, a SIM card must be able to function in any GSM phone.

The specific functionality of GSM SIM cards is laid down in ETSI standards. The original ETSI standard, GSM 11.11, allows the operator to place files on the SIM that can be edited using the phone and provides a means by which the phone can authenticate itself to the operator on the network. More recent (Java Card) SIM cards may also comprise the so-called SIM Toolkit.

- ETSI: SIM (GSM 11.11): This standard (see [5]) describes the authentication mechanisms that the SIM employs to gain access to the network. The SIM uses a symmetric key that cannot be extracted. GSM 11.11 also describes a simple, static ISO 7816 file system (the standard lists the files that need to be present). Access to files may be restricted in several ways, for instance readable by the user only after entering a PIN or readable only by the Mobile Operator (MO) after authentication. This file system can be used by the operator to store properties of the SIM that can be read and processed by the GSM phone. Some files are uniquely targeted at the user, such as the phonebook.
- ETSI: SIM Application Toolkit (GSM 11.14): This standard (see [6]) defines a framework on top of Java Card for so-called SIM Toolkit applets. These Java Card applets exist outside of the GSM 11.11 file system, causing older mobile handsets to ignore the applets on a GSM 11.14 SIM. Newer handsets will regularly poll applets, allowing a SIM Toolkit applet to (1) add GUI menus and (2) respond to events, such as reception of certain SMS text messages or selection of specific menu items by the user. The SIM Toolkit applets are managed by Global Platform, which explicitly allows management over-the-air (OTA) by the mobile operator. Most mobile phones place the menus generated by the SIM Toolkit applets in a “Extra” or “Tools” menu. This has three major advantages in user-friendliness: firstly, the MO can perform remote applications management (install new versions, etc.); secondly, SIM applications can respond

to messages from the network without disturbing the user, and thirdly, a SIM application can use the handset's GUI, and such an application takes priority over applications that are not located on the SIM (such as MIDlets⁹).

- Following the GSM age (the so-called 2nd generation mobile phones), we are currently in the UMTS age (the so-called 3rd generation mobile phones). GSM 11.11 and GSM 11.14 have UMTS equivalents published by 3GPP that describe the USIM and USAT respectively.

It is unclear what percentage of SIM cards in current use can employ the SIM Toolkit applications to support Mobile PKI. Although the greater majority of distributed SIM cards support the SIM application toolkit, not all SIM cards provide hardware support for cryptography.

2.2.3 PKI standards

In addition to the smartcard and mobile phone standards, much of the technology related to authentication and digital signatures has been standardised. Strong authentication and advanced digital signatures require cryptographic processing and therefore involve key management. Public Key cryptography is indispensable to keep key management manageable. Describing key management is outside the scope of this report. In practice public key cryptography always employs X.509 certificates, following the ITU-T and RSA Security's Public Key Cryptography Standards (PKCS).

So-called PKI tokens are sometimes used because the client-side environment cannot always be trusted: the user as well as the software on the user's platform may be malevolent. PKI tokens are devices that can be used for cryptographic calculations for authentication and digital signatures. A PKI token will contain a private key in protected memory - so it can be used in transactions but cannot be extracted - and certificates that link the identity of the user to the corresponding public key. A Certificate Authority (CA) signs these certificates. The CA's signature can be checked by anyone using the CA's root certificate. To prevent unauthorised persons from using a stolen PKI token a PIN code is usually required for using the token for cryptographic computations.

A PKI-based solution for digital signatures can also be used for authentication. A one-time message (which will never be repeated) is then generated by the server; the client is requested to digitally sign this message. The resulting signature along with the public key certificate are then checked by the server to establish that the client has access to the corresponding private key, and therefore holds the PKI token issued to that client.

An overview of RSA Security's PKCS standards:

- PKCS #1 Public Key cryptography (RSA).
- PKCS #6 Certificates (is deprecated and has been replaced by X.509 version 3).
- PKCS #7 Cryptographic message syntax. This describes the result of a digital signature (so-called 'signed data').
- PKCS #9 Selected attributes. This describes aspects of the content of the 'signed data' structure.
- PKCS #11 Crypto tokens. This describes the operation of PKI tokens with respect to the interface with client software running on the PC.

⁹ MIDlets are Java ME applications that are installed on the handset (and not the SIM card). See <http://java.sun.com/javame/index.jsp>.

- PKCS #15 Crypto token information format. This describes the content of the token, specifically certificates, the selection of a key if a token contains several private keys, etc.

It should be noted that Mobile PKI solutions generally do not employ PKCS#11 and higher, as the token used for Mobile PKI has no direct interaction with the client-side, i.e., the user's PC.

2.2.4 Standards for electronic signatures

Several countries, including the Netherlands¹⁰, introduced legislation based on a European directive from 1999 [2]. Article 2 of this directive defines a number of requirements that a so-called advanced electronic signature must meet. An advanced electronic signature is required:

- to be uniquely linked to the signatory;
- to be capable of uniquely identifying the signatory;
- to be created using means that the signatory can maintain under his sole control;
- to be linked to the data to which it relates in such a manner that any subsequent change of the data is detectable.

Additionally the directive states the requirements for a so-called qualified certificate. These relate not so much to the technology used (X.509 is the de facto standard) as to the requirements for the CA. The requirements refer to attributes embedded in the certificate, such as name and location of the CA, the valid uses of the certificate (signing, authentication, encryption, ...), validity dates, etc.

According to the ETSI Mobile PKI standards, electronic signatures are also classified within the European Electronic Signature Standardization Initiative (EESSI) of the International Communications and Technology Standards Board (ICTSB). This initiative distinguishes between general electronic signatures (which cannot be processed automatically but have legal status), qualified electronic signatures (that as a consequence of the technical implementation should have at least the same legal status as a written signature) and an enhanced electronic signature with the same properties as the European directive's advanced electronic signature. The EESSI tasks appear to have been taken over by ETSI-ESI¹¹. The terminology from the European directive now appears to be leading.

The Common Criteria protection profile for Secure Signature Creation Devices (described in [3]) also refer to "qualified" certificates as well as "qualified electronic signatures".

Mobile PKI solutions have the ambition to implement advanced signature devices that allow the creation of signatures using qualified certificates.

2.2.5 Mobile Signature Service (MSS)

Mobile PKI is described extensively in Chapter 3. This section is therefore restricted to an overview of the relevant standards with a short description.

¹⁰ See <http://www.e-overheid.nl/e-overheid-2.0/live/binaries/e-overheid/juridisch/wet-elektronische-handtekening.pdf>

¹¹ See <http://portal.etsi.org/esi/el-sign.asp>

- ETSI 102 203 – “Business functional requirements”.

This presents a non-technical introduction to Mobile PKI explaining the concepts of the other ETSI documents in advance. It also describes a large number of application scenarios, and contains a justification of the design criteria and a possible way to assign the various roles.

- ETSI 102 204 – “Web service interface”.

This describes the syntax of the messages exchanged between the MSSP (the server) and the Application Providers (APs). This is the most concrete of the ETSI documents describing Mobile PKI.

- ETSI 102 206 – “Security framework”.

This comprises a framework of security requirements for a Mobile PKI solution. Unfortunately this is technology-agnostic and therefore abstract. These requirements are the foundation for Chapter 4.

- ETSI 102 207 – “Roaming”.

The ETSI standards for Mobile PKI support roaming at the transaction level whereby transactions are passed on through a chain of linked MSSPs. This does not apply to the current pilot and is therefore ignored in this report.

- WPKI document [16].

This document is more recent than the ETSI standards. The document claims to have minimal dependencies on the ETSI standards. It is less abstract than the ETSI standards. Its status is unclear (not a standard).

- Common Criteria protection profiles for SSCDs.

Software used for security purposes can be certified in an independent process in order to compare the quality of the suppliers of these products. The so-called Common Criteria (CC) are standards for this certification. Part of the CC certification is the selection of a so-called protection profile (PP). Such a PP (see [3]) exists for devices that are used for creating electronic signatures (so-called Secure Signature Creation Devices, SSCDs). The PP for SSCD systematically presents the requirements for an SSCD. It also describes the environment in which an SSCD is used.

The ETSI standards for Mobile PKI are “technology-agnostic” and thus provide little detail (with the exception of the SOAP specification for the traffic between AP and MSSP). They do comprise many requirements, including security requirements. It is striking that concepts such as “SIM card” and “mobile phone” are hardly mentioned as such; a distinction is made between the bare Signature Creation Device (in practice this is the SIM Toolkit application on the SIM card) and the Signature Creation Application (in practice this is the mobile phone).

2.3 Conclusion

Mobile PKI technology is based on the basic ingredients that have been around since 2001 and are technically mature and standardised. Mobile PKI does require reasonably advanced (and possibly more expensive) SIMs, however. There may have been a lack of business opportunity for mobile operators to introduce such SIMs up till now. However,

the issues regarding the conditions for a large-scale implementation of Mobile PKI in the Netherlands are beyond the scope of this report.

The use of open standards ensures that an identity provider or application provider such as SURFnet can adopt the technology and integrate the solutions with its own systems with relative ease.

Recent years (2007-2009) have shown an increased adoption of Mobile PKI, especially with respect to Valimo's solutions. A large number of pilots have been undertaken and Mobile PKI is already being put to actual use by banks to provide services (for instance in Turkey, Scandinavia and the Baltic states). SIM card manufacturers and the mobile operators, the major parties that play a part in Mobile PKI, appear to engage in partnerships to bring Mobile PKI to the market.

3 Architecture

This chapter describes the architecture of Mobile PKI solutions in general as well as the specific Valimo solution. We present the registration and authentication (and signing) processes of Mobile PKI at a higher abstract level. In the case of the Valimo solution we describe the authentication process in more detail, based on the experiences from the small-scale pilot (as much as could be determined).

3.1 Processes

Every identity management process comprises a number of stages: creation, registration (provisioning of identities, enrolment of users, issuing certificates), use of identities, revocation (recall, retraction) of identities.

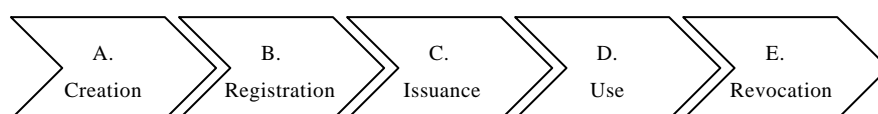


Figure 3: Lifecycle for means of authentication

In actual practice all sorts of variants may exist, but these stages can generally be distinguished.

Apart from the identity of the user, the *means* of authentication (such as the smartcard) have a life cycle of their own. The life cycle consists of stages such as: creation, personalisation, issuing, use, termination. Sometimes the creation substage is also taken to include the *development* of hardware and software. This happens, for instance, in security evaluation in the context of the Common Criteria: protection profile [3] mentions the following stages in the lifecycle in Figure 3 on page 8: Design, Fabrication, Initialization, Personalization, Usage, Destruction. This report is restricted to the registration and the use of means of authentication. The other stages may differ from case to case.

The Registration stage links an identifier (usually a name or a number) to an end-user and provides the end-user with credentials (in this case an application on a SIM card containing the private key). The identifiers are included in the user’s public key certificate.

During the Use stage, sessions are established between users, APs and the Identity Provider (IdP). Each session can be divided into stages: sign on, authentication, authorisation, use, and sign off.

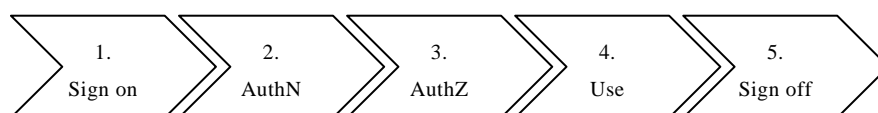


Figure 4: Stages in an authentication session

The two sets of stages have a completely different timescale. We will consider both and they will be reflected in the architecture/implementation. They are also the foundation for the security analysis in Chapter 4.

3.2 Roles

Each of the stages in the process involves different parties in different roles:

- The end-user;
- The mobile operator (the MO);
- The manufacturer of the means of authentication (the SIM) and the corresponding back-end systems (the MSSP), for instance Valimo;
- The IdP, for instance SURFnet;
- The certificate authority (CA), for instance Diginotar;
- The application provider (AP), for instance an educational institution connected to SURFnet, in some scenarios this may be SURFnet itself.

Please note that the various responsibilities, discussed in more detail in the two processes below, may be assigned to various of the parties involved. For instance, managing the attributes that belong to the identities may be assigned exclusively to the IdP or may be shared across AP, IdP, CA, MO and the end-user. Similarly, checking the authenticity (or integrity) of transactions may be allowed for all or may be restricted to a subset of these parties.

3.3 Registration process

Registration (also referred to as provisioning or enrolment) is the process in which the identity of an end-user is initially recorded in the system. We assume that an end-user already has a relation with an MO and possibly also with other parties such as the IdP. These parties have already stored attributes pertaining to the end-user, using a specific identifier as an index. These identities must be linked in the databases of the various parties involved and the CA has to create credentials (generate a key pair with the private key stored on the SIM and the public part in the MSSP). Several links need to be made between the identifiers under which a user is known to the various parties.

The diagram in Figure 5, based on the descriptions in [8], [10] and [16], provides an example of the infrastructure that may have been employed for the pilot.

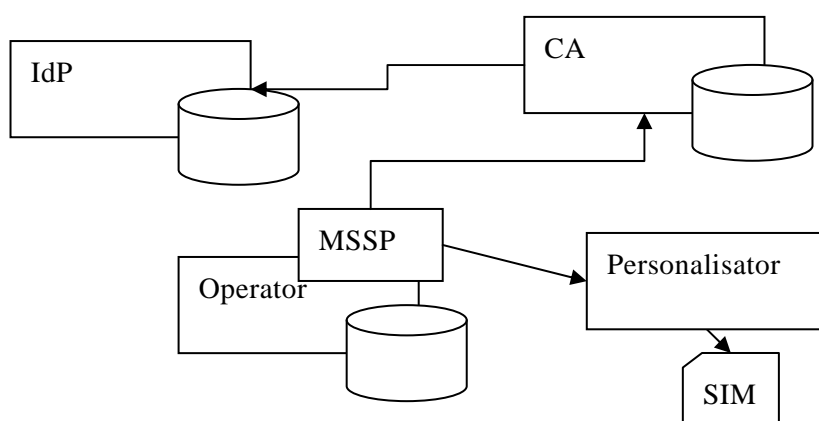


Figure 5: Parties and messages during a (possible) registration process.

Variations on this infrastructure are possible. The assumption is that both the IdP and the MO already know the user. The user administration of the MO is used by the CA to create certificates. These certificates will include the SIM's International Mobile Subscriber Identity Number (IMSI) as an attribute. The public key that is comprised in the certificates must correspond to the private key as it appears in the SIM application

toolkit applet. There are a number of ways to achieve this, the first being the most probable:

- The key pair is generated in the SIM: The card, including the application, is issued to the user; the application generates the keypair and sends a Certificate Signing Request (CSR) to the CA (by way of the mobile operator). The CA signs the certificate and the public part is stored in the MSSP.
- The CA generates the keypair: The CA generates the keypair and the private key is sent securely to the personalisator. The SIM containing the application and the keys is issued to the user. The CA signs the certificate that is stored in the MSSP.
- The MSSP generates the keypair: The MSSP generates the keypair and sends the CSR to the CA. The CA signs the certificate, this is stored in the MSSP and the mobile operator sends the private key to the personalisator.

In all cases the application on the SIM contains the private key which, if proper procedure is followed, is not stored elsewhere during the process. The public key certificate always remains in the MSSP (although in principle it may also be stored in the application on the SIM). The IdP only has to store the identifier that refers to the certificate and requires an online connection to the MSSP.

The role of the personalisator (often the manufacturer of the SIM card), who loads the applications onto the SIM card on behalf of the MO, may be taken on by the MO itself because the SIM Toolkit API supports over-the-air application management. This means that the Mobile PKI application may be installed and personalised post-issuance, i.e., after the SIM card has been issued and placed in the user's mobile phone.

3.4 Authentication process

Authentication is the process in which the end-user proves to the IdP that he holds the credentials that were issued to him. In this case that is done by signing a message that was generated by the IdP using the private key that is present on the SIM. Authentication takes place at the request of the end-user when attempting to interact with a service provided by an AP where the AP requires confirmation of the identity of the end-user. Often the communication of stored attributes by the IdP to the AP (the so-called attribute release) is subsumed as part of the authentication process.

Figure 6 (again based on the descriptions in [8], [10], [16] and on a Valimo marketing presentation¹²) represents the Mobile PKI authentication process. The protocol involves all parties. In this case, it is initiated by the end-user who connects to the AP using a PC. Because authentication is required, the AP redirects the user to the IdP. The IdP connects to the MSSP which retrieves the user's credentials, specifically the IMSI number, and checks with the CA if the corresponding end-user certificate is still valid. If that is the case, then the IdP generates a message containing a random challenge and sends this to the MSSP. The MSSP now sends an SMS text message to the telephone number of the user by way of the MO (who is able to look up the telephone number (MSISDN) that belongs to a specific IMSI). The SMS text message contains a request to sign the message generated by the IdP. The SMS text message is not meant to be seen or read by the end-user, but should be processed by the SIM (a service SMS message). The Mobile Valimo VMAC applet on the SIM signs the challenge after the user has been requested to enter his PIN code in the GUI of the handset. The AP receives the signature by way of

¹² PowerPoint presentation by Erkki Saharanta, Valimo Wireless Ltd, available at <http://www.oasis-open.org/events/forum/2006/slides/saharanta.ppt>.

the MO and MSSP and is then able to act upon this (for instance by retrieving further attributes).

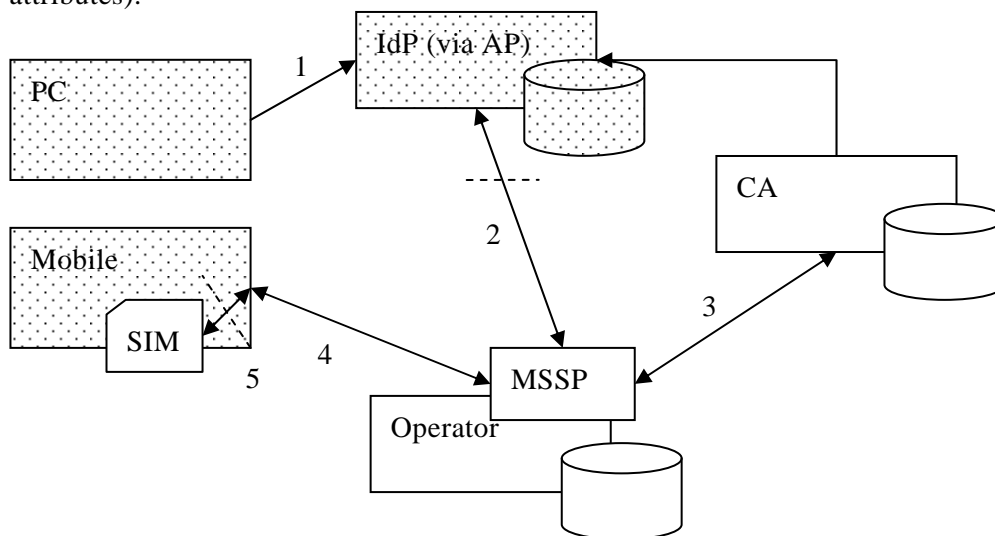


Figure 6: Parties and messages during the authentication (or signing) process.

The IdP takes the following steps to check the response to its challenge message that it receives from the SIM (by way of the MSSP):

1. The message contains the public end-user certificate that corresponds to the private key in the SIM. The IdP uses the root certificate of the CA to check that the end-user certificate has been signed with the private key that corresponds to the CA's root certificate.
2. The message also comprises a PKCS#7 structure which, if all is well, contains the original challenge. The IdP checks if this is really the case. The IdP also checks if this structure was signed using the private key corresponding to the end-user certificate.
3. The IdP regularly communicates with the CA to determine if the certificates are still valid¹³.

The APs may be allowed access to the evidence (challenge and response) proving the identity.

The dotted areas in the figure indicate the components under the control of the authors of this report during the pilot. Section 3.5 provides a brief technical analysis of what happens at interface 2 (the exchanges between MSSP and AP) and interface 5 (the traffic between MO and SIM).

The figure positions the MSSP with the MO. Alternatively, the MSSP could be hosted by the IdP, the CA, or Valimo, the manufacturer of the means of authentication. It seems most obvious however, that the MO hosts the MSSP. Our part of the SURFnet pilot employed an MSSP hosted by Valimo.

3.5 Technical analysis of various interfaces

Although the ETSI standards are agnostic about the technology that is employed (page 14 of [10]), we can achieve a reasonable insight into how Valimo implements the set of

¹³ Either by retrieving a Certificate Revocation List or by using the Online Certificate Status Protocol (OCSP).

requirements by subjecting the Valimo solution to a number of tests in the authentication scenario. We analyse the traffic between IdP and MSSP (interface 2 in Figure 6) and the traffic between the mobile telephone and the SIM card (interface 5 in Figure 6). After all, we have these two interfaces at our disposal during the pilot. These analyses provide a more detailed view of the security checks that are to be performed during a transaction.

3.5.1 Interface between AP and MSSP on the internet

The IdP sends an XML message (a SOAP call according to [9]) to the MSSP over interface 2 containing the ‘data to display’ and a request for signing the ‘data to sign’. An example of such a request is included in Appendix A.

The IdP is then authenticated (in the pilot this was done using a password) and the MSSP sends an SMS text message to the end-user’s SIM who then signs the ‘data to sign’. The MSSP responds on interface 2 with an XML message (a SOAP response according to [9]). An example of such a response is included in Appendix B.

This response contains a base64-encoded PKCS#7 signed data structure. This structure has three parts. Firstly, the PKCS#7 structure contains a collection of so-called signed attributes. Three such attributes exist:

- 1.2.840.113549.1.9.3 (ContentType) → 1.2.840.113549.1.7.1 (data)
- 1.2.840.113549.1.9.25.3 (randomNonce) → 30B18CD97E66EB29
- 1.2.840.113549.1.9.4 (messageDigest) →
D32C2CE05DD77BEF3687B40D361532FD4030716D

The first attribute indicates the type of the signed content. The second attribute ensures that old transactions cannot be reused. The last attribute turns out to be a SHA1 hash of the ‘signed data’ exactly as it was submitted.

Secondly, the PKCS#7 structure contains the end-user’s public key certificate:

Certificate:

Data:

Version: 3 (0x2)

Serial Number:

90:28:79:d8:34:75:9e:83:24:df:02:8c:41:d5:e2:1a

Signature Algorithm: sha1WithRSAEncryption

Issuer: C=fi, O=valimo, OU=RSADevpCA

Validity

Not Before: Jun 16 15:23:41 2009 GMT

Not After : Jun 16 15:23:41 2010 GMT

Subject: C=fi, O=Valimo, OU=nonrep, CN=8935801080084400490F

Subject Public Key Info:

Public Key Algorithm: rsaEncryption

RSA Public Key: (1024 bit)

Modulus (1024 bit):

00:92:b8:58:e4:50:af:27:c7:bd:40:55:59:df:c8:9e:14:7f:b9:f0:6e:05:68:ed:b4:a4:33:49:3d:23:

57:91:af:e6:4e:b2:bc:aa:ff:55:c4:91:2b:19:24:9f:52:66:c4:32:5f:60:c1:38:04:6b:57:9d:7c:81:

07:8d:c1:ce:24:cb:81:c0:1a:4f:4d:a1:65:e4:a5:34:00:ec:3f:6d:d2:1f:94:2a:9c:c1:18:6e:4b:01:

ee:ef:0e:d2:0c:e0:96:ce:14:31:13:a3:d3:78:91:5a:8c:f3:d0:71:cb:6d:5a:75:0d:e0:39:fe:30:6f:

2b:cd:8e:96:9e:e9:72:05:29

Exponent: 65537 (0x10001)

X509v3 extensions:

X509v3 CRL Distribution Points:

DirName:/C=fi/O=valimo/OU=RSADevpCA

Signature Algorithm: sha1WithRSAEncryption

55:29:2f:5a:c1:60:02:dc:a9:3b:61:e4:55:aa:e5:d7:0e:3a:07:b0:41:0b:fc:62:a3:19:32:3e:7f:be:a0:77:2a:5f:ce:b4:ee:5f:36:ba:29:a9:0e:da:7b:e4:1f:8d:79:de:0b:9e:c8:ab:d3:25:a0:2b:08:07:83:34:c1:a4:7b:b0:ee:be:bd:1b:74:b4:4e:2a:5c:09:84:2e:a3:e9:90:16:c2:1c:43:b9:9d:80:4d:6d:9a:47:50:56:35:1d:ba:0d:d9:4f:b7:db:62:69:65:40:89:85:9c:20:62:87:d4:15:26:5b:7a:83:e0:83:48:81:b9:68:61:77:09:72

Note that the name in the subject of the certificate comprises the SIM's IMSI number: /C=fi/O=Valimo/OU=nonrep/CN=8935801080084400490F. The Valimo test root certificate is used to determine that this included certificate was indeed signed by Valimo.

Thirdly, the PKCS#7 structure contains the actual signature covering the attributes (an RSA encryption of the SHA1 hash of the signed attributes):

5FC6951CA01DBE0C1B4C6457B04855B58931214034540016C496260452EE8EE5894B51F11FE195DED55330729101F2638C0BE8E28B145B8ED02FF85BD1753BB238D058510357B6A4CAE5AD2FAFF9268365EFB8C5ADA37467820262B52767C1EAFABA75F051B5171743BC1CD0C759AB2CBC1B574153C5ED819832C7653D38B82

In the pilot the signature covering the signed attributes turned out to be incorrect. This was most probably caused by an incorrectly configured MSSP.

3.5.2 Interface between handset and SIM

The communication between the mobile phone and the SIM card can be analysed by placing a so-called "Season Interface"¹⁴ between the contacts of the SIM and the card reader in the handset. The big question here is if and how end-to-end security (secure messaging) is established between the VMAC applet and the MSSP. A proprietary high-level communication protocol is generally used at this interface, although it must be based on the GSM 11.14 standard (see [6]) and is subject to the ETSI Mobile PKI standards (see [10]). Unfortunately it was not possible to complete this experiment successfully within the scope of this project.

3.6 Conclusion

In a scenario as used in the pilot, and as it typically would be employed by SURFfederatie, Mobile PKI uses two channels in communicating with the user: one channel through the user's PC and the second channel through the user's mobile phone. Authentication of the user is initiated over the PC channel and confirmed by checking whether the user has the corresponding keys in a SIM Toolkit application on the SIM card by way of the GSM channel.

A number of parties (MO, IdP/AP, CA) are involved in the authentication process. Each of these parties manages a part of the user's identity (an identifier with a number of attributes). The architecture is flexible, which allows for many variations in the configuration. A different configuration means that the responsibilities are assigned to the stakeholders in a different way. However, the MO has a special role because it is the owner of both the SIM and the SIM Toolkit application installed on it. This means that Mobile PKI can be used only if the MO collaborates, in contrast to SMS-OTP, for instance, where the MO only needs to pass on data.

¹⁴ See http://en.wikipedia.org/wiki/Pirate_decryption#Terminology_and_Definitions.

4 Security analysis

Two possible approaches can be taken to determine the security of Mobile PKI solutions. Firstly, the ETSI standards that describe Mobile PKI comprise a framework of security requirements (see [10]). This document was the starting point for this chapter and several of the requirements will be discussed in section 4.2. Secondly, there is also a Common Criteria protection profile (see [3]) for mobile signature devices that also formulates requirements for Mobile PKI solutions that are to be certified at the EAL4+ level. Section 4.3 outlines where and how these requirements differ from the ETSI standard requirements.

However, this chapter starts with an ad-hoc security analysis in which the authors have attempted to approach the specific Valimo solution from the viewpoint of an attacker during the pilot.

4.1 An ad-hoc threat analysis

As indicated in chapter 3, a Mobile PKI solution consists of several components. The most important components are: a server hosted by the MO (the MSSP), the user's mobile phone (the SCA) and the SIM Toolkit application on the SIM in the mobile phone (the SCD). The user's identity is determined by the user's SIM card. Each of these components store data that deserve protection: the so-called assets. We identify the following assets based on chapter 2 (how does PKI work, how does Mobile PKI work) and chapter 3 (the architecture of Mobile PKI).

- Private keys that are stored on the SIM card.

These private keys are stored in the SIM memory and are used during transactions to convince the MSSP of the authenticity of the SIM card. With proper security measures these keys cannot be extracted, not even by the end-user.

- The SIM itself.

A stolen SIM card may be used by others to conduct transactions while posing as the user. This requires that the attackers have the user's PIN code. This could be obtained later under false pretences or could possibly be guessed.

- PIN code.

An obtained PIN code is useful only in combination with a stolen SIM card. A properly secured SIM allows an attacker three guesses at most before the PIN check blocks the SIM.

- Integrity of the server and underlying systems.

APs connect to the MSSP over a network and this requires that the MSSP is online and therefore can be reached by attackers. Typically this asset will be protected by requiring mutual authentication between the MSSP and APs. This also comprises the privacy related attacks that reveal the online behaviour of users (against the wishes of the users) and attacks against the registration procedures (having the SIM cards of unsuspecting users blocked, registering non-existent users).

The assets in a Mobile PKI solution are protected by a number of security measures: the SIM card is a smartcard that is able to resist all kinds of attacks, that is blocked when an incorrect PIN code is entered too many times, etc.

We attempt to build as complete a picture as possible of potential attacks and their impacts by approaching the situation from the position of an attacker. We assume a standard authentication scenario (see Figure 6) in which a user is trying to access a service provided by an AP. The impact of the attack is given a qualification of either [OK] or [BEWARE].

- Passive online attacker
 - An online attack is made significantly more difficult by using two channels (to the PC over internet, to the mobile phone over the MO's network). [OK]
 - The traffic between PC and IdP/AP is generally secured by SSL and this means that a passive attacker can hardly do any harm. The traffic between IdP/AP and MSSP is preceded by authentication of the IdP/AP using a username/password combination (at least during the pilot). This does not provide sufficient security. Mutual authentication based on certificates over an SSL link would be an improvement. [OK]
 - The GSM standard describes securing the traffic between the mobile telephone and MSSP (by way of the MO) using algorithm A5, but recently the strength of this algorithm has been called into question . [BEWARE]¹⁵

It is probable that the SIM Toolkit application first establishes a secure channel before the application opens communications over the unsecured GSM channel (as the Global Platform standard provides options for this). [OK]
- Active online attacker
 - The use of two channels is also an impediment to an active online attacker. [OK]
 - An active attacker could attempt to seduce a user to visit a fake AP through a phishing attack over the PC channel. However, because an AP can only communicate with the MSSP after successful authentication, the fake AP will not be able to initiate a transaction that involves the user's mobile phone. [OK]

However, a naive user might be seduced into entering his PIN code over the PC channel. [BEWARE]
 - Obviously, the MMSP and the AP server must be sufficiently secured against the attacks that originate on the network. This is a standard security issue. [OK]
- Attacker with access to the PC
 - An attacker with physical access to the PC still has no access to the user's mobile phone and therefore cannot pose as the user to conduct transactions. [OK]

¹⁵ See, for instance, <http://reflexor.com/trac/a51>.

- Similarly, installed malware on the user's PC does not pose an actual threat. [OK]
- However, installed malware may be used to launch a so-called “man in the browser” attack¹⁶, in which the unsuspecting user is made to believe that he is engaging in a (low-value) transaction with the AP that he selected, while in actual fact a trojan in the browser posing as the user is conducting a (high-value) transaction with another AP. [BEWARE]
- Any installed malware might also record which APs are visited by the user (this poses a privacy problem). This is not related to the authentication solution.
- Attacker with access to the mobile phone
 - An attacker with physical access to the SIM cannot obtain the private key(s) and PIN code (under the assumption that the smartcard platform is sufficiently secured). This is a standard security issue. [OK]
 - An attacker who steals (a mobile phone with) a SIM does not know the PIN code for the Mobile PKI application (nor the global PIN code of the SIM card if this is enabled). [OK]

In general, people quickly notice that their mobile phone has been stolen, more quickly than other PKI tokens. The user's certificate can then be revoked at the MSSP. [OK]

 - It is unclear if malware installed on the mobile phone (for instance a Java ME, Symbian or Windows Mobile application) can disrupt the functionality of the SIM Toolkit application (both with respect to the user interface as well as receiving/sending SMS text messages) and in that way could give the user the impression that the Mobile PKI PIN is requested, or that a transaction has or has not taken place. [BEWARE]
 - The threat of malware on the mobile platform is not as real as it is on PCs, [OK] but this may change as mobile phones are used for more and more services. [BEWARE]
- Attacker with access to the servers (an insider)
 - An (employee at an) AP could be in the position to locally modify or substitute transactions and make a user sign a (high-value) transaction while the user is led to believe he is signing another (low-value) transaction. [BEWARE]
 - An (employee at an) AP could be in the position to implement a “mafia in the middle”¹⁷ attack and lure users to his AP (for instance through phishing). In a “mafia in the middle” attack an unsuspecting user thinks he is conducting a (low-value) transaction while an intermediate party poses as the user in conducting a (high-value) transaction at another AP. [BEWARE]

Mobile PKI appears to be quite a secure solution, apart from a number of standard security issues regarding the SIM smartcard that is used.

¹⁶ See, for instance, http://en.wikipedia.org/wiki/Man_in_the_Browser.

¹⁷ See Chapter 2 of Ross Anderson's *Security Engineering*.

Under certain circumstances (malware on the end-user's PC, insider attack at the AP) it may be possible for an attacker to mount a "man in the middle" attack. The user will then be made to sign a transaction that is different from the intended transaction. Such an attack can be recognised by end-users if they scrutinise the Data To Be Signed (DTBS) that is displayed on the handset. This must correspond to the transaction that is being conducted. This requires that APs construct the DTBS in such a way that users understand what the transaction that they are agreeing to entails.

Mobile PKI requires that the MSSP and the APs mutually authenticate each other in order to determine that they are communicating with the right party. In the case of the SURFnet pilot, the MSSP only communicates with the SURFnet IdP. The APs that are connected to the SURFnet IdP have a trusted relationship with the IdP. Mobile PKI can be used by these APs as an additional authentication process, for instance for step-up authentication.

4.2 ETSI security requirements

ETSI 102 206 [10] presents a framework of requirements. The document describes requirements for the whole of the Mobile Signature Creation Service and for the individual components that are part of this service: the SIM card, the mobile phone, and the MSSP server (requirements are classified as either MSCS, MSCD, MSCA and MSSP). The requirements are formulated at a relatively abstract level, which allows providers of Mobile PKI ample freedom in making choices regarding implementation.

It would go too far to go through the entire list of requirements in this chapter. This chapter therefore only presents a selection of the requirements, in the expectation that this provides the reader with sufficient insight into the status of the security.

For example, general requirements for the Mobile PKI service include:

- *The DTBS shall contain a Signer's Document. (MSCS1)*
This requirement is not as much a requirement for the Valimo solution as a requirement for the IdP/AP. The IdP/AP is responsible for compiling the SOAP request that is submitted to the MSSP.
- *The DTBS shall contain the Signer's Certificate related to the Signature Creation Data that the MSCD uses to generate the Mobile Signature and intended by the Signer. (MSCS2)*
Several systems must be properly attuned to each other in order to meet this requirement: the AP has to indicate which user is involved and the CA has to provide the proper certificate, namely the certificate that belongs to the private key in the user's SIM.
- *The MSCA, MSSP and MSCD shall maintain the confidentiality of the DTBS components, DTBSF, DTBSR, Mobile Signature, and SDO. (MSCS8)*
This requirement will be met if the SIM Toolkit applet establishes a secure channel to the MSSP using a shared key.
- *The System shall ensure that the DTBS components used to create the DTBSF and DTBSR are the same as those presented to the Signer during the representation process and that they are identical to those signed by him. (MSCS9)*

This requirement refers to the consistency in the solution implemented by the IdP/AP, MSSP and the SIM Toolkit application.

An example of the requirements for the SIM card (including Valimo's application) are:

- *The MSCD shall provide unambiguous detection of physical tampering that might compromise the security functions. (MSCD8)*

This requirement entails that the smartcard platform on which the SIM is based provides some resistance against attacks that attempt to retrieve the private keys from the SIM using hardware-based methods.

An example of a requirement for the mobile phone is:

- *A Trusted Path for the transaction of SAD to the SSCD shall be provided through the MSCA. (MSCA15)*

This is taken to mean that a GSM handset must have a secured keypad to ensure that a the PIN that is typed in can securely reach the SIM card. For instance, any installed key logger on the mobile phone should not be able to impact the interactions between the SIM Toolkit application and the user. This is quite an extensive requirement that may not satisfied by all manufacturers of handsets.

An example of a requirement for the MSSP server is:

- *The MSCA and the MSSP shall mutually authenticate each other, to assure the Signer that the MSSP and this particular Mobile Signature request can be trusted. (MSSP10)*

The ETSI requirements cover the majority of the threats. The Valimo solution appears to meet the ETSI requirements in most aspects. The authors of this report do not have relevant information for those aspects where it is not clear if the Valimo solution meets the requirements.

4.3 Security requirements from the Common Criteria protection profile

A presentation of the protection profile [3] will help to make the security analysis more complete, even though it cannot be established with complete certainty that the Valimo SIM has been subjected to the Common Criteria standard tests for this protection profile.

A Common Criteria protection profile describes the requirements that are imposed on a product (the so-called Target of Evaluation, or TOE) and the environment in which the TOE occurs.

The protection profile commences with a list of assets. It has been reproduced in Table 2. The protection profile subsequently systematically lists several assumptions about the TOE and its environment: the threats that are relevant to these assets, and the security policies of the organisation such as the use of qualified certificates. Other examples are that only the signatory can use the TOE and that in practice data can be signed only once.

1. SCD: private key used to perform an electronic signature operation (confidentiality of the SCD must be maintained).
2. SVD: public key linked to the SCD and used to perform an electronic signature verification (integrity of the SVD when it is exported must be maintained).
3. DTBS and DTBS-representation: set of data, or its representation which is intended to be signed (Their integrity must be maintained).
4. VAD: PIN code or biometrics data entered by the End-user to perform a signature operation (confidentiality and authenticity of the VAD as needed by the authentication method employed)
5. RAD: Reference PIN code or biometrics authentication reference used to identify and authenticate the End-user (integrity and confidentiality of RAD must be maintained)
6. Signature-creation function of the SSCD using the SCD: (The quality of the function must be maintained so that it can participate to the legal validity of electronic signatures)
7. Electronic signature: (Unforgeability of electronic signatures must be assured).

Table 2: Assets of a Signature Creation Device according to the Common Criteria protection profile.

A systematic overview of the security aims for the TOE and its environment are then based on this threat analysis and these assumptions. This is subsequently used for formulating a long list of requirements for the TOE and its environment. Many of the requirements in [3] are similar to the ETSI requirements described in section 4.2. One may assume that the authors of the ETSI standard may have based their work in part on this protection profile.

However, there are a number of aspects in which [3] differs from the ETSI standard. Firstly, the protection profile (and the Common Criteria in general) address a much larger part of the lifecycle of the TOE, such as whether the TOE is manufactured and configured in a responsible manner. Secondly, the protection profile makes a more general distinction between TOE (the SCD, i.e. the SIM card) and the environment of the TOE (the SCA, the MSSP). Though [3] also imposes requirements on the environment, the emphasis is on the TOE. Thirdly, a number of very specific requirements are included that refer to the physical characteristics of the SIM card. Advanced attacks such as ‘differential power attack’ are mentioned specifically, whereas ETSI remains somewhat more abstract.

4.4 Conclusion

Mobile PKI is a very secure solution that meets the extensive programme of requirements (as far as we can determine) as formulated by ETSI. The solution is much more secure than username/password combinations because it involves a hardware token that meets stringent requirements (Common Criteria EAL4+ certified).

However, a “man in the middle” attack by a malevolent AP is always possible if a user cannot recognise the real IdP/AP by the DTBS displayed on his handset.

5 Application for higher education

In this chapter we consider the impact of Mobile PKI technology for SURFnet and higher education in general, including potential practical applications in the short term. We place an emphasis on use in SURFfederatie, but we also list some other possible uses.

5.1 Present situation: username/password

The *de facto* most widely used means of authentication in higher education is a username/password combination. This is also the most widely used means of authentication in SURFfederatie. It is not suited for services that require a higher level of security, since a username/password combination is very vulnerable to phishing and is easily shared with colleagues. Experts seem to agree that username/password combinations will disappear in time as a sole authentication method, although there appears to be little consensus regarding the timeframe in which this will happen. Advantages of username/password combinations are of course that they are cheap (e.g. no costs relating to hardware or communication) and that people are used to them. However, the costs of resetting passwords (when people forget them, for instance) can be high if this is to be done securely.

5.2 Criteria for the use of Mobile PKI

The trade-off between security and user-friendliness must be taken into proper consideration when introducing a safer means of authentication such as Mobile PKI. The advantage of Mobile PKI is that it is not only more secure than a username/password combination, but that it also is reasonably user-friendly.

The following criteria can be used to select services that are suited for a Mobile PKI introduction in the short term:

1. *High level of security* – For instance, services that are vulnerable to fraud, or involve privacy sensitive information, or may have a large detrimental impact on the integrity of systems or networks require a more secure means of authentication.
2. *Low number of users* – If only a relatively small number of users is given access to the service, then the costs will be less of a problem.
3. *For employees* – Mobile PKI will only require collaboration with one mobile operator for access for employees of an organisation that has a corporate contract with a single mobile operator. In contrast, a service for all students in the Netherlands will require the cooperation of all mobile operators.

5.3 Potential services for using Mobile PKI

In meetings during the project the following services were named as candidates for a short term introduction of Mobile PKI for authentication. These have been grouped as services for research, for education and for ICT/network.

Research:

Grid and e-science authentication: Access to grid and e-science resources typically uses PKI, with all the known disadvantages of distributing certificates.

Setting up lightpaths: The high costs of this service require a high level of authentication.

Education:

Access to exam results and student administration in general: The privacy sensitive nature of the data requires a high level of authentication.

Exam certificates at connected institutions: This service is typically very sensitive to fraud. It often, if not always, uses paper.

ICT/network:

Requesting certificates: This currently employs a USB PKI token but would be more user-friendly if employing a mobile phone. Specifically, this would no longer require the installation of software on the PC.

DNS registration: An effort has been initiated for this service to use SMS One-Time-Password as step-up authentication. Mobile PKI would be both more user-friendly and more secure than SMS OTP.

Account management at institutions: The Mobile PKI identity may be used to manage provisioning and password resetting of accounts at the connected educational institutions. This saves on costs for the helpdesk with respect to password reset, and it is more user-friendly and more secure.

VPN authentication: This requires a high level of authentication due to the security requirements (for the network and for the identity of the users).

6 Conclusions and recommendations

The question arises of how Mobile PKI compares to other solutions. What are the Unique Selling Points for Mobile PKI?

- Mobile PKI uses a “something you have” token, specifically the SIM card in the mobile phone. This overcomes many problems that are associated with, for instance, simple username/password authentication. Phishing attacks, for example, are made a thing of the past.
- As with a TAN list, Mobile PKI uses an external channel. Mobile PKI has many advantages with respect to user-friendliness because most users have a mobile phone with them but few carry a list of TAN codes.
- As with SMS-OTP, Mobile PKI uses an external channel, i.e., the user’s mobile phone. Mobile PKI has the advantage that it does not require the user to type in a code from the mobile phone on the PC. The PIN code the user must enter on the mobile phone is always the same.
- An ‘OTP token with display’ (for instance a banking token) also requires that the user types in codes from the handset on the PC. A user may forget such a token, but hardly ever forgets his mobile phone. The latter disadvantage does not hold if the OTP token is a SIM Toolkit application running on the SIM card, instead of a separate hardware token.
- Users are more likely to forget their USB PKI token or PKI smartcard than they are likely to forget their mobile phone. Also, some tokens require the user to enter a PIN code using the keyboard of a PC, which cannot be trusted to be secure (consider key loggers and other malware) and they require installed hardware (a card reader) and/or software (drivers, middleware).

This comparison only considers authentication solutions. Properly speaking, only the last solution, the USB PKI token, can be compared to Mobile PKI because it allows digital signing.

Please note that the over-the-air capabilities of the SIM Toolkit API deliver additional advantages of Mobile PKI over traditional USB PKI tokens and smartcards, because this allows for a flexible migration path from simple (unqualified) certificates to qualified certificates. The mobile operator in fact provides a secure connection to the SIM card that allows updates to be performed post-issuance.

The fact that it depends on the MO is also the weakness of Mobile PKI. Introduction is possible only in collaboration with an MO. A heterogeneous group of users (with contracts with different MOs) can only make use of Mobile PKI if all MOs cooperate.

Based on the preceding remarks and the conclusions in the various chapters (2.3, 3.6, 4.4) we conclude:

- Mobile PKI technology is based on the standard components that have been around since 2001 and are technically mature and standardised. The use of open standards ensures that IdP/APs such as SURFnet can adopt the technology with relative ease. However, the technology requires quite advanced (and therefore expensive) SIMs. The mobile operators, who own the SIMs, play a key part in the implementation. The introduction of Mobile PKI at a national level is possible only with the support of all mobile operators.

This report does not answer the question why mobile PKI has not yet been introduced on a large scale in the Netherlands. Some progress can be discerned in recent years (2007-2009). Many pilots have been announced or undertaken. Mobile PKI has also been deployed for banking services, for instance in Turkey, Scandinavia and the Baltic states. Valimo, the provider of the technology employed in this pilot, features relatively often in press releases and has set up partnerships with all large SIM manufacturers and many European mobile operators.

- The Mobile PKI architecture is very flexible as a result of standardisation. This allows for many variations in the configuration. The mobile operator plays an important part in every configuration variant because it manages the access to the SIM card.
- The security of Mobile PKI is quite sufficient. ETSI has formulated an exhaustive programme of requirements for Mobile PKI. The implementation by Valimo meets these requirements (as far as the authors of this report have been able to determine). Mobile PKI is a much stronger form of authentication than username/password combinations because it gives the SIM card (which adheres to the stringent Common Criteria requirements) an essential task. The only threat scenario, a “mafia in the middle” attack executed by an untrusted application provider, or a “man in the browser”, is possible only if the users do not pay proper attention.
- Regarding the application of Mobile PKI for SURFnet, it seems that the security is unnecessarily strong for many of the current applications. Depending on the costs this does not need be an impediment, as the solution is user-friendly. The solution does have merits for some target groups (requiring a high level of security and involving a low number of employees). It seems advisable to experiment with the solution for these target groups and to delay further introduction until Mobile PKI technology is more widely adopted in the Netherlands and more insight is gained into the costs for support by all three mobile operators. In the meantime, a more extensive use of SMS one-time-passwords may be considered for step-up authentication or password reset.

Appendix A - Example request to the MSSP

This XML document is an example of a request such as the IdP (generally speaking an AP) might send to the MSSP. The password and telephone numbers have been blanked out.

```
<?xml version="1.0" encoding="UTF-8"?>
<s:Envelope xmlns:s="http://schemas.xmlsoap.org/soap/envelope/">
  <s:Body xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema">
    <MSS_Signature xmlns="http://uri.etsi.org/TS102204/v1.1.2#">
      <MSS_SignatureReq MinorVersion="1" MajorVersion="1" MessagingMode="synch" TimeOut="180">
        <AP_Info AP_ID="http://mssp.valimo.com/diginotar" AP_TransID="_x0032_20091007T160109"
        AP_PWD="*****" Instant="2009-10-07T16:01:09.+02:00" AP_URL=""/>
          <MSSP_Info>
            <MSSP_ID>
              <DNSName/>
            </MSSP_ID>
          </MSSP_Info>
          <MobileUser>
            <MSISDN>+350000000000</MSISDN>
          </MobileUser>
          <DataToBeSigned MimeType="text/plain" Encoding="UTF-8">data to sign</DataToBeSigned>
          <DataToBeDisplayed MimeType="text/plain" Encoding="UTF-8">data to display</DataToBeDisplayed>
          <SignatureProfile>
            <mssURI>http://mss.valimo.com/DiginotarDS</mssURI>
          </SignatureProfile>
          <KeyReference>
            <CertificateIssuerDN>C=fi, O=Valimo, OU=RSADevpCA</CertificateIssuerDN>
          </KeyReference>
        </MSS_SignatureReq>
      </MSS_Signature>
    </s:Body>
  </s:Envelope>
```


References

- [1] 3G Americas, Identity Management – Overview of Standards & Technologies for Mobile and Fixed Internet, January 2009
- [2] European Parliament and Council, Directive 1999/93/EC, Official Journal of the European Communities, December 13th 1999
- [3] CEN/ISSS, E-SIGN workshop - expert group F, Common Criteria EAL4+ Protection Profile - Secure Signature-Creation Device Type 3, Versie 1.05, 25 July 2001
- [4] ETSI SR 002 176 Algorithms and Parameters for Secure Electronic Signatures
- [5] ETSI, TS 100 977 v8.10.0 (2003-09), Digital cellular telecommunications system (Phase 2+); Specification of the Subscriber Identity Module – Mobile Equipment (SIM-ME) interface (3GPP TS 11.11 version 8.10.0 Release 1999), 2003
- [6] ETSI, TS 101 267 v8.13.0 (2003-03), Digital cellular telecommunications system (Phase 2+); Specification of the SIM Application Toolkit (SAT) for the Subscriber Identity Module – Mobile Equipment (SIM-ME) interface (3GPP TS 11.14 version 8.13.0 Release 1999), 2003
- [7] ETSI, TR 101 903 XAdES
- [8] ETSI, TR 102 203 v1.1.1 – business functional requirements
- [9] ETSI, TR 102 204 v1.1.4 – web service interface
- [10] ETSI, TR 102 206 v1.1.3 – security framework
- [11] ETSI, TR 102 207 v1.1.3 – roaming
- [12] GlobalPlatform Card Specification Version 2.2, available at <http://www.globalplatform.org/>, March 2006
- [13] ISO/IEC, Identification cards – Integrated circuit cards Part 4: Organization, security and commands for interchange, ISO/IEC 7816-4: 2005
- [14] Sun, Java Card Platform Specification 2.2.2, available at <http://java.sun.com/javacard/specs.html>.
- [15] Elena Trichina, Konstantin Hypponen, Marko Hassinen, SIM-enabled Open Mobile Payment System Base don Nation-wide PKI, In proc. Secure Electronic Business Processes, 355 – 366, ISSE/Vieweg, 2007
- [16] WPKI Projectgroup, WPKI Main Specification, Rev. 2.2 (and appendices), available at <http://www.wпки.net/>, May 14th 2009

