Master's Thesis

The Virtual SIM

— *A Feasibility Study* —

By

Marc Richarme

In collaboration with:

DTU

NOKIA
Connecting People

Copenhagen,
21 April 2008

Preface

This Master's thesis has been prepared at the Department for Informatics and Mathematical Modelling (IMM) at the Technical University of Denmark (DTU), in partial fulfilment of the requirements for the degree of Master of Science in Engineering.

This thesis has been prepared over six months and the workload amounts to thirty-five ECTS credits.

Acknowledgements

I would like to thank Nokia Danmark A/S for having me as a project student. This unique opportunity to be in an inspiring and dynamic environment with access to people and information central to the problem has been the main factor in making this thesis possible.

I wish to express my sincere appreciation to my supervisors, Henning Olesen, Peter Anglov, Peter Vestergaard and Jens-Ole Madsen for their support and guidance during the preparation of this thesis. Special thanks also go to Jais Agertoft for giving me the opportunity to prepare my thesis with Nokia.

I would also like to thank the following people for sharing their knowledge, and making it possible to get a broad understanding of the different perspectives portrayed in this report:

Antti Kiiveri, Valtteri Niemi, Lauri Paatero, Jukka Parkkinen, Jens Benner, Mikael Buchholtz, Tjaard Meier, Palle Staffeldt, Henrik Jørgensen, Terhi Pere (all from Nokia); Peter Aaage (unspecified smart card vendor); Ross Campbell (TeliaSonera); Claus Rasmussen (Cloakware); Stefan Kaliner, Martin Froels (both from T-Mobile).

Finally, thanks goes out to everybody else from Nokia who has helped me in various ways.

Copenhagen, April 2008

Marc Richarme

Abstract

This report demonstrates the feasibility of replacing the physical SIM card – the subscriber authentication token used in every GSM and UMTS terminal – by a virtual counterpart, the 'Virtual SIM'. The SIM card is a critical component for mobile operator as it plays a key role in billing subscribers for their use of the network, and in addition protects against various types of fraud. Furthermore, operators use the SIM card as a platform for the deployment of advanced network- and user services. This means that a Virtual SIM scheme must satisfy many challenging requirements regarding security, functionality and operational feasibility. Furthermore, in order for such a system to gain the acceptance of terminal manufacturers, it must also be low-cost and preferably have other advantages justifying the extra effort and risk incurred by moving the SIM functions to the terminal.

The chosen approach leverages the secure processor already existing in many mobile terminals to provide the required level of protection to the critical SIM functions. The system also largely builds on technologies already present in most existing mobile phones to provide platforms for supplementary applications and for remote management.

An over-the-air provisioning system makes it possible to sell subscriptions in the form of numeric codes to be entered in the terminal, which is compatible with all existing subscription sales channels. In addition the system makes it possible for operators to implement web-based subscription portals that users can access directly from their mobile terminal, even if it doesn't already have a working subscription. This allows existing users to get their subscription into new phones and new users to buy a subscription directly from their mobile terminal, and use it immediately after.

The proposed scheme has several advantage of the physical SIM card. In the terminal side some advantages are saved board space, increased the durability of phones, increased standby time, better SIM lock, and even a prevention mechanism again device counterfeiting. For operators, the Virtual SIM effects cost savings and opens the door to several new ways of selling subscriptions.

Contents

Glossary

This glossary is meant as a quick reference for the reader, and not as comprehensive definitions of the terms. Source: Nokia termbank ([1]), Wikipedia English version ([2]), and own work ([3]).

Baseband 5 security (BB5)

Security architecture used in new Nokia devices. BB5 security provides boot integrity checking ensuring platform security, and also provides a secure execution environment contained within the chip, allowing the execution of small protected applications (PAs) isolated from the outside world.[3]

Churn rate

Proportion of subscribers who leave an operator in a given time period. This is a measure of how well an operator retains subscribers, and can be seen as an indicator of the competition on the market, but many factors are involved.[2]

Global System for Mobile Communications (GSM)

Digital system for mobile communications used worldwide.[1]

International Mobile Equipment Identity (IMEI)

Identity with which the mobile station can be uniquely identified. The IMEI serves as the serial number of the device.[1]

International Mobile Subscriber Identity (IMSI)

World-wide unique subscription identifier stored on the SIM card. The IMSI serves as a key to derive subscriber information such as directory number(s) from the home location register (HLR).[1]

Mobile Network Operator (MNO)

Telecommunications company which provides network services and maintains a telecommunication network. Note: This term is used interchangeably with the terms 'operator' and 'mobile operator'. In the Unites States, an MNO is called a 'carrier'.[1]

Mobile Virtual Network Operator (MVNO)

Network operator that appears in all significant respects to the customer as a fully-fledged mobile operator, yet has neither its own radio network nor its own licence, but instead pays for access to at least one of the existing operators' networks.[1]

Open Mobile Alliance (OMA)

Organisation that acts as a mobile industry standards forum aiming at interoperable mobile services across geographic areas, network operators, and mobile terminals and at an open standards-based framework that permits services in a multi-vendor environment.[1]

Over-the-air technology (OTA)

A technology that enables the operator to transfer data over the air to terminals and remote sites. OTA technology can be used, for example, to update the contents of data fields in the SIM card or to download applications, such as ringing tones or games, remotely to a wireless device.[1]

Provisioning

This refers to the setting up of new services for an existing subscriber of a mobile phone network. This can include provisioning of device settings required for the use of services such as GPRS, MMS, Instant Messaging, etc., or it can be application provisioning, where an application is sent to *the handset or to the SIM card. Provisioning is often achieved using over-the-air (OTA) technology, e.g. using SMS messages.*[2,3]

Public Land Mobile Network (PLMN)

Mobile network for the specific purpose of providing land mobile communication services to the public. GSM is an example of public land mobile network system.[1,3]

SIM Lock

Restriction applied to a mobile handset subsidised by a mobile operator, in order to prevent the user from using it with a subscription from another operator.[2]

Subscriber Identity Module (SIM)

Security module that is inserted into a piece of mobile equipment for subscriber identification and other security related information.[1]

Universal Integrated Circuit Card (UICC)

Multi-application smart card platform used in mobile telecommunication networks. The term UICC denotes the non-technology specific card platform, which can contain applications specific for a certain use, e.g. an USIM application in the case of UMTS.[3]

Universal Mobile Telecommunications System (UMTS)

Third generation mobile communication system based on WCDMA. UMTS can be considered the successor of GSM, and it has been designed to be backwards-compatible with GSM and reuses much of the existing network infrastructure.[1,3]

Universal Subscriber Identity Module (USIM)

Security module for UMTS. As opposed to the SIM card, this term denotes the logical application performing the SIM functionality, and not the physical card itself (see UICC).[3]

Vanilla handset

Handset destined to be sold independently of any operator, and which has no operator customization or network/SIM lock.[3]

Virtual Subscriber Identity Module (VSIM)

An implementation of the (U)SIM functionality inside the terminal device instead of in a separate physical module (UICC). In this work, the term VSIM will be used both to denote the concept as a whole, and to denote the piece of data transferred from the operator to the terminal to get the subscription running. The latter use maintains the analogy with the SIM card, as a subscription is activated on the handset by the transfer of a VSIM, which can be compared to the insertion of a SIM card.[3]

Abbreviations

AuC	Authentication Centre
BB5	Baseband 5 (security) *(see glossary)*
BSC	Base Station Controller
BTS	Base Transceiver Station
DM	Device Management *(see glossary)*
DRM	Digital Rights Management
EIR	Equipment Identity Register
GMSC	Gateway MSC
GSM	Global System for Mobile Communications *(see glossary)*
HLR	Home Location Register
IMEI	International Mobile Equipment Identity *(see glossary)*
IMSI	International Mobile Subscriber Identity *(see glossary)*
ME	Mobile Equipment *(terminal excluding the SIM card)*
MNO	Mobile Network Operator *(see glossary)*
MS	Mobile Station *(terminal including the SIM card)*
MSC	Mobile services Switching Centre
MVNO	Mobile Virtual Network Operator *(see glossary)*
ObC	OnBoard Credentials
OMA	Open Mobile Alliance *(see glossary)*
OMC	Operation and Maintenance Centre
OTA	Over-The-Air *(see glossary)*
PKC	Public Key Cryptography
PKI	Public Key Infrastructure
PLMN	Public Land Mobile Network *(see glossary)*
POTS	Plain Old Telephone Service

PSTN	Public Switched Telephone Network
SE	Secure Element
SIM	Subscriber Identity Module *(see glossary)*
UMTS	Universal Mobile Telecommunications System *(see glossary)*
VLR	Visitor Location Register
VSIM	Virtual Subscriber Identity Module *(see glossary)*

Chapter 1

Introduction

The SIM card is a ubiquitous authentication token that is present in all GSM and UMTS handsets. In this thesis, the possibility of replacing the physical SIM card by a virtual counterpart is examined.

According to the report of the first SIM Expert Group (SIMEG) meeting in January 1988, “the SIM is the physically secured module which contains the IMSI, an authentication algorithm, the authentication key and other (security related) information and functions. The basic function of the SIM is to authenticate the subscriber identity in order to prevent misuse of the MS (Mobile Station) and the network” [1]. Since then, the SIM card has evolved far beyond its original purpose and is now capable of controlling various aspects of the phone’s behaviour and providing value added functionality to the end-user [2]. Future SIM capabilities, such as the USB interface, device management server, smart card web server, and DRM, to name only some, are increasing the SIM’s ability to compete against differentiating features in the terminal, and against the terminal’s brand [3]. An alternative to the physical SIM card, specifically a ‘Software SIM’ or ‘Virtual SIM’[1] integrated in the device, would be one way to stop this development.

The GSM standards were established in such a way that a removable smart card was to be used to store subscriber information, the authentication algorithm and the secret subscriber key used to connect to the mobile network. During the initial work on the SIM card, SIMEG considered the possibility of having a ‘fixed’ SIM, but it was rejected for the following reasons: First, there was no viable way to securely store the secret algorithm and subscriber key in the terminal. Second, there was the problem of how to load (and subsequently replace) the key and algorithm in the terminal [4, 5]. Finally, a ‘fixed’ SIM could be a commercial barrier with respect to the possibility of free trade with mobile equipment [6]. Indeed, in GSM’s early days the interoperability of terminals and networks in the system, facilitated partly by the use of the removable SIM card, allowed competition to flourish, resulting in lower prices and leading to mass-adoption of GSM [7].

[1] The term ‘Virtual SIM’ will be used in this report, since the concept involves much more than merely a software implementation of the SIM functions.

Many high-end mobile phones sold today have built-in security hardware which offers a level of tamper-resistance comparable to that of smart cards and which could potentially solve the problem of protecting the authentication algorithm and the key. Furthermore, progress in the field of cryptography might make provisioning schemes that were not realizable in the eighties possible. For instance, public key cryptography (PKC) was considered for GSM security. Two of the main reasons it wasn't adopted were that implementations were immature at the time, and that longer keys were required than with symmetric cryptography [8]. Today, PKC is widely used, and key length is much less of an issue due to the technological development of execution platforms and algorithms. With this in mind, it is not unconceivable that a well-designed provisioning scheme could support an open terminal market and overcome the commercial concerns raised.

For mobile operators, the use of the Virtual SIM could result in cost savings on account of the SIM cards and the logistics necessary for their distribution to end-users. The SIM industry has revenues of over two billion euros (2006) [3], which come from mobile operators. It seems clear that both mobile operators and device manufacturers have something to gain from finding an alternative to the SIM card.

However, the move from the physical SIM to the Virtual SIM impacts the established system on many levels. First, the SIM vendors could become obsolete and disappear from the SIM ecosystem (at least in their present form), which implies some level of changes to the existing processes. Furthermore, mobile operators rely on the SIM card to authenticate users, and thereby ensure that only legitimate subscribers can use the mobile network and that they are properly billed, which is an extremely business-critical function. Operators are free to choose which SIM vendor(s) to trust with this responsibility, and SIM vendors are experts in security and have security certifications for their products and processes. In the Virtual SIM model, operators will have to trust device manufacturers, since (in an open terminal market) they do not control the devices accessing their networks. This also raises liability issues due to the potential situation where a failure in the security of a terminal would result in abuse of the mobile network and financial loss for the operator.

The questions which remain are whether the obstacles can be overcome, and whether the benefits from having a Virtual SIM outweigh the possible disadvantages.

1.1 Rationale for this study

Although the physical SIM card has contributed to the success of GSM, this does not necessarily imply that an alternative could not fill its role, and the technological evolution of mobile terminals is removing the original barriers.

Since July 2007, a technical specification group (TSG) in 3GPP (the standardization body responsible for GSM and UMTS) has been investigating the feasibility of “Remote management of USIM application on M2M Equipment” (TR 33.812 [9]), which is a concept similar to the Virtual SIM, but solely for the use in the machine-to-machine communication domain, thus eliminating many of the constraints present in the consumer segment. While, the M2M industry has some needs that cannot be met by regular SIM cards, a potential Virtual SIM standard developed for this purpose could also be made to cover normal consumer terminals, but this is not addressed in the 3GPP study.

In December 2007, Motorola proposed that the same TSG should “consider whether the time is right to start work on replacing the current hardware SIM card with a secure downloadable version that is stored and runs directly in a secure environment on the UE", i.e. for 'normal' handsets [10]. This proposal was, however, rejected by mobile operators.

Clearly, the concept of a Virtual SIM is something that industry players need to take a position on, and an analysis of the technical and business aspects is needed, which is what this thesis aims to provide.

1.2 Research questions

Qualitative studies are usually based on research questions instead of hypotheses, [11] and this convention will also be used here. Werner and Schoepfle suggest an approach with a primary question defining the overall goal of the study followed by sub-questions narrowing the focus without constraining it [12]. Using this approach, the primary research question of this thesis is:

- Should the idea of using a Virtual SIM for the GSM/UMTS networks be pursued?

The following sub-questions describe the main focus areas of the thesis:

1. What are the success criteria?
2. Which technical approach should be used?
3. What are the advantages and disadvantages compared to the status quo?
4. Is the concept economically sound?
5. Would the introduction of the Virtual SIM be an advantage for Nokia?

1.3 Scope

The Virtual SIM concept could be reused for other current and future access technologies, but in this report, focus will mainly be on 2G/3G mobile telephony, specifically GSM and UMTS. Specifically, the technical analysis will only aim at describing how the Virtual SIM fits into these systems. Furthermore, the goal of the technical analysis is to prove the feasibility of the concept on a high level, and the details of an actual implementation are outside the scope of this study. The study focuses of the use of Virtual SIMs in mobile handsets, although other application areas are possible, e.g. machine-to-machine communication devices.

1.4 Empirical methods

This study is based on information obtained through an extensive literature study comprising academic articles, books, technical standards, and working documents from standardization bodies. Furthermore, information has been obtained through interviews and discussions with technical experts as well as people with general knowledge and experience about the topics covered, including people working for mobile operators, smart card vendors and terminal vendors. Some of the conducted interviews were semi-structured and conducted in a formal manner (notes and transcripts available in appendix A), but a lot of the knowledge gained has been so through informal discussions.

The study was carried out at Nokia in the team responsible for the SIM card, which has made it possible to get an understanding of the actual state of the industry that could never have been acquired solely through interviews and/or a literature study. On the downside, this has made it difficult to attribute some bits of information to a specific source.

1.5 Organisation

This report is organized into seven chapters. The first chapter presents the concept of the Virtual SIM and explains the motivation behind this study. Chapter 2 gives the reader an introduction to the SIM card and the environment in which it operates, both in terms of technologies and of the industry landscape. Chapter 3 gives an analysis of the challenges and requirements the Virtual SIM has to take into account. Chapter 4 presents a viable technical solution that takes into account the identified requirements, and in Chapter 5, the feasibility of the proposed solution is assessed. Finally, chapter 6 puts the Virtual SIM concept into a broader perspective, and chapter 7 concludes on the finding of the report.

Chapter 2

Background

The purpose of this chapter is to give the reader the required background knowledge to understand the topics discussed in the rest of the report. However, a certain level of prior technical understanding is assumed. The first section aims at introducing GSM and UMTS systems and the role of the SIM card. The second gives a more comprehensive introduction to the functions of the SIM, and the last section gives an overview of the mobile industry and its current development, which is important in order to understand the implications of the Virtual SIM.

2.1 GSM and UMTS architecture

The GSM architecture can be considered as consisting of four subsystems, as shown in Figure 2.1:

- The mobile station (MS) is in the users' possession and consists of the mobile equipment (ME), i.e. the handset, and of the SIM card.
- The base station subsystem (BSS) constitutes the access-part of the network and consists of base transceiver stations (BTS) and base station controllers (BSC), the latter implementing the 'intelligence' behind the BTSs, i.e. radio channel allocations, handovers.
- The network subsystem (NSS) constitutes the core of the network. It consists of mobile switching centres (MSC) routing voice traffic within the operators network, as well as gateway MSCs interfacing with other telephone networks. For packet data connections, such as GPRS, an entity called the servicing GPRS support node (SGSN) acts as a router within the network, and the gateway GPRS support node (GGSN) routes traffic to the outside world (none of these is shown in Figure 2.1). The home location register (HLR) is essentially a database of mobile subscribers, and the associated authentication centre (AuC) holds the secret key for each subscriber and handles the authentication procedure which will be described in the following section. Supplementing the HLR, one or more visitor location registers (VLRs) store temporary information regarding other operators' users roaming on the network.

- The operation subsystem (OSS) is used to monitor, diagnose and manage the network, functions which are handled by the operation and maintenance centre (OMC). Finally, the equipment identity register EIR is a database of mobile terminal identities, which can be used to block stolen terminals from accessing the network.

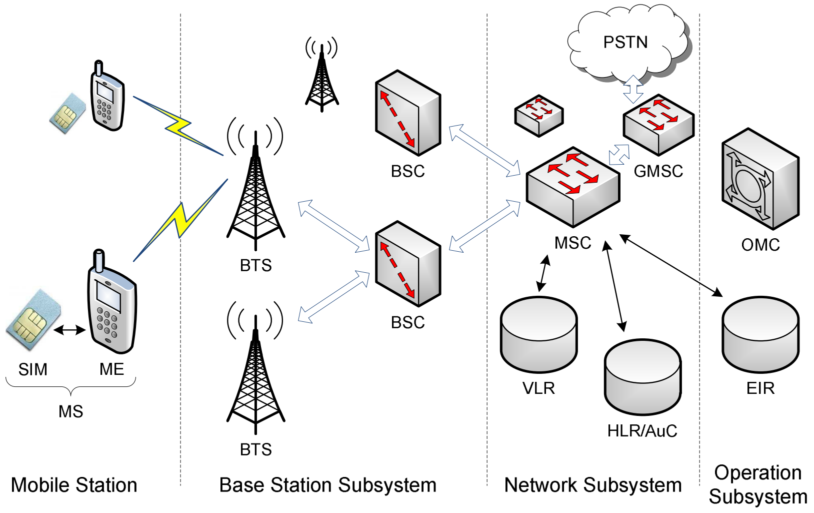


Figure 2.1: Overview of the GSM architecture. Source: adapted from [13, 14].

When roaming, users are said to connect to a ‘visited network’ or ‘serving network’, while their own operator is referred to as the ‘home network’.

A more detailed description of the GSM architecture can be found in the several books, e.g. [13] or [14], or in the standards (GSM TS 03.02).

UMTS largely reuses the GSM architecture, the major differences being in the BSS part. For the purpose of this thesis the differences are insignificant. A detailed description of UMTS can be found in e.g. [15].

2.2 GSM and UMTS Security

The GSM system was designed with the simple security goal of providing a degree of protection of the radio path approximately equal to that provided in the fixed network [8]. This resulted in the following security requirements for the GSM system [16]:

- Subscriber identity authentication. This protects the network from unauthorised use.
- Subscriber identity confidentiality. This provides protection against the tracing of a user’s location by listening to exchanges on the radio interface.
- User data confidentiality across the radio interface. This protects the user’s connection orientated data against eavesdroppers on the radio interface.
- Connectionless user data confidentiality across the radio interface. This protects user information sent in connectionless packet mode in a signalling channel from eavesdropping on the radio interface.
- Signalling information element confidentiality across the radio interface. This protects selected fields in signalling messages from eavesdropping on the radio interface.

To meet these goals, cryptography is used to authenticate the subscriber and (optionally) to encrypt the data passing over the radio interface.

2.2.1 Authentication in GSM

In GSM and UMTS, subscriber authentication is based on the SIM card, which is a hardware token placed into any mobile terminal. Each SIM contains an identifier for the subscriber (known as the IMSI) and a secret symmetric key called the subscriber key (or Ki, in short) used to authenticate the subscriber. This key is only shared with the authentication center (AuC) of the mobile operator that issued the SIM.

The IMSI (International Mobile Subscriber Identity) is a globally unique identifier. Its structure is shown in Figure 2.2: It is composed of a tree digit long mobile country code (MCC), a two or three digit long mobile network code (MNC), specifying which operator the subscriber belongs to within the given country, and finally a mobile subscriber identity number (MSIN), identifying the specific subscriber within that operator's user base.

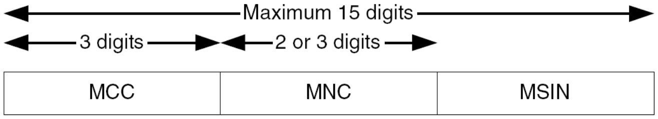


Figure 2.2. Structure of the IMSI. Source: 3GPP TS 23.003 [17].

When a subscriber wants to access the mobile network, the IMSI from the SIM card is sent to the corresponding operator's AuC, which finds the Ki corresponding to that subscriber in a database. In order to grant access to the network, a check is made to verify that the SIM also knows the corresponding Ki. To this end, the AuC generates a random challenge and encrypts it using Ki as key. The challenge is then sent to the SIM card which is requested to perform the same calculation and to send back the result to the network. The network can then check if the two results match. If so, the Ki in the SIM is known to be the same as that stored in the AuC, and the subscriber is considered to be authenticated and is granted access to the network.

This scheme ensures that the subscriber key never leaves the SIM or the AuC, which is especially important when a subscriber is roaming in a network other than his 'home' network. In this case, the 'visited' network routes the authentication request to the AuC of the home network which then follows the same procedure as described above. Only in this case, it is the visited network's responsibility to check whether the results from the SIM and the AuC match. Thus, the Ki doesn't need to be disclosed in order to authenticate the user, and this procedure can be performed even if the link between the SIM and the AuC isn't trusted.

In fact, this scheme authenticates the SIM card rather than the user. In order to tie the SIM to a subscriber, a secret 4 to 8 digit personal identification number (PIN) normally needs to be entered by the user in order to unlock the SIM, although this is optional. When the PIN is used, this can be seen as a two-factor authentication scheme. To prevent attempts to guess the PIN, SIM cards have a limit of 3 wrong tries before it is blocked. If this happens by accident, users can unblock it using an 8 digit personal unblocking key (PUK), a procedure which can be attempted 10 times before the SIM is irreversibly blocked.

The SIM can be an anonymous or pseudonymous credential, as there is no technical requirement to combine its information with any personal information of the actual subscriber [18]. For post-paid subscription, billing information is obviously required, but pre-paid subscriptions can be entirely anonymous, and often are.

Figure 2.3 illustrated the SIM-based authentication process in more details: The AuC first generates a random number RAND. Based on the IMSI, the AuC looks up the subscriber-specific key, Ki, in its database and uses it with algorithm called A3 to encrypt RAND, thereby forming the number SRES. The random number RAND is sent to the terminal, which passes it on to the SIM card. The SIM card also contains the A3 algorithm and the Ki for that subscriber, and the same calculation is performed. The resulting SRES is then passed back to the network and compared with the one calculated by the AuC. If they match, this proves that the SIM card contains the correct subscriber key (and algorithm), which is considered a proof its authenticity.

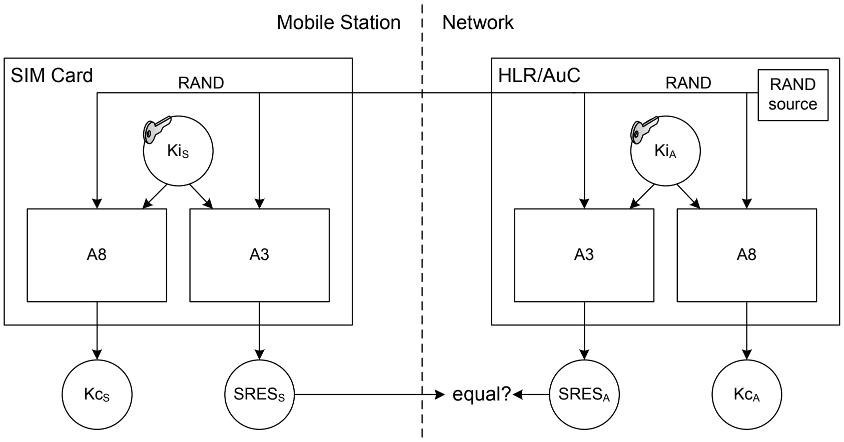


Figure 2.3: Challenge-response authentication in GSM. Source: based on description in [13].

It is clear that the subscriber key is the cornerstone of GSM security, and one of the main functions of the SIM card is to protect its secrecy.

2.2.2 Encryption in GSM

Another algorithm, called A8, is also run in the AuC and the SIM during the authentication process. The output number is normally referred to as Kc, and is used as a session key for encrypting the data traffic, if encryption is enabled for the network (which is not necessarily the case). The bulk encryption is not performed by the SIM card, and instead, Kc is passed to the terminal, which then takes care of (de)ciphering data sent to/from the terminal using an algorithm called A5. On the network side, the encryption/decryption is normally handled by the BTS[2].

In fact, when the user is authenticated, the AuC generates an ‘authentication triplet’ consisting of (RAND, SRES, Kc) which can be used by the network to authenticate the subscriber and encrypt traffic without any further need to contact the AuC. This is particularly useful for roaming users, where the AuC for a given user may be located half way across the world, and there may be significant communication delays. In this case, the AuC is typically requested to generate several authentication triplets, which makes it possible for the visited network to authenticate the user more than once without contacting the home operator.

[2] Thus, the user data is only encrypted on the radio-interface, and not within the network.

2.2.3 UMTS Security

In UMTS, the biggest addition to the GSM security model is mutual authentication. That is, the SIM also verifies that the terminal is connected to a legitimate serving network, which is useful for preventing 'fake base station' attacks to which GSM is vulnerable[3]. To this end, a number called AUTN is sent to the SIM card during authentication, alongside the RAND number. AUTN is calculated by the network and consists of a message authentication code, MAC, and an (encrypted) sequence number, SEQ, which is incremented every time an authentication is attempted. Based on the RAND value and its Ki, the SIM card can calculate whether the correct MAC has been sent by the network, i.e., whether the network is also in possession of the secret key. The SIM card also maintains a corresponding sequence counter, which is used to verify the freshness of the authentication request, in order to prevent replay attacks.

Another omission of GSM security is integrity protection of signalling messages. This is addressed by UMTS by the use of message authentication codes attached to these messages. These codes are generated based on an integrity key, IK, which is derived from the subscriber key. Just as with the Kc key (which has been renamed CK in UMTS), the serving base station must know the value of IK in order to verify the traffic from the user. Thus, in UMTS systems, instead of authentication triplets the AuC generates authentication quintets consisting of (RAND, SRES, CK, IK, AUTN). An algorithm called MILENAGE is used to generate and verify these numbers. This algorithm can be seen as the successor of A3/A8.

For a detailed description of UMTS security, see [19] or [20, chapter 6].

2.2.4 Algorithms

The A3/A8 algorithm pair (which is normally implemented as one algorithm) can be specified by each operator, since it is implemented in the SIM card which is issued by each operator, and the AuC (which is also operator-specific). On the other hand, the A5 algorithm is defined globally, since it must be implemented in every base station and mobile terminal.

It may be desirable for operators to use a custom algorithm for A3/A8 for security reasons. The GSM Association originally specified an example algorithm for A3/A8 called COMP128 (or COMP128-1), but this algorithm was broken in 1998 by Berkeley researchers [21], making it possible to clone SIM cards. Operators who used a custom algorithm were not vulnerable to this attack [22] (which doesn't necessarily imply that their algorithms weren't vulnerable, but merely that the specific attack targeted the widely used example algorithm).

To address this issue, the GSM Association has released two successors to the algorithm, called COMP128-2 and COMP128-3. The COMP-128 algorithms have never been officially publicized, which was without doubt one of the main reasons that the mentioned vulnerability was not found earlier (e.g. before the algorithm was being used). Instead, many operators now use GSM-MILENAGE, a modified version of the MILENAGE algorithm [23]. An overview of all GSM related algorithms can be found on GSMA's web site [24]. As with A3/A8, operators are also free to create their own implementation of MILENAGE instead of using the example algorithm. Many operators stick to the example algorithm (e.g. TeliaSonera [25]). However, some operators

[3] In fact, it is still possible to create fake base stations in UMTS which relay messages between the network and the subscriber, but it is not possible to obtain the content of the messages, to modify them, or to insert spurious messages in the data stream, so the attack is essentially limited to an advanced denial of service, which could just as well be achieved by radio jamming.

do use a custom version. T-Mobile, for instance, uses a modified version of the example algorithm [22].

The bulk encryption algorithm in UMTS (the successor of A5) is called KASUMI.

2.3 SIM primer

The SIM card plays a central role in the mobile station in the GSM and UMTS systems. Indeed, the SIM card is what binds a mobile subscription to the handset, and without it, the mobile ecosystem where the handset and the mobile subscription are two separate entities would not exist. This section will give an introduction to the SIM card and how it assumes this important role.

From the user's point of view, the SIM can be considered the physical incarnation of the subscription. It takes the form of a smart card, which is a small, tamper-resistant integrated circuit, embedded on a plastic card. It is removable from the handset, making it possible for the user to migrate the subscription to a different handset by simply removing the card from the old device and placing it onto the new one. It is also possible to change the subscription used with a given handset by replacing the SIM card. However, many mobile operators sell subsidised handsets, where the user is contractually bound to a particular subscription for a certain period of time[4], and where the handset is locked to the particular operator's network, so that it cannot freely be used with any SIM card [26]. This is called "subsidy lock", "network lock" or, most commonly, "SIM lock".

2.3.1 Basic functions

The SIM card's single most important function is to identify and authenticate the subscriber towards the mobile network. In addition to subscriber authentication, the SIM card also has a range of other functions. These functions can be categorized into four general areas, namely 'security', 'data storage', 'management functions' and 'supplementary applications'. These are depicted in Figure 2.4.

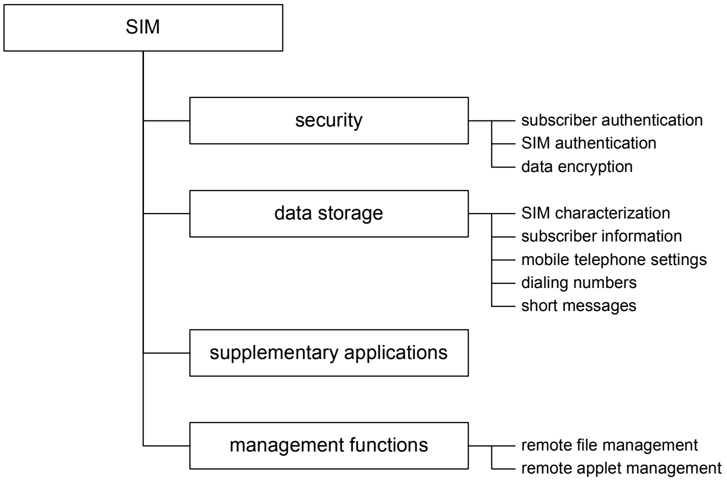


Figure 2.4: Classification of the basic functions of the SIM card. Source: inspired by fig. 13.10 in [27].

[4] Typically up to 24 month [26].

2.3.1.1 Security

The security-related features of the SIM card include the ability to authenticate the subscriber and to provide encryption of traffic on the radio path. Indeed, one of the key strength of the SIM is its tamper-resistance, which allows it to securely store the subscriber identity key (Ki) and to run the authentication algorithm in an environment protected from attackers. SIM cards (and smart cards in general) go to extensive lengths to provide these security features, as will be described in section 2.3.2.

2.3.1.2 Data storage

The SIM card features a file system storing various pieces of data permanently attached to the mobile subscriber, and various network-related parameters. It also holds data that can be dynamically updated by the handset, as well as user data, such as address book entries and short text messages. The SIM card may also hold non-standard files for internal use, but of course, the terminal won't normally be aware of their presence. The file system is organized hierarchically with a distinction between directory files (DF) and elementary files (EF). Elementary files can be of three different types, as shown in table Table 2.1.

The existence of a SIM file does not necessarily mean it can be accessed or overridden. Each file has an associated access control list defining how it can be used, depending on the authorization level of the user. Authorization is granted by the use of PIN codes, and the end-user is normally given two different PINs: one for normal use and one for administrative use. The second one (called PIN2) can for example be used to restrict the SIM to allow only calls to a specific set of destination numbers during normal use (fixed dialing numbers). Furthermore, there is a default mode active when no PIN has been entered yet, and five additional administrative levels in the SIM cards requiring other codes. These are used during card manufacturing and personalization to write data that cannot subsequently be altered. They are also used by network operators to alter SIM files that cannot be modified by the end-user.

Table 2.1: SIM file types. Source: [2].

Linear Fixed	**Transparent**	**Cyclic**
Many records, all are the same length.	A single block of data.	Many records, all are the same length.
Last Record does not wrap to first record.		Last Record wraps to first record.
Used mainly by the Phonebook.	Used for most files.	Used for Last Number Dialled.

A few concrete examples of SIM files are shown in Table 2.2 below. Some of the functions they fulfill are [2]: SMS settings and storage, last dialed numbers, phonebook, and roaming list. The latter (also known as the 'PLMN list') is a prioritized list of mobile networks used to determine which network the terminal should connect to if the home network cannot be reached.

Table 2.2: A few examples of SIM files and a description of their function. Source: 3GPP TS 31.102 [28].

Filename	Content / description
EF_{IMSI}	IMSI
	This EF contains the International Mobile Subscriber Identity (IMSI).
EF_{Keys}	Ciphering and Integrity Keys
	This EF contains the ciphering key CK, the integrity key IK and the key set identifier KSI.
EF_{HPPLMN}	Higher Priority PLMN search period

	This EF contains the interval of time between searches for a higher priority PLMN.
EF_{FPLMN}	Forbidden PLMNs
	This EF is read by the ME as part of the USIM initialization procedure and indicates PLMNs which the UE shall not automatically attempt to access. A PLMN is written to the EF if a network rejects a Location Update with the cause "PLMN not allowed".
EF_{LOCI}	Location Information (LAI+TMSI)
	This EF contains the following Location Information: Temporary Mobile Subscriber Identity (TMSI); Location Area Information (LAI); Location update status.
EF_{AD}	Administrative Data
	This EF contains information concerning the mode of operation according to the type of USIM, e.g. normal, type approval, cell testing, or manufacturer specific. It also provides an indication of whether some ME features should be activated during normal operation as well as information about the length of the MNC, which is part of the International Mobile Subscriber Identity (IMSI).
EF_{ECC}	Emergency Call Codes
	This EF contains emergency call codes.
$\text{EF}_{\text{NETPAR}}$	Network Parameters
	This EF contains information concerning the cell frequencies.

2.3.1.3 Supplementary applications

The SIM Application Toolkit (often abbreviated STK or (U)SAT) provides a range of powerful features to the SIM card, enabling it to interact actively with the terminal, the user and the network. This makes it possible to extend the role of the SIM card far beyond that of a basic network authentication token, and allows it to provide various value added services controlled by the network operator. Table 2.3 gives an overview of the most important features supported by the SIM Toolkit.

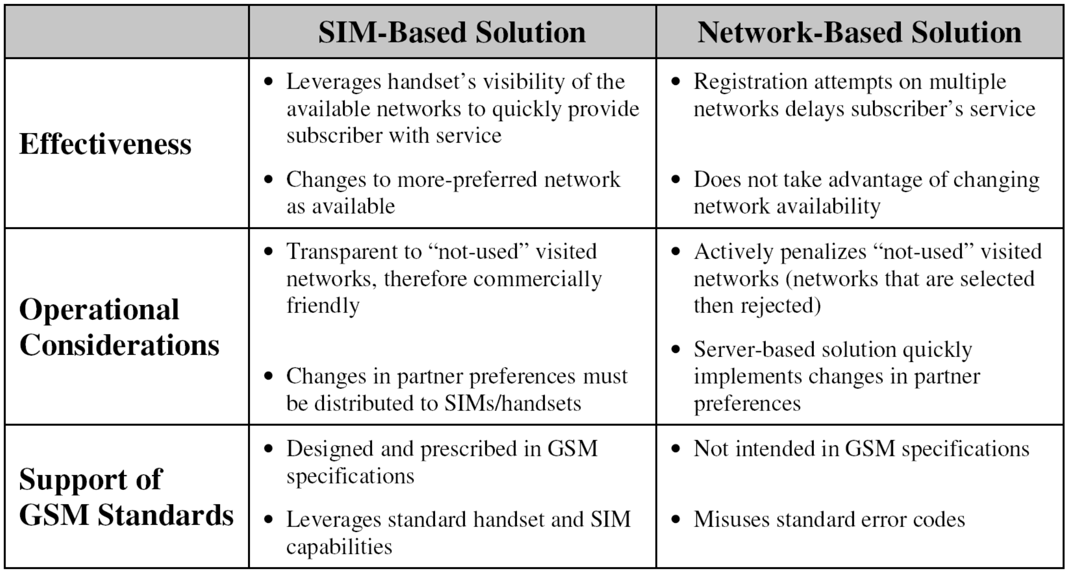
Table 2.3: Features of the SIM Toolkit. Source: 3GPP TS 31.111 [29]

Basic user interaction	**Communication**	**Provide local information**
• Terminal menu item • Selection menu • Display text message • Get user input • Show text on the idle screen • Play audio tone • Launch browser	• Send SMS • Send SS request • Send USSD string • Set up call • Send DTMF • Bearer Independent Protocol (BIP) • Data download to the SIM card by means of SMS messages	• Location information (mobile country code, mobile network code, location area code, and cell ID of the current serving cell) • Terminal identity information • Network measurement results • Current access technology
Event notifications	**Miscellaneous**	**Call Control**
• Call connected/disconnected • Location changed • User activity • Idle screen available • Language changed	• Request terminal to refresh SIM card configuration • Control polling rate • Run AT command • Timers	• Intercept voice calls, SS and USSD operations, and SMS messages • SIM card can allow, bar or modify a request

Nevertheless, the possibilities for user-interaction provided by the SIM card are very limited by today's standards, and although many operators have an 'operator menu' powered by the SIM Toolkit, they are rarely used, and many instead rely on WAP pages to provide services and infotainment to the user [25].

Many operators are using the SIM toolkit in what is called a 'smart roaming' applet which ensures that the terminal connects to the preferred roaming partner, beyond what is possible with the PLMN list. This type of SIM Toolkit application is very popular with operators [25], and Figure 2.5 shows some of its advantages. Operators also use the SIM Toolkit to implement differentiating offers such as Vodafone's ZuHause, where subscribers get flat-rate telephony within their home area.

Another important use of the SIM toolkit is as an enabler for the management functions described in the next section.



	SIM-Based Solution	Network-Based Solution
Effectiveness	• Leverages handset's visibility of the available networks to quickly provide subscriber with service • Changes to more-preferred network as available	• Registration attempts on multiple networks delays subscriber's service • Does not take advantage of changing network availability
Operational Considerations	• Transparent to "not-used" visited networks, therefore commercially friendly • Changes in partner preferences must be distributed to SIMs/handsets	• Actively penalizes "not-used" visited networks (networks that are selected then rejected) • Server-based solution quickly implements changes in partner preferences
Support of GSM Standards	• Designed and prescribed in GSM specifications • Leverages standard handset and SIM capabilities	• Not intended in GSM specifications • Misuses standard error codes

Figure 2.5: Advantages to SIM Toolkit-based 'smart roaming', as highlighted by a whitepaper promoting such a solution. Source: SmartTrust [30].

The SIM Toolkit specification itself only defines the protocol and commands; how applications using it are implemented on the card is up to card vendors. However, most modern SIM cards use a Java-based application platform called Java Card. This allows operators to develop application that are compatible with SIM cards from different vendors, and makes it easier for third parties to develop services involving the SIM card. The latest version of this platform, called Java Card 3, was released in March 2008. It removes many of the limitations of previous versions of Java Card, and supports many features found in the standard edition (SE) of Java in addition to new Java Card specific features. For instance, it supports multi-threading, advanced connectivity (TCP/IP, HTTP, HTTPS), and – in addition to the classical applet-based programming model – a web application model [31].

2.3.1.4 Management functions

In most modern SIM cards, remote management (also known as OTA management or OTA provisioning[5]) functions are available that can be used by the operator to remotely modify SIM files, and to provision new services and toolkit applications to the user[6]. This is usually achieved over-the-air using short messages, although it is technically possible to manage the cards over a packet-oriented data connection. The remote management functions are actually made possible by the SIM Application Toolkit, which gives the SIM the necessary communication capabilities.

[5] Not to be confused with OTA provisioning of a Virtual SIM subscription; this will be discussed later.

[6] In the industry jargon, this is known as Remote File Management (RFM) and Remote Applet Management (RAM), respectively.

According to Ross Campbell [25], most management operations are small updates to files, which only take a single SMS to carry out. The most common types of management operations are updates to the PLMN list (roaming agreements change frequently, and this list must be updated so that roaming subscribers use the networks with the lowest cost [2]), re-branding of the operator name (e.g. due to a merger), and updates to the SIM phonebook (e.g. if the operator's call-centre changes its number). Sending a SIM toolkit application over SMS is a lengthy procedure, taking up to 10 minutes per subscriber.

Most operators have an OTA platform which is used to manage the content of subscribers' SIM cards, and when requested, to dispatch updates (files and applications) to SIM cards in the field. OTA platforms are available from most SIM card vendors, but at least one widely used platform if offered by a third party (SmartTrust).

In the future, the use of remote management functions is likely to increase, especially for operators aiming to provide a wide range of value-added services to their subscribers. As new services are developed, operators will want to make them available to existing subscribers, and to this end, remote management is a key enabler.

2.3.2 Smart card security

The raison-d'être of the smart card is its ability to secure whatever programs and data are stored inside it. This is reflected in the physical, electrical and software design of the cards as well as in the production and personalization processes used.

The interface of smart cards is logically protected by their operating system and program code, which enforce file access policies and other security policies (e.g. PIN code check). However, this is far from enough, due to the fact that smart cards usually need to protect long-term secrets, which could potentially be extracted from the SIM in a number of physical attacks that bypass the legitimate interface. An overview of these attacks will be given here, and for more details the interested reader is referred to Michael Tunstall's excellent article on the subject [32], which also has additional references to the attacks described.

Three categories of attacks are normally considered:

'Invasive attacks' require the microprocessor in a smart card to be removed and directly attacked through physical means. In theory, this type of attack can compromise any secure microprocessor, but they are typically very expensive and time-consuming. For example, probes could be placed on bus lines on the chip in order to derive secret data sent through them (which requires a protective layer to be penetrated, as will be described shortly). In the extreme case, a focused ion beam could be used to destroy or create tracks on the chip surface.

'Semi-invasive attacks' require the surface of the chip to be exposed. An attacker then tries to compromise the security of the microprocessor without directly modifying it. For example, electromagnetic emanations could be picked up using a special probe, again making it possible to derive secret information. An alternative is to inject faults into the microprocessors logic (by changing opcodes) or data using laser or white light. For instance this could be used to make the chip bypass security checks or dump its memory, including secret keys.

'Non-invasive attacks' seek to derive information without modifying the smart card (i.e. the plastic remains intact). Side-channel analysis, where the attacker attempts to derive secrets from information that leaks during the computation of a command, falls into this category. For instance, differences in the time it takes to run a cryptographic algorithm for different inputs can sometimes be used to derive information about the key. The same goes for the power consumption of the chip, which could either be analysed for patterns or be used statistically to validate hypotheses about the cryptographic key (differential power analysis). Non-invasive fault-injection attacks can also be performed, including variations in power supply, variations in the external clock, extremes of temperature, and electromagnetic flux.

Countermeasures include anomaly sensors stopping the chip if it is exposed to conditionsoutside their expected operating ranges (e.g. voltage, temperature, light). Important functional blocks are randomized to make reverse engineering difficult, and the chip surface is covered by a metal layer preventing the chip's features to be identified. Side-channel analysis can be prevented by making the execution time of algorithms constant, by introducing random delays, by randomizing the execution order of operation, and by randomizing the data input to algorithms and subsequently un-randomizing the result. To protect against fault-injection attacks, various integrity checks and redundancy can be used, along with some of the randomization countermeasures, which make it difficult to time such an attack. Smart cards are also protected against such attacks as cutting the power to the chip before it gets the chance to decrement the PIN attempt counter.

2.3.3 Evolution

The smart card architecture has evolved from a monolithic model, where a small change in the software required a whole new IC mask, to a virtual-machine based architecture, where applications can be installed even after the card has been issued [33]. This evolution is illustrated in Figure 2.6. Current SIM cards are based on both 3$^{\text{rd}}$ and 4$^{\text{th}}$ generation technology, with 3$^{\text{rd}}$ generation cards dominating the low-end market, and 4$^{\text{th}}$ generation used in high-end markets. These high-end SIM card are based on Java Card technology, which makes it possible to re-use applications on cards from different vendors, and which significantly reduces the time it takes to develop and test new applications.

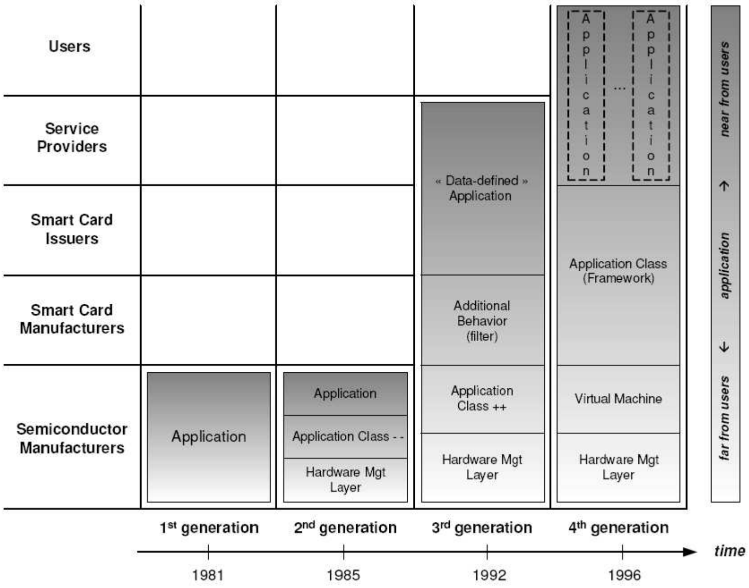


Figure 2.6: Smart card software architecture evolution. Source: [33]

Some clarification regarding the terminology used for the SIM card is needed. The whole story is described in detail in [5], but to summarize, the term "SIM card" was used in the GSM specifications created by ETSI until 1998, where standards development organizations in

Europe, Japan, Korea and the USA founded the Third Generation Partnership Project (3GPP) to jointly develop a third generation (3G) mobile communication system. In the same process, it was decided that a new working group, TSG-T3, should be the central focus point for all next generation telecommunication smart cards. The original SIM card specifications were split into two to support the concept of a multi-application card. One application-independent part described the physical and electrical characteristics of the card, as well as the basic communication protocols and logical functionality. This basic card platform became known as the Universal Integrated Circuit Card, or simply the UICC. The other half of the original specifications was formulated as a 'USIM application' running on top of the UICC platform, and was specific to the forthcoming UMTS system. This separation can be illustrated as follows:

GSM: SIM = physical card + GSM 'application'

UMTS: UICC = physical card and basic logical functionality
+ USIM = UMTS application on a UICC

More than one application can reside on the UICC. This could for example be a SIM application for GSM, an USIM (Universal Subscriber Identity Module) application for UMTS, or an R-UIM application for CDMA. These applications can be present on the card side by side, such that the same UICC can work in a both a GSM phone, an UMTS phone, or a CDMA phone. Applications not related to network authentication can also be installed on the card, such as for example an electronic purse application, a ticketing application or a credit card application, as illustrated in Figure 2.7.

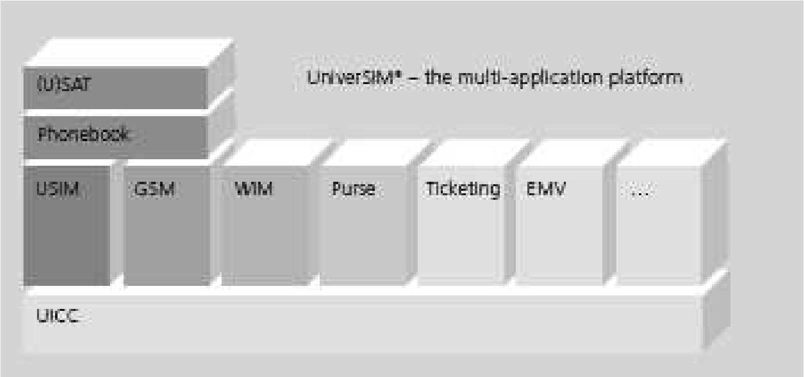


Figure 2.7: Multi-application UICC. Source: G&D UniverSIM whitepaper [34].

Interestingly, a major topic at the first meetings of T3 was whether the USIM was a removable module. "GSM delegates considered this to be a matter of fact. This view was, however, not shared by all delegates. Not everybody considered an open terminal market to be an advantage for the operator. The issue was finally resolved at the plenary meetings of TSG-T, TSG-SA3 and TSG-SA in Fort Lauderdale, in March 1999, where it was agreed that the 'USIM is a removable hardware module like the SIM is for GSM'." [5] Stated in other words, some mobile operators were interested in having a non-removable SIM for the purpose of binding the terminal to a predefined subscription or operator.

For the sake of simplicity, this thesis will mostly use the terms 'SIM card' or simply 'the SIM' in their classical sense — referring to both the physical card and the GSM or UMTS application. In cases where the distinction is relevant, the terms 'UICC' and 'USIM application' will be used.

2.3.4 Current developments

The SIM card is currently gaining a whole range of new capabilities. A new high-speed USB interface between the SIM and the terminal has recently been standardized [35], which could become the enabler for a range of advanced use cases. For instance it paves the way for flash-based high capacity (HC) SIM cards featuring from megabytes to gigabytes of memory [36]. This could be an enabler for various operator-services: "In addition to allowing portable storage of multimedia content such as videos, pictures and music, HC SIMs are well suited for securely storing DRM-managed content, rights tokens and digital certificates" [36].

The USB is also likely to pave the way for true IP connectivity to the SIM card. Currently, the SIM can use the Bearer Independent Protocol to establish a TCP connection with the operator, but true IP connectivity would mean that applications on the phone would be able to access IP-based servers on the SIM card.

Another new feature that is being developed for the SIM cards is the so-called Smart Card Web Server (SCWS). According to Telecom Italia, SCWS enables the following [37]:

- to use WAP contents in local and/or without coverage
- instant data portability between terminals
- address book and multimedia SIM services
- local management of the tariff profile
- remote administration of SIM contents and services on the part of the Operator
- secure management (HTTPS) of personal information
- management of certificates for authentication procedures (PKI)
- access to SIM contents through the PC browser

It is also a convenient way to create SIM-services based on protocol working on top of HTTP. Any service on the terminal that uses such a protocol now or in the future could be directed to use the SIM card as server, without technical barriers: "The SCWS is a pragmatic specification that leverages the mature and widely used HTTP protocol to enable a range of solutions such as on-SIM portals, NFC, just-in-time customisation and DRM" [38].

Finally, the SIM will also be a central element in the deployment of NFC for mobile phones, since many of the use cases put forward for NFC (e.g. ticketing, micro-payments, etc.) require a security element. The SIM is in an excellent position to fill this role, and the standardization in this regard is progressing, with the adoption of the Single Wire Protocol for the communication between the SIM and the NFC modem, and with the recent adoption of the Host Controller Interface [39].

2.4 Industry landscape

The primary actors involved in the value system of the SIM card are depicted in Figure 2.8. The IC vendors manufacture the silicon chips, which are sold to SIM vendors. In turn, SIM vendors package the chips into operator-branded plastic cards, add the operating system and other software, and personalize each card with operator- and card-specific data, including secret keys. The SIM cards are bought by operators and are passed on to their end-users so that they can gain access to the operators' networks. SIM- and device-management vendors sell various management solutions to operators.

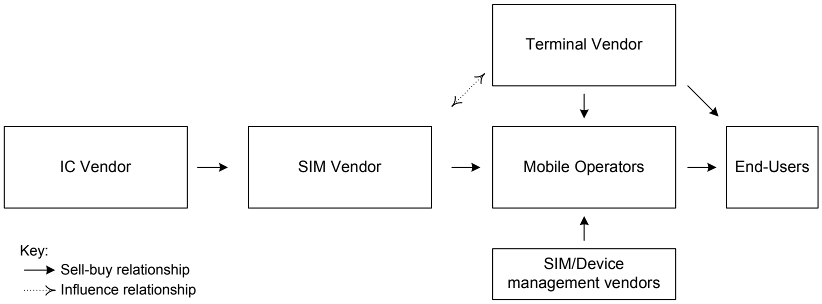


Figure 2.8: Value system of the SIM card. Source: own work.

Device vendors are not involved in the SIM supply chain, but are still included in the depicted value system, since a mutual influence relationship exists between them and the SIM industry in general. Indeed, most SIM features would be useless without support for them from the handsets, meaning that for most new SIM features to become usable, terminal vendors must first ship devices supporting them, which they are not necessarily prepared to do. However, SIM features are governed by standardization bodies (ETSI and 3GPP), which puts some pressure on terminal manufacturers to implement them in order to comply with the standards. SIM vendors, terminal vendors and mobile operators all play an important role in the standardization of SIM features. Finally, the large mobile operators, being the biggest customers of terminal vendors, also have some power over the SIM (and other) features supported by terminals. This is also a motivating factor for terminal vendors to comply with new standards.

2.4.1 Mobile operators' value systems

Up until now, 2G and 3G digital cellular systems have been based on a top-down vertically-integrated model [40], where mobile operators provide all aspects of the mobile experience, from network access to service platforms, all integrated in a seamless way [41]. This can partly be explained by the high investment needed to license the spectrums and establish the network infrastructure required by those technologies, but the use of the SIM card also plays a key role in preserving this structure due to its ability to function as an exclusive platform for service provisioning, and as a control point. Although today's phones have web-browsers and execution environments, mostly making it possible for third party services to be installed on the terminal without cutting a deal with the mobile operators, this has not always been the case. For a long time, the operator-controlled SIM was the only viable approach to providing services beyond simple voice- and text-messaging interactions. The days of walled gardens are not entirely gone, though. For example Vodafone was heavily criticised in 2007 for blocking VoIP applications on their phones [42].

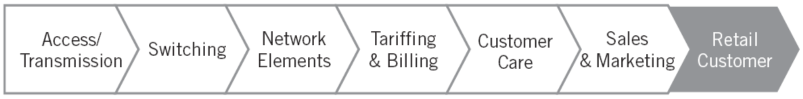


Figure 2.9: The early MNO (2G) value chain. Source: [43].

The vertically-integrated approach stands in contrast to the model used for internet communications, where the internet service provider may provide an e-mail account, a web-portal and perhaps some other services, but the bulk of value for the end-user lies in services made available by other parties through the connection.

2.4.1.1 The broadened value system

As the mobile industry has matured, so have the product offerings, which are becoming much more complex than simple voice services. For example, "some operators, appreciating the role of handsets as drivers of consumer choice, have pushed into earlier and greater control over device and interface design through partnerships with Far Eastern-branded vendors (Vodafone and Sharp); original device manufacturers (O2 and HTC XDA); and operating system sponsors (Orange, and Microsoft SPV2)" [43]. Some operators put a lot of energy into developing mobile portals (e.g. Vodafone's Live!, TMobile's T-Zones, and Orange's Orangeworld) through which they sell various value-added services and products, such as music, ringtones, games, etc. This has resulted in a broadened value chain [43], as illustrated in Figure 2.10.

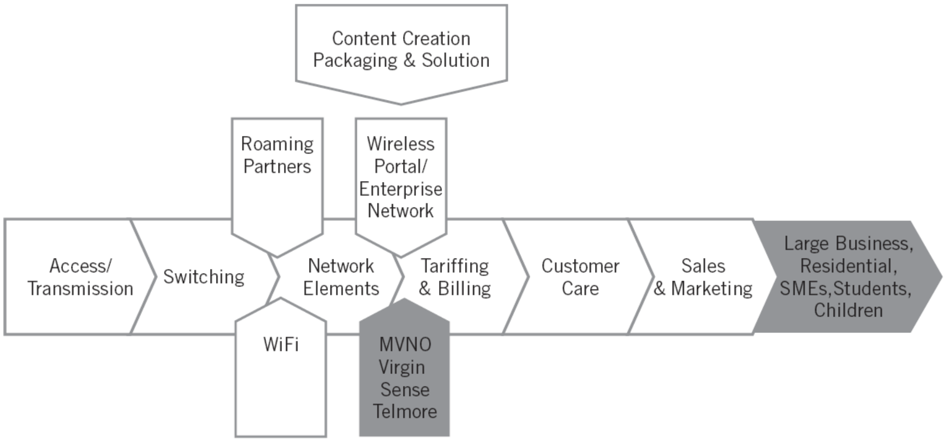


Figure 2.10: Current MNO value chain (2G & 3G). Source: [43].

2.4.1.2 Virtual operators

Some newer market entrants have a radically different strategy than traditional MNOs, as they don't seek to establish their own network infrastructure and don't have their own spectrum allocations, but instead make business agreements with traditional MNOs in order to use their network infrastructure. Different models are used for these agreements: 'Pure' MVNOs own a part of their core network infrastructure, especially the HLR, which gives them a lot of flexibility and enables them to offer highly differentiated services. Other MVNO's strategies are to own as little infrastructure as possible, and only provide basic voice and data services [43]. Pure MVNOs operate on the same basis as an international roaming partner towards the incumbent operator.

2.4.1.3 Current developments

Constant technological advances result in mobile terminals' capabilities becoming comparable with PCs, which has resulted in an increased demand to access the same internet services. This is posing a serious challenge to mobile operators' vertically integrated model. In the long run, this could result in a commodisation of access services, forcing operators to adapt, either by expanding their value chain to encompassing advanced services, or by adopting a bit pipe strategy, where the core business is to provide simple network connectivity [44], as is the case for Internet Service Providers today.

Another factor threatening the vertical integration of mobile services is the possible advent of ad-hoc access networks [45, 46, 47]. In this model, the user obtains network-connectivity by whatever access networks are locally available, which can span the whole range of short- and long-range access technologies. This would decouple the services from the network access part, which would focus on providing service-independent (all-IP), seamless, ubiquitous connectivity. This type of scenario is commonly referred to as 'beyond 3G' or '4G'[7]. It requires fundamentally different business models from those seen today [46], and will require a horizontal value structure.

So how does this relate to the Virtual SIM?

As mentioned, the SIM card contributes to upholding the vertically integrated market structure. An alternative may well catalyze the evolution towards any of the above mentioned scenarios: commodisation of mobile access, operators as bit pipes, ad-hoc networks, beyond 3G, etc. Depending on the strategic plans and visions of a given player, this may or may not be a promising prospect, which may affect the pros/cons ratio of the Virtual SIM seen from that player's perspective. While this should be kept in mind, one conclusion drawn from the conducted interviews and various informal discussions, is that the perception of the future and whether something is an opportunity or a threat, is very subjective and can vary wildly from organization to organization and even between people of the same organization. One example of this is from the interview with Stefan Kaliner, head of the UICC department of T-Mobile, who is of the perception that the SIM card could advantageously be replaced by some successor technology [22]. His superior, on the other hand, is of the firm belief that the physical SIM is key asset for operators, which should be preserved in the future.

[7] Although these terms, especially '4G', are often being used to mean different things, for instance LTE-technology, or its future successor, is sometimes being referred to as '4G'.

Chapter 3

Analysis

3.1 Adoption of the Virtual SIM

The process by which a new idea, product or process spreads within and across economies is called “technological diffusion” [48]. Understanding the factors influencing the adoption of innovative technologies, such as the Virtual SIM, is essential in order to design successful systems. The following two subsections will present a theoretical framework for technological diffusion, and the remainder of the section will discuss the factors that might influence the adoption of the Virtual SIM by the involved stakeholders.

3.1.1 Everett Rogers’ diffusion of innovations theory

The original concept of technological diffusion was put forward by Beal and Bohlen [49] and further promoted by Rogers [50] in his “diffusion of innovations” theory. This theory takes a sociological approach, where adopters are classified into the following categories depending on the time of adoption relative to a normal distribution (see Figure 3.1):

- **Innovator**: Venturesome (not a change agent) — ahead of his/her time.
- **Early Adopter**: Respectable (not necessarily a good change agent) — willing to try the innovation before it has proved to be useful.
- **Early Majority**: Deliberate
- **Late Majority**: Sceptical — even after others embrace the innovation
- **Laggards**: Traditional — don’t see a need to change.

Roger argues that the adoption decision is governed by the model shown in Figure 3.2. Although the terminology used targets applications in the field of sociology, the model can readily be generalized, and is in fact the most generally used model within the field of innovation diffusion [51].

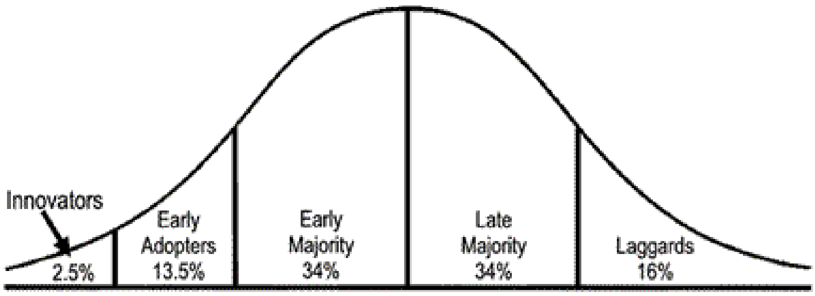


Figure 3.1: Bell curve illustrating the diffusion of innovation. Source: [50].

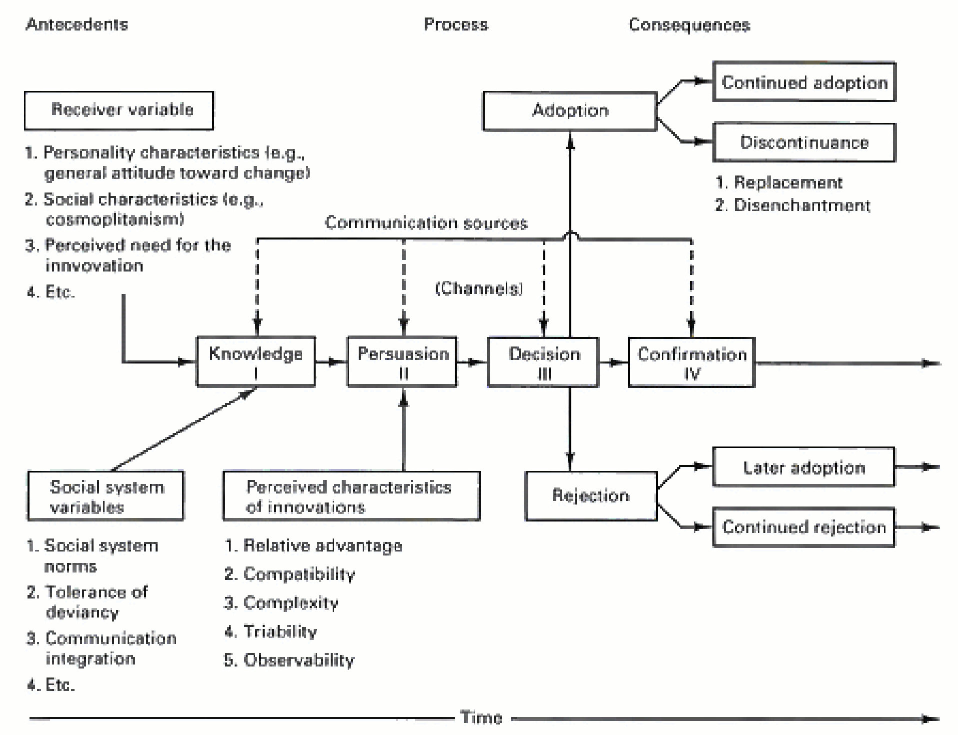


Figure 3.2: Everett Roger's innovation decision process. Source: [50].

The model takes into account a priori knowledge such as personality traits of the decision-maker (for a company these variables would be of a strategic nature) as well as the nature of the network ('social system' in Roger's terms) within which the decision-maker operates. In the model, the decision process itself is influenced by the a priori disposition of the subject, as well as by characteristics of the innovation. The following five characteristics are highlighted by Rogers [50, 52]:

- **Relative advantage:** The innovation is better than the status quo.
- **Compatibility:** The innovation is compatible with current values and practices.
- **Low complexity:** The innovation is not difficult to understand and use.
- **Triability:** The innovation can be tried out on a partial or temporary basis.
- **Observability:** The impact of the innovation is noticeable to other potential adopters.

Two other relevant characteristics described in the literature [52] will also be considered:

- **Low cost:** The less expensive the innovation the more likely it is that it will be quickly adopted and implemented.
- **Profitability:** The level of profit to be gained from adoption of the innovation.

Generally, innovations with positive characteristics are more attractive and will diffuse more rapidly than those with less favourable characteristics [50]. However, some make the distinction between primary and secondary characteristics, where the former are inherent to the innovation, whereas the latter are specific to the decision-maker in question [53]. Values for primary characteristics have been assessed based on logical inferences about the innovation or by relying on expert judgements. Values for secondary characteristics can be inferred from objective features of the organization, and can also be captured by soliciting the perceptions of key informants [51]. This model is adequate for studying organizations' reactions to innovation, and will be used to assess the feasibility and acceptance of the Virtual SIM by mobile network operators and by device manufacturers.

In his famous book titled 'Crossing The Chasm' [54], Geoffrey Moore argues that there is a large barrier (a chasm) between the adoption of a product or technology by early adopters (visionaries) and by the early majority (pragmatists). He attributes this to the fact that visionaries and pragmatists have very different expectations. However, this only applies to 'discontinuous innovations' that force a significant change of behaviour by the customer; continuous innovation is still best described by the original technology adoption lifecycle [55]. The Virtual SIM could be classified as a discontinuous innovation, since it radically changes the SIM industry; however, the essence of Moore's theory is that discontinuous innovations will have a hard time being adopted beyond the original niche market targeted by the technology. This notion of niche markets mostly applies to end-users, and cannot readily be applied to an industry group such as mobile operators. For this reason, Moore's model will not be considered in this context.

3.1.2 Technology acceptance model

How potential adopters perceive an innovation is a key determinant of adoption [50, 52]. Innovation perception is a function of the innovative technology and of the subject observing it, and can operate on two levels, resulting in two different types of adoption. When focus is on the organization decision to adopt, it is the perception of leaders and key decision makers that matters and their decision will result in the organization's formal adoption of the innovation (or not) [51]. However, after formal adoption, end-users of the technology are often relatively free to choose whether to use an innovation and how [56]. Thus, for many types of innovation, a key element is the acceptance of the technology by its intended users, which is driven by the individual perceptions of an innovation [57]. These user perceptions are usually not motivated by the classical innovation characteristics. Instead Davis et al. propose the Technology Acceptance Model, with the following innovation characteristics favouring end-users' acceptance of technology change [58, 59]:

- **Usefulness**
- **Ease of use**

This will be our framework for evaluating end-users' acceptance of the system.

3.1.3 Adoption by device manufacturers

As with any innovation, the Virtual SIM (VSIM) will follow some adoption curve as it replaces the SIM card. The SIM card adoption will decrease from its current 100% for GSM/UMTS and will at some point reach 0%, assuming it is completely replaced by the VSIM technology. During the transition, however, the VSIM is likely to coexist with the SIM, both in terms of support from network operators, and in devices: If a mobile operator stops issuing SIM cards for people with 'old' phones not supporting the VSIM, he will probably lose this customer group; and if device makers do not produce dual SIM/VSIM phones, while there is still a high demand for SIM-capable devices due to operators not yet supporting the VSIM, either two variants of each model must be produced, or the manufacturer will lose market shares to vendors supporting those markets, neither of which is likely to be acceptable. Initially, device manufacturers will probably coordinate the launch of VSIM-capable devices with the first operators' support for VSIM in their networks. However, while this doesn't force other operators to jump on the bandwagon, device manufacturers ship the same products globally (more or less), and scale effects are likely to make it more profitable to implement dual SIM/VSIM functionality on all new devices, rather than targeting the segment with special products (even if this means that such phones are used in some areas with slow VSIM diffusion, and only the SIM functionality will be used). This situation can be compared with 3G phones still being used on 2G networks, where several of their advanced features cannot be used. Thus, the predicted scenario is that device manufacturers will start supporting the VSIM over the range of devices in their device-portfolios more quickly than the operators will support the system in their networks. Thus, the availability of VSIM functionality for end-users will largely depend on the rate of adoption by operators.

3.1.4 Adoption by end-users

The above discussion makes it clear that there are two concurrent adoption curves. One is the support for the technology by the operators, which can be measured as the percentage of operators offering VSIM-based subscriptions. The other is the adoption by end-users, i.e. the proportion of subscriptions using a VSIM. Figure 3.3 illustrates how such adoption curves might look. It should be noted that these are purely speculative.

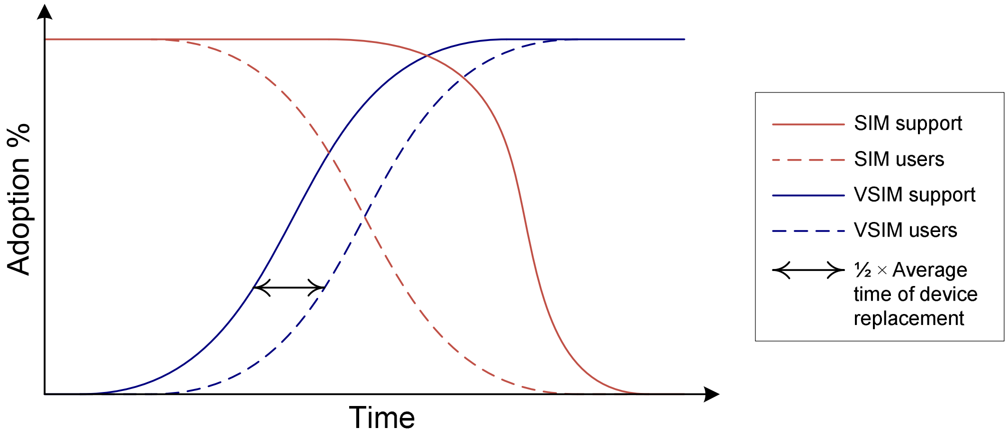


Figure 3.3: Speculative illustration of the adoption by mobile operators (solid lines) and by end-users (dotted lines). Source: own work; see description in text for further details.

The dotted lines in Figure 3.3 represent the percentage of subscribers using a SIM card (red) or a VSIM (blue). The curves are symmetric, and their sum is always 100%, since a VSIM and a SIM card will not be issued simultaneously for the same subscription. Users' adoption of the VSIM is assumed to be time shifted compared to the support for the VSIM in networks and devices. This is because a prerequisite to end-users using the VSIM is that both their mobile terminal and the network support the technology. Disregarding the second-hand market and assuming that users adopt the VSIM as soon as possible, this time shift will be the average time of device replacement, divided by two. However, when given the choice (in terms of device and network support), there is a risk that users will 'stick' to the SIM card technology they are used to, which would result in a lower adoption rate than the one depicted. This effect is undesirable, since all stakeholders have an economic incentive to minimize the transition period. To counter it, the VSIM should have some advantages for the end users, such as lower cost or new features, which should be advertised by the operator. In the technology adoption model (described in section 3.1.2), this corresponds to increasing the *usefulness* of the innovation, thereby increasing the incentive for change. For instance, operators could demand a fee each time a SIM card is issued, which users could avoid by choosing the VSIM. From the user's point of view, the most important criteria for adopting the VSIM are *ease of use* and *usefulness* [58]. This means that the VSIM should preferably not be more complicated to use than the normal SIM card. Furthermore, it should not restrict what the user's possibilities compared to the SIM card. Appendix B.1 contains a set of usage scenarios, showing how the usability of the proposed scheme compares to that of the SIM card. Initially, users might fear that the usefulness of their device will be restricted, and therefore be reluctant to change. A transition period with dual SIM/VSIM devices would give users a chance to get accustomed to the new technology, without feeling forced into a less advantageous position.

3.1.5 Adoption by operators

3.1.5.1 Incumbent operators

The solid blue curve in Figure 3.3 illustrates how support for the VSIM will be slow at first, driven by a few 'innovators' and subsequently by 'early adopters'. Once these first operators can prove to the industry that the concept if feasible, it is likely that "as more information and experience accumulate it becomes less of a risk to begin using it. Competitive pressures mount and 'bandwagon' effects occur. Where the profitability of using the innovation is very difficult to estimate, the mere fact that a large proportion of its competitors have introduced it may prompt a firm to consider it more favourably" [60:p137]. This effect will eventually drive the rest of the operators to adopt the VSIM, and at some point, when most devices in circulation support the VSIM and when it has been adopted by a majority of the user-base, operators will stop supporting the old SIM card, as illustrated by the fall of the solid red curve in Figure 3.3 (note that the sum of these two curves is not necessarily 100%, since operators are likely to support both technologies concurrently during the transition period). Innovators might be operators seeking to differentiate their product offerings in the highly competitive market, or small operators targeting niche markets, for instance MVNOs.

If users who turn on a subscription-less phone are able to get a list of operators from whom they can download a VSIM subscription (a possibility that will be described in section 4.6.2), this may be an additional incentive for operators to support this system quickly, since otherwise it may be more convenient for users to get a subscription from a competitor.

3.1.5.2 Innovating newcomers

Innovators might also be newcomers to the market seeing an opportunity in offering a VSIM-only service to a niche market. Typically, this would be an MVNO, since they avoid large investments in infrastructure, and thus can afford to take more risks. A requirement is that the MNO providing network access would support the VSIM in his infrastructure. With such an ally, the MVNO newcomer could create an entirely 'virtual' network by selling subscriptions over the internet and providing users with VSIMS. There would be no physical stock or stores, and essentially the whole business could reside on a few servers. This would allow cost savings compared to a normally operating MVNO, and thus lower prices, which could capture a niche part of the market. As mentioned earlier, it would not be possible for an incumbent MNO or MVNO to stop supporting the physical SIM card right away, since this would mean losing subscribers. For large incumbent MNOs, such newcomers are unlikely to be a threat. On the contrary, it would allow them to observe if VSIM has the expected potential without taking any risk, and if the low-end segment is targeted, this may also be an opportunity for the incumbent operator to capture market shares from competitors. This pattern has been seen before. One of the first MVNOs, Telmore, quickly captured a significant part of the low-end segment with prices incumbent operators couldn't compete with. Telmore was later acquired by TDC, the incumbent MNO which provided it with wholesale network access in the first place. TDC thereby re-captured the lost market share and, most importantly, the market shares its competitors lost to Telmore [43, 61].

3.1.5.3 Factors influencing operators' acceptance

Innovations are adopted primarily on the basis of some expected benefit, but compatibility and complexity also influence the likelihood of adoption [52, 50]. Thus, it can be expected that the compatibility of the virtual SIM system with the existing infrastructure and processes will have a positive impact on its success, and conversely that the complexity of the system and procedures introduced will have a negative impact. Most important is the compatibility with the existing systems. Big differences in operational and logistic processes would result in a significant financial penalty for operators during the SIM/VSIM transition period. This aspect is discussed further in section 3.2.1. Furthermore, new processes require the training of staff, which is also a costly affair (proportionally to the extent of the differences between the systems and processes), especially if people working in retail stores are affected.

3.1.6 The standardization process

The goal of GSM was to create a standard that would meet everyone's demands. Of course, there are always conflicts of interest, but the mobile industry was still in its infancy, and the system was designed with the interests of the industry as a whole in mind. The spirit was that decisions should be made unanimously, and most of them were. Today, the industry is firmly established, and GSM has proved to be very successful. However, this means that the various stakeholders participating in the standardization meetings are primarily concerned with their own immediate business interests when decisions are to be made. Since these stakeholders have diverging interests, it is not uncommon that there are large disagreements between two or more camps, and it is more often necessary to resort to a vote[8].

[8] No citation is available regarding this consideration, but the information has been confirmed by Jens-Ole Madsen, Nokia.

If it is agreed that a Virtual SIM system will be standardized, it is likely that different schemes will be proposed (both at a conceptual level and with concrete technological choices), each serving different stakeholders' interests. This introduces the strategic risk that the final system will not have the advantages originally hoped for, or even that the system becomes a strategic misstep. This risk is present for all involved stakeholders; most notably for mobile operators and device manufacturers (obviously, for SIM manufacturers the entire concept is one big threat).

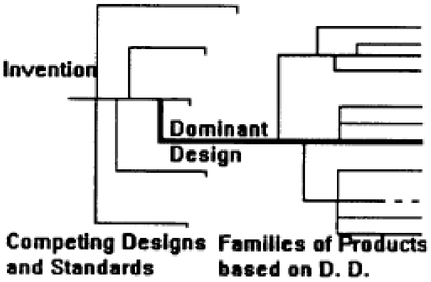

Invention
Dominant Design
Competing Designs and Standards
Families of Products based on D. D.

Figure 3.4: Structural patterns of technological change. Source: [62]

A parallel can be drawn to the competition between different technological standards in the open market (as opposed to here, where a standardization body is involved, and the technological choice is made before the technology is put on the market). A classical example is the choice of VHS over Sony's Betamax. The main difference here is that with the involvement of a standardization body, the decision is made by the industry players, as opposed to the end-users. Many different models have been derived to describe such technological change [63]. The model in Figure 3.4, proposed by Susan Sanderson [62], suggests that a number of designs and standards compete to become the *dominant design*, and that only the one that becomes dominant is continued. Subsequently changes are introduced to this design to fit various needs, and at some point, one of these derived designs may become dominant. This model seems to apply well to the standardization process that can be observed in organizations such as ETSI and 3GPP. The fact that, in this model, the outcome of the original technological choice is also the basis for future evolutions, underlines the importance for stakeholders to proactively promote the adoption of the technology that serves their interests best; once a less advantageous technology becomes dominant, stakeholders might have to accept the consequence of this outcome for a long time.

On the other hand, in this specific situation, one might argue that since a working system (the SIM card) is already in place, and since a considerable effort is required from all large stakeholders (mobile operators and terminal manufacturers) in order to deploy the VSIM, a system that doesn't serve the interests of both groups is very unlikely to see the light of the day, even if it passes the standardization process. This is in itself unlikely, since a majority vote of 71% is necessary for the final proposal to be accepted, and that SIM vendors are unlikely to welcome such a proposal. In the event that the original VSIM concept undergoes so many changes that SIM vendors might accept it (in which case it probably isn't advantageous to handset vendors), the above point still remains valid.

It is the opinion of the author that the threat discussed above (that a proposal to standardize a VSIM system might have a disadvantageous outcome for terminal manufacturers) is negligible. However, these considerations do highlight the importance of devising a scheme that is advantageous to both operators and device manufacturers, in order to increase the probability of successful standardization and subsequent real-life deployment: a win-win business model is needed.

3.2 Operator requirements

3.2.1 SIM logistics

In the context of this study, SIM card logistics are important for two major reasons: First, the process of packaging and distributing personalized SIM card to retail stores is a major cost item for operators (for Telia Denmark, this amounts to approximately 1€ per card [25]), and secondly, making radical changes in established procedures would significantly increase the acquisition cost for the Virtual SIM for operators, and should thus be avoided. On the other hand, changes are desirable if they result in cost savings in the long run.

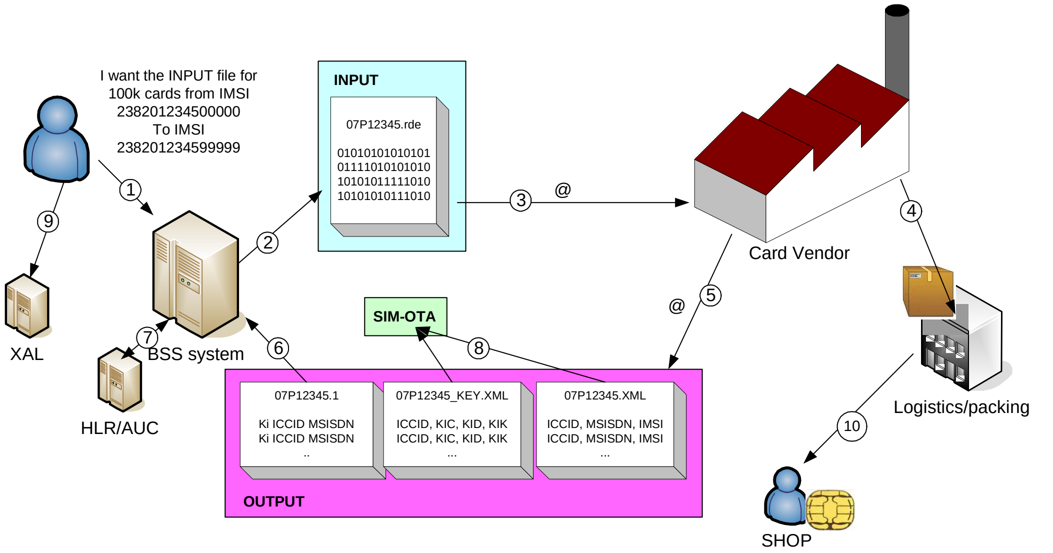


Figure 3.5: SIM card ordering and provisioning at TeliaSonera Denmark. Source: [64]

The SIM card ordering and provisioning process for pre-paid cards at TeliaSonera Denmark is illustrated in Figure 3.5, and consists of the following steps [64, 25]:

1. When SIM card stocks run low in retail stores, the logistics personnel file a request for a new batch of SIMs by means of the business support system (BSS). A batch is typically 50,000 or 100,000 cards. The request includes the type of SIM card desired from a set of "profiles", previously agreed with the SIM vendor, specifying the electrical and graphical features of the card. The request also specifies a range of IMSI numbers to be used for the cards.

2. IMSIs and ICCIDs are allocated in the relevant databases, and an electronic file is generated containing the request data: card profile, IMSI and ICCID ranges.

3. The request file is sent to the card vendor over an electronic channel, and the vendors start personalizing the cards. The cards are typically already in stock, and are only lacking the final personalization stage. Personalization includes customizing the SIM files containing the ICCID, IMSI, and for pre-paid cards, the MSISDN, as well as generating a Ki and OTA keys for each card.

4. The finished batch of personalized cards is sent to another company, responsible for the packaging and distribution of the SIM card. This involves putting the card in a physical package, which also contains terms-of-use documents and the like. The packages also include the PIN/PUK codes and stickers with a bar code specifying the ICCIDs/IMSIs of the contained cards. The cards are then distributed to retail shops.
5. The card vendor uses another electronic channel to send back an encrypted set of data files to the operator containing the personalization data (ICCID, IMSI, MSISDN, Ki, OTA keys, etc.) of the cards.
6. Data regarding the cards (ICCID, IMSI, MSISDN, Ki, etc.) are loaded into the BSS.
7. Card data are loaded from the BSS to the HLR/AuC.
8. OTA related data files (including OTA keys) received from the card vendor, are loaded into the OTA servers.
9. Data are loaded into the logistics/ERP system.
10. Cards are sold by retailers. Post-paid subscriptions are activated and assigned an MSISDN when the subscription is sold — this is managed entirely on the network side, and doesn't involve the SIM card. Pre-paid subscriptions have a pre-assigned MSISDN, and are pre-activated.

Some network operators let the SIM card vendors handle the packaging and distribution step (e.g. T-Mobile [22]), but this is not important in the context of this study.

With the introduction of the Virtual SIM, these processes will indisputably change. This change is wanted, since there is the opportunity to save a lot of costs related to the SIM card. The cards proper are one large source of expenses, but equally important is the cost of the logistics associated with distributing SIM cards to retail stores and managing them in subscriber databases. Nevertheless, change also costs money in terms of new equipment and staff training, and it is therefore desirable to have the option to remain compatible with existing procedures and equipment. The extent of change required in this area largely depends on how the VSIM provisioning system is designed.

3.2.2 Retail channels

The distribution channels for SIM cards vary from market to market, but the following general retail models have been identified:

- Post-paid model: SIM cards are sold stored either owned by operators or affiliated with them. Subscriptions are often bundled with a subsidised.
- Pre-paid model: SIM card are sold in general-purpose stores, such as super-markets, gas stations, etc. Subsidised phones are less common.
- American model: Subscriptions are always bundled with a terminal.
- Internet model: Subscriptions are sold on the internet and the SIM is mailed to the user. Usually not sold with subsidised phones.
- Third-world model: Subscriptions are sold through ad-hoc channels, such as by market vendors, etc.

Operators must be able to sell Virtual SIM based subscriptions through all of the currently used sales channels [65, 25], and the approach taken for provisioning Virtual SIMs to the terminal must reflect this: Either, the chosen provisioning method should be usable in all these scenarios, or the system should be flexible enough to allow different methods of provisioning Virtual SIMs that can be applied to different retail conditions.

3.2.3 Security aspects and trust model

One of the main barriers to the endorsement of the Virtual SIM concept by mobile operators is their reluctance to give away the subscriber key [66]. This is understandable, since its secrecy is essential for being able to bill subscribers for their consumption. Thus, a key requirement for operators is that the terminal must be able to protect the subscriber key adequately.

One of the merits of the SIM card is that it is entirely under the control of the operators. They specify the hardware and software requirements and have a contract with the vendors, which solves liability problems, etc. A fixed SIM scheme could take the same approach, allowing operators to provide a secure element that would be embedded in the terminal during production [22]. Of course, this has major drawbacks, since terminals would be forever bound to a single operator, and since it complicates the logistics of terminal manufacturing immensely. Also it would not be a ‘real’ Virtual SIM scheme, since it would merely be an embedded physical SIM.

Instead, in a Virtual SIM scheme, terminal vendors control the hardware platform of the SIM. This requires some kind of trust model. Operators will need some assurance that a given terminal will be able to protect the subscriber key, before they issue a Virtual SIM to the terminal. This aspect is one of the keys to the success of the Virtual SIM.

3.2.3.1 Authentication algorithm

In the current GSM and UMTS systems, operators are free to choose their own authentication algorithm, as described in section 2.2.4. The main reason for this is that if the example algorithm provided by the GSM Association is found to be vulnerable, operators with another algorithm won’t be affected. Also, if the amount of people using a specific algorithm becomes smaller, the motivation and resources put into attacking that algorithm will also diminish.

Although many operators use the example algorithms [25], at least some of the large operators, such as T-Mobile and Vodafone, use their own algorithms, which are often implemented as a modified version of the standard example [22].

Nevertheless, the value of having custom algorithms is debatable, in the light of the very limited consequences of the breach of COMP128. A Virtual SIM system could be accepted without the possibility of having custom algorithms if it has got other sufficiently large advantages, but from the operator’s point of view this would make it a less attractive solution [22, 66].

For operators using a custom algorithm, an important requirement is that this algorithm remains secret. If the algorithm is disclosed, most of the point in having a custom algorithm in the first place is lost. In fact, this would probably lead to a situation where operators using a custom algorithm are potentially more vulnerable to attacks, since that algorithm would have undergone less scrutiny than the standard MILENAGE algorithm.

3.2.4 Functionality requirements

As mentioned in section 2.3, operators use the SIM for various other purposes than merely authenticating subscribers. For instance, it allows the remote updating of files such as network parameters and the execution of toolkit applications making it possible to implement a wide spectrum of non-standard functions. Operators are not likely to want to relinquish the control these features give, and a Virtual SIM scheme should take this into account. Specifically, it should provide the means to remotely manage SIM files and to execute operator-applications that should at least have the same features as the SIM Toolkit.

3.2.5 Subscriber retention

The Virtual SIM could potentially make it very easy for subscribers to change operators, which would increase churn rates for the whole industry. This is very undesirable for operators, and therefore the Virtual SIM must incorporate a SIM lock mechanism, preventing a subsidised phone to be used with another operator before the contract runs out. SIM lock mechanisms used today usually check if the IMSI of the SIM is within a predefined range[9] before allowing the SIM to be used. There is no reason why the exact same implementation should not be used with the VSIM also.

Fraud threats related to the SIM lock will be described further in section 3.5.1.3.

3.3 Terminal vendor requirements

The shift from a SIM card owned by operators to a Virtual SIM implemented in the terminal could raise concerns that the cost of terminals will increase, since they need to incorporate new functionality. The move to the Virtual SIM does not in itself create new value for the terminal – new SIM functions will only benefit operators, and the terminal itself only really benefits from being able to access the network, which is already possible with the physical SIM. This means that the manufacturing cost of a terminal supporting the Virtual SIM should not be greater than that of an equivalent terminal using a regular SIM card. A huge quantity of mobile phones is produced every year (over 1.15 billion in 2007 [67]), and even a small increase in production cost has a big impact. It must be noted that production cost not only includes the bill of materials but also the complexity of the manufacturing process and the manufacturing time per unit.

Furthermore, the use of the Virtual SIM should not make the logistics significantly more complicated. Today, the terminal manufacturing process is independent of which mobile operator the produced terminals will be used with (except for the case where terminals are customized for an operator, but this is a different matter).

3.4 End-user requirements

As mentioned in section 3.1.2, the main adoption criteria for end-users are usefulness and ease of use. These two aspects can be considered in the context of common usage scenarios related to the SIM, such as:

- *Borrowed phone:* Is it possible for a user to borrow a phone and use his own subscription? With the SIM card, this is possible if the borrowed phone isn't SIM locked.
- *Borrowed subscription:* Is it possible for a user to borrow a friend's SIM and use it in his own phone? Also possible with the SIM card, unless the phone is SIM locked.

[9] This is a highly simplified view. SIM lock mechanisms are terminal vendor dependent, but can usually check several IMSI ranges as well as the value of other SIM files.

- *Transfer phonebook contacts or messages between terminals:* Is it possible to transfer personal data between phones easily? The SIM card can be used for this purpose.
- *Multiple subscriptions:* Can a phone hold more than one subscription? This is for example useful for someone having a personal subscription as well as a subscription paid by his employer. Also relevant in developing countries, where many people share one terminal but each has his own subscription. In almost all cases, mobile phones can only hold one SIM card. It is, however, possible to use more than one subscription per phone by swapping SIM cards, assuming the terminal isn't SIM locked.
- *Multiple device ownership:* Can the same subscription be used with more than one terminal? Currently, it is necessary to physically move the SIM card between terminals.

One can also consider the ease of use aspect when acquiring a subscription. With the SIM it is necessary to physically place the SIM of a new subscription in the phone, which can be difficult for users who are not so tech-savvy. And even for users who are, figuring out how to open the phone for the first time can sometimes be quite a challenge – even with the manual. On the other hand, once the SIM is placed into the terminal, access to the network normally requires no additional configuration.

3.5 Security

A controversial change such as the Virtual SIM will probably meet a lot of resistance, especially from SIM card vendors, who are likely to point out security as a reason why SIM cards are indispensable. The change to the Virtual SIM requires a new approach to SIM security, and this shift is likely to be challenged. This section gives an analysis of the options when it comes to security.

3.5.1 Mobile fraud

In a perfect world, nobody would try to cheat, and there would be no fraud, but history shows that this is not the case. The ultimate goal of the security measures discussed in this section is to protect systems from abuse and fraud, in order to preserve stakeholders' assets and to defend end-users from being defrauded.

Fraud is a major concern for mobile operators, as it can have a negative impact on the bottom line as well as on brand image. Two types of fraud-related loss can be distinguished: 'hard' and 'soft' loss [68]. Soft loss is a theoretical figure derived from lost revenue due to illegal use of the system, which assumes that the illegal user would have paid for the service used without permission. Hard loss is about real money that the operator has to pay to someone else. Most frauds result in both types of losses with varying proportions. For instance, shoplifting results in a hard loss corresponding to the purchase cost of the items stolen, and a soft loss corresponding to the gross profit that would have been made if the shoplifter had instead bought the items. Another way to distinguish between them is that hard loss can be accurately measured (whether in monetary value or otherwise[10]), whereas soft loss can only be expressed as a hypothetical maximum of lost revenues. Clearly, hard loss is much more serious to a business than soft loss.

[10] Hard loss can also be incurred on intangible assets. For instance, if counterfeiting activities lead to a deterioration of the company's image, this loss of brand-value is a hard loss, whereas the lost revenue from people knowingly buying a counterfeit product is difficult to determine, since most of those would probably not buy the product at its full price.

In the case of telecoms, a typical example is that a person has placed fraudulent calls for some amount: The hard loss incurred by the operator comes from the actual cost of routing the calls. If the person is calling a number served by the same operator, the hard loss will be very small, whereas in the case of calls placed to international numbers, the operator will be billed for the traffic, which may result in a substantial hard loss. Similarly, if fraudulent mobile calls are placed while roaming abroad, roaming partners will bill the operator for the traffic.

According to the Communications Fraud Control Association, loss due to fraud in the communications sector was in the range of 54 to 60 billion dollars globally in 2003 [69], so it is clearly a real concern for operators (although the cited article does not specify anything about how this number was obtained and whether it includes soft loss).

3.5.1.1 Cloning fraud

If a fraudster is able to identify himself towards the network using the identity of an existing subscriber, the services used will be billed to that person. This is damaging to the operator since he will have to pay for the fraudster's communications (assuming they are noticed by the legitimate user). Moreover this can be a major annoyance for his customers and may damage his reputation. This type of fraud was very common in the days of analogue systems, where no special security measures were present. Subscriber identification numbers could be obtained by monitoring the radio communication between a terminal and the network passively, or by requesting them from terminals actively using fake base stations [70]. These identifiers were then sold on the black market or used in call-selling operations, where phone services were stolen and resold cheaply [71]. The challenge-response authentication introduced in second generation mobile systems (GSM in Europe and IS-41 in the United States) largely eliminated the problem, since physical access to the terminal was now needed in order to get access to the secret shared key used for subscriber authentication. In addition, in the GSM system, this key is stored on a tamper-resistant smart-card (the SIM), further increasing the barrier. In either case, subscription cloning now has a very poor cost/benefit ratio and is virtually non-existent [72]. This assumption is confirmed by the fact that even though Marc Briceno, Ian Goldberg and Dave Wagner made SIM cloning possible in 1998, with a few vulnerable SIM cards still in circulation today, this did not result in subscription cloning fraud becoming significant [71] – simply because physical access to the victim's terminal (or SIM card) is still needed, and thus, other types of fraud are much easier to carry out, and more profitable.

Nevertheless, according to Eugene Bergen Henegouwen, executive vice president and managing director for Syniverse, a player in the mobile security space, SIM cloning is still being used for fraudulent activities: "The SIM copying business is a big business. It's not an individual copying SIM cards. It's illegal organizations making money by selling free calls to people. It's a growing concern but hard to get a handle on it." [69]. Even though the quoted article is recent, it is hard to conceive how this information can be accurate: A part from the vulnerability in the original COMP128 algorithm (discussed in section 2.2.4) found in 1998, there has been no publicly disclosed way to clone SIM cards, and vulnerable cards have not been produced for many years and few are still left in the field.

Other possible motivations for cloning fraud are: identity theft and eavesdropping (if the subscriber key is known, radio communications can potentially be decrypted). But also in these cases, the need to be in physical possession of the terminal in order to extract the necessary information is a major hindrance to the practical feasibility of such attacks.

3.5.1.2 Subscription fraud

This type of fraud occurs when a subscription is acquired through normal means, but services are not paid for. Operators can normally monitor subscriber's consumptions, and block a subscription if extraordinary patterns are detected, or if the credit of a prepaid subscription drops to zero. This limits the opportunities for abuse. However, when a subscriber is roaming, the home operator cannot normally know the consumption of a subscriber until the visited operator sends back the bill, which can take up to several days, although this delay has been shortened to 24 hours for call charges exceeding a certain amount [72]. This still gives abusers plenty of time to resell calls massively before the subscription is disabled. This type of fraud is sometimes called 'roaming fraud'. In practice, it is often carried out by specialized call reselling centres, where GSM's conference call capability is used to connect two other parties. The SIM card can then leave the conference and set up a new simultaneous one to two other numbers, and so on, everything being billed to a subscription that won't be paid.

To counter this type of fraud, operators are often monitoring their subscribers' usage for suspicious patterns, and if such patterns are detected the subscription is blocked. This is of course limited by the reporting delay when roaming, but operators are deploying counter-measures against this as well, with a standard known as 'Near Real Time Roaming Data Exchange", and this is expected to reduce roaming fraud by as much as 90% [69].

3.5.1.3 Subsidy fraud

Subsidy fraud takes advantage of the cheap handsets offered by many operators in conjunction with a new subscription. New subscriptions are established (e.g. with fraudulent credit cards) but not paid for, and the handsets are sold cheaply abroad with a profit. Fraudsters can get their hands on large quantities of handsets in this fashion, e.g., by pretending to be small enterprises. The main counter-measure available to prevent this is the SIM-lock mechanism, which ties phones to be sold by operators to a specific range of SIM cards, so that they can't be used on others' networks. However, until now, fraudsters have been able to defeat even the most advanced SIM lock mechanisms on the market, and subsidy fraud is a very profitable business and a big cost for operators.

3.5.2 Level of hardware protection

One of the biggest technical challenges of moving the SIM into the terminal is to provide the required protection for its security-critical functions, most notably the authentication algorithm and subscriber key. It must be assumed that a potential attacker has physical access to the device. The normal approach in such situations is to physically protect the inner workings of a system and provide a limited set of interfaces to the outside environment that can only be used to interact with the system in legitimate ways. For instance, an automated teller machine is essentially a PC running inside a safe, where a screen/keypad interface typically allows users to interact with the machine in predefined ways. With unrestricted access to the computer inside the ATM, it would be possible to instruct the machine to dispense cash without the corresponding amount being debited from an account. Similarly, the smart card provides a physically secure environment (as described in section 2.3.2) where the secret subscriber key can safely be stored and where the operator's algorithm can run without being tampered with. The electrical interface to the smart card offers a restricted way of interacting with the system, allowing only certain 'legal' operations to be performed.

3.5.2.1 Secure element

A secure element refers to a dedicated chip package containing a microprocessor, RAM, ROM and non-volatile memory (EEPROM or flash-memory). This can either be in the form of a removable card (i.e. a smart card), or it can be in a more traditional chip package, embedded in the terminal. A secure element can be programmable, so if it is used to protect the VSIM, it could be possible to change subscription in the field. An example of such a chip is the Smart-MX manufactured by NXP Semiconductors, which was used by Nokia in the 'Nokia 6131 NFC' phone to provide security for NFC services from multiple third parties [73]. Smart cards have proved their worth in fields with high security requirements, e.g. telecommunications (SIM cards) and banking (credit cards). While it is theoretically possible to compromise their security, given enough resources, smart cards have extensive protection mechanisms making such attacks extremely expensive and time-consuming, as described in section 2.3.2.

3.5.2.2 Secure processor

Normally, a processor reads/writes data and instructions from/to external memories, which makes it possible to tamper with the chip's operation with physical access to the electrical interfaces connecting these components. But assuming that an adversary cannot penetrate the CPU itself, it might be possible to perform private computation, by making sure that instructions and data are always in encrypted form outside of the chip, and by cryptographically verifying the integrity of code and data as they are read from external memories [74]. Some research projects explore this idea, e.g. XOM [75] and AEGIS [76], and commercial products already implement the basic concept of a general purpose processor containing a logical 'secure environment' that can be used to run security-critical code.

For instance, Nokia uses a design called Baseband 5 (BB5) in its new medium- to high-end handsets, which features such a secure environment (and eventually, this system will be used throughout the whole product line). Motorolas 3GPP proposal to consider the introduction of a SoftSIM [10] also indicates that (some of) their terminals have a secure environment. Furthermore, commercial equivalents of this technology are available from Texas Instruments under the name M-Shield [77, 78], and from ARM, which licenses a secure extension to its processor architecture called TrustZone [79].

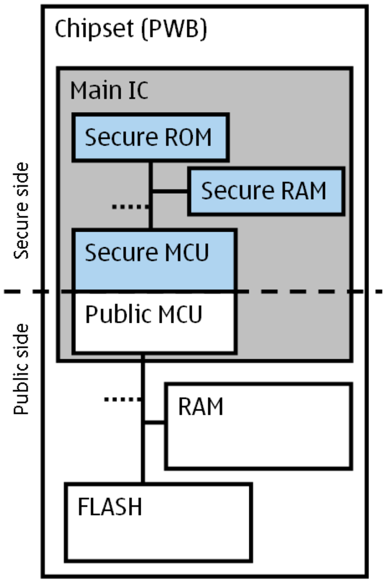


Figure 3.6: Illustration of the Baseband 5 memory architecture, where the secure environment has access to secure resources not available to the public side. Source: Internal Nokia slides.

In these products, the secure environment is implemented as an additional secure processor mode, which provides additional privileges compared to the normal mode(s) used by the OS. These additional privileges include access to secure RAM and ROM physically located inside the chip, as well as access to other hardware resources reserved for use by the secure environment, e.g. secure interrupts. The processor is designed such that code running in normal modes cannot interfere with or gain access to code and data in the secure environment.

Such a secure environment can be used to establish a trusted computing base for the rest of the system using integrity & digital certificate checks. But more importantly (in this case) is the fact that code running in the secure environment is protected from interference from normal software, and that it never leaves the physical chip. It would thus be necessary to open the chip package and access the silicon die in order to spy on or modify data in the secure environment. This makes it a good candidate for protecting the operator's algorithm and the secret subscriber key.

The tamper-resistance provided by secure processors is arguably smaller than that of secure elements, partly because they commonly lack the special tamper-protection mechanisms that are standard in secure elements (see section 2.3.2), and partly because of the fact that the interface between the secure environment and the rest of the system is often not as simple as is the case with a secure element which has a specialized physical interface. The more complex interface is an inherent characteristic that differentiates secure processors from secure elements, but there is no reason why a secure processor could not feature some or all of the tamper-protection mechanisms found in secure elements – other than cost, that is. On the other hand, secure processors significantly raise the bar on tamper-protection compared to the software-only methods discussed in the next section, due to the fact that, theoretically, a physical attack is required in order to compromise the system. Such an attack could for example involve removing the chip package in order to analyze the chip features with a microscope. Such attacks are likely to render the chip unusable (or at least the product in which it is used), which would significantly raise the cost of the attack, and thereby reduce the scenarios in which such an attack is profitable. Case in point, if a phone's SIM-lock is based on a secure processor, destroying the processor (and thus the phone) in order unlock it is not a profitable attack. Conversely, if the phone contained long-term secrets that could be used to unlock all other devices of the same model, it would probably be worthwhile to sacrifice one to be able to unlock all others.

3.5.2.3 Trusted platform module

The term Trusted Platform Module (TPM) can refer to both a specification [80] authored by the Trusted Computing Group (TCG), and to implementations of this specification. TPM refers to a secure chip added to a computing system in order to augment it with a root of trust, which can be used for various purposes. TPM modules offer various facilities such as checking program and data integrity, secure key storage, remote attestation and sealed storage. Remote attestation creates a (digitally signed) summary of the state of a system, which can be sent to a remote entity. This entity can then determine if the system has been manipulated with or if it can be trusted. Sealed storage allow applications to store encrypted data, that can only be decrypted if the TPM releases the associated key, which it only does to the same application that encrypted the data.

The TPM specification can be implemented by any vendor, but implementations are tested for compliance by the TCG [81]. A mobile version of the TPM concept has recently been standardized under the name Mobile Trusted Module (MTM) [82]. MTM modules differ from TPMs in the following ways [83]:

a) The concept of secure boot is introduced, where the boot sequence is not only integrity checked, but also aborted if the verification fails.

b) As opposed to TPM, an MTM implementation does not need to be in hardware, but can build on an existing security architecture in a device.

c) The reference architecture takes into account the possibility of several parallel MTM instances in the same device.

The following possible approaches can be for the implementation of a MTM module [84]:

a) A specialised MTM chip.

b) A TPM 1.1 or 1.2 chips and some extra layer in software to implement extra commands.

c) Another HW chip bound to the platform and running an MTM application amongst others.

d) SW MTM running in a virtualised engine with the virtualisation environment protected by an underlying HW MTM.

e) SW MTM running in a CPU chip.

Trusted Platform Modules are mostly designed to protect a system from software based attacks, and do not provide a secure execution environment. For instance, when an application has passed integrity checks and is allowed access to a cryptographic key, the key is made available to the application in the main memory. Thus, an attacker with physical access could recover the key on the memory bus or directly from RAM using e.g. a cold boot attack [85]. Furthermore, the fact that MTM modules are flexible about their implementation means that the use of an MTM-based system does not in itself guarantee the underlying level of security.

3.5.2.4 Software protection

This type of protection relies solely on software mechanisms to guard the executing code and data from tampering or disclosure – no special security-hardware is involved. Such approaches are common in the PC world, for example where software running on legacy hardware is used to enforce access restrictions on copyrighted digital media files. In this case, the files are stored in an encrypted format, and only the official player or viewer, which will enforce the access restrictions, contains the secret key or algorithm required to decrypt them. These programs must be protected from attempts to bypass the restrictions or to extract the secrets they contain. This is not unlike the situation where a Virtual SIM scheme was designed to protect the operator's secrets without hardware-based protection methods: an attacker might try to build his own replica of the system which would leak the secrets, or he might attempt to extract the secrets from a genuine system.

History shows that such systems are notoriously insecure [86, 87, 88]. An interviewee pointed to white-box cryptography as a potential software protection method [89]: Chow et al. have proposed methods for implementing the DES and the AES block ciphers [90, 91] in such a way that it becomes hard to extract the embedded secret key in a white-box attack context, in which the attacker has full access to the implementation and its execution environment. Several attacks have been published affecting these methods [88, 92, 93, 94, 95]. While these implementations can still provide an effective protection of the key, and the task of extracting a key can be complicated considerably [96], software-attacks are inexpensive and non-invasive, and it would probably be unwise to trust these methods to provide a high level of security for some time into the future.

Another approach called Pioneer is presented in [97], which provides a pure software-based method of remote execution attestation (that is, a trusted remote system can get a proof that an executable has run unmodified on an untrusted system). However, it requires the remote system to have detailed knowledge of the hardware architecture on which the executable runs. This is probably unrealistic in our usage scenario. Furthermore, as stated in the paper, it assumes that an attacker cannot directly manipulate the memory or the hardware, which would be possible with physical access to the device.

3.5.2.5 Summary

Table 3.1 summarizes the protection provided by the various platform security schemes described above. Secure device identity refers to the ability to prevent the unique identifier of the device from being changed. Secure storage is the ability to provide a system where stored data (e.g. keys) can be protected from being recovered in plaintext both by unauthorized applications (or the operating system) on the device and by someone with direct access to the physical memory, if there is no access control/tamper-protection on the (physical) memory module. A scheme providing a secure execution environment allows custom code to be executed entirely inside a tamper-proof device, such that e.g. keys manipulated by the code cannot be extracted from memory and that the intended control flow of the code cannot be modified. Protection from software attacks refers to whether the system can protect critical code and data from being disclosed/altered by malicious software running on the system. Finally, protection from simple and advanced physical attacks refers to the level of protection of code and data (whether it is stored in permanent memory or being executed) from situations where an attacker has physical access to the device. Invasive and semi-invasive attacks (see section 2.3.2) are classified as advanced.

Table 3.1: Summary of protection provided by different platform security schemes. 'depends' means that the protection level provided depends on the way the MTM module is implemented.

	software-only	TPM	MTM	secure processor	secure element
Secure device identity	no	yes	depends	yes	yes
Secure storage	no	yes	depends	yes	yes
Secure execution environment	no	no	no	yes	yes
Protection from software attacks	yes	yes	yes	yes	yes
Protection from simple physical attacks	no	no	no	yes	yes
Protection from advanced physical attacks	no	no	no	no	yes

The security of MTM largely depends on the chosen implementation, but it can at most reach the same security level as a TPM (for the aspects considered here).

Figure 3.7 below summarises the possible approaches, and shows how the choice is a compromise between security, flexibility and cost.

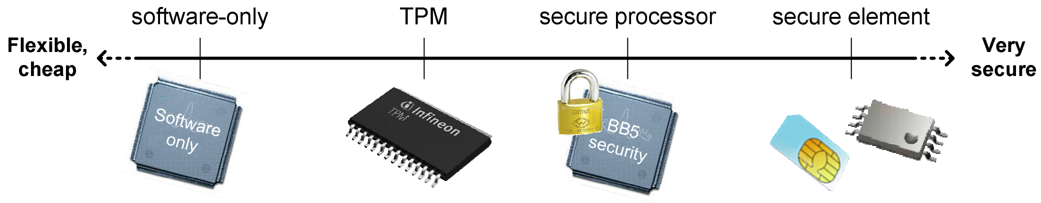


Figure 3.7: Scale showing the different levels of hardware security considered for protecting the Virtual SIM. Whether the TPM is cheaper than the secure processor depends on the use, but they have been put in this order due to their respective levels of security. The MTM system is omitted, since the security depends on the actual implementation used. Source: own work.

3.5.3 On-Board Credentials

Regardless of which hardware protection scheme is chosen, a system for storing and facilitating the use of Virtual SIM credentials and algorithms securely is needed. A team at the Nokia Research Center (NRC) has developed a platform for handling user credentials called 'OnBoard Credentials' (ObCs) [98]. This platform leverages general-purpose security hardware, such as secure processors or TPM/MTM modules (or a dedicated secure element, although this option isn't mentioned in the NRC paper), to provide a high level of security for the credentials while also featuring a high degree of flexibility.

The fact that the ObC platform was designed to reuse existing general-purpose security hardware means that it is cheap to implement in devices already featuring such hardware. It also means that the system must take into account the constraints of these hardware platforms. For example, in secure processors with on-chip memory, the resources available for the OnBoard Credentials are very limited – in the order of tens of kilobytes of RAM and hundreds of kilobytes of ROM.

The rest of this section will give a short introduction to ObCs, as they will play an important role in the Virtual SIM system described in the following chapter. For an in-depth description of the ObC platform, see [98].

In the ObC paradigm a credential consists of *secret data* such as keys, and an algorithm that operates on this data, known as a *credential program* (also referred to as *ObC secrets* and *ObC programs*). ObC programs are isolated from each other and cannot normally access each others' secrets, although it is possible to have a group of programs share a piece of secret data (for instance the same service provider could have two versions of an algorithm operating on the same key). Figure 3.8 shows the ObC architecture when implemented using a secure environment.

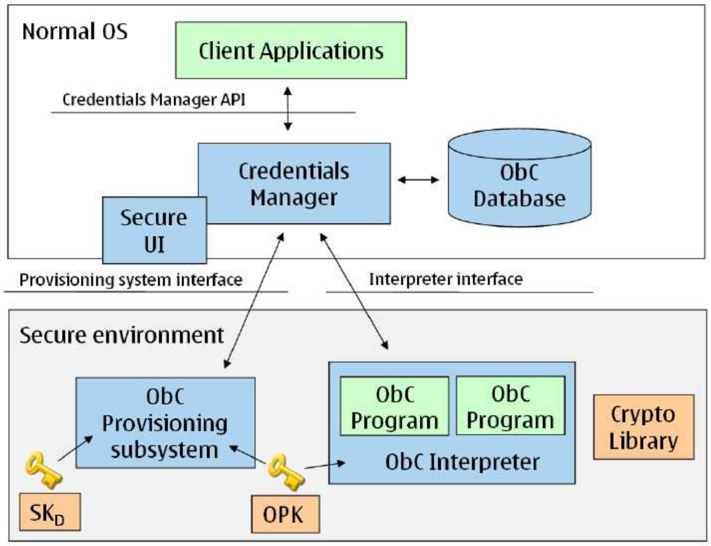


Figure 3.8: On-Board Credentials architecture. Source: [98].

3.5.3.1 Provisioning

The provisioning subsystem is used to transfer a credential (ObC programs and secrets) securely from a service provider to the client system. It relies on the assumption that a device-specific public/private key pair exists where the private key is available only within the secure environment of the device, and that the service provider knows the public key. The provisioning system works by sending a collection of encrypted and integrity protected packages to the client. Figure 3.9 shows the packages used for provisioning a credential consisting of two secrets and one program. The system is based on the notion of families, which is a group of secrets or programs. A family is established using the ObC/Init package, which contains a family-specific root key and is encrypted using the public key of the device. In the example in Figure 3.9, two families are established, each with its own root key: one for secrets and one for programs. The root key for the 'secrets' family (RK_S) is then used to encrypt two ObC/Xfer packages containing each of the two secrets to be transferred. Similarly, the root key for programs (RK_P) is used to encrypt an ObC/Xfer package containing the credential program. Finally an ObC/Endorse message is created which is used to instruct the device that a certain program should be allowed to use a family of secrets. The endorse message contains a hash of the program, and is encrypted with the root key of the 'secrets' family.

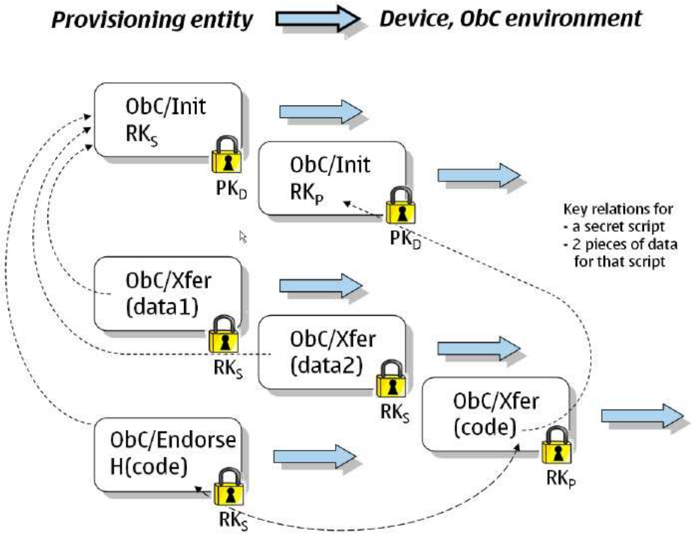


Figure 3.9: ObC provisioning packages and key relations. Source: [98].

The provisioning packages are self-contained when it comes to security, and can be transferred using any secure or insecure mechanism.

3.5.3.2 Interpreter

The ObC programs are run by a byte-code interpreter in order to isolate them in order to prevent them from interfering with other programs running in the secure environment, and in order to facilitate interoperability between terminals. Due to the memory restrictions of the secure environment, a memory-efficient byte-code interpreter is needed, which a programming language called Lua (version 2.4) [99] was chosen for ObC programs. A part from the small footprint of the interpreter, some other important advantages of Lua are that it is portable (the interpreter is written in ANSI C), it has a liberal license, the source code is freely available so that a certain version of Lua can be used for a specific purpose without the need to incorporate new features of the language, and additionally, the language is simple and execution is fast [100]. Additionally, compared to Java, Lua has the advantage that it's license-free.

An example of a Lua script written for the ObC interpreter is given in Appendix C.2. The example is an implementation of the MILENAGE algorithm for UMTS authentication (see section 2.2.4).

3.5.3.3 Credential manager

The credential manager is a software interface to the ObC system, providing access from the untrusted side of the environment to both the provisioning and interpreter subsystems in the secure environment. It provides an API for creating, managing and using credentials, and also supports PIN-based access control. Credentials can be created by passing the provisioning packages described earlier to functions in this API. Other functions are available to invoke credential programs. The details of these procedures are outside the scope of this report.

3.6 Prior work

3.6.1 Mobile trusted platform approach

Kasper et al. have proposed a virtual SIM system [101] based on the Mobile Trusted Module (MTM) specified by the Trusted Computing Group (TCG) [82], which is the mobile equivalent of the Trusted Platform Module. While the use of the Trusted Computing Group's proposed architecture has some advantages, such as an interesting separation of the trust domains in the terminal and not having to define a new foundation for the security of the terminal, their proposed system has some major drawbacks, when it comes to real-world use of the system. The provisioning procedure is illustrated in Figure 3.10 and Figure 3.11. Apart from its complexity, it has a few other problems making it conflict with the actual needs of such a system. First, the point of sale of the subscription needs to communicate with the mobile operator in order to activate the subscription once it is sold to the user, which is a problem when subscriptions are not sold in operator-owned stores, e.g. supermarkets, gas stations or other general purpose store. In those cases the stores would need to have a secure connection to all the operators for whom they sell subscriptions. Furthermore stores need special equipment, since they are required to interface with the handset, which means that staff needs to be trained in using it. This is highly unlikely to happen in non-specialized stores. Subscriptions cannot be pre-activated, since the activation step transfers the user's certificate to the operator.

Once, the user has bought the subscription in a point of sale, the virtual SIM must be rolled out to the terminal (i.e. downloaded). The paper fails to discuss how this should happen, but simply states that "the mobile device is able to access the registration service provided by MNO over some kind of channel. For instance, this service is implementable as a network teleservice or internet download service."

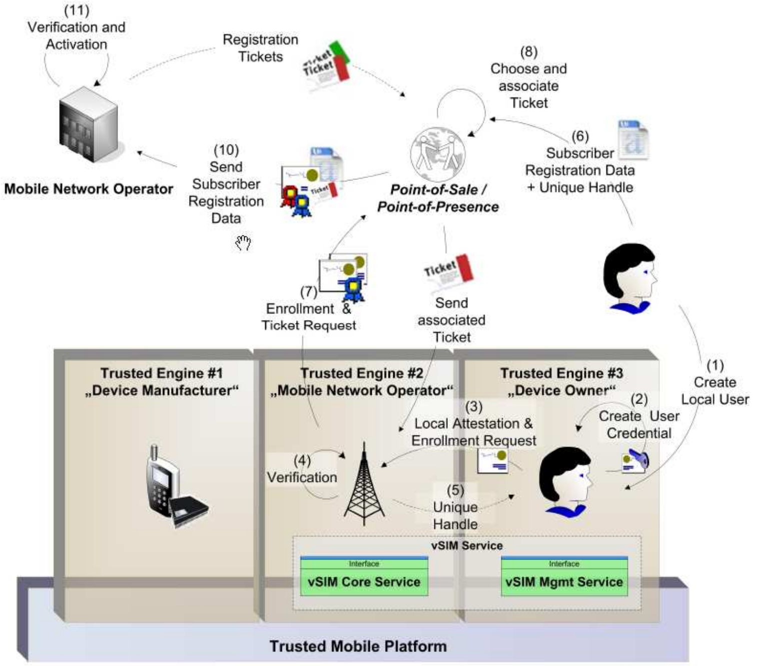


Figure 3.10: Model for "Subscriber Registration and Enrolment". Source: [101].

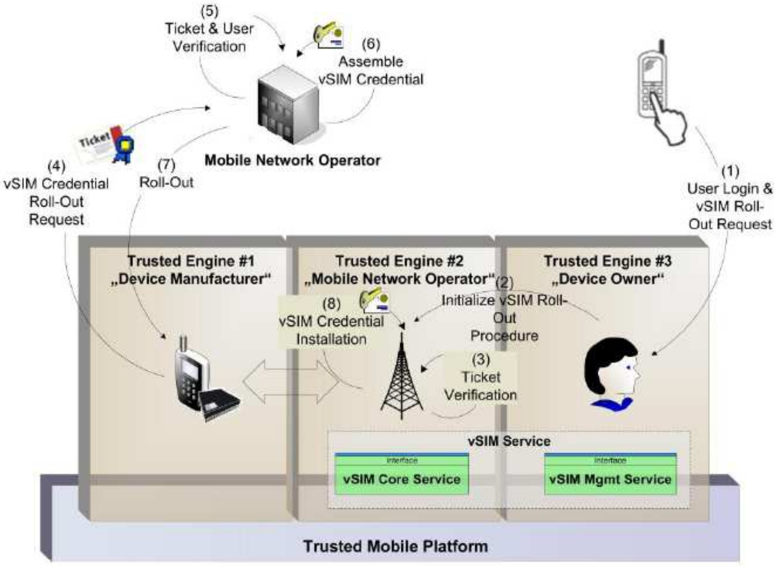


Figure 3.11: Model for "vSIM Credential Roll-Out". Source: [101].

Chapter 4

System design

4.1 The challenge

Besides its core functions of subscriber identification and authentication, the SIM also has numerous other capabilities. In many cases there is a certain overlap with functions in the terminal (e.g. message storage and phonebook), and in other cases, the reason for having a specific function in the SIM card is its security capabilities. With the Virtual SIM being managed by the terminal, the arguments for having duplicate functionality vanishes, and the control over these various functions should be placed either on the terminal or on the VSIM side, as illustrated in Figure 4.1.

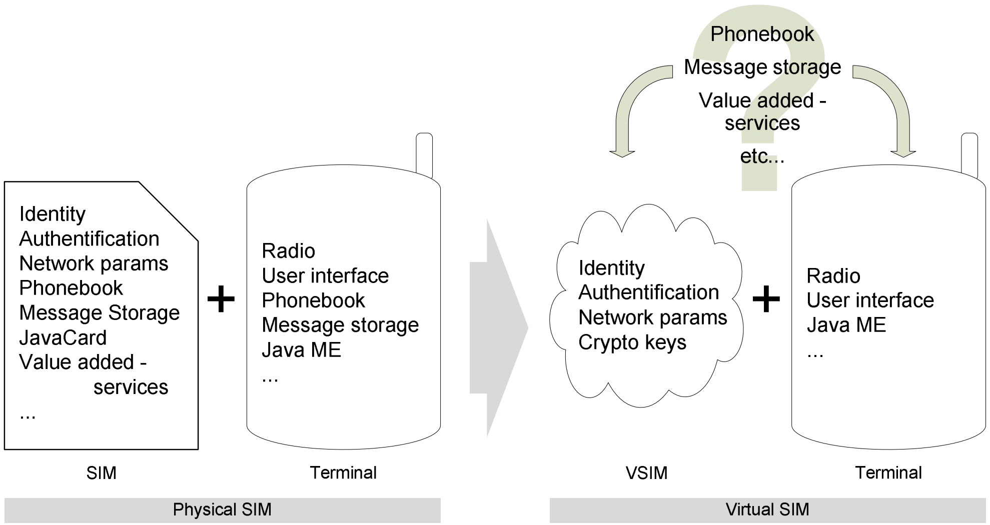


Figure 4.1: With the introduction of the Virtual SIM, the control over some features could potentially lie with the VSIM or the terminal, while some key functions such as subscriber identification an authentication naturally lie with the VSIM. Source: own work.

If the decision was entirely up to operators, they would probably want to provide all the important functions themselves, and the terminal would be a shell providing CPU power, man-machine interface, and network access. Conversely, terminal vendors would probably want operators to merely provide network access, and not try to control aspects of the user experience. Obviously, these 'ideal' situations cannot satisfy both parties, and a compromise must be found that at least addresses the minimum requirements of each stakeholder.

4.2 Overview

As mentioned, one of the key challenges with the Virtual SIM is that operators need to relinquish control over the subscriber keys, and trust device vendors to protect them. For this to work, operators need some sort of assurance that when they issue a VSIM to a terminal, the terminal will indeed be capable of protecting the secret key. The chosen solution to this issue is to give operators the freedom to decide whether a given terminal is trusted or not. This approach gives terminal vendors an incentive to protect VSIMs adequately, since operators might otherwise decide to boycott a specific device or vendor. In reality, this is probably not how operators would act, unless a certain terminal was critically flawed. Instead, new terminals would undergo a conformity test by an independent body, where the security of the device would be ascertained. After passing the test, the device would be considered trusted by operators. Of course, this presupposes that operators have a reliable method of determining the identity of a terminal before issuing it a VSIM, so that rogue terminals cannot spoof their identity and pretend to have passed the conformity test. This aspect will be discussed in section 4.6.4.

The security of the proposed Virtual SIM scheme is based on the use of a secure processor (see section 3.5.2.2), although some vendors may choose to implement this in a separate secure element. This choice satisfies several of the requirements discussed in the previous chapters: The secret key and algorithm remain within a tamper-proof environment, and in addition the use of a secure processor means that it can be implemented at a negligible cost.

In section 2.3, the SIM card was described as having four basic types of features, 'security', 'data storage', 'management functions' and 'supplementary applications', which must be implemented in the terminal. Furthermore, with the VSIM, there is no token that can simply be placed into the handset to give it a subscription. Thus, another essential feature is a provisioning mechanism allowing the transfer of a subscription to a 'blank' handset. These features are summarized in Figure 4.2. The essence of the Virtual SIM is that these features should be implemented by the underlying device and network, but controlled by data provided, in order to provide a standardized platform on top of which different mobile network operators can deploy their Virtual SIMs.

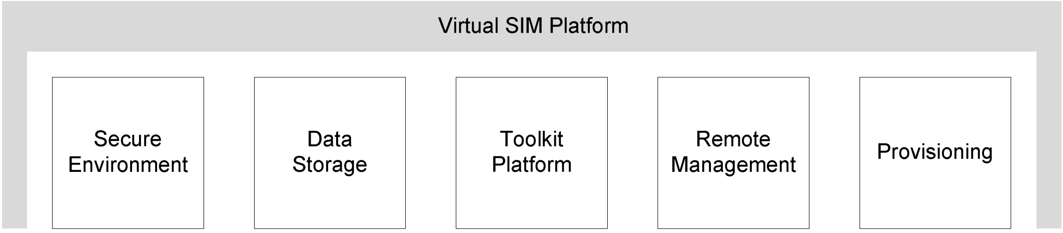


Figure 4.2: The five basic functions of the Virtual SIM platform. Source: own work.

4.3 Basic SIM functions

The OnBoard Credentials platform described in section 3.5.3 provides a good foundation for the authentication functions of the VSIM. However, it only fulfills part of the functionality required for the basic SIM functions. In addition to the authentication algorithm, a core function of the Virtual SIM platform is to handle the SIM file system. As described in section 2.3.1.2, SIM files have various access restrictions applying to them which the Virtual SIM should also be able to enforce. Several approaches can be considered:

a) SIM file access is controlled in secure mode. This is only possible if the SIM file system is entirely managed inside the secure environment.

b) PIN check is handled in secure mode, but file access control is delegated to a public mode driver.

c) PIN check and access control is handled in public mode.

A part from cryptographic keys, which in the case of the VSIM would be stored by the ObC platform, and not in the file system, SIM files are used by the terminal. If the terminal is compromised, it could simply ignore the contents of SIM files. Thus, letting a public mode driver handle the SIM file system does not create any security threats compared with the situation where it is managed by a physical SIM card, and there is no reason to waste resources requiring the secure mode to manage the SIM file system. On the other hand, it might be desirable to prevent access to SIM files if the user hasn't entered his PIN code, in order to protect personal data in case a terminal is stolen. For these reasons, the second approach has been chosen: letting the terminal's operating system manage the SIM file system, but only allowing this after the PIN code has been checked by the secure mode.

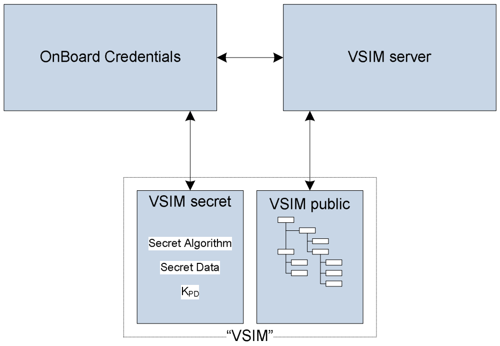


Figure 4.3: Virtual SIM architecture for the basic SIM functions. Source: own work.

The architecture drawing in Figure 4.3 shows how this is implemented in the terminal. The architecture is built around two major components. The OnBoard Credentials part is implemented in the secure mode of the processor and is responsible for handling the secret subscriber key and the operator's algorithm. The VSIM server is the interface towards the rest of the terminal, and also manages the SIM file system. It handles SIM requests from the terminal, and determines how they should be carried out. File requests, for instance, are handled by the VSIM server itself, whereas requests to perform PIN checks or GSM authentication are redirected to the ObC module in secure mode.

The VSIM data is stored in two separate parts. The secret part contains the GSM/UMTS authentication algorithm(s) and the related secret data (i.e. the subscriber key, and possible sequence number counters for UMTS mutual authentication, etc. – this is up to the implementer of the algorithm). In addition it contains a randomly generated key used to encrypt the public part of the VSIM containing the SIM file system. When the correct PIN code is entered, the OnBoard Credentials platform will grants access to the VSIM, which will make the key available to a VSIM server on the public side. The VSIM server then has access to read/write the SIM file system, and is responsible for enforcing access policies on SIM files.

As mentioned, the VSIM server acts as an interface for the Virtual SIM towards the rest of the terminal. How this is implemented is up to the individual terminal vendors to decide. One approach would be for the Virtual SIM server to emulate a SIM card at the APDU level[11], as illustrated in Figure 4.4. This would allow the rest of the terminal to be indifferent as to whether a physical SIM card or a Virtual SIM is being used. This would make it simple to have hybrid terminals supporting both systems during a transitional period. After this, the interface to the physical SIM card could be removed, and the APDU interface could possibly be dropped altogether.

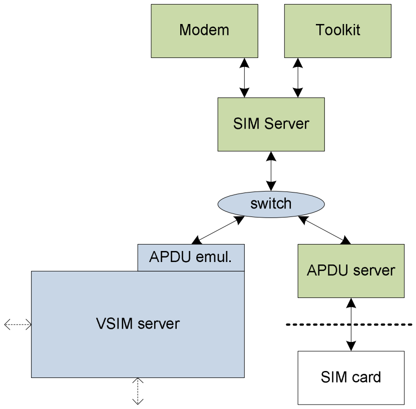


Figure 4.4: Virtual SIM architecture coexisting with the physical SIM. The VSIM Server emulates APDUs and implements the standard SIM card commands. A switch selects whether the terminal communicates with the VSIM or the physical SIM. Green boxes represent components that already exist in terminals, and blue boxes are components introduced with the Virtual SIM. Source: own work.

4.4 Toolkit platform

Although Lua scripts are used for credential programs (e.g. the GSM authentication algorithm), the Lua environment provided by the ObC platform is probably too limited to function also as a full-fledged SIM Application Toolkit platform. Instead, the Java ME [102] environment present in most modern phones could be used to provide a modern, interoperable platform for operator applications. As previously describes, it would be very challenging to introduce such a new platform unless it had at least the same capabilities as the one currently used. Thus, Java ME applets would need to have access to an API implementing the SIM Toolkit functions listed in Table 2.3 on page 12. This API (the 'VSIM API') should include an interface to the modem in order to implement call control function, network measurements and communication functions (e.g. data download, bearer independent protocol). In addition, the VSIM API should have some mechanism giving access to the SIM file system. This architecture is illustrated in Figure 4.5.

Another feature that should be supported is the interaction between Java and the OnBoard Credentials platform. Service provider applets should be able to interact with ObC programs belonging to the same service provider, in order to make it possible for to create secure Java ME based services.

[11] APDUs are data packages in the high-level protocol used for communicating with smart cards, see e.g. [27].

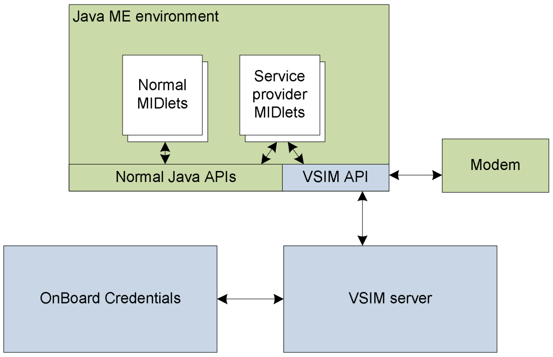


Figure 4.5: Proposed architecture for a VSIM API for Java ME. Source: own work.

4.4.1 Service provider privileges

Powerful mechanisms such as call control should not be made available to any Java MIDlet[12] in the terminal. Also access to SIM files should be restricted to the MIDlets belonging to the service provider (SP) that issued the given SIM. Thus, SP applications need different permissions than regular applications, and to accommodate the possibility of having multiple VSIMs, there should be a mechanism to associate MIDlets with VSIMs so that access restrictions can be enforced.

Most of the solution to this problem already exists in the form of the MIDP 2.0[13] security model [103, 104]. This specification introduces the concept of trusted MIDlets which are digitally signed. When the MIDlet is installed to the terminal its signature is verified against appropriate root certificates, thereby authenticating the signer. These can either be stored in the device itself, on the Wireless Identity Module (WIM), or the (U)SIM. MIDlets are assigned a protection domain, which is determined based on the root certificate used for signing it. Protection domains define a set of permissions granted to MIDlets, which in turn enables or disables access to various APIs (for instance, only some trusted applications should have the right to make phone calls, access phonebook entries, etc.). There are two sets of permissions: 'allowed' and 'user'. The latter asks for confirmation by the user when an application wishes to use a restricted API. Mobile phones (GSM and UMTS) have at least the following protection domains:

- 'Manufacturer Domain' is the most powerful domain and, allows unconditional access to all APIs.
- 'Operator Domain' contains applications that are signed by the mobile operator. This domain is used when a signed MIDlet is successfully verified against a root certificate on the WIM or (U)SIM. On the (U)SIM, the operator certificate(s) are stored in a SIM elementary file. The operator domain sets all possible permission to 'allowed'. The specification takes into account the case where the SIM card is changed, see [103:ch8]. In this case, MIDlets in the operator domain whose signatures no longer match a root certificate on the new SIM must be disabled by the terminal.

[12] MIDlets are Java ME programs

[13] Mobile Information Device Profile (MIDP) is a specification for the use of java on embedded devices and is part of the Java ME framework.

- ‘Third-Party Domain’ is intended for applications signed by other entities than the terminal manufacturer and the operator. All permissions are set to ‘user’, and the user must thus approve access to all protected API calls.
- ‘Untrusted Domain’ contains all the MIDlets which are not signed.

There are some security concerns with this concept of protection domains due to some of the problems of PKI systems [105, 104]. However, these concerns only apply to the third-party domain, since in the manufacturer and operator domains each organization can define their own root keys and do not need to rely on a public PKI. In fact, the concept of protection domains is almost ideally suited for ‘Toolkit MIDlets’, and the current Java ME standards can be used as-is to implement this functionality by the addition of the VSIM API. This API would, however, need to keep track of which VSIM a currently running operator MIDlet is associated to, in order to properly perform VSIM-specific operations (e.g. SIM file access). This is relevant in the case where multiple VSIMs are installed (even though only one of them is active at a time).

4.5 Remote management

Mobile operators should be able to remotely manage SIM data in the terminal, as discussed in the previous chapters. The OMA Device Management (OMA DM) platform [106] can be leveraged to provide this functionality. Device management refers to technologies that allow third parties to configure mobile devices. Terminals supporting OMA DM contain a device management tree, which at its leaves contains management objects representing e.g. settings that can be accessed and manipulated. The management tree is structured such that different entities can have access to only a subset of the possible settings, and various methods of access control enforce these rights [107]. Figure 4.6 shows such a setup where the mobile operator and the device vendor each control their respective management domains.

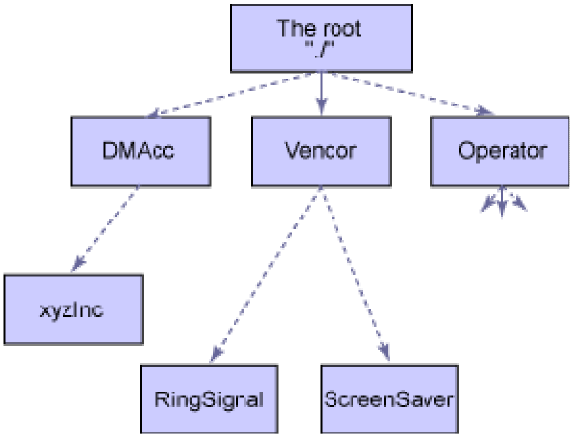


Figure 4.6: Example OMA DM management tree. Source: [108].

In the case of the Virtual SIM, SIM files and ObC programs and secrets could be managed under a sub-tree controlled by the operator. The VSIM would contain credentials giving the operator’s management server access to the corresponding branch in the management tree.

OMA DM is a widely accepted standard that is already deployed in many mobile terminals, and using OMA DM as a foundation for the remote management of the Virtual SIM also has the advantage that most of the stakeholders are already familiar with the technology.

4.6 Provisioning

One of the critical aspects of a Virtual SIM system is how to get a subscription into a device when a physical token isn't used. Provisioning refers to the process of transferring an operator's VSIM into the terminal, in order to bring it to a state where it is able to communicate using the current GSM/UMTS system. The term 'VSIM' is used to denote the operator-specific subscription data. We will assume a minimal set consisting of the IMSI, the Ki, and the authentication, although more information could be included (for instance, all mandatory SIM files plus files required in order to establish a data connection).

To complicate matters further, the operator must be able to ascertain whether the target device can be trusted, as explained in section 3.2.3. Consequently, the operator must have proof of the identity of the device in order to be able to make this decision, and the subscription information must be transferred in a way that ensures that only the intended device receives the information, in order to prevent an attacker eavesdropping on the communication channel to duplicate the VSIM.

4.6.1 Provisioning channels

Many possible channels (methods of transferring the VSIM) can be used to provision the terminal, each with their strengths and weaknesses. Which method to adopt largely depends on the usage scenarios the system is designed for. One approach may be better when subscriptions are bundled with the terminal, whereas another may better suit an open terminal market where users must be able to interchange subscriptions and/or terminals easily. In this section, a description is given of the methods considered in this study as well as their respective strengths and weaknesses.

4.6.1.1 Secure token

The VSIM can be distributed to the user in a secure hardware token. In this case, 'secure token' refers to a secure physical device, such as a smart card or a secure memory card, and not an authentication token such as SecurID. The role of the secure token is to evaluate whether or not a terminal it is inserted into is trusted to store the VSIM. In the affirmative case, the VSIM is transferred to the device, and the secure token can be removed. The SIM card is also a secure token, the difference being that the SIM card must remain in the device (and does not evaluate the terminal or copy protected data). This scheme has a few advantages over the status quo with the SIM card. For example the phone can have multiple VSIMs, and if a secure memory card is used, the connector could be used or for other purposes once the VSIM has been copied to the terminal. However, these are generic VSIM advantages, which also apply to the other provisioning schemes described here.

Advantages	Disadvantages
• Current retail methods can be reused. • Memory card slots can be used for other purposes once the VSIM has been copied to the phone.	• Does not have cost advantages over the SIM card.

4.6.1.2 Provisioning in factory

The VSIM can be transferred to the device during the manufacturing process. This way, operators do not need to bother about the supply of physical SIMs. The process is also simplified at retail stores, since users get a working phone, and do not need to go through the trouble of inserting a SIM card. For obvious reasons, this is only feasible for devices sold by operators, as opposed to vanilla phones, which are sold directly to end-users without a bundled subscription. Furthermore, this method does not by itself provide any way of changing the subscription in the device subsequently. Some operators might find this prospect appealing, but it is probably not good for the industry as a whole, and certainly not for the end-user [22]. The VSIM transfer mechanisms are easy to secure, since they are only used in controlled environments (terminal production facilities).

Advantages	Disadvantages
• No change needed in access network.	• Does not work for vanilla handsets.
• Phone works out of the box.	• Does not include a solution for changing subscription later on.
• Operator can control in which devices VSIM is used.	• Operator must allocate IMSI numbers a long time before the subscription is put to use, which is wasted inventory.
• Device manufacturers can sell this as a service to operators.	
• Easy to make secure.	

4.6.1.3 Provisioning using a local connection

The VSIM can also be transferred to the phone using its local connectivity interfaces. For example, this could be a cable/USB connection, Bluetooth, infrared, memory card, or any other transfer method. The provisioning system can be made very flexible and could easily be adapted to many kinds of situations. It does, however, require that the VSIM can be properly encrypted, since it will be transferred over an insecure channel. Furthermore, the operator needs a way to determine whether the terminal the VSIM ends up in can provide adequate protection measures, and that it is not a ‘malicious device’, built specifically to compromise the VSIM (this is also the case for the other provisioning methods, except of course factory provisioning).

The transfer could be done either at a retailer with the right equipment, or by the end-user who could download a subscription from the internet to his home computer and from there transfer it to the phone. This is also the main disadvantage of the scheme: Advanced equipment (e.g. a computer) and internet connectivity is a necessity, and users must be trained in using it. For some retail channels, this will be impossible or expensive to implement, e.g. gas stations or super-markets, and it will also be impossible to use in developing countries, where such advanced equipment and connectivity are rarely available. Furthermore, the complexity involved would probably result in an increased need for customer care and support.

Advantages	Disadvantages
• No change needed in access network.	• Requires computer equipment in all retail stores.
• Subscriptions could be downloaded from the internet and installed by the end-user.	• Requires training of retail personnel.
	• Impossible to make a ‘low-tech’ sale, e.g. in gas stations or for in developing countries.
	• Transmission security and device trust issues (see section 4.6.4).

4.6.1.4 Over-the-air provisioning

Most of the above mentioned problems could be solved if the user could simply turn on his phone, and download the subscription he wants directly from the mobile network. However, the current standards do not permit a device without a subscription to do anything else than placing an emergency call and receive cell broadcast messages. The access network would need to be changed to allow a subscription-less device to connect in order to download a VSIM, and such changes raise the acquisition cost of a VSIM system for operators significantly. Furthermore, this also raises the following question: Should it be possible to download a VSIM from one operator while roaming on another operator's network?

Essentially, there are three possible approaches to this roaming issue:

a. VSIMs can only be downloaded from the operator's own network. In this case the user must be inside the coverage area of the operator he wants to download a VSIM from.

b. VSIMs can be downloaded while the user is roaming, assuming the serving operator's network supports OTA provisioning, and that the serving operator and the 'home' operator (the one the user wants to download a VSIM from) have a special roaming agreement covering VSIM transfer.

c. VSIMs can be downloaded when connected to a visited operator, without requiring special support from the visited operator's network, assuming that a normal roaming agreement exists.

The last option requires the VSIM provisioning system to somehow use the existing GSM authentication scheme in a non-standard way, in order to signal to the desired operator that a terminal without a subscription wishes to connect. The feasibility of this option has been investigated, including whether it would be possible to use non-unique IMSIs for provisioning purposes along with a default Ki, or whether devices could be issued a 'preliminary IMSI' to be used for provisioning. Unfortunately, no viable solution was found, and in any case it would be an abuse of the existing design, which is usually a bad idea. Furthermore, any such solution would be problematic due to the shortage of IMSIs. Thus, the rest of this section will focus on the first two options.

In order to download a VSIM from an operator, the minimum amount of information required from the user is the identity of the operator. In GSM, operators are identified by a three-digit *mobile country code* (MCC) and a two- or three-digit *mobile network code* (MNC) [17]. The combined number is known as the *PLMN code*[14]. Once this code is known, it is possible for the device to establish a communication channel to a given operator, as will be described in the next section. This channel can then be used to download a VSIM.

The user could choose the desired PLMN code in three ways:

a. The terminal could contain a list of all supported operators for the user to choose from. In fact, mobile devices already have such a list, which is used for manual network selection. An updated list is regularly distributed to device manufacturers, but it is not kept up to date by the device itself [13:p448]. While this is not a problem for the manual selection feature, which is rarely used and is in no way business critical for operators, it would be a major problem for new VSIM adopters who wouldn't be listed in existing devices.

b. The PLMN code could simply be typed into the terminal by the user. Five to six digits are not too difficult for users to remember, especially since the MCC part would quickly become familiar to users of a given country. This option is especially useful in the situation where subscriptions are sold in a store or over the internet (as is currently the case with all subscriptions), since the user could be issued the PLMN code through the

[14] These five to six digits also constitute the first part of the IMSI of any subscriber associated with the given operator, the remaining digits identifying the specific subscriber.

(existing) sales channel. In this situation, the PLMN number could be followed by additional digits, which could serve as proof of purchase of the subscription allowing it to be downloaded without further user interaction. This possibility will be described in the next section.

c. The terminal could scan for visible networks and the user be presented with a list of those supporting OTA provisioning. This could potentially make it very easy and convenient for users to get a new subscription, even if they are not close to a retail store. Considering that pre-paid subscription can often be bought in e.g. gas stations and super-markets, operators could make this process even easier for the user by allowing the subscription to be acquired right from the terminal. Note that compared to the other two options, this one does not permit subscriptions to be downloaded while the user is not under the desired operator's coverage area.

Note that these options are not mutually exclusive and a combination of them would give a versatile system. For example, users in their home country could find the desired operator by a network scan, but when abroad, they would enter the PLMN code or select the operator from a list.

One of the challenges of OTA provisioning is how to handle MVNOs. Pure MVNOs (see section 2.4.1.2) have their own distinct PLMN code, but a terminal scanning for networks (option c) would only detect operators who own the network infrastructure, since only their identity is broadcasted. This would result in incumbent MNOs having a competitive advantage over MVNOs, which may become a regulative problem in some countries. A way to handle this is for the MNO to broadcast a list of the PLMN identities of MVNOs which have agreements with the serving MNO. The mobile terminal would include this additional information in the list of found operators presented to the user. With such a system, operators could also have agreements to broadcast the identity of some international roaming partners as a service to their users or as publicity.

MVNOs that piggyback on incumbent operators' PLMN identity, won't even be addressable using the above scheme. To work around this, an extended PLMN code could be used to also address the MVNO served by an MNO. An extra digit or two in the entered code could be used to differentiate between users requesting a VSIM from the incumbent MNO and users requesting a VSIM from one of the serviced MVNOs. Other network entities or the terminals don't need to know the details of such a scheme, and it is entirely up to the serving operator to decide how additional digits in the code are interpreted. Such extra digits could also be included in the above mentioned list of served operators broadcasted by base stations, which would allow terminals to find non-pure MVNOs as well.

Advantages	Disadvantages
• Existing retail stores and processes need no change. • Low operating cost. • Convenient for users.	• Changes to the access network required (high acquisition cost). • Transmission security and device trust issues (see section 4.6.4).

4.6.2 Functional description of OTA provisioning

OTA provisioning seems to be one of the most versatile and powerful possibilities of the Virtual SIM and for this reason, a possible approach will be elaborated further in this section.

From a functional perspective, the scheme would work as such: Handsets would have a ‘Subscription’ menu with the structure shown in Figure 4.7.

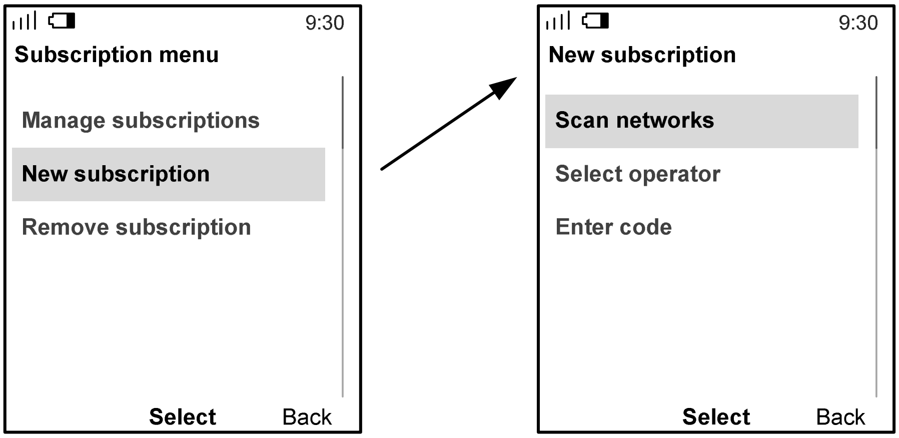


Figure 4.7: Concept: subscription menu structure. Source: own work.

The first and the last menu items (manage subscriptions and remove subscription) are used to manage subscriptions already in the terminal (if multiple subscriptions are supported). The ‘New subscription’ menu is the one related to provisioning. As can be seen, there are three possible ways for the user to acquire a subscription. He can either let the terminal scan for available networks, select an operator from a pre-programmed list on the terminal, or he can enter a code obtained from an operator.

If the user chooses to scan for local networks, the phone scans for network operators in the normal fashion, except that it also listens for an additional piece of information on the broadcast channel of found frequencies that indicates if an operator support virtual SIM transfers and what additional MVNOs (or roaming partners) it serves. Such a scan would take about 10-20 seconds[15]. The user is then presented with a list of found operators supporting VSIM transfers.

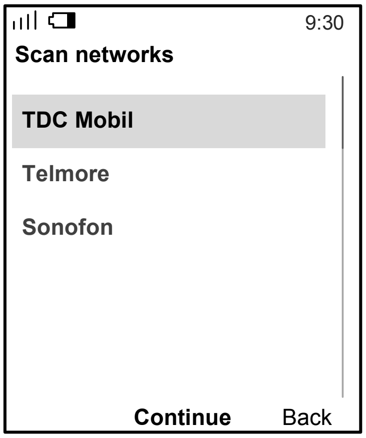


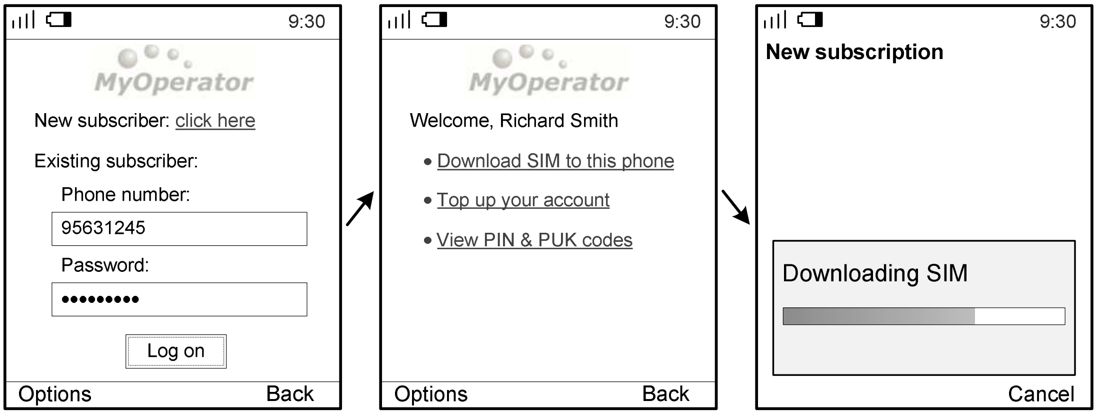
Figure 4.8: Concept: Result of network scan. Source: own work.

The mobile connects to the selected network using a *GPRS-Attach* message. Normally, this message is given the IMSI from the SIM card as a parameter, but in this case, the terminal’s

[15] Roughly the delay experienced when the ‘Manual network selection’ option is activated on a GSM phone.

IMEI will be used instead. While this violates the current standards (which need to be adapted), the signalling messages are actually designed to support taking the IMEI as a parameter instead of the IMSI: In the voice domain, this is used to establish an emergency call without a SIM card being present in the terminal. Assuming the network is designed to support this feature, this type of GPRS Attach will succeed and allow the terminal to establish a specific type of packet data context, which will be limited only to allow connections to a special 'subscription server'. This server could be addressed using a special operator-specific DNS name formed from the MCC and MNC codes, e.g. 'https://mss.01.238.plmn/'[16]; this name could then be resolved by the serving network to the desired address. The terminal would open a web browser to that address and establish a HTTPS connection to this subscription server.

The user now has a functional connection to the operator of his choice, and what happens from this point on is largely up to the individual operator. For instance, the user could be given the choice between buying a new subscription and downloading a VSIM for an existing subscription. If the first option is chosen, the user is taken to a page where he can enter his credit-card details, and so on. The second option could allow the user to enter a username and a password, taking him to a subscription management page. This page could allow him to download a new VSIM, and could potentially also serve other functions (note that the user would be able to access this page from any phone, with or without a subscription – some interesting services could result from this!)



New subscriber: click here

Existing subscriber:

Phone number: 95631245

Password: ●●●●●●●●●

Welcome, Richard Smith

- Download SIM to this phone
- Top up your account
- View PIN & PUK codes

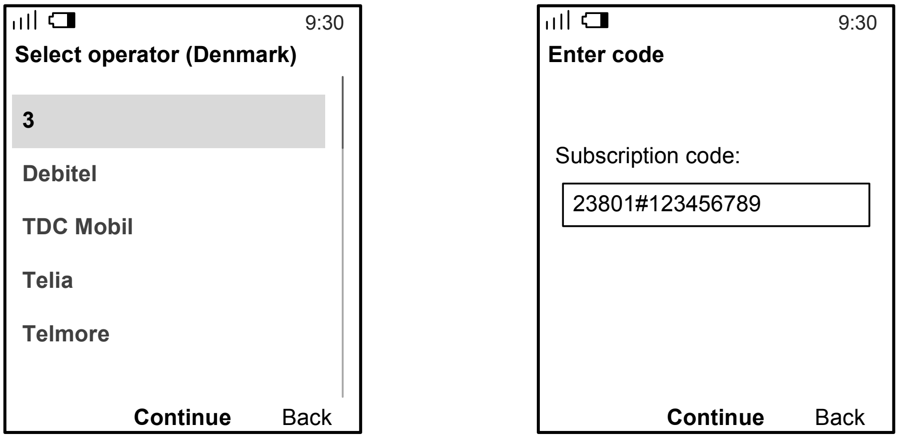
New subscription

Downloading SIM

Figure 4.9: Concept: pages served from an operator's subscription server. Left: the first page allows new subscribers to register, or existing subscribers to log in. Middle: example of what a welcome page for an existing subscriber could look like. Right: a VSIM is downloaded from the subscription server to the terminal. Source: own work.

The 'choose operator' option (in Figure 4.7) works in much the same way, but instead of scanning for networks, the terminal presents the user with a predefined list of operators, as illustrated in figure Figure 4.10 (left). Of course, this list will only contain the operators known to the terminal manufacturer during its production. This also makes it possible to select an operator while outside his coverage area or while roaming. As mentioned in the previous section, special roaming agreements would be required for this to work. If no such agreement was present, the connection would be rejected (or possibly the user would be redirected to the competitor's subscription server).

[16] The first part stands for 'mobile subscription service', the next part (01) is the MNC, and the following (238) is the MCC. The suffix, 'plmn' signifies that the domain name refers to a mobile (PLMN) network. This scheme more or less follows the convention used for reverse DNS lookup [130]. The example given here would refer to the subscription server of the Danish operator TDC Mobil.



Subscription code: 23801#123456789

Figure 4.10: Concept: Alternative methods of initializing the provisioning connection. Left: using a predefined list of operators (in this example, the user has chosen to view operators from Denmark in a previous screen). Right: using a subscription code provided by the operator. Source: own work.

If the user instead chooses the 'enter code' option, he is presented with the screen shown in Figure 4.10 (right), where a decimal code can be entered. As a minimum, this code consists of a number identifying the operator. This could conveniently be a concatenation of the Mobile Country Code and the Mobile Network Code (see section 2.2.1), which uniquely identifies every operator in the world. These two numbers are 5 or 6 digits long in total.

If the entered code doesn't contain other digits, this method works just as if the user has selected the specified operator from the predefined list mentioned above. However, the code can also be longer, in which case, the significance of the remaining digits is operator specific, and the number would automatically be passed to the operator's subscription server, e.g. as a HTTP header or as POST data. Operators could for instance use this facility when selling subscriptions in regular stores, where the subscription would simply consist of a code to be entered in the terminal, handed to the user on a slip of paper (wrapped in some operator-branded package, of course). The operator's subscription server would detect this and allow the user to download the VSIM directly and conveniently. Note that this 'paper slip' way of selling subscriptions is compatible with the current retail practices of the SIM card, and would also work in low-tech situations (i.e. emerging markets). Another use for the extra digits could be for promotional codes. Again, operators would be free to implement such schemes without interoperability issues, since the additional digits are interpreted by the operator's subscription server, and not by the terminal.

The use of additional digits raises the problem of how the terminal should determine whether the MCC + MNC part of the code consists of 5 or 6 digits. Indeed, while all MCC codes are 3 digits long, the length of the MNC depends on the MCC (and in theory it is possible to have both 2 and 3 digit MNC codes within one MCC, although the standard recommends against this [17]). The terminal needs to know which part of the code it should use to refer to the network. In the case of the IMSI, a special file on the SIM card is used to help the terminal determine the length of the MNC. To work around this problem, the user could be forced to enter the subscription code in two parts (PLMN code + additional digits) separated by a delimiter (e.g. the hash sign, '#') or in two distinct fields. The former is illustrated in Figure 4.10 (right).

This ends the functional description of the provisioning scheme. But before moving on, a cautionary note about security should be given. As described, this scheme is extremely vulnerable to man-in-the-middle (MITM) attacks. An attacker could set up a fake base station and either act as a rogue operator with bogus identifiers, hoping that users would search for operators and randomly choose him, or alternatively broadcast the MCC and the MNC of an existing operator in order to capture their current or prospective customers. The operators could then redirect users to their own 'subscription server', thereby performing a phishing attack. This could be used for a number of purposes: The attacker could trick the user into entering his credit card details, or if the user entered a 'paper slip code', this could be intercepted and abused by the attacker while presenting the user with an error message. A third option would be to intercept the username and password of subscribers logging on to their account using the method described previously. In fact, an attacker could carry out all of these attacks simultaneously! Obviously this is unacceptable, and this is where the DNS scheme proposed above comes in handy. Operators could be assigned SSL/TLS server certificates based on their specific DNS name (which is derived from the MCC and MNC values). One or more trusted certificate authorities could issue such certificates to legitimate operators. The certificate authorities' root keys would be placed in all produced phones, which could then check whether they were connecting to a legitimate operator, and whether the operator is indeed the one requested (in this way serving operator's wouldn't be able to perform MITM attacks). Note that the mutual authentication present in UMTS does not solve the mentioned problem, since no pre-shared key is present, so no UMTS authentication can take place.

4.6.3 Roaming agreements for VSIM transfer

As mentioned, agreements between operators regarding the VSIM are a prerequisite for being able to download VSIMs while roaming. There is a difference between this and normal roaming situations, though, since the subscribers cannot be authenticated before being granted access (to the subscription server), so in practice, everyone would be able to establish an expensive (international) connection to an operator. Since connection is only allowed to that specific server, the abuse-cases are limited, but malicious persons (or competitors) could use this to generate large amounts of international traffic to an operator that the targeted operator would have to pay for. This type of attack can, however, be stopped by blocking the attacker's EMEI.

4.6.4 Provisioning security

To meet the security requirements, the provisioning system must meet the following goals:

- The operator must be able to ascertain the trustworthiness of the device.
- If the operator decides to issue a VSIM to that device, it must not be used on another device.

The trustworthiness of the device can be determined in several ways. For example, Trusted Platform Modules have a feature called 'remote attestation' that can be used by remote parties to obtain a measure of the integrity of the platform [109]. This feature is very useful on platforms that are inherently untrusted, such as PC's or in the vSIM architecture proposed by Kasper et al. (section 3.6.1), but if the platform itself is tamper-resistant, as will be the case for the VSIM system proposed here, it is redundant[17]. However, the integrity of the platform is not enough. If the platform has been designed in an insecure way, either accidentally or by malice, integrity is not worth much. What is really needed is a way for the operator to identify the platform reliably, such that he has the ability to decide whether a specific platform can be

[17] While the whole mobile device cannot be considered a trusted platform, the security-critical functions of the VSIM will be managed entirely within a tamper-resistant unit.

trusted or not. In other words, the device must be authenticated by the operator. This is not to say that verifying the integrity of the platform, operating system, and other software is not useful. But how this is done can be left to the device manufacturer, and as long as the operator has a means reliably determine the identity of the device reliably, he will also be able to decide whether it provides adequate security.

4.6.4.1 Solution using public key cryptography

One way of achieving these goals is to use a public-key infrastructure, as is illustrated in Figure 4.11. The root of trust in this model is a central certificate authority (CA) recognised by all parties. For example, this role could be held by the GSM Association, or by one of the existing commercial certificate authorities (e.g., VeriSign). This CA issues certificates to all operators, which they can use to sign the VSIM files they generate, so that terminals can check that the files have not been compromised in some way during the transmission, and that the sender is a legitimate operator. The CA would also issue certificates to terminal vendors, which would in turn be used to certify device-specific public keys along with the device's IMEI. The corresponding private keys would be stored in the protected memory of the terminals, so that they could only be used inside their secure environment. When requesting a VSIM, the terminal would sign the request with its private key, and send its public key certificate along with it to the operator. The operator would then check the validity of the certificate, thereby obtaining a guarantee that if he encrypts a VSIM with the given public key, only the device with the IMEI specified in the certificate will have the corresponding private key. The operator is then able to make a decision as to whether the device should be trusted based on the received IMEI. If he decides the device is ok, the operator encrypts the VSIM with the public key, and signs it with his own private key. The VSIM can then only be decrypted inside the secure environment of the intended device, and the device can check the origin of the SIM, thereby closing the loop.

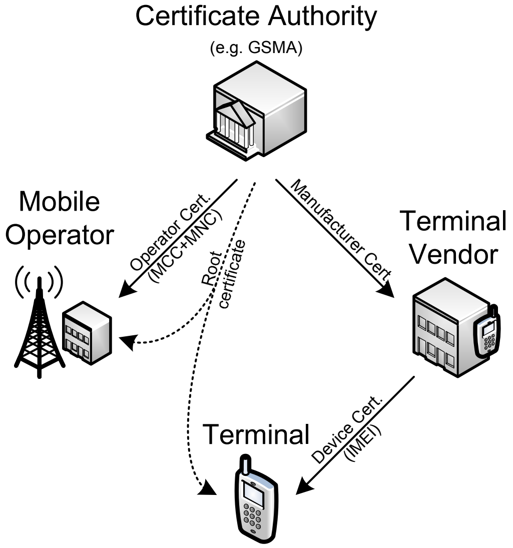


Figure 4.11: Public Key Infrastructure for provisioning security. Source: own work.

The same operator-certificate could potentially be used to establish the HTTPS connection for the provisioning process described in section 4.6.2, but it is probably a better idea to separate these domains. The terminal's private key should definitely not be used outside of the secure environment of the processor.

The downside of this approach is that the terminal and the operator need to exchange certificates, which places some overhead on the communication between them. With packet data connections, this isn't much of a problem; however, it makes it difficult to use the system using other low-bandwidth bearers, e.g. SMS. In emerging markets, GPRS etc. is not necessarily available.

4.6.4.2 Solution using identity-based cryptography

An alternative is to use identity-based cryptography (IBC), in which any identifying information can be used in lieu of a public key, thereby eliminating the need for certificates. Adi Shamir proposed the concept in 1985 [110], and demonstrated an ID-based scheme for digital signatures, but was only able to construct one for encryption. In 2001 the problem of identity-based encryption (IBE) was solved independently by Boneh and Franklin [111] and Cocks [112]. Subsequently a lot of research has been done on the subject (a 2004 survey can be found in [113]).

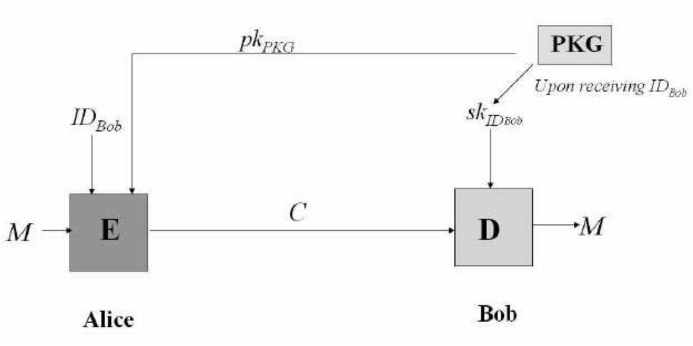


Figure 4.12: Identity-based encryption. Source: [113].

Figure 4.12 illustrates how IBE works. Alice wants to transmit an encrypted message to Bob. They use a trusted third party called the private key generator (PKG), who has a public key pair. The receiver identifies himself to the PKG who generates Bob's private key ($\mathrm{sk}_{\mathrm{IDbob}}$) as a function of the secret PKG key and Bob's identity ($\mathrm{ID}_{\mathrm{bob}}$). Using Bob's identity and the public PKG key, Alice can encrypt a message to Bob. Upon receiving the encrypted message, Bob uses his private key to decrypt it.

This scheme can also be extended to support a hierarchical PKG structure, as demonstrated in [114], where the responsibility to generate private keys can be delegated for a subset of identities.

Using hierarchical IBE (HIBE), terminal authentication and VSIM encryption could work like this: A top-level PKG, trusted by all parties, generates private keys for all terminal manufacturers. Their identity consists of an IMEI range (more than one key could be assigned to each manufacturer, e.g. one for each phone model). During production, the terminal manufacturer uses his private key to generate a device-specific private key for each terminal based on its IMEI, and store it securely in the terminal's secure environment. During normal use, when the terminal requests a VSIM, it sends along the IMEI to the operator (and nothing else); note that the user could also communicate the IMEI to the operator without the terminal being involved. Based on the received IMEI, the operator can make a decision about whether the VSIM should be issued. If yes, it is encrypted using a key derived from the public key data of the top-level PKG and the IMEI, and the encrypted VSIM is sent to the terminal. If the terminal lies to the operator about its IMEI, it won't be able to decrypt the received VSIM, since it does not have the required private key corresponding to the IMEI used in the VSIM encryption.

The advantage of this scheme is that the operator only needs to know about the terminal's IMEI in order to generate the VSIM. This means that a request for a VSIM could easily fit inside an SMS message or an USSD string, or could be communicated to the operator by the user himself (something which is not possible if a digital certificate has to be attached to the message).

An intrinsic disadvantage of identity-based encryption is that a secure channel is needed between the PKG and the receiver (Bob) in order to transfer his private key. For the proposed use, this is not a problem, since the private key would be transferred during terminal production, where a safe channel is assumed to be present. Another IBE concern is that the PKG functions as a key-escrow, meaning that this entity is able to recover the private keys it has generated at a later time (certificate-based cryptography doesn't have this problem since the certificate authority only needs the public key in order to generate a certificate). Again, in this case, it is not a problem, as the PKG is located in the terminal production facilities, and the terminal vendor is assumed to be trusted (this is the basic assumption for the whole Virtual SIM system).

The main disadvantage to this scheme is that identity-based cryptography is such a new concept, especially with the conservative approach usually taken when it comes to cryptography. This means that an IBC-based system is not likely to receive a lot of support. (On the other hand, Rijndael was adopted as the basis for the MILENAGE algorithm before it won the AES contest, so perhaps, the mobile phone industry is more open to 'innovative' cryptographic concepts than other sectors).

4.7 Summary

The Virtual SIM scheme proposed in this chapter largely relies on the reuse of existing technologies, to provide the necessary functionality. This reduces the cost of the system and the perceived risk taken by adopting it.

The complete terminal architecture for the Virtual SIM is shown in Figure 4.13.

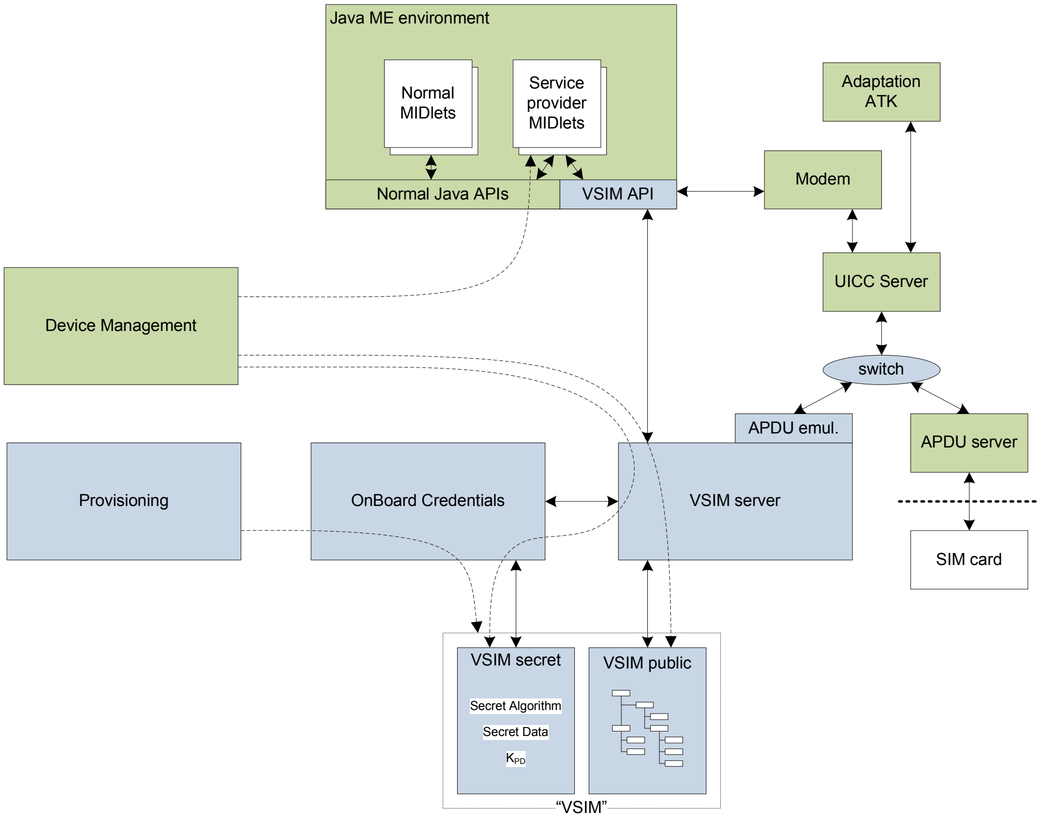


Figure 4.13: Overview the terminal architecture for the Virtual SIM. Source: own work.

Chapter 5

Feasibility Assessment

This chapter aims at giving an evaluation of the feasibility of the proposed Virtual SIM scheme. The technical feasibility has been demonstrated in the previous chapter, so the first section of this chapter will focus on whether the Virtual SIM system can realistically be implemented. Following this, the operational and economic feasibility will be addressed.

5.1 Technical feasibility

5.1.1 Availability of secure platforms

The Virtual SIM scheme assumes the availability of a secure platform in the mobile terminal. Specifically, this could be a secure element or a secure processor. While secure processors are at least used in some devices from Motorola (as they point out in [10]) and Nokia, they are far from being used universally. However, low-end phones also evolve, and since secure processors, when used, are the foundation for important functions such as platform integrity, prevention of IMEI spoofing, SIM lock, etc., they are likely to be adopted in these types of phones as well in a few years.

5.1.2 Availability of Java ME in terminals

Some of the arguments for using Java ME as platform for a SIM Toolkit (see section 4.4) were that it provides a modern, interoperable platform. But the most important factor is that it's already present in most devices, which reduces the cost and overhead of the Virtual SIM. However, current entry-level terminals do not support Java, and this could be a problem, especially if emerging markets are to be targeted. Figure 5.1 shows the forecasted evolution of mobile phone price segments. As can bee seen, the category of devices with basic features (such as Java) is beginning to replace the 'basic phone' category. In the long run, almost all phones will probably support this basic technology.

Finally, the absence of Java ME in a terminal does not imply that it will not be able to accommodate a Virtual SIM. It simply means that the terminal will not support extended operator applications, just as with current phones that do not support the SIM Toolkit (there are some).

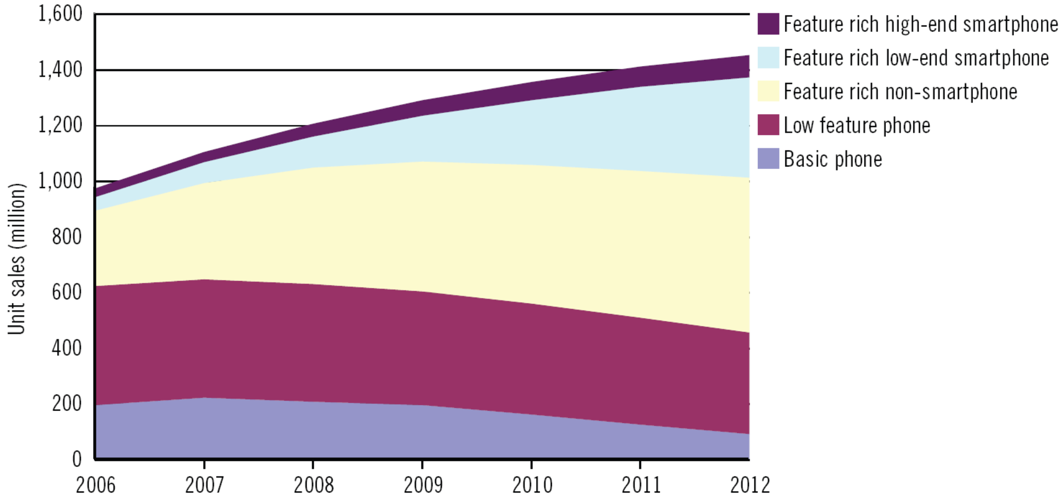


Figure 5.1: Global mobile handset sales by technology segmentation. Source: Informa Telecoms & Media [115].

5.1.3 Security assessment

According to Smith [116], “Risk in any context is the sum of threats (those events which cause harm), vulnerabilities (the openness of an enterprise to the threats) and asset value (the worth of the asset in danger). Increase any of these factors and the risk increases; decrease any, and the risk decreases”. Although this definition is formulated with enterprise strategy in mind, it is valid in other contexts as well. For any system, the security measures – which are put in place to reduce the vulnerabilities – should be adapted to the value of the assets they are designed to protect. Assuming the threats remain the same for a given scenario, a corollary to Smith’s axiom is that the goal of security measures is to increase the cost of exploiting vulnerabilities in a system above the value of the protected assets. In other words, attacks should be made unprofitable.

5.1.3.1 Effect of the Virtual SIM on fraud threats

The description of mobile fraud schemes given in section 3.5.1 shows that the SIM card plays a role in subscription cloning (because the SIM card must leak the secret key of the subscriber for cloning to be possible) as well as in subsidy fraud (because the fraud is made possible by removing the intended SIM card from a subsidized phone, and putting in another one). Figure 5.2 shows a threat diagram (following the guidelines of the CORAS framework for security analysis [117]) for these cases. The open padlocks represent potential vulnerabilities that could make the depicted threats possible. So how does the Virtual SIM affect these threats?

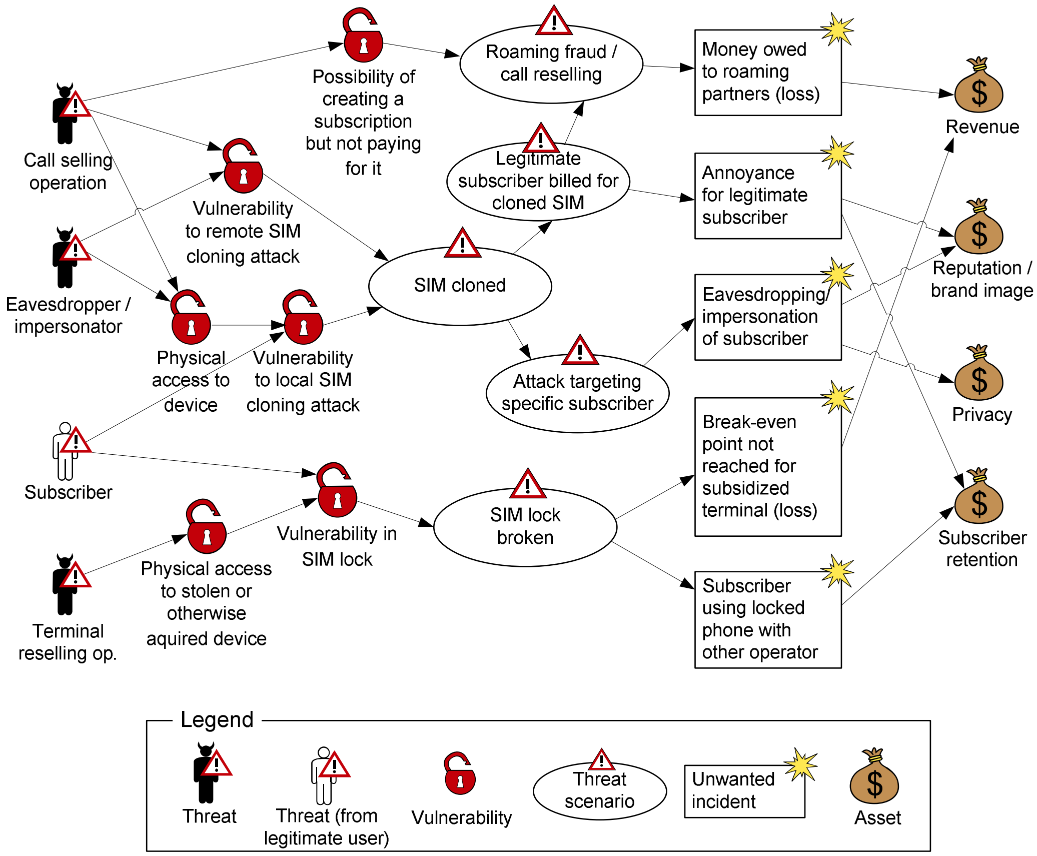

Call selling operation
Eavesdropper / impersonator
Subscriber
Terminal reselling op.
Possibility of creating a subscription but not paying for it
Vulnerability to remote SIM cloning attack
Physical access to device
Vulnerability to local SIM cloning attack
Physical access to stolen or otherwise aquired device
Vulnerability in SIM lock
Roaming fraud / call reselling
Legitimate subscriber billed for cloned SIM
SIM cloned
Attack targeting specific subscriber
SIM lock broken
Money owed to roaming partners (loss)
Annoyance for legitimate subscriber
Eavesdropping/ impersonation of subscriber
Break-even point not reached for subsidized terminal (loss)
Subscriber using locked phone with other operator
Revenue
Reputation / brand image
Privacy
Subscriber retention
Legend
Threat
Threat (from legitimate user)
Vulnerability
Threat scenario
Unwanted incident
Asset

Figure 5.2: Threat diagram for SIM-related threats based on pictograms described in the CORAS framework for security analysis [117]. The pictograms for threats (represented by persons) are not used entirely according to the guidelines of the framework, since the white person is supposed to represent an 'accidental threat' in opposition to a 'deliberate threat' depicted by the black person. In our case, the two different pictograms are used for deliberate threats to distinguish between legitimate users attacking the system for convenience (white) and attackers abusing the system for profit (black). Source: own work.

First of all, the responsibility for protecting the subscription key moves from the SIM card to the terminal. Due to the networked nature of mobile phones, both remote and local vulnerabilities could potentially serve as attack vectors to compromise the secrecy of this key. There is a big difference between these two cases in terms of the consequences of vulnerabilities. Remote attacks are much more critical, since they are usually low-risk and can be performed on a large scale. They are also very unlikely, since vulnerabilities need to be present in several distinct protection layers (application, operating system and secure mode) in order for them to be exploitable for the purpose of cloning the SIM. Local attacks require physical access to the terminal, which means that the terminal must first be stolen or borrowed (the attack could be carried out by e.g. a repair technician or a 'friend'), or the attacker must be the owner of the terminal, i.e., the user. If the terminal is stolen, the user will notice and block his subscription. If the key is extracted while the terminal is in someone else's possession, it will usually be possible to determine who the culprit is, either by the user or during a criminal investigation. If it is the user himself who extracts the key, he could potentially use his subscription on more than one device, but he would be paying for the subscription usage himself, so there is no economic loss for the operator. In either case, cellular networks are centrally monitored by fraud-detection systems, so any attempt to use a subscription from multiple terminals simultaneously is likely to be detected quickly, and the subscription will be blocked in response. With the proposed Virtual SIM system, the subscriber key is protected by the secure environment of the terminal. As with any such system, this protection can be circumvented in the presence of design flaws or software bugs in the code running in the secure environment, but assuming a well-designed and well-tested system, attackers will have to revert to attacking the secure environment itself[18]. Ideally, compromising the security in this way would involve invasive attacks which are likely to be too expensive compared to the monetary gain that can be obtained through them. In reality, some of the non-invasive attacks described in section 2.3.2, such as side-channel attacks might be applicable, unless secure processors (and software running within them) are adequately protected against this – which is not currently the case, at least in commercial products. Although this is probably not a showstopper for the Virtual SIM, it is an area that should be the target of further investigations and improvements.

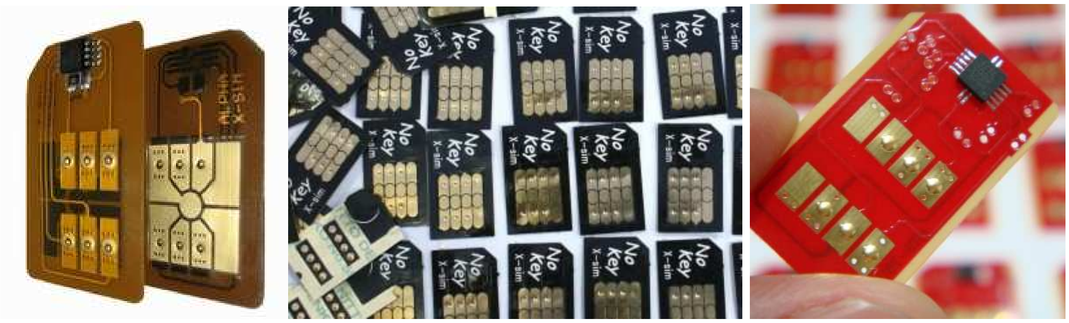


Figure 5.3: Examples of commercial SIM-Lock bypassing devices designed to fool the terminal into thinking the SIM has a valid IMSI. These 'X-SIM' devices are placed between the SIM card and the terminal's SIM connector intercepting and altering the communications. Source: internet bulletin board (GSM Forum).

Currently, subsidy fraud is prevented by the use of the SIM lock which is implemented by the terminal. The Virtual SIM doesn't change this. However, the absence of a physical module that can be moved makes it more difficult to cheat the protection mechanisms put in place by the terminal. For instance, a recent attack against the SIM lock involves intercepting the communication between the SIM card and the terminal, in order to make the terminal believe that the SIM card has an IMSI that passes the SIM lock check. It utilizes the fact that the terminal doesn't send the IMSI to the network if the TMSI stored on the card can be used. Commercially, this concept is sold as 'X-SIM' devices, which take the form of microcontroller-equipped 'stickers' placed between the (unauthorized) SIM card and the SIM connector in the terminal, see e.g. [118]. Figure 5.3 shows a few examples of commercially available X-SIMs. This problem could be avoided if the SIM Lock protection is integrated with the Virtual SIM security mechanisms.

5.1.3.2 Security of public-side VSIM functions

In the proposed design, SIM file handling and the equivalent of toolkit applications would be running outside of the secure execution environment, as opposed to the current practice where these functions are protected by the SIM card. Instead, these functions rely on the security of the terminal's operating system, and in the case of toolkit application on the Java ME implementation. Thus, they are not suited to handle security-critical functions, which might at first sight seem like a significant issue. However, the question is: what really constitutes security-critical functions? Subscriber authentication is one, but this is handled on the secure

[18] The SIM-lock mechanism on recent Nokia phones is based on such a secure processor, and the few known attacks to this system are due to design flaws (which have subsequently been corrected), and not to the secure environment being compromised.

side by the ObC platform. In the case of SIM files used by the terminal, their integrity is important to the correct functioning of the mobile service, but if they are compromised this would not result in a critical breach of security resulting in monetary loss for the operator. The closest one comes to this is if the PLMN list is modified to select sub-optimal roaming networks. In fact, modifying SIM files is currently straightforward using man-in-the-middle attacks between the SIM card and the terminal, as described in the previous section. The Virtual SIM would, in fact, provide a better protection against manipulation of SIM files. If in the future a secure communication channel is introduced between the SIM card and the terminal, in the end, it is still the terminal that uses the content of the files, and this is just as susceptible to a security break of the terminal as would be the case for the Virtual SIM.

This observation also extends to toolkit applications. The current communication channel between the SIM card and the terminal is in no way secure, so SIM Toolkit commands cannot be relied on for security-critical applications. Take for instance the previously mentioned ZuHause service, where subscribers get flat-rate telephony within a home area. This could be implemented by letting the SIM card determine whether or not the terminal is inside the home area using network measurements performed by the terminal, and then send the results back to the network. However, this approach would be vulnerable to an attacker faking the communication of network measurements, e.g. to fool the network into thinking the terminal is always within the home area. Instead, such functionality needs to be implemented in the network is any reasonable level of confidence is needed. On the other hand, SIM measurements could be used to indicate to the user whether or not he is in the home zone – this is not a security critical function.

In the end, all functions of the SIM card rely on the correct operation of the terminal, even authentication – but in this case the security requirement is that the key and algorithm must not be disclosed, which is possible with the SIM card as well as the VSIM. For supplementary applications, true security-critical function could be implemented as credential programs interfacing with a Java ME application. Thus, the differences between the SIM card and the Virtual SIM regarding what runs in a tamper-resistant environment have no effect on security.

5.1.3.3 Security of protected-side VSIM functions

The network authentication functions running in the secure execution environment are protected against software flaws in phone applications as well as in the operating system. However, vulnerabilities in code running in secure mode could potentially expose the VSIM secrets. Thus, it is primordial that such code is free of bugs. Fortunately, there is only a limited amount of code running with this privilege level. Thus, through testing and possibly certification is not unrealistic.

The secure processor also provides protection against basic physical attacks. High-budget attacks are still feasible, but not profitable. The biggest threat lies with the possibility of non-invasive side-channel attacks. If such attacks were possible, the VSIM could be compromised (i.e., cloned) at a low cost. There are many protection mechanisms against such attacks, as described in section 2.3.2, and many are low-cost software protections. A thorough investigation of the possibility of such attacks is outside the scope of this report, and should be performed on a case-by-case basis. It is, however, highly recommended that terminal manufacturers conduct such analyses and implement the required countermeasures.

5.2 Operational feasibility

5.2.1 Compatibility with current practices

Naturally, the SIM logistics and processes would change with the introduction of the Virtual SIM, since physical SIM card would no longer need to be distributed. In this case, the change is an advantage due to cost saving. Of course, the VSIM would still need to be “ordered” either automatically in the case of OTA subscription purchases or manually in the case of paper-slip codes. In the latter case, the process of managing stock and distribution could remain the same as today, except that the associated logistics would be simplified since the whole process of ordering, personalizing and packaging SIM card would be avoided – only the printed material needs to be distributed to retail stores.

Another important point is, that IMSIs only need to be allocated to subscribers the moment they download the subscription to the phone (for the first time). This has two advantages: First, IMSIs are sometimes a scarce resource, and having piles of unused IMSIs waiting in retail stored is suboptimal. Second, many operators pay their HLR vendor (partly) based on the number of entries in the database. Thus, money can be saved by delaying the creation of the subscription till the very last moment.

The proposed OTA provisioning is compatible with all the distribution methods described in section 3.2.2: Prepaid subscription can be acquired in the same manner as today by substituting the physical SIM card by a subscription code provided by the operator (and possibly printed and packaged for branding purposes). This method is also valid for subscriptions sold by street vendors, kiosks, gas stations, super-markets, over the internet, etc. Subsidized terminals can be locked to an operator in the same way they are today. In this case, the subscription can either be installed by operator’s staff, by the end-user (using the paper-slip method), or potentially by the terminal manufacturer. In the latter case, either the VSIM itself could be flashed to the terminal during the labelling/personalization phase or alternatively, a subscription code could be embedded in the device, which would automatically be used to download a subscription the first time the phone is started. The last option required the least coordination between terminal vendors and operators, since the terminal vendor could simply generate random subscription codes himself and then send the list of valid codes to the operator along with the terminals. This also has the advantage of delayed IMSI allocation described above.

The method for OTA purchasing of subscriptions is not comparable to any other retail channel available today. However, there is no clash in processes due to the highly autonomous functioning of this method. At most, support personnel would need to be trained in the system.

On the terminal manufacturer’s side, the Virtual SIM affects the production in one major way: The need to generate per-device asymmetric keys. This is already being done for devices supporting DRM, so most production facilities will support this or could be extended to support this with the existing expertise from the DRM case. This does however require a lot of CPU (at least in the case of RSA), and the process would approximately take 2.5 seconds per phone. Thus, manufacturing sites would probably need to get upgrades in terms of key-generating computing power.

5.2.2 User experience

As mentioned in section 3.1.2, the characteristics deemed to influence the adoption by end users are usefulness and ease of use. In a few cases, the Virtual SIM arguably decreases the easy of use compared to the SIM card. This is for instance the case when the SIM card is used to transfer contacts and messages from one terminal to another. With the Virtual SIM, this is not inherently supported. However, in the vast majority of cases, the Virtual SIM improves both the usefulness and the ease of use of the SIM system. This is demonstrated in Appendix B.1, which compares how different use cases would work with the SIM card and the Virtual SIM.

5.3 Economic feasibility

This section will evaluate whether the concept of the Virtual SIM is economically sound for terminal manufacturers and mobile operators.

5.3.1 Terminal manufacturers

5.3.1.1 Bill of materials

Once past the transition period with hybrid SIM card / Virtual SIM implementations, it is assumed that support for physical SIM cards is completely removed from mobile terminals. This will result in savings on the bill of materials (BOM): The SIM connector currently amounts to around 0.10–0.15 euros on the BOM of terminals plus potentially some additional discrete components. Appendix B.4.2 presents a set of calculations of the impact of this cost saving on the total BOM as well as on the profit margin of terminal vendors, based on cost estimates for terminals in five different cost categories for the years 2006–2012. These calculations show that for high-end phones, the cost savings are negligible but for the two lowest categories, ‘basic’ and ‘low feature’, the SIM-related components amount to 0.93% and 0.55% of the BOM, respectively, in year 2012. In these cases, removing support for the physical SIM card would result in increases of the margin of device vendors of 5.65% and 3.41%, according to the estimates.

In Nokia, secure processors are expected to become part of all classes of terminals at some point in the future. Thus, the Virtual SIM would not increase the cost of the terminal.

5.3.1.2 Power consumption

SIM cards consume only a small fraction of the total power consumption of terminal when they are active. When idle, terminals enter a sleep mode where the used current is only a few milliamperes, and SIM cards can also be put into sleep. However, the SIM Toolkit enables SIM to issue pro-active commands, which are implemented by letting the terminal to poll the SIM card at regular intervals. The SIM card can decide the duration of this interval or even disable polling entirely, however many SIM cards do not alter the default polling interval. In this case, Nokia phones poll cards for commands approximately every 25 seconds – even when idle. This requires terminals to power up while communicating with the card, which brings the current consumption from a few milliamperes to a value ranging from tens to over a hundred milliamperes during this interval. Even if the duration of this is only a fraction of a second, the high polling frequency means that this negatively affects the standby time of terminals. Detailed power measurements made on the Nokia 6500 (see Appendix B.3) have shown that avoiding this polling would increase the standby time by approximately 3.75%. For more advanced terminals, such as smartphones, consuming more power, estimates have indicated that around 10% extra standby time could be gained, which could be achieved by using the Virtual SIM.

5.3.1.3 Warranty costs

The SIM card relies on a mechanical connector in the phone, which is naturally prone to wear and failure. Furthermore, the electrical signals to and from the SIM card are filtered, which is typically done using passive electronics integrated in a single component, a part which is also prone to failure. A breakdown of the total expenses incurred by Nokia in 2007 due to warranty repairs related to the SIM card can be found in Appendix B.2. These numbers do not include all costs, but the total cost of SIM-related warranty repairs is estimated to be at least 2.5 million euros.

5.3.2 Mobile operators

The OTA provisioning scheme requires modifications to the network, as described in section 4.6.1. Both the access and the core networks only requires software updated to existing components in order to support this, which is not too expensive (it was not possible to get an estimate of this cost). Additionally, some central servers are required for VSIM provisioning. These items are one-time acquisition costs, that should be compensated for by cost saving / additional revenues generated by the use of the Virtual SIM.

Operators have a metric for the cost of getting a new subscriber called the subscriber acquisition cost (SAC), which includes the cost of the SIM card, as well as the marketing cost, the cost of subsidising a handset, etc. SIM card can cost anywhere between a half and tens of euros. Operators with highly vertically integrated value chains usually focus on post-paid subscribers which generate high average revenue per user (ARPU). They also invest a larger SAC in capturing such users, partly because they will tend to use more expensive SIM cards that can support their service offerings, but certainly also because subsidised mobile handsets are usually offered with new subscriptions, which constitutes a large part of the SAC. In this case, the coast of the SIM card is much less significant due to the overall high SAC. This stands in contrast to smaller operators with a high percentage of prepaid subscribers. These will typically seek to minimize the cost of the SIM, due to the higher churn rate and the lower ARPU. Also MVNOs usually have fewer expenses than incumbent operators, and in some cases, e.g., subscriptions sold over the internet, the SIM card constitutes a comparatively large part of the SAC. In either case the Virtual SIM results in cost savings, and is worth the initial investment.

The overall average cost of the SIM card is around 1 euro, and on top of this comes the cost of packaging and shipping, which also amounts to approximately 1 euro [25] (see also Appendix B.4.1). Worldwide, this translates to revenues of approximately 2.2 billion euros [3] that can be saved by operators on a yearly basis by using the Virtual SIM instead of the SIM card.

Chapter 6

Discussion

6.1 Extended VSIM use cases

The Virtual SIM opens the door to some interesting use cases that are not possible with the physical SIM.

6.1.1 Multiple subscriptions per device

The Virtual SIM makes it possible to have more than one subscription on the phone without the extra hardware needed to accommodate multiple SIM cards. In theory, more than one subscription could be active at the same time, but in practice, this would require the radio components of the phone to be duplicated, which would increase the cost and power consumption of the handset. Another approach is to have a system where calls to inactive subscription are automatically forwarded to the active one. Call forwarding is supported by all mobile networks, but unfortunately, this is not the case for SMS messages. This problem can be addressed in three ways. Either it is accepted that an 'active subscription' is the only one that can receive text messages. However, delivery of text messages times out after typically four days, and the user would run the risk of losing important messages if the subscription is not activated regularly. Another approach would be to cycle between subscriptions regularly in order to download SMS messages. This is feasible, but could result in the desired subscription being unavailable when needed by the user. This approach could also be implemented in such a way that the terminal would only check for messages for inactive subscriptions when requested by the user. The last approach consists of having some intelligence in the network that would known which subscriptions were present on a phone and which one was active, and redirect messages or at least notifications of awaiting messages to the active subscription. For such a SMS delivery system to work between subscriptions from different operators, additional changes are required to the current network signalling protocols and inter-operator agreements, and this is probably not realistic in practice.

Multiple subscriptions are of interest, since the user could change between them at will, and for instance easily change to a local operator when in a foreign country to avoid paying roaming charges. There is no theoretical limit to the number of VSIMs the terminal could contain, and if this concept is taken a step further, the phone could detect which subscription was the cheapest under a given set of conditions, and either make suggestions to the user as to which operator to choose, or simply select the cheapest one automatically.

For obvious reasons, operators in developed markets are not keen on this kind of concept. First, it would make it more convenient for users to avoid international roaming, thereby damaging operators' roaming revenues, which constitute a large part of their total revenue stream (in 2006 between 10% and 18% according to a study by Evalueserve [119], while Informa estimates this to be more than 26%! [61]). Secondly, it would probably increase churn, since the user is less bound to one operator. It would also shift the role of operators into 'connectivity providers' or 'bit pipes', and while some may argue that this is unavoidable (see section 2.4.1.2), it is certainly not something operators would want to encourage. Since the Virtual SIM would put this kind of system within reach, operators are likely to perceive this as a major threat. There is not much one can do about this, since the Virtual SIM inherently removes the physical constraint to having multiple SIMs (which is not the same as saying it will necessarily happen in practice if the VISM is introduced). As any change, this could also be an opportunity for operators who adapt and embrace the new competition model.

Another opportunity is to make it possible to have separate 'personal' and 'work' subscriptions on the same terminal – without the hassle of switching SIM cards.

Another variant of the multiple subscription use case is where the concept is used in emerging markets, where not everyone can afford to own a personal mobile phone. In this case, a group of people could share a phone, but have their own VSIM stored on it, thus saving the hassle to change SIM cards.

6.1.2 Multiple device ownership

A different possibility that the Virtual SIM facilitates is multiple device ownership, which is when one person actively uses more than one mobile phone. For example, one could have a work phone and a personal phone, or fashion aware people might have different phones to match different styles. Of course, this is possible today by switching SIM cards, but that is not something most people would want to do on a daily basis. Theoretically, it would be possible to issue more than one SIM card for the same subscription, but this is not commonly seen, probably due to the cost of issuing and managing more than one SIM per subscription. The Virtual SIM could make this service possible at a low cost. This use case is especially interesting for terminal vendors, since it would increase the demand for their products.

6.1.3 Device counterfeiting prevention

Counterfeiting of mobile terminals is becoming an increasing problem for terminal vendors, as the quality of the replicas is becoming so high that average consumers could confuse them with genuine terminals. For example high-quality Chinese counterfeits of the Nokia N93i [120] and N95 [121] phones have been spotted on the black market.

Due to the proposed provisioning scheme, the Virtual SIM could mitigate this problem: If handset manufacturers need to be accredited in order to be part of the provisioning PKI (or HIBE infrastructure), counterfeit terminals would simply not be able to download a VSIM subscription, since the device needs to be authenticated in the process. Even if a counterfeiter somehow managed to sign terminals, either using a certified signing key or by extracting the private key from an original device and using it to create counterfeit units with identical IMEIs, this operation could easily be stopped in cooperation with operators. Either by blocking a specific IMEI or by blocking all terminals signed with a compromised key. This would even mean that terminal replicas already in the field would be unable to download new subscriptions.

6.2 Authentication for other services

6.2.1 Ad-hoc network authentication

Concurrently with traditional PLMN networks, other mobile access technologies are emerging. For instance, WLANs have become very common for home use and for public hot spots, and WiMAX is emerging as a possible alternative to UMTS. In the future, the number of ways a mobile terminal can get network connectivity is likely to increase [45, 122, 46], and there is a possibility that network access will be established in an ad-hoc fashion. An ad-hoc mobile network is a transient network formed dynamically by a collection of wireless mobile nodes that doesn't use existing network infrastructure or a centralized organization [47].

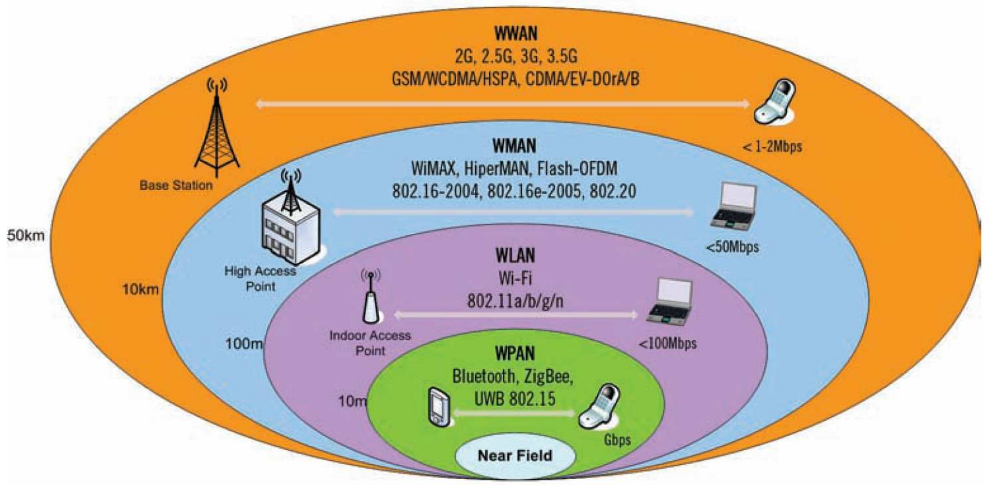


Figure 6.1: Mobile and wireless broadband standards. Source: Informa Telecoms & Media [115].

The Virtual SIM is flexible enough to deal with different types of network authentication, different authentication credentials, and different service providers without the necessity for them to have a prior business agreement, as would be the case if they needed to share a SIM card.

6.2.2 User-centric identity management

The term 'identity management' covers a broad range of meanings, as described in [18]: On one end of the spectrum it denotes various schemes to manage a diversity of identities that pile up for a single user. This type of identity management helps users get their work done without having to deal with too many identifiers and corresponding credentials, e.g. passwords. This is sometimes known as single sign-on (SSO) or unified user management. On the other end of the spectrum, there are approaches that aim at protecting users' privacy and reduce the danger of usage profiles. This type of identity management helps users maintain a set of different identities or pseudonyms that they can use in different contexts.

In the identity management literature, a real-world person or organization is known as an entity. An entity may have several identities, each of which is unique within an 'identity domain'. These identities are referred to using identifiers (commonly there is a one to one relationship between identifiers and identities, in which case the separation between these terms is somewhat blurred in common use). Figure 6.2 illustrates these correspondences.

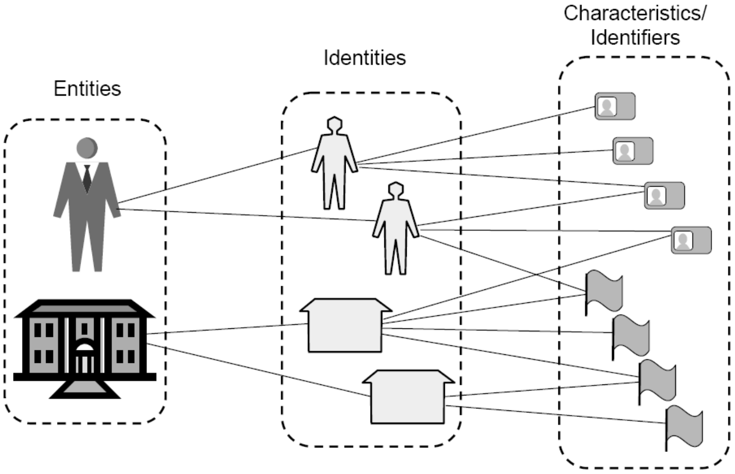


Figure 6.2: Correspondence between entities, identities and characteristics/identifiers. Source: [123].

There are several identity management models in common use [123]:

Isolated user identity: The service provider issues both identifiers and credentials to users. In this case, a user gets a separate unique identifier from each service provider he transacts with. Many registration-based internet portals are examples of this.

Federated user identity: In this case, a group of service providers establishes multilateral agreements, which will enable service providers within the group to recognize user identifiers and entitlements from the other providers. Such a group is known as a 'federated domain'. The user still holds separate identifiers for each service provider, but does not necessarily need to know or possess them all, as a single identifier and credentials is sufficient to access all service providers in the federated domain. There are several standards for identity federation, including the OASIS Security Assertion Markup Language (SAML) [124] and the Liberty Alliance framework [125].

Centralised user identity: In this model, a single identifier and credential provider is used by all service providers, either exclusively or in addition to other identifier and credentials providers. This type of model can be implemented in several different ways, e.g. the common identifier model, the meta-identifier model, and the single sign-on (SSO) model. See [123] for more details on these.

Jøsang and Pope [123] argue that while the motivation for federated and centralised identity management models is to simplify the user experience, there will never be a single identity domain for all service providers. Instead, they propose a system where a personal authentication device (PAD) is used to store the various identities and credentials of the user, as illustrated in Figure 6.3. The user authenticates himself towards the PAD using a PIN code, and the PAD can in turn authenticate the user towards various services providers. This can be seen as a form of virtual single sign-on, which also has the advantage that legacy identity management models (e.g. username/password authentication for web services) can remain unchanged, i.e. legacy services don't need to be aware of the fact that a PAD is being used for authentication. Such a scheme falls under the category of 'user-centric identity management' [123, 126].

Security-wise, this has another advantage, since an identity stored in the PAD can also hold information identifying the service provider it belongs to. This would solve some of the problems that exist with e.g. HTTPS, where it is often the user's responsibility to check the identity of the service provider. "Not everybody understands what are the consequences [*sic*] of giving the "Yes" answer to the certificate related questions during HTTPS initialization" [104], and not everybody checks if the URL displayed in the location bar really belongs to the expected service provider. This flaw in usability is a big threat to the security of the protocol, since users often neglect to do this, which is one of the reasons why phishing attacks are so common. A PAD could automatically refuse to authenticate the user if the entity requesting a credential is not the service provider that issued it.

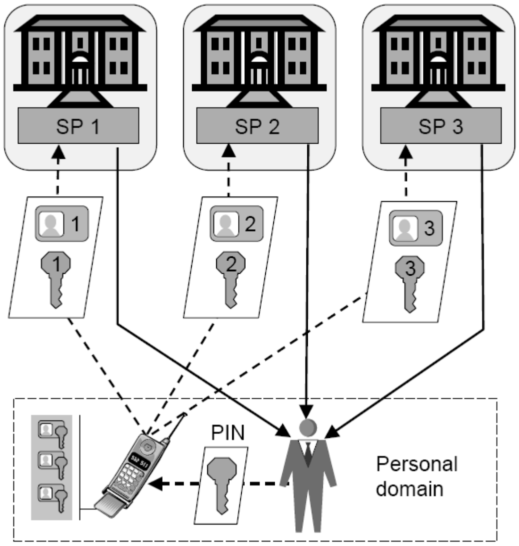


Figure 6.3: User-centric identity model. Source: [123].

The Virtual SIM system described in this report can be extended to handle other types of service providers than mobile operators, thereby functioning as a PAD. If the Virtual SIM system were to fill this role, even more powerful services could be provided, since service providers could potentially be allowed to issue specialized credentials that could include some logic (in LUA script) running inside the secure processor

6.2.3 Third-party secure services

Currently, third parties wishing to create mobile services requiring some level of protection against physical attacks (e.g. ticketing, e-purse, loyalty programs), need to make a business agreement with operators in order to be able to leverage its security features. A Java API, called the Security and Trust Services API [127] (JSR-177), enables Java applications to communication with a secure element, such as the SIM card, which could enable such applications. Indeed, service providers could make an agreement with a mobile operator to place a specialized application on the SIM card that would provide the security for a certain application. This approach has two drawbacks. First, the service provider must enter in a business relation with an operator, which can complicate the revenue model for the service. Second, the distribution of the service is limited to users having a subscription with supported operators. In effect, this means that such services can only be advertised as part of an offering from an operator.

The Virtual SIM scheme could be extended to handle the generic case of independent service providers taking advantage of the secure environment in the terminal to provide advanced secure services. Such services could be distributed to any user with a supporting mobile terminal, regardless of the used operator or the brand of his terminal.

Such a feature is essentially possible using simply the OnBoard Credentials framework. However, some of the elements proposed for the Virtual SIM could facilitate the creation of rich services. For instance the interface between the credentials program and Java ME used by the VSIM could be extended to allow interaction between non-SIM credentials and MIDlets in the third-party domain associated with the credential.

6.3 Strategic considerations for Nokia

As mentioned in the introduction, new features of SIM cards, such as high capacity, USB interface, web services, NFC, etc., threaten to increase the competitiveness of services delivered by the SIM platform compared to those present on the terminal. Sometimes, such services are in direct competition, such as with personal information management (PIM), such as the phonebook. In other cases, the SIM is used as a vehicle to promote the brand of the operator, which inevitably diminishes the brand-presence of the terminal vendor. In fact, SIM vendors and operators have a common strategic advantage in taking value (services, brand, etc.) out of the terminal [3]: Operators have an interest in creating operator loyalty instead of device loyalty. For SIM vendors, the more responsibility and value is held by the SIM, the more expensive SIM cards and SIM applications can be sold to operators.

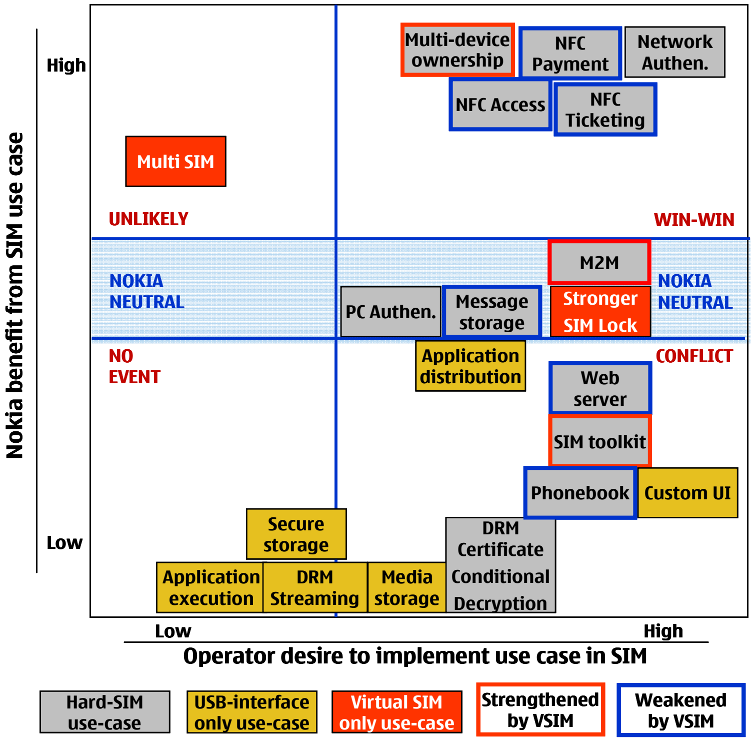


Figure 6.4: Map of SIM use cases according to operator wishes and Nokia benefit. Source: SIM card use cases and their position in the chart have been obtained from a Nokia strategy presentation [3]; Colour classification (USB-only/strengthened/weakened) and positioning of the use cases 'Multi SIM', 'Multi-device ownership', 'Stronger SIM Lock' and 'M2M': own work.

Seen from a strategic point of view focused on Nokia, the Virtual SIM can be seen as a way to stop this development of the SIM card, and restore its former role as a simple network authentication device. Even though the proposed Virtual SIM scheme allows operators to create supplementary applications in Java ME, this is nothing compared to the potential use cases that could arise as a consequence of the new SIM technologies.

Figure 6.4 shows a range of SIM card use cases that have been identified in [3]. They are positioned according to their benefit for operators and Nokia: the further to the bottom they are placed, the more detrimental they are to Nokia. The use cases have been categorized according to which ones are invalidated by the VSIM (yellow, and to a lesser extent blue) and which ones are facilitated (red). In general, the Virtual SIM prevents many of the undesirable use cases (yellow).

The Virtual SIM also redefines the way operators influence the technological development of terminals. Often, the current modus operandi is that operators and smart-card vendors have enough votes in the standardization bodies to push through new SIM-centric features without the approval of terminal vendors, and vendors are subsequently forced to implement the features due to the bargaining power of the operators. The Virtual SIM removes SIM vendors from this game, which means that new technologies can be introduced based on a closer cooperation between operators and terminal vendors. Operators will not lose power due to this shift, as their bargaining power remains intact, but since new technologies are not influenced by the interests of SIM vendors, they are more likely to result in win-win scenarios for operators and terminal vendors.

6.4 Limitations

The use of interviews for gathering information has some limitations:

- Using interviews increases the probability of studying perceptions rather than facts.
- Different people within the same organizations have radically different opinions.
- It is difficult to determine whether interviewees tell ‘the whole story’ or whether they are withholding information due to strategy/confidentiality issues — potentially they could have a hidden agenda and could be manipulating the interviews to put things under a certain light, although this is rather unlikely, and probably hasn’t been the case.

Thus, it is very important for the interviewer to correctly construe the information obtained based on the background of the context, the person being interviewed, the tone of his voice, and other indicators. This necessary interpretation step also means that the obtained information is coloured by the interviewer himself.

A possible way to reduce the subjectivity would be to use techniques involving many subjects. The Delphi method is one such technique [128, 129]. This method aims at obtaining forecasts from a group of experts. The participants first answer a questionnaire and include comments to the answers. A moderator collects the responses, merges them and filters out irrelevant information. The result is given back to the experts for further comments. The experts are allowed to modify their earlier statements and responses at any time. This process is repeated a number of times, until consensus is reached. The experts remain anonymous during the whole process.

This Delphi method could be used to collect information in a way that is much less influenced by personal biases, spontaneous ideas, etc. The main disadvantage is that it is time-consuming and that a suitable group of experts must be found that are willing to sacrifice the required time.

6.5 Future work

The study presented in this report has focused on shaping the abstract idea of a Virtual SIM into a well-defined concept with a feasible overall design. Before the Virtual SIM becomes a reality, stakeholders need to be convinced that the concept is sound. A proof of concept is an effective method of proving that a concept is technically feasible and demonstrating its advantages, which is the first step on the way. Given that enough organizations support the concept, all stakeholders need to agree on its exact form and create the necessary standard.

6.5.1 Proof of concept

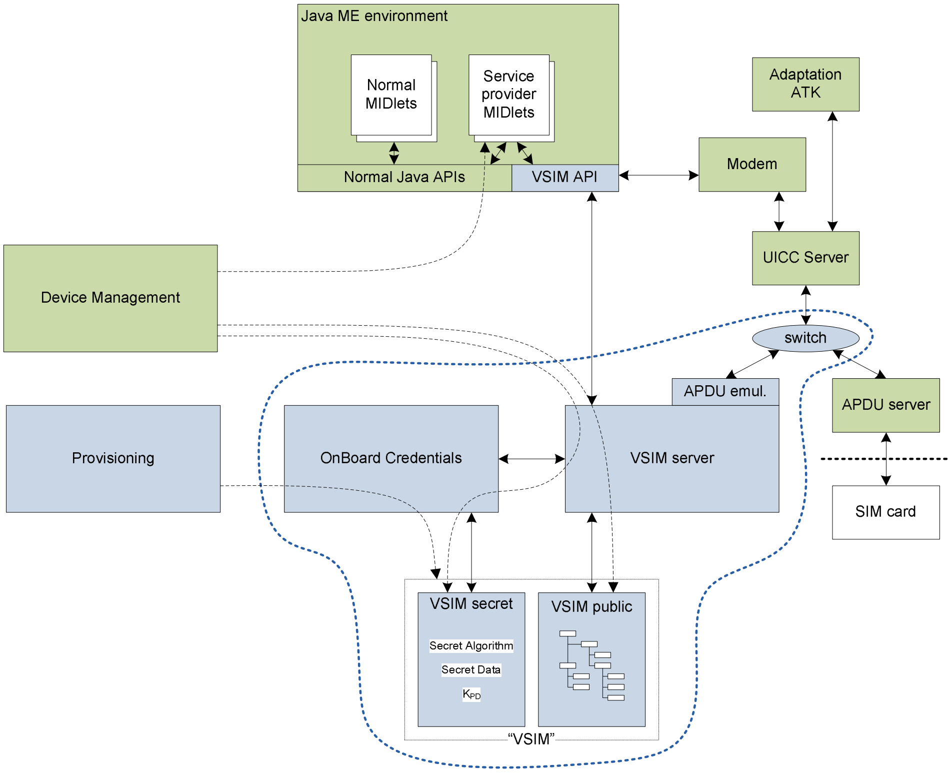


Figure 6.5: The dotted line defines the scope of the first VSIM proof-of-concept being implemented at Nokia. Source: own work.

The first step to be taken is to develop a basic proof-of-concept, which is an important tool in selling the idea. The following approach could be taken:

1. Implement the basic VSIM functionality on the terminal:

    d. Implement a SIM/APDU emulator on the phone, including a SIM file system emulator.

   e. Leverage the On-board Credentials (ObC) scheme to provide the security functions of the Virtual SIM.

   f. Implement a provisioning subsystem that integrates with the ObC subsystem and the emulated SIM file system. This system takes an encrypted VSIM file as input, decrypts in the secure mode, re-encrypts the sensitive data for the ObC subsystem, and stores the non-sensitive data in an emulated SIM file system.

   g. Create a 'VSIM manager' that handles multiple VSIMs and provides a user interface to the provisioning system, allowing the addition and removal of VSIMs.

2. Implement an OTA provisioning scheme working on a test network:

   a. Implement support for OTA provisioning on the terminal.

   b. Implement support for OTA provisioning in the test network.

   c. Create a simple HTTPS-based 'subscription server' connected to the test network, which can generate test VSIMs.

3. Add a Java ME API to the terminal implementation, as described in section 4.4.

4. Extend the Device Management server to support the Virtual SIM, as described in section 4.5.

A team in Nokia has begun working on a proof of concept incorporating the first level of functionality defined above. The scope of this project is shown in Figure 6.5 below.

6.5.2 Standardization efforts required

While the Virtual SIM system has been designed to reuse as many existing technologies as reasonable, the proposed system still requires quite a lot of work in standardization. Specifically, standardization efforts are required for the following items:

- *Terminal implementation:* The security architecture of the terminal needs to be specified, including the role of various components, the interfaces, and implementation guidelines. Furthermore the format of the VSIM files used for provisioning must be determined, including the format of provisioned SIM files, ObC secrets and ObC algorithms. The latter needs a standard for the format of LUA scripts for ObC, including an API specification.

- *Trust model:* Some set of agreements need to be put in place between operators and terminal vendors regarding security and conformance testing of terminals, etc. The interoperability of VSIM credentials on different devices is an extremely important aspect of the system.

- *Bootstrapping:* Modifications to the access and core network as well as inter-network signaling required in order to accommodate the VSIM need to be standardized. Moreover, a PKI or Hierarchical IBE infrastructure needs to be put in place, which involves all terminal manufacturers and operators.

- *OMA-DM extensions*: The proposed remote management scheme requires the standardization of extensions to the OMA-DM protocol.

- *Java ME VSIM 'Toolkit' API:* The VSIM API giving access to 'SIM Toolkit'-like functionality in Java ME for privileged applets needs to be specified.

Chapter 7

Conclusion

Moving from the use of a physical SIM card to a Virtual SIM system has far-reaching implications for the GSM and UMTS industry. It affects fundamental aspects ranging from the way in which operator-centric features of the terminal are standardized and implemented to the way operators sell subscriptions to end-users. On one hand, this means that the Virtual SIM is an opportunity to reshape the mobile world, but on the other it raises concerns that the Virtual SIM effects changes that are unfavourable for some of the stakeholders. Thus, the main success criteria for the virtual SIM are that it can provide the same basic functions as the SIM (e.g. protect operators against fraud, provide a platform for operator-centric services, and support an open terminal architecture), and that, in addition to this, it has clear advantages over the physical SIM cards for all involved stakeholders.

The Virtual SIM scheme proposed in this report uses a secure processor in the terminal, that provides a high enough level of protection to make fraud schemes based on SIM cloning unprofitable. At the same time, this approach does not significantly add to the cost of devices that already use such a processor (which is at least the case for modern Nokia and Motorola phones). The system leverages the ubiquitous Java ME environment to provide an interoperable platform for value-added services, and reuses the OMA DM system to enable remote management of the Virtual SIM. Additionally, a provisioning scheme is proposed, which, in its general form, can use any communication channel between the terminal and the operator to transfer a subscription to the terminal, while allowing the operator to determine whether the terminal can be trusted to protect the secret subscription data. Building on this, an OTA provisioning mechanism is proposed, which allows users to easily download a subscription directly from their terminal, and which supports operators' existing sales channels and systems. This OTA scheme also makes it possible for operators to provide new types of services to their users, and gives them a new channel through which they can market and sell subscription plans.

The proposed Virtual SIM system has several advantages compared to the use of a physical SIM, and solves a whole range of problems that currently apply. It removes the mechanical and electric design restriction imposed by the SIM card, increases the standby time, increases mechanical reliability thereby reducing costs of warranty repairs, removes the problem of man-in-the-middle attacks between the SIM and the terminal, improves the security of SIM-lock mechanisms, and could help preventing mobile phone counterfeiting. Operators also benefit from these advantages, and additionally, the Virtual SIM can reduce the subscriber acquisition cost, and greatly simplify their logistics. Furthermore, it could facilitate usage scenarios such as multiple device ownership, and multiple subscriptions per device.

As a final note, many of the people interviewed during the course of this study seem to believe that the SIM card is a technology that will eventually be replaced by an alternative better suited for the future of mobile communications. Leading the development of the Virtual SIM could be an opportunity for Nokia to shape the future and reinforce its position as the market leader.

References

Note: All web sources accessed 20-04-2008.

[1] "Report of the 1st SIMEG meeting," tech. rep., SIMEG 28/88, The Hague, January 1988.

[2] K. E. Mayes and T. Evans, "Smart cards for mobile communications," in *Smart Cards, Tokens, Security and Applications* (K. E. Mayes and K. Markantonakis, eds.), ch. 4, pp. 85–113, Springer US, 2008.

[3] A. Celen, "SIM Strategy," internal powerpoint presentation, Nokia, 15 March 2007.

[4] "Report of the 2nd SIMEG meeting," tech. rep., SIMEG 43/88, Paris, March 1988.

[5] K. Vedder, "The subscriber identity module: Past, present and future," in *GSM and UMTS: The Creation of Global Mobile Communication* (F. Hillebrand, ed.), ch. 13, pp. 341–369, John Wiley and Sons, 2002.

[6] "Letter from MoU-BARG meeting to SIMEG," tech. rep., SIMEG 12/88, 1988.

[7] G. Schmitt, "The Contribution of the GSM Association to the Building of GSM and UMTS; Section 2: God Send Mobiles, the Experiences of an Operator Pioneer," in *GSM and UMTS: The Creation of Global Mobile Communication* (F. Hillebrand, ed.), ch. 21, pp. 490–494, John Wiley and Sons, 2002.

[8] M. Walker and T. Wright, "Security," in *GSM and UMTS: The Creation of Global Mobile Communication* (F. Hillebrand, ed.), ch. 15, pp. 385–406, John Wiley and Sons, 2002.

[9] "Feasibility study on remote management of USIM application on M2M equipment," 3GPP TR 33.812 v0.2.2, March 2008. `http://www.3gpp.org/ftp/Specs/archive/33_series/33.812/`.

[10] Motorola, "Introduction of SoftSIM into 3GPP," in *Technical Specification Group Services and System Aspects, Meeting #38*, TSGS#38(07)0768, (Cancun, Mexico), 3GPP, 03–06 December 2007. Copy of this document enclosed in Appendix C.1.

[11] J. Creswell, *Research Design: Qualitative and Quantitative Approaches.* Thousand Oaks, CA: Sage Publications, 2nd ed., 2003.

[12] O. Werner and G. M. Schoepfle, *Systematic Fieldwork: Vol. I Foundations of Ethnography and Interviewing.* London: Sage Publications, 1987.

[13] M. Mouly and M.-B. Pautet, *The GSM System for Mobile Communications.* Telecom Publishing, 1992.

[14] S. Redl, M. K. Weber, and M. Oliphant, *An Introduction to GSM.* Artech House, 1995.

[15] F. Muratore, *UMTS: Mobile Communications for the Future.* John Wiley and Sons, 2001.

[16] ETSI, “Digital cellular telecommunications system (phase 2); security aspects.” GSM 02.09 V4.5.1, ETS 300 506. Available at `http://www.3gpp.org.`

[17] 3GPP, “Numbering, addressing and identification, Rel. 7.” 3GPP TS 23.003 V7.0.0. Available at `http://www.3gpp.org.`

[18] K. Rannenberg, “Identity management in mobile cellular networks and related applications,” *Information Security Technical Report*, vol. 9, no. 1, pp. 77–85, 2004.

[19] V. Niemi and K. Nyberg, *UMTS Security*. John Wiley and Sons, 2003.

[20] P. Chandra, *Bulletproof Wireless Security: GSM, UMTS, 802.11, and Ad Hoc Security*. Newnes, June 2005.

[21] I. Goldberg, M. Briceno, and D. Wagner, “Gsm cloning,” April 1998. `http://www.isaac.cs.berkeley.edu/isaac/gsm.html.`

[22] Meeting with Stefan Kaliner, Head of UICC Development, T-Mobile, 10 January 2008. See notes in Appendix A.

[23] 3GPP, “Specification of the GSM-MILENAGE Algorithms: An example algorithm set for the GSM Authentication and Key Generation functions A3 and A8, Rel. 7.” 3GPP TS 55.205 V7.0.0. Available at `http://www.3gpp.org.`

[24] GSM Association, “GSM security algorithms.” `http://www.gsmworld.com/using/algorithms/index.shtml.`

[25] Meeting with Ross Campbell, System Manager, TeliaSonera, 7 January 2008. See notes in Appendix A.

[26] M. Tallberg, “Bundling of handset and subscription.” Helsinki University of Technology, Networking Laboratory, 2004. `http://keskus.hut.fi/opetus/s38042/s04/Presentations/13102004_Tallberg/Tallberg_paper.pdf.`

[27] W. Rankl and W. Effing, *Smart Card Handbook*. John Wiley and Sons, 3rd ed., November 2003.

[28] 3GPP, “Characteristics of the USIM application, Rel. 7.” 3GPP TS 31.102 V7.0.0 . Available at `http://www.3gpp.org.`

[29] 3GPP, “Universal Subscriber Identity Module (USIM) Application Toolkit (USAT), Rel. 7.” 3GPP TS 31.111 V7.0.0. Available at `http://www.3gpp.org.`

[30] SmartTrust, “Steering of roaming technologies.” white paper, 2006.

[31] Sun Microsystems, “Runtime Environment Specification - Java Card Platform, Version 3.0 - Connected Edition,” March 2008. available from: `http://java.sun.com/products/javacard/3.0/.`

[32] M. Tunstall, “Smart card security,” in *Smart Cards, Tokens, Security and Applications* (K. M. Mayes and K. Markantonakis, eds.), ch. 9, pp. 195–228, Springer, 2007.

[33] D. Deville, A. Galland, G. Grimaud, and S. Jean, “Smart Card Operating Systems: Past, Present and Future,” in *the 5th USENIX/NordU Conference*, (Vasteras, Sweden), February 2003.

[34] Giesecke & Devrient, “UniverSIM – securing the 3rd generation,” whitepaper, 2004. `http://www.gdaus.com.au/pdf/Telecommunications/UniverSIM.pdf.`

[35] ETSI, “Smart Cards; UICC-Terminal interface; Characteristics of the USB interface, Rel. 7.” ETSI TS 102 600 V7.1.0.

[36] A. Constantinou and F. Benlamlih, “High capacity SIMs.” Whitepaper, Informa telecoms & media, 2006.

[37] Telecom Italia, “Smart Card Web Server.” web page. `http://c5.telecomitalia.com/default.aspx?idPage=476.`

[38] A. Constantinou, "The sim card evolution: finally, a breakthrough?." VisionMobile blog, 19 March 2008. `http://www.visionmobile.com/blog/2008/03/the-sim-card-evolution-finally-a-breakthrough/`.

[39] J. Walko, "ETSI okays HCI spec for NFC in handsets." web article, 29 February 2008. EE Times Europe, `http://www.eetasia.com/ART_8800507332_499495_NT_5f6cb6ff.HTM`.

[40] W. Lehr and L. W. McKnight, "Wireless internet access: 3G vs. WiFi?," *Telecommunications Policy*, vol. 27, no. 5, pp. 351–370, June 2003.

[41] O. Martikainen, "Complementarities creating substitutes – possible paths from 3G towards 4G and ad-hoc networks.," *CTIF workshop on Beyond 3G/4G*, 2005.

[42] B. White, "Nokia N95 sees crippling by UK carriers." Web article, 20 April 2007. `http://www.engadgetmobile.com/2007/04/20/nokia-n95-sees-crippling-by-uk-carriers/`.

[43] J. L. Anderson and B. Williams, "Unbundling the mobile value chain," *Business Strategy Review*, vol. 15, pp. 51–57, 2004.

[44] K. Palletvuori, "Changing MVAS environment," in *Proceedings of the Research Seminar on Telecommunications Business* (S. Luukkainen, ed.), Helsinki University of Technology, 2003.

[45] B. Evans and K. Baughan, "Visions of 4G," *Electronics & Communication Engineering Journal*, vol. 12, no. 6, pp. 293–303, December 2000.

[46] I. Armuelles, T. Robles, I. Ganchev, M. O'droma, M. Siebert, and M. Siebert, "On ad hoc networks in the 4G integration process," in *The Third Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc)*, 2004.

[47] I. Chlamtac, M. Conti, and J. J. N. Liu, "Mobile ad hoc networking: imperatives and challenges," *Ad Hoc Networks*, vol. 1, no. 1, pp. 13–64, July 2003.

[48] P. L. Stoneman, "Technological diffusion : The viewpoint of economic theory," *Richerche Economiche*, vol. 40, pp. 585–606, 1986.

[49] G. M. Beal and J. M. Bohlen, "The diffusion process," *Increasing Understanding of Public Problems and Policies*, pp. 111–121, 1956.

[50] E. M. Rogers, *Diffusion of Innovations*. New York, NY: The Free Press, 4th ed., 1995.

[51] R. G. Fichman, "The diffusion and assimilation of information technology innovations," in *Framing the Domains of IT Management: Projecting the Future Through the Past* (R. W. Zmud, ed.), pp. 105–127, Cincinnati, OH: Pinnaflex Education Resources, 2000.

[52] L. Tornatzky and K. Klein, "Innovation characteristics and innovation adoption-implementation: A meta-analysis of findings," *IEEE Transactions on Engineering Management*, vol. 29, no. 1, pp. 28–45, 1982.

[53] G. W. Downs and L. B. Mohr, "Conceptual issues in the study of innovation," *Administrative Science Quarterly*, vol. 21, no. 4, pp. 700–714, December 1976.

[54] G. A. Moore, *Crossing the Chasm: Marketing and Selling High-Tech Products to Mainstream Customers*. New York: Harper Business, 1992.

[55] W. Schirtzinger, "Ten reasons high-tech companies fail," in *Proceedings of the University of Washington Computer Fair Conference*, publication year unknown. `http://www.hightechstrategies.com/10_reasons.html`.

[56] D. Leonard-Barton and I. Deschamps, "Managerial influence in the implementation of new technology," *Management Science*, vol. 34, no. 10, pp. 1252–1265, October 1988.

[57] R. E. Kraut, R. E. Rice, C. Cool, and R. S. Fish, "Varieties of social influence: The role of utility and norms in the success of a new communication medium," *Organization Science*, vol. 9, no. 4, pp. 437–453, July 1998.

[58] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly*, vol. 13, no. 3, pp. 319–340, 1989.

[59] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, "User acceptance of computer technology: A comparison of two theoretical models," *Management Science*, vol. 35, no. 8, pp. 982–1003, August 1989.

[60] E. Mansfield, *Industrial Research and Technological Innovation*. New York: Norton, 1968.

[61] Informa, "Mobile industry outlook." Informa telecoms & media, 2006.

[62] S. W. Sanderson, "Managing generational change: Product families approach to design." School of Management, Rensselaer Polytechnic Institute, no date.

[63] E. Ehrnberg, "On the definition and measurement of technological discontinuities," *Technovation*, vol. 15, no. 7, pp. 437–452, September 1995.

[64] R. Campbell, "General SIM-OTA presentation," internal powerpoint presentation, TeliaSonera Denmark, November 2007.

[65] Meeting with Jens Benner, Business Development Manager, Nokia, 16 October 2007. See notes in Appendix A.

[66] Meeting with Peter Aage, employee at unspecified smart card vendor, 5 October 2007. See notes in Appendix A.

[67] "Gartner says worldwide mobile phone sales increased 16 per cent in 2007." Gartner press release, February 2008. available online: `http://www.gartner.com/it/page.jsp?id=612207.`

[68] J. Hynninen, "Experiences in mobile phone fraud," seminar on network security. report tik-110.501, Helsinki University of Technology, Department of Computer Science and Engineering, 2000.

[69] C. Kuhl, "Fraud creeps into wireless psyche." WirelessWeek, November 2007. `http://www.wirelessweek.com/Article-Fraud-Creeps-into-Wireless.aspx.`

[70] M. ho Shin, J. Ma, A. Mishra, and W. Arbaugh, "Wireless network security and interworking," in *Proceedings of the IEEE*, vol. 94, February 2006.

[71] K. Mitnick, "Telecom system security," in *Security Engineering: A Guide to Building Dependable Distributed Systems* (R. J. Anderson, ed.), ch. 17, John Wiley and Sons, 2001.

[72] R. J.-P., C. J.-C., L. J.-L., and K. W., "Mobile phone fraud – are gsm networks secure?," *Computer Fraud & Security*, vol. 1996, pp. 11–18(8), November 1996.

[73] J. Yoshida, "Nokia unveils NFC phones for U.S. carriers at CES." EETimes.com web article, January 2007. `http://www.eetimes.com/conf/ces/showArticle.jhtml?articleID=196802444.`

[74] S. Smith and J. Marchesini, "Hardware-based security," in *The Craft of System Security*, ch. 16, Addison Wesley Professional, November 2007.

[75] D. Lie, C. Thekkath, M. Mitchell, P. Lincoln, D. Boneh, J. Mitchell, and M. Horowitz, "Architectural support for copy and tamper resistant software," *SIGPLAN Not.*, vol. 35, no. 11, pp. 168–177, 2000.

[76] G. E. Suh, D. Clarke, B. Gassend, M. van Dijk, and S. Devadas, "Aegis: architecture for tamper-evident and tamper-resistant processing," in *ICS '03: Proceedings of the 17th annual international conference on Supercomputing*, (New York, NY, USA), pp. 160–171, ACM, 2003.

[77] H. Sundaresan, "Omap platform security features." Texas Instruments whitepaper, July 2003. `http://focus.ti.com/pdfs/vf/wireless/platformsecuritywp.pdf.`

[78] J. Srage and J. Azema, "M-Shield mobile security technology." Texas Instruments whitepaper, 2005. `http://focus.ti.com/pdfs/wtbu/ti_mshield_whitepaper.pdf.`

[79] T. Alves and D. Felton, "TrustZone - integrated hardware and software security." ARM whitepaper, July 2004. `http://www.arm.com/pdfs/TZ_Whitepaper.pdf.`

[80] Trusted Computing Group, “Trusted platform module (tpm) specifications.” available online. `https://www.trustedcomputinggroup.org/specs/TPM/`.

[81] S. Kinney, *Trusted Platform Module Basics: Using TPM in Embedded Systems*. Newnes, 2006.

[82] Trusted Computing Group, “Mobile phone specifications.” available online. `https://www.trustedcomputinggroup.org/specs/mobilephone`.

[83] J.-E. Ekberg and M. Kylänpää, “Mobile Trusted Module (MTM) - an introduction,” tech. rep., Nokia Research Center, November 2007. NRC-TR-2007-015, available online: `http://research.nokia.com/files/NRCTR2007015.pdf`.

[84] Trusted Computing Group, “Mobile Trusted Module Specification FAQ.” available online, June 2007. `https://www.trustedcomputinggroup.org/specs/mobilephone/MTM_Specification_Technical_FAQ_062007.pdf`.

[85] J. A. Halderman, S. D. Schoen, N. Heninger, W. Clarkson, W. Paul, J. A. Calandrino, A. J. Feldman, J. Appelbaum, and E. W. Felten, “Lest we remember: Cold boot attacks on encryption keys.” Center for Information Technology Policy, Princeton University, April 2008. available online: `http://citp.princeton.edu/memory/`.

[86] P. C. van Oorschot, A. Somayaji, and G. Wurster, “Hardware-assisted circumvention of self-hashing software tamper resistance,” *IEEE Transactions on Dependable and Secure Computing*, vol. 02, no. 2, pp. 82–92, 2005.

[87] M. Madou, B. Anckaert, B. D. Sutter, and K. D. Bosschere, “Hybrid static-dynamic attacks against software protection mechanisms,” in *DRM '05: Proceedings of the 5th ACM workshop on Digital rights management*, (New York, NY, USA), pp. 75–82, ACM, 2005.

[88] M. Jacob, D. Boneh, and E. Felten, “Attacking an obfuscated cipher by injecting faults,” in *Proceedings of ACM CCS-9 Workshop DRM* (J. Feigenbaum, ed.), vol. 2696, pp. 16–31, 2002.

[89] Meeting with Claus Rasmussen, Field Application Security Engineer, Cloakware, 29 December 2007. See notes in Appendix A.

[90] S. Chow, P. Eisen, H. Johnson, and P. C. van Oorschot, “A white-box des implementation for drm applications,,” in *Proceedings of ACM CCS-9 Workshop DRM* (J. Feigenbaum, ed.), vol. 2696, pp. 1–15, 2002.

[91] S. Chow, P. Eisen, H. Johnson, and P. C. van Oorschot, “White-box cryptography and an aes implementation,” in *Proccedings of SAC'02* (K. Nyberg and H. M. Heys, eds.), vol. 2595, pp. 250–270, 2003.

[92] O. Billet, H. Gilbert, and C. Ech-Chatbi, “Cryptanalysis of a white box aes implementation,” in *Proceedings of the 11th Annual Workshop on Selected Areas in Cryptography*, pp. 227–240, 2004.

[93] H. E. Link and W. D. Neumann, “Clarifying obfuscation: Improving the security of white-box des,” in *ITCC '05: Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'05) - Volume I*, (Washington, DC, USA), pp. 679–684, IEEE Computer Society, 2005.

[94] L. Goubin, J.-M. Masereel, and M. Quisquater, “Cryptanalysis of white box des implementations,” in *Proceedings of the 14th Annual Workshop on Selected Areas*, 2007.

[95] B. Wyseur, W. Michiels, P. Gorissen, and B. Preneel, “Cryptanalysis of white-box des implementations with arbitrary external encodings,” in *Proceedings of the 14th Annual Workshop on Selected Areas in Cryptography*, pp. 264–277, 2007.

[96] W. Michiels and P. Gorissen, “Mechanism for software tamper resistance: an application of white-box cryptography,” in *DRM '07: Proceedings of the 2007 ACM workshop on Digital Rights Management*, (New York, NY, USA), pp. 82–89, ACM, 2007.

[97] A. Seshadri, M. Luk, E. Shi, A. Perrig, L. van Doorn, and P. Khosla, “Pioneer: verifying code integrity and enforcing untampered code execution on legacy systems,” in *SOSP '05: Proceedings of*

the twentieth ACM symposium on Operating systems principles, (New York, NY, USA), pp. 1–16, ACM Press, 2005.

[98] J.-E. Ekberg, N. Asokan, K. Kostiainen, P. Eronen, A. Rantala, and A. Sharma, "OnBoard Credentials Platform, Design and Implementation," tech. rep., Nokia Research Center, January 2008. NRC-TR-2008-001, available online: `http://research.nokia.com/files/NRCTR2008001.pdf.`

[99] "The programming language lua." `http://www.lua.org.`

[100] R. Ierusalimschy, L. H. de Figueiredo, and W. Celes, "The Evolution of Lua." available online: `http://www.tecgraf.puc-rio.br/~lhf/ftp/doc/hopl.pdf.`

[101] M. Kasper, N. Kuntze, and A. U. Schmidt, "Subscriber authentication in cellular networks with trusted virtual SIMs," in *Proceedings of the 10th International Conference on Advanced Communication Technology*, (Phoenix Park, Korea), IEEE, February 2008. http://andreas.schmidt.novalyst.de/docs/icact2008-2008302-Subscriber-Au thentication-with-Trusted%20vSIMs-final-paper.pdf.

[102] Sun Microsystems, "The Java ME Platform." `http://java.sun.com/javame/index.jsp.`

[103] "Mobile Information Device Profile Specification 2.0 (JSR-118)," 2002. available online: `http://jcp.org/aboutJava/communityprocess/final/jsr118/index.html.`

[104] O. Kolsi and T. Virtanen, "MIDP 2.0 security enhancements," in *HICSS '04: Proceedings of the Proceedings of the 37th Annual Hawaii International Conference on System Sciences (HICSS'04) - Track 9*, (Washington, DC, USA), p. 90287.3, IEEE Computer Society, 2004.

[105] C. Ellison and B. Schneier, "Ten risks of PKI: What you're not being told about public key infrastructure," *Computer Security Journal*, vol. 16, no. 1, pp. 1–7, 2000.

[106] Open Mobile Alliance, "Oma device management v1.2." Available online: `http://www.openmobilealliance.org/technical/release_program/dm_v1_2.aspx.`

[107] Open Mobile Alliance, "Oma device management security v1.2," February 2007. available online: `http://www.openmobilealliance.org/technical/release_program/dm_v1_2.aspx.`

[108] S. Lin, S. Jiang, H. Lin, and J. Liu, "An introduction to oma device management," October 2006. Available online: `http://www.ibm.com/developerworks/library/wi-oma/index.html.`

[109] T. C. Group, "TCG specification architecture overview." `https://www.trustedcomputinggroup.org/groups/TCG_1_4_Architecture_Overview.pdf.`

[110] A. Shamir, "Identity-based cryptosystems and signature schemes," in *Proceedings of CRYPTO 84 on Advances in cryptology*, (New York, NY, USA), pp. 47–53, Springer-Verlag New York, Inc., 1985.

[111] D. Boneh and M. K. Franklin, "Identity-based encryption from the weil pairing," in *CRYPTO '01: Proceedings of the 21st Annual International Cryptology Conference on Advances in Cryptology*, (London, UK), pp. 213–229, Springer-Verlag, 2001.

[112] C. Cocks, "An identity based encryption scheme based on quadratic residues," in *Proceedings of the 8th IMA International Conference on Cryptography and Coding*, (London, UK), pp. 360–363, Springer-Verlag, 2001.

[113] J. Baek, J. Newmarch, R. S-Naini, , and W. S. A, "Survey of identity-based cryptography," in *AUUG*, 2004.

[114] C. Gentry and A. Silverberg, "Hierarchical id-based cryptography," in *ASIACRYPT '02: Proceedings of the 8th International Conference on the Theory and Application of Cryptology and Information Security*, (London, UK), pp. 548–566, Springer-Verlag, 2002.

[115] Informa, "Future Mobile Handsets, 9th Edition." Informa telecoms & media, 2007.

[116] M. R. Smith, *Commonsense Computer Security: Your Practical Guide to Information Protection.* McGraw-Hill Book Co., 1993.

[117] F. Braber, I. Hogganvik, M. S. Lund, K. St, and F. Vraalsen, "Model-based security analysis in seven steps — a guided tour to the coras method," *BT Technology Journal*, vol. 25, no. 1, pp. 101–117, 2007.

[118] D. Belic, "No-key Nokia X-SIM unlocks Nokia phones to work on any network; Warranty stays intact." web article, 27 February 2008. `http://www.intomobile.com/2008/02/27/no-key-nokia-x-sim-unlocks-nokia-phones-to-work-on-any-network-warranty-stays-intact.html.`

[119] FierceMobileContent, "EU launches inquiry into mobile data roaming fees," 17 January 2008. `http://www.fiercemobilecontent.com/story/eu-launches-inquiry-mobile-data-roaming-fees/2008-01-17.`

[120] "Cloning (r)evolution: Surprising good looking N93i clone!?." Symbian Freak web article, 22 September 2007. `http://www.symbian-freak.com/news/007/09/surprising_good_looking_n93i_clone.htm.`

[121] "Cloning (r)evolution: Surprising good looking N95 clone!?." Symbian Freak web article, 2 January 2008. `http://www.symbian-freak.com/news/008/01/surprising_good_looking_n95_clone.htm.`

[122] J.-Z. Sun, J. Sauvola, and D. Howie, "Features in future: 4G visions from a technical perspective," *Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE*, vol. 6, pp. 3533–3537 vol.6, 2001.

[123] A. Jøsang and S. Pope, "User centric identity management," *AusCERT Conference*, 2005.

[124] OASIS, "Conformance Requirements for the OASIS Security Assertion Markup Language (SAML) V2.0," committee draft, Organization for the Advancement of Structured Information Standards, 15 January 2005.

[125] Liberty Alliance, "Liberty ID-FF Architecture Overview." Version: 1.2-errata-v1.0. Liberty Alliance Project, 2005.

[126] A. Bhargav-Spantzel, J. Camenisch, T. Gross, and D. Sommer, "User centricity: a taxonomy and open issues," in *DIM '06: Proceedings of the second ACM workshop on Digital identity management*, (New York, NY, USA), pp. 1–10, ACM, 2006.

[127] JSR 177 Expert Group, "Security and Trust Services API (SATSA), JSR 177," 2003. Available online: `http://java.sun.com.`

[128] G. Rowe and G. Wright, "Expert opinions in forecasting: Role of the delphi technique," in *Principles of Forecasting: A Handbook of Researchers and Practitioners* (J. Armstrong, ed.), Boston: Kluwer Academic Publishers, 2000.

[129] K. C. Green, J. S. Armstrong, and A. Graefe, "Methods to elicit forecasts from groups: Delphi and prediction markets compared," MPRA Paper no. 4999, Monash University Business and Economic Forecasting Unit, 7 November 2007.

[130] P. Mockapetris, "Domain names - implementation and specification." RFC-1035, November 1987.

Appendix A

Interviews & meetings

A.1 Antti Kiiveri, 27-08-2007

Subject:	Virtual SIM		
Date:	27-08-2007	Location:	Nokia, Copenhagen
Type:	Interview	Output:	Transcript
Participants:	Antti Kiiveri, Senior Architect, Security, Nokia Marc Richarme, Student, DTU		

MR: Are you working with BB5?

AK: Yes, I've been working with BB5 from day one. Basically, I have been building the architecture.

MR: What are you using BB5 for? Of course there is SIM lock, but are there any other features that are using BB5?

AK: DRM is one. There is a slideset I could show you if you are intersted in the features we are using BB5 for.

[...]

MR: What level of security is BB5 designed to achieve?

AK: Basically the ideas behind is thet some day we could do DRM, but the three basic ideas were IMEI protection, SIM lock and software authenticity. Then we were thinking: Could this be something that we could use for DRM. And also basically all the use cases that we wanted in later were kind of easier.

MR: Because you had a solid foundation

AK: Yes. So the SIM lock and DRM they are hard enough for most of the things. So then the question would be how secure the hardware actually is. And I think that it is a very good question, and we don't have a very good answer for that.

MR: I think Mikael told me that you had a company investigating it, trying to crack it, and came up with a solution where you could peal off the upper layer off, and with a microscope read the key of the chip. Is that true?

AK: Yes, that is doable.

MR: Is that the most advanced or best hack you've found until now?

AK: Maybe. That was kind of overkill in the sense that we kind of accepted that that is something that can be done, but since it is very difficult and you need to do it for every device you want to hack. You can't kind of copy the attack very easily so that you break one, and the the rest of the devices would be piece of cake, but you need to do the whole attack for every device. For us it was kind of acceptable level.

MR: So you have accepted this threat. Do you know if they have the same problem with smart cards?

AK: I'm not actually a very good expert on that. We used to have... after we have been designing that bb5 for a while, there was one guy Lauri Paatero. he joined Nokia. He has a history of smart card security. He's been in one Finish company who was doing smart cards. [...] So he came in and brought all this knowledge [...] and fullfilled the details one we started to need those.

MR: Ok. The hack by Dejan, the SIM unlock hack, that only affects the SIM lock part and is not a basic flaw in the bb5 security, right?

AK: Yes

MR: Ok, so a part from the microscope thing, there is no security flaw.. until now?

AK: Yes.. But since he has been doing some so good analysis on the system, I guess that a similar attack can be targeting for example software SIM or whatever.

MR: Ok. What's that attack.. how does it work. You said something about chaning the value of a register.

AK: I didn't, so..

MR: Ok.. somebody said..

AK: Somebody said. Yes, that would be the idea, that..

MR: Can you explain the basic idea behind the attack?

AK: I think that, the idea is that at a certain moment they changing the.. or actually they are directing the answer to a different register that it is supposed to go.

MR: I'm not sure I understand what answer? [...]

AK: There would be also kind of instructions.. code instruction. And I guess that their are chaning one instruction which is supposed to write values to a register, let's say A.

MR: In the memory which is outside the BB5 chip?

AK: It's kind of CPU registers which is inside the.. Its one of the ARM processor registers. So it's supposed to write the register, let's say A, but it writes to A, because of this hardware hack.

MR: And these registers are inside the secure part of the chip?.. the a and b registers that are modified?

AK: That I don't know. I guess that would be the case.

MR: But the communication on the bus when the bit is changed? What is the chip communicating with?.. the memory or another ASIC, or?

AK: Memory bus.

MR: Ok, it's the instruction memory that's actually altered on the way to the chip dynamically.

AK: Yes, it must be instruction i think.

MR: So in principle, you can change any instruction?

AK: There are actually guys in the next room that can actually give the right details to you if you are interested. Janne Takala for example knows that better than I do.

MR: Ok.. Have you though about software SIM or is that not your domain.

AK: We have been discussing about that earlier also

MR: As a use case for BB5 or in general?

AK: In general.. but of course also for BB5,but we didn't really go deep in that area. [...]

MR: You didn't go deep into it.. why?

AK: Umm.. I guess that we had something else also by that time.

MR: So it's just a matter of resources, and not that the area wasn't interesting.

AK: Yes.. it would be very interesting in the sense that if we can put all of the smart card vendors out of business [laughs].

MR: Do you know what the production steps are? There is a labeling step involved with every device. I guess that all the data is stored in a secure server in the production facility.

AK: Yes

MR: Is there any.. I know that with smart cards there are some standards where the secure area is.. only special personnel can go in there and they have all these security procedures which are certified, etc. Do you know if you have some similar..

AK: We have the same without the certification.

MR: Ok, so it's basically the same procedures as with smart cards, security-wise.

AK: Yes

MR: Why haven't you been certified. Because you don't have to prove anything to anyone, or?

AK: Yes, more or less so.

MR: But is you wanted to, you could

AK: Yes

MR: I don't know if you can tell me some more about the more advanced features of BB5. What is the security based on? Is it the same system as TPM for instance?

AK: No in the sense that.. or basically we have now this mobile TGC workgroup, which has basically define the TPM for mobile phones, which is called something else, I don't remember. But basically they have the same now, and I guess it is out now. It has been published during the summer. And Nokia has been also involved in that area. And the target in that work was that we would be able to utilitze BB5 security system for creating kind of a terminal internal TPM. So this would be one use case. There might be slight differences.. if you read the specification. There might be some kind of things that we are not quiet able to fulfill with BB5 now. Mostly because this guy left the standardization workgroup when he was doing spec version 0.9. And then without this guy they upgraded it from 0.9 to 1.0, and there was those changes done without Nokia intercepting it. But that would be one use case for BB5 system also. [...] Currently I'm not very big fan og TGC anyhow, because I don't really understand kind of the added value of it. Because we basically have all the same things with the.. or most of the same things with BB5 already. Of course there are interoperability stuff which would be good, but I don't know what would be the value if all the device would for example boot the same way. Because at least the mobile spec doesn't really say anything about this in the interoperability or something like that. And we don't have any kind of infrastructure supporting for example that you can from the network you can somehow ask values from a mobile device, such as "are you now ok or not".

MR: About the software SIM.. In the GSM (and I think also in the UMTS) speciifcation you have the A3 and A8 algorithms which can be defined by the operator.. so in order to put the software SIM in the phone, there are two options: Either the operator could download their specific algorithms to the phone in order to make the authentication etc., or some fixed algorithm could be used. But if the operators were to download their own algorithms alongside the SIM authentication keys etc., do you think that would be possible to execute in a secure way for instance inside BB5. Is there a system allowing to authenticate code or instruction, which can be downloaded from the outside and the executed without disclosing the algorithms to the outside world.

AK: The is just the raw authentication. There is no system for kind of allowing that by the application to be executed in the BB5 security subsystem. So that would be that we would need to implement those algorithms by ourselves, or that we should have some kind of interpreter in the BB5 subsystem. That would kind of protect the secure environment itself.

MR: [...] It is imaginable to make a byte-code interpreter or something running in secure mode.

AK: Yes, the ... have already the interpreter ready.

MR: Is it Java based, or?

AK: It is based on LUA.

[...]

A.2 Peter Aage, 05-10-2007

Subject:	Virtual SIM		
Date:	05-10-2007	Location:	Nokia, Copenhagen
Type:	Interview	Output:	Transcript
Participants:	Peter Aage, M.Sc. in Computer Science, employee at a smart-card vendor Peter Anglov, M.Sc. in Engineering, Danish Defence Marc Richarme, Student, DTU		

MR: I 3GPP's security work group har de kommet med et forslag til et Work Item, som handler om, at man i forbindelse med M2M overvejer muligheden med at fjerne det fysiske SIM kort og lægge applikationen ind i terminalerne direkte, for det giver en masse fordele i forbindelse med M2M. Og målet med mit projekt det er at finde ud af, om det giver mening - også for Nokia - og om det er noget, Nokia skal beskæftige sig med. Og så kigge på løsninger på problemet og at finde ud af, hvad der bedst kan betale sig. Så jeg vil gerne interview dig om input til det.

MR: Har du lyst til at give en kort præsentation af dig selv, for the record?

PA: Altså sådan hele historien, eller?

MR: Bare hvad der har med branchen at gøre.

PA: Som sagt så har jeg arbejdet i $5\frac{1}{2}$ år hos Gemplus, og i den periode der lavede jeg en hel del projekter med OTA opdatering af SIM kort, altså typisk sådan noget med.. altså ja, praktisk så installerede jeg platforme, men jeg kørte også nogle kampagner hvor jeg fx rebrandede en hel operatørs portefølje af kunder, ikke. Så det var sådan noget med.. altså vi sendte for eksempel nede hos Orange i Holland, som skiftede fra DutchTone til Orange, der sendte vi 2 millioner SMS'er for at opdatere et par filer i SIM kortet. Så det var.. ja altså det er med de mekanismer som man nu kan i SIM kortene. Det er sådan noget halv-proprietært noget.. det er blevet lidt mere åbent i forbindelse med USIM. Men altså, det var den slags projekter jeg var involveret i, og det er, klart at jeg i hele min karriere faktisk har været involveret i projekter som både var rent telekom-relaterede men også sådan en blanding af forskellige projekter. Altså i Gemplus arbejdede vi med den her Mobilix betalingsløsning. Jeg var ikke decideret involveret i det, men jeg var alligevel lidt på sidelinien; jeg arbejdede i firmaet på det tidspunkt. Nu er jeg primært inden for bankkort; specialist på området - men jeg ved udmærket hvad der foregår på SIM siden også.

MR: Som jeg nævnte, så er 3GPP i gang med at lave en undersøgelse, som handler om at flytte SIM applikationen fra SIM-kortet ind i telefonen i forbindelse med M2M. Er det noget du har hørt om før? - er det noget nyt for dig?

PA: Nej, det er ikke nyt. Fordi der er mange i M2M, der har kigget efter det her. For det er klart at, at integrere et SIM kort i sådan en lille embedded device er en udfordring. Og det er specielt en udfordring når man skal bruge udstyret i sådan nogle harske miljøer, fordi det er virkelig en dårlig connector, der sidder i telefonen. Og selvfølgelig skal man også lige implementere en T=0 protokol, og vide hvordan kortet kommunikerer. Altså, det ville være nemmere at have sådan en eller anden SpeedBus eller noget andet, man bare kunne integrere med sin mikroprocessor, som man gør med alt muligt andet.

MR: Hvorfor tror du så ikke, at det er sket endnu?

PA: Operatørerne er jo ikke specielt glade for at give sådan en Ki til nogen som helst - altså den nøgle der skal bruges i SIM-kortet. Og derfor er men også nød til at forsikre operatøren om, at hvis man får oplyst nøgler og den slags, så skal de også blive gemt på en sikker måde. Og jeg tror i virkeligheden at det der er udfordringen og grunden til det ikke er blevet gjort endnu det er at man ikke har nogen sikker måde at loade de her chips, eller hvad man nu har tænkt sig at bruge i stedet for. At der ikke rigtig er nogen standardiseret måde at flytte data over i en embedded device, som måske havde den SIM-funktionalitet.

MR: Så det der mangler, siger du, det er, at den måde de bliver overført på, at man over for operatørerne kan vise, at det er sikkert.

PA: Det er i hvert fald en af tingene. Jeg vil ikke sige at det er det eneste der mangler. Men det er klart, at man skal jo altså hele tiden skal kigge tilbage på, hvorfor indføre man overhovedet SIM. Det har man jo gjort fordi, at man da man startede med GSM ikke havde noget sikkert sted at gemme - altså der var flere forskellige grunde, ikke, men - man havde ikke noget sikkert sted at gemme de her nøgler i telefonen. Jo så kunne man putte en mikroprocessor ind, hvor de var brændt i et eller andet sted fra og så havde man så låst telefonen til en specifik abonnent, eller til et eller andet uhensigtsmæssigt. Hvis man ligesom skal bevæge sig ud, så skal man kigge på, hvad kan man lave af sådan en åben infrastruktur, hvor man ligesom kan flytte de her nøgler rundt på en sikker måde. Og jeg tror en af udfordringerne det er netop det.. altså der findes jo masser af nøgleudvekslingsformater. Og jeg ved da også at man i router-miljøer og sådan noget, allerede har routere med et certifikat i, der kunne gøre.. Altså, du kunne få routerens identitet, og så ud fra identiteten kunne du bare kryptere nøglerne ned til den router, og så er man nogenlunde sikker på, at de blev gemt fornuftigt. Men altså, det kan man jo stadigvæk ikke.. fordi du krypterer dem til routeren, så kan routeren jo stadig behandle dem på en usikker måde.

MR: Så man mangler også et bevis for, at der hvor nøglerne ender, at det også.. at de så ikke kan komme ud derfra.

PA: Ja præcis. Og hvis man skal gå helt over i betalingsverdenen, så har man noget der hedder PCI-certificering, som jo er sådan en certificering mod at du ikke stort set kan komme ind nogen steder og lave side-channel attacks, eller sådan nogle ting. Så derfor, altså det ville nok være overkill at lave en PCI-certificering. Men, du bliver nødt til at forsikre din operatør om at de her nøgler bliver gemt et sikkert sted, fordi det er virkelig operatørens ejendom. Hvis operatøren ikke ved hvor hans nøgler er, så risikerer han en masse svindel i systemerne. Og det er klart, at det er man bare nødt til at overbevise operatøren om, at det er sikkert.

MR: Så du mener også, at tiderne har ændret sig siden GSM blev indført.. at det nok er mere realistisk nu?

PA: Ja, fordi nu begynder du at have sådan noget trusted platform chips, og sådan nogen ting, der begynder at komme i PC'erne, og man kunne måske forestille sig, at det også kom i mobilerne, på et eller andet niveau.

MR: Og det gør det jo faktisk også fordi den gruppe der har indført trusted platforms på PC'er, lige har standardiseret en mobil pendant.

MR: Hvis nu, at det viser sig at det kan lade sig gøre i M2M terminaler, det her. Hvorfor skulle det så ikke kunne lade sig gøre i almindelige telefoner?

PA: Ja det var også nærmest mit spørgsmål til dig. Fordi, altså, hvorfor kigger man kun på M2M. Fordi i virkeligheden, hvis man laver infrastrukturen til det her - og det er jo nødt til at være billigt, for det skal sådan noget M2M være - hvorfor skulle man så ikke kunne bruge det i en almindelig mobiltelefon.

MR: Du ser ikke nogle specielle forhindringer?

PA: Nej bestemt ikke. Altså igen, man må tilbage til hvorfor har man indført et SIM kort? Og det er simpelthen mangel på et sikkert sted at gemme nøglerne. Og hvis det er til stede i en terminal, jamen så kan man lige så godt bruge den.

MR: Hvad ser du af fordele og muligheder? Altså, du siger, at du har overvejet det her før... Ser du nogen specielle fordele ved at have det i normale mobiltelefoner frem for at have et SIM kort.

PA: Altså, det er da klart at noget af det der er den meget store udfordring for en operatør, det er jo det her med at han skal skifte SIM-kortet ud. Men samtidig er det jo også en loyalitetsfaktor det der med, at det skal jo også helst være en lille smule besværligt at churne til andre selskaber. Så det er sådan både og. Der er fordele og ulemper. Jeg vil sige, at hvis man skal bruge et virtuelt SIM, så skal man i hvert fald også kunne lave en eller anden form for SIM-lock for at kunne fastholde abonnenten. Det er klart, at der er selvfølgelig også den ulempe, at hvis man har en eller anden specifik funktionalitet i SIM kortet, noget form for microbrowser eller.. Altså, der findes nogle forskellige teknologier, hvor SIM-kortet det står for eksekvering af nogle cardlets eller sådan noget, det vil man jo ligesom ikke kunne flytte med over i et virtuelt SIM, men jeg vil så til det sige, at det er jo ikke noget, der i dag er specielt udbredt. Altså der er ikke mange i dag i Danmark, der bruger SIM-browsing og den slags teknologier. Ikke engang WAP er specielt udbredt.

MR: Nu snakker du om den funktionalitet der er i SIM kortet. Jeg har sådan et skema her, hvor jeg har prøvet at skrive op hvilke funktionaliteter der er. Har du en ide om, hvor vigtigt det vil være for de forskellige funktionaliteter at blive overført til telefonen, hvis nu man skulle afskaffe SIM-kortet.

PA: Altså den første den er jo ligesom obligatorisk: Autentificering op mod netværket [...]. Jamen altså telefonbogen det kan man jo stille spørgsmålstegn ved, fordi du jo i dag allerede har funktioner til at udveksle din telefonbog mellem telefoner. Altså, hvorfor skal den i det hele taget ligge i SIM-kortet, det kan man stille spørgsmålstegn ved. Forså altså, jeg kan jo eksportere min telefonbog som Windows adressekartotek til en anden telefon. Så det vil jeg ikke sige er noget must. SIM Application Toolkit, altså det er jo noget man *kan* bruge til specielt.. hvis der nu er ting i telefonen, som man af en eller anden grund har brug for at putte noget ekstra sikkerhed på.. eller, hvad skal man sige, en applikation du har brug for at putte noget ekstra sikkerhed på - f.eks. den her Mobilix betalingsapplikation - jamen så kan man implementere i det i SIM-kortet ved hjælp af SAT. Men hvis du i terminalen har ekstra sikkerhedsfunktionalitet, så du på en sikker måde kan bruge nogle nøgler i telefonen, altså til f.eks. at signere en betalingstransaktion, så er det heller ikke noget must. Og jeg vil også sige, at den type applikation er jo ikke noget, der er specielt udbredt i dag. Altså, der er ikke nogen her i Danmark, der egentligt bruger SIM Toolkit Applikationer til noget specielt fornuftigt. Det er i hvert fald min konklusion. "Roaming Control", der går jeg ud fra at du refererer til PLMN-listen i telefonen?

MR: Ja, samt at der er nogle operatører, som vist ændrer den dynamisk.

PA: Ja, det er der rigtig mange, der gør. Fordi det er klart, at den måde man som operatør laver roaming på, det er, at man laver nogle roaming agreements med nogle forskellige operatører rundt om i verden. Og når man laver nye agreements så er det klart, så skal man have de her operatørers numre ind i PLMN-listen. Og det er faktisk en vigtig funktionalitet at have. Men spørgsmålet er om den skal ligge i SIM-kortet. Fordi som det er i dag, der er det jo rent faktisk sådan at brugeren jo selv kan gå ind og rette i den her, hvis han har den rigtige telefon. Ikke at der er ret mange, der gør det. Men altså, man kunne lige så godt.. Det ville da være mere oplagt at have en liste, som man signerede og pushede ned i telefonen en gang imellem. Altså hvorfor skal den helt ned i SIM-kortet. Det er der ikke nogen specielt sikkerhedsmæssig begrundelse for. Altså, man kunne lige så godt arbejde med nogle handset profiler, og man har jo også handset provisioning systemer i dag. Så der er ikke nogen sikkerhedsmæssig argumentation for at have den i SIM-kortet. Det er bare en liste du læser op. Så er der det her "hele filsystemet". Altså der ligger jo nogle funktioner, men der er jo ikke nogen voldsom argumentation for, at have det i et kort. Der er ikke ret mange af de her filer, der ligger i filsystemet, som har nogen speciel sikkerhed, man ikke kan implementere i en terminal. Så er der noget WIM. Jeg går ud fra at det er sådan noget PKI-funktionalitet. Og det er klat, at grunden til at man har WIM funktionalitet i SIM-kortet det er igen, fordi man mangler den sikre platform i terminalen, og hvis man har det, så er der ikke nogen grund til at lægge den i SIM-kortet. "ISIM" ved jeg ikke, hvad der dækker over.

MR: Det er til IP Multimedia Subsystem, IMS, som bliver sådan en IP baseret.. fremtidens IP netværk eller hvad man skal kalde det, som også er blevet standardiseret her for nylig. Man begynder vist allerede at lave SIM-kort med det, tror jeg.. Det er noget IP identitet.

PA: Ja, men igen, hvis det er identitet, og du har en trusted platform, du på en eller anden måde kan gøre det i, så er det ikke hensigtsmæssigt at gøre det i SIM-kortet. "JavaCard Environment".. det er jo lækkert fordi du kan downloade applets i kortet og ny funktionalitet. Men hvis du har et tilsvarende miljø i din terminal, hvor du kan downloade en signeret applet, og sikre at den kun gør nogle bestemte ting, jamen så kan man selvfølgelig lige så godt have det i terminalen.

MR: Du sagde noget med microbrowsers. Er det noget, som operatørerne ville synes var vigtigt at have i telefonen?

PA: Altså igen, jeg mener, at TDC har det i mange af deres SIM-kort.. eller havde i hvert fald. Og jeg ved også, at det ikke bliver brugt. Altså, det er jo en information on demand application, man typisk bruger den slags til. Og problemet er, at det er for dårligt integreret i telefonen. Det ligger for langt nede i nogle menuer, for at man gider at bruge det.

PAN: Er der nogle andre ting i SIM-kortet, som der ikke er nævn, der interessant for SIM-kortet, eller ikke interessant?

PA: Ja, men altså der er selvfølgelig nogle funktioner omkring... at man kan jo lave det der hedder fixed dialing numbers med sådan en PIN2, så du kan slå den over til kun at ringe til at antal bestemte numre og den slags, som også er nogle sikkerhedsfunktioner, der ikke er ret mange, der bruger. Og grunden til der ikke er ret mange der bruger det, det er, at der ikke er nogen sikker kommunikation mellem telefonen og kortet. Og dvs.. at hvis man bare er en lille smule smart, så går man bare ind og sætter en eller anden spy eller et eller andet ind i mellem, der ligesom får SIM-kortet til at opføre sig anderledes, fordi det er der ikke nogen der kan opdage. Så der er ikke rigtig nogen sikkerhed i det. Og så ligger der selvfølgelig nogle - nu ved jeg faktisk ikke hvor meget det bliver brugt - altså der er jo nogle funktioner, hvor telefonen viser operatørens navn og sådan noget ud fra nogle data - men det er alt sammen noget data, der bliver læst op af kortet... som i virkeligheden godt kunne ligge et andet sted.

MR: Man kunne godt forestille sig at filsystemet det lå.. eller de relevante filer lå i telefonen og kunne opdateres af operatøren.

PA: Ja, altså, man kunne jo i princippet bare have en applet, der gjorde det samme. Altså i java-kort der er SIM applikationen jo en applet der bliver downloaded. Og der er ikke nogen der siger, at den behøver at ligger i et SIM-kort. Det er for at man har et ensartet, sikkert miljø til at eksekvere sin applet i.

MR: Kan jeg få dig til at sætte nogle tal på hvor vigtigt du synes det er?

[...]

PA: Jeg spekulerer lidt på hvordan man ligesom skal gøre det op..

MR: Spørgsmålet er om det skal være sådan en operatør-telefonbog, eller..

PA: Der ligger jo f.eks. nogle service dialing number også i nogle kort. Altså, man har en bestemt telefonbog, hvor du kan lægge nogle opratørafhængige numre ned i. Så derfor er der jo flere typer af telefonbøger. Så hvis man tager den mest almindelige "ADN" telefonbog, der ligger i SIM-kortet. Ja, altså så.. jeg tror ikke engang.. jeg bruger ikke min, fordi de her entries er for korte, ikke.. jeg mener ikke den er specielt brugbar som den ligger i SIM-kortet. SIM Application Toolkit.. hvis man kigger i andre lande.. altså jeg ved der er nogle lande, hvor man.. altså jeg mener i Saudi Arabien og sådan nogle steder, har de nogle browsing applikationer i kortet. Så det kan godt være en vigtig ting at tage med. I hvert fald funktionaliteten. Roaming control, den mener jeg også er crucial at have.. Den der [entire UICC file system] er nok en 2'er. Altså, funktionaliteten er selvfølgelig vigtig, men det er ikke vigtigt at have i et SIM-kort.

[...]

PAN: Hvis du skulle lave kravspecifikationen til en virtuel SIM enhed, hvilke ting skulle så være med i forhold til hvad der er i dag?

PA: Jamen jeg tror sådan set jeg ville tage udgangspunkt i, hvad der er i dag. Der er jo en grund til at man har de forskellige datafiler i et SIM-kort.

[...]

PAN: Hvis du skulle være mere præcis og sige, "det skulle kun være den, den, den, den", altså virkelig præcist skære igennem.

PA: Ja man er jo nødt til at kigge på hvordan operatøren rent marketingmæssigt bruger de her forskellige ting. Altså, der er jo nogen der bruger meget krudt på at lægge de rigtige numre i. Altså at man kan ringe til et servicenummer hvis man har et problem. Andre har det overhovedet ikke. Og det er derfor at sådan noget som et par entries i telefonbogen, der allerede ligger der når man får sit kort, kan være interessant. Og i det hele taget hele pakketeringen af.. når man køber et taletidskort, hvad får man så i hånden.. hvad ville man så skulle få i hånden som et virtuelt SIM? [...] I dag, når man går ned og køber et taletidskort, får man jo noget i hånden, og det er klart, at der er jo nogen marketingmekanismer, som man også er nødt til at overveje lidt. Altså hvis man kigger på det rent teknisk, jamen så er den første her jo.. den skal være der, det er jo derfor SIM-kortet er der. Igen, telefonbogen er ikke vigtig, som sådan, hvis den i øvrigt er i telefonen. Og hvis man køber en telefon sammen med et kort, så vil den også være personaliseret med operatørens forskellige ting og sager - GPRS settings og alle de her ting. Men altså, det er klart, at man er jo nødt til at vide sådan lidt mere specifikt hvor meget der egentligt bliver brugt af de her funktioner. Er SIM Toolkit Applikationer noget, som man satser på i Nokia i fremtiden, ikke? Altså der er jo Ericsson-telefoner, der ikke understøtter det.

MR: Det er i høj grad et spørgsmål om, hvad operatørerne kræver, fordi det er i høj grad dem, som laver kravspecifikationerne. -- Ved du noget om sikkerheder omkring autentificeringsmekanismerne i SIM-kortene.

PA: Ja, altså, jeg kan ikke A8/A3 algoritmen by heart, men jeg ved nogenlunde..

MR: Jeg prøver at fiske efter.. hvis man har et virtuelt SIM-kort.. sådan som specifikationerne er nu, så kan operatørerne selv definere deres custom-algoritme. Er det noget, som du vil mene er nødvendig at bibeholde, hvis man vil flytte SIM-kortet over i telefonen, eller kan man forestille sig at det er nok at lave en fast algoritme - f.eks. MILENAGE, som bliver brugt i UMTS, og som har nogle forskellige parametre, som kan customizes. Tror du at operatørerne, at det er meget vigtigt for dem, at de har den mulighed for selv at vælge algoritme helt frit, eller at de kan stilles tilfredse med at kunne vælge nogle parametre, eller om det bare er nok at have en fast algoritme.

PA: Altså, jeg mener at Orange har deres egen algoritme, og så der jo den her HMAC, som vist nok er en Nokia-opfindelse, som jo også bliver brugt i det, der tidligere hed Orange i danmark, som nu er Telia. Altså de brugt det jo også før det egentligt blev udbredt. Men jeg tror der er nogen operatører, der vil være interesseret i at have deres egne algoritmer. Men altså, som du selv antyder, er det jo ikke det fleste der gør det - det er de færreste - men der er nogen, der fokuserer på at have deres egne algoritmer.

MR: Tror du at de ville give afkald på det hvis de fik nok ud af det? Tror du de ville være villige til at give afkald på det?

PA: Det er et udmærket spørgsmål. Jeg ved ikke hvad det er der gør at man vælger at definere sin egen algoritme. Jeg tror det er noget med at det skal være svært.. det skal være endnu sværere at hacke ind på nettet.. Jeg kan ikke svære dig på om det er et must.

PAN: Er det relevant? Er det et relevant spørgsmål?

PA: Ja det er det jo, fordi det er jo også noget af det, der.. Det er jo en af de ting som man vælger når man implementerer et nyt SIM-kort. For kan have noget at gøre med en HLR, om det kun kan noget specifikt. Og det kan jo også have noget at gøre med.. at hvis du vil lave tingene.. eller terminalerne bagud-kompatible.. altså du skal jo være opmærksom på at man skal sikkert skifte sit HLR, hvis ikke man understøtter de eksisterende algoritmer. Så derfor så får du jo nok et problem, hvis ikke du kan lægge hvad som helt i af underlige algoritmer. Fordi det jo i dag er noget man kan gøre åbent, som det passer en. Så jeg vil nok tilråde at man havde nogle valgmuligheder. Også fordi, at hvis man kigger på det sådan rent historist, så have man jo den har COMP128 security by obscurity algoritme, som jo blev hacket på et tidspunkt. Og sådan vil det jo være fremadrettet. Så derfor er du jo nødt til at kunne skifte dem ud.

[about subscription provisioning...]

PA: [...] Det der selvfølgelig vil være kampen det vil være, hvilken operatør er det, der dukker først op på listen, fordi det er jo det, alle vil være interesseret i. Og hvordan sikrer man at valget bliver tilfældigt. Altså når jeg starter min device op, hvad dukker der så op på skærmen?

MR: Det er i tilfælde af at det er almindelige telefoner..

PA: Ja.. Men det er jo også et eller andet form for valg der skal foretages hvis det er machine to machine. Altså, hvis man nu forestiller sig at den skal startes automatisk op.

[...]

MR: Du sagde noget med at det bliver et problem at finde ud af hvordan selve telefonen bliver garanteret over for operatøren.. at den vil beskytte hans hemmelighed. Kunne man forestille sig, at telefonerne bliver certificeret, og via sådan et certifikat, som telefonerne indeholdt, så blev checket inden operatøren sendte SIM-kort ned til dem.

PA: Ja, altså man har jo også.. altså for kortene i dag har du jo sikkerhedscertificeringer. Altså, man opererer med sådan nogen IRL certificeringer eller common criteria certificeringer af forskellige dimser. Altså i smart card verdenen er det typisk noget IRL-certificering, hvor man validerer, så vidt jeg husker.. man har en eller anden profil, hvor man siger at den opfylder det og det, og så bliver det valideret at den rent faktisk gør det.

MR: [...] Hvem er det, som så står for at kontrollere det?

PA: Der er jo nogle forskellige instanser, som certificerer den slags. Jeg ved ikke om det der, men altså.. for eksempel sådan nogen som TNO i Holland, det er også sådan nogen, der lave PCI certificeringen. Og sådan nogen som T-Systems i Tyskland, jeg tror også de laver IRL, men jeg er ikke sikker på det. Det er sådan nogen laboratorier, der går ind og kigger på software og hardware og kigger på om de mener at det er så sikkert, som man selv påstår, at det er. Der har også tidligere været nogen andre - jeg kan ikke huske hvad de hedder - hvor man også faktisk går ind og certificerer hele processen omkring fremstillingen af de devices, således at man ikke kommer og får certificeret en eller anden platform man har lavet, og så ligger source-koden der hjemme i en eller anden papircontainer ude på parkeringspladsen.

MR: Det er vel også det, der bliver brugt i smart card industrien

PA: Ja. ... Det lyder fornuftigt. Jeg kan ikke lige gennemskue om der er nogen decideret åbne uhensigtsmæssigheder der. Altså der er sådan.. den måde man bør lave den slags på.. man kan jo lave en analogi til hvordan man distribuerer digitale signaturer. Det er lidt samme koncept vi er ude i. Og der skal man jo bare sikre sig, at hvis man har behov for at vide hvem det er man distribuerer identiteten til, så skal man på en eller anden måde sikre sig, at man ved det.

PAN: Jeg tror jeg faktisk jeg har set et eksempel, hvor det er man har brugt.. devicen har et certifikat.. altså har et nøglepar, hvor fabrikationsenheden har certificeret det certifikat som et engangscertifikat. Og første gang enheden henvender sig får den så udstedt et nyt certifikat, afhængigt af hvem der skal være operatør, for eksempel. Det var en mulighed.

MR: Jeg har et andet lille skema her, hvis du vil kigge på det; omkring hvordan det her subscription management skal foregå. Og hvad for noget der er vigtigt, og hvad for noget som måske er mindre vigtigt.

PA: "Users can buy the device independently of an operator"

MR: Hvis man for eksempel tænker på en entry-level telefon eller et handset eller sådan noget. Er det så vigtigt at brugeren kan købe den i en forretning, hvor den ikke er tilknyttet en operatør fra starten af, eller ville det være acceptabelt at man kun kunne købe den slags telefoner gennem operatører.. hvor de så kunne have deres subscription lagt på telefonen fra start af.

PA: Altså, hvis man kun kan købe den med et fast abonnement i, vil man så efterfølgende kunne skifte operatør.

MR: Ja, det antager det her spørgsmål, ja.

PA: Fordi der er klart, at der er jo et ønske, og det er jo klart at, hvad skal man sige.. hvis man skal leve op til de almindelige regler fra konkurrencemyndigheden, så skal der være en hvis frihed. Altså så kan man jo ikke låse en.. [...] Men altså man kan sige der er jo.. der er kommet et større marked for nye telefoner der ikke kommer fra en operatør. Det har der jo ikke været specielt meget tidligere. Men altså.. hvis man kigger på device, tænker man sig så at man måske kan købe den fra flere forskellige operatører.. men at man så vælger hvor man kan købe den.

MR: Ja, lige som der er nu, men telefoner, som er sponsoreret.

PA: Ja. ... Vigtigheden er det ud fra... ?

MR: For succesen af der her system med virtuelle SIM-kort.

PA: For succesen. Så det er ikke set ud fra operatørens eller brugerens synspunkt.

MR: Nej, men det vil jo også være et kriterium for succesen at de to parter synes at det er en god ide, så..

PA: Ja. Nå, men du kan jo sige, at operatøren vil altid synes noget andet end abonnenten, fordi operatøren vil jo selvfølgelig helst have at man ikke kan churne til et andet selskab.

MR: Så de vil synes det er en god ide.

PA: De vil synes det er en rigtig god ide, at man ikke kan skifte.. Men brugeren vil jo helst kunne skifte nærmest dagen efter at kan har købt. ... Altså, det er jo et marketing spørgsmål. Det er jo ikke sådan rigtig teknisk, eller hvad kan man sige. Det er rigtig svært for mig at svare på.

PAN: Er det er relevant spørgsmål?

PA: Ja, men det er jo relevant at have en ide om det, kan man sige, men det er nok nærmere...

MR: Vil det være vigtigt for brugeren? [...] Vil brugeren nægte en telefon, hvis den kun bliver solgt hos operatører?

PA: Ja.. men samtidig, hvis du siger at det er en eller anden Nokia-telefon, som alle operatører sælger, ikke, så kan man bare gå til en anden operatør, hvis man er skide hamrende utilfreds med TDC, eller hvis man står i deres debitor-kartotek, eller sådan et eller andet. Så hvis jeg skal sige noget så er det måske vigtigt.. "somewhat important". Der er et svært spørgsmål.. et rent marketingspørgsmål.

[...]

PA: "The end user can change operator during the life of the product", det mener jeg er "very important". Fordi det er noget folk virkelig gør, altså. Men hvis du spørger operatørerne, så er det nok noget andet ikke.

PA: "During the life of the product, the operator can be changed remotely, while the device is in the field"

[...]

PA: Jeg har skrevet her, at jeg mener at det er vigtigt, at man kan skifte operatøren remotely. Fordi, det skal jo helst være nemmere end at skifte et SIM-kort.. det skal jo også være en af fiduserne ved det her. Men jeg mener at den her SIM-lock funktionalitet, den skal på en eller anden måde være der. Men derefter, så skal det være nemt at skifte operatør.

PA: "The operator can be changed remotely, even when in situations where the subscription has stopped working". Altså, hvad er det specielt for nogle situationer der tænkes på?

MR: Det her spørgsmål det er nok mere ment til M2M.. Hvis nu man har to millioner biler ude i marken og ens kontrakt løber ud, eller operatøren går fallit, eller der sker et eller andet, som gør at man har brug for at skifte SIM-kortene. I det her tilfælde med biler, så vil det nok være ret vigtigt, at man kan skifte dem uden at alle bilerne skal indkaldes til eftersyn. Men mere generelt, tror du så at det er sådan en.. kun i det her tilfælde.. eller er det noget man generelt kunne forestille sig var nyttigt?

PA: Ja som sagt, jeg mener man skal have mulighed for det, og det er sådan set uanset om abonnementet virker eller ikke virker. Altså, man skal jo have en frihed til at skift over. Altså hvis man går helt galt af operatøren og skal i retten, eller et eller andet, på grund af at man ikke er enig i telefonregningen, så skal man jo have sin frihed til at flytte over til noget andet. Det er jo også det man.. i dag der kan du jo bare hive SIM-kortet ud og skift ud med noget andet. Og det er klart at du jo have en langt større fleksibilitet hvis.. ja ok, selvfølgelig, hvis du har en SIM-lock, så kan du selvfølgelig ikke gøre det. Altså, det er en fleksibilitet men er nødt til at have. Man kan jo sige at et virtuelt SIM skal jo ikke have begrænsninger i forhold til et almindeligt SIM. Så jeg synes det er vigtigt.

PA: "The operator can be changed remotely, even when the device is not under network coverage of the new operator". Altså, det er jo sådan et spørgsmål om at man tror på at alle operatører vil tilslutte sig denne her løsning med at have sådan en signallerings-båndbredde til den slags. Altså hvis man tror på at det er muligt, så er man også nødt til at tro på at det her kan lade sig gøre. Så hvis det her setup ligesom skal fungere, så skal det jo også være muligt.

[...]

PA: "The device has a backup subscription that can be used in the event that the normal mobile subscription stops working". Altså det her med backup subscription det er jo ikke noget man bruger i dag. Og med det setup der er beskrevet, så vil man jo til hver en tid kunne flytte over på et nyt, hvis der er et problem. Så det mener jeg ikke er vigtigt.

MR: Et andet spørgsmål kunne være: Er det vigtig at man kan mere end et subscription i telefonen, som man kan skifte imellem.

PA: Altså, jeg kender godt operatørernes svar til det: Det er ikke noget der er voldsomt interessant. Fordi, det gør jo at man vil have mulighed for at switche sin trafik, ligesom man gør på fastnet. Det vil de absolut ikke være interesserede i. Men en anden ting er at som det virker i dag, så kan man faktisk ikke lave en GSM device, der altid virker. Fordi, du kan ikke spejle en GSM forbindelse. Du kan ikke have to SIM kort der har det samme nummer. Og det er jo et klart problem. Men det er jo spørgsmålet om ikke det vil være løst.. det er jo en anden problematik man også kan overveje: Om man kan downloade samme identitet til flere devices, og om de må være på nettet samtidigt. Der er jo nogen der tillader det, og andre gør ikke. Og grunden til at de ikke gør det er, at man betaler licens til ens HLR-leverandør og sådan noget per antal brugere. Og jeg tror nok, at hvis man putter to IMSI Ki'er på, så vil det være to brugere, selv om det er det samme. Sådan er der nogen, der har det. Og det kan være det, der er begrænsningen; at det simpelt hen er et spørgsmål om penge, at man ikke vil gøre det. Men det er noget af det, man skal overveje. ... Men det er da et interessant spørgsmål. Men altså, jeg mener igen, altså.. at hvis man har et åbent system, så vil du altid kunne downloade en ny subscription, hvis det første ikke virker..

PAN: Er der nogen ting der mangler, man kunne spørge om.

PA: Til en kravspecifikation? Altså det her er jo nogle meget bløde spørgsmål kan man sige? Du er sådan i opstartsfasen? ...

PAN: Hvad kunne et hårdt spørgsmål være?

PA: Jamen det mere det rent tekniske om hvordan man sikrer sig at.. Altså.. i dag har man jo eksempelvis ikke noget.. kan man ikke kommunikere sikkert med et SIM kort. Bør man kunne det? Bør man kunne lave en sikker session til et virtuelt SIM-kort af en eller anden grund.

PAN: Jeg syntes også du nævnte på et tidspunkt at man havde mulighed for at lave et eller andet microchannel helt nede på chippen.. så man kunne opdatere et eller andet langt fra. Noget med at lave en krypteringstunnel fra chippen til en eller anden central device.

PA: Altså, SIM-kortet har ikke indbygget sådan en secure messaging funktionalitet. Men omvendt har de jo OTA opdateringsmulighed. Men altså, det er nogle ting man er nødt til at tage med. Altså OTA opdateringer det er et must og have med. Fordi, der kan bare være nogle ting man ønsker at ændre, altså hvis man har lavet en.. Ja der er jo nogle der går ind og retter i koden i SIM kortene ved OTA opdateringer.. Hvis man finder ud af at der er en adgangsparameter, der er sat forkert i kortet, kan man gå ind og opdatere det. Men altså, det er jo selvfølgelig vigtigt at gå ind og overveje, hvad skal man have i et virtuelt SIM-kort: Er det bare data, eller er det også noget funktionalitet. Fordi hvis man ønsker at kunne tilbyde noget af det samme som SIM toolkit applikationer, jamen så er man nødt til at overveje måske at have et eller andet sikker Java miljø at afvikle tingene i, som kan afvikle en eller anden form for signeret applet man kan downloade i handsettet.

[...]

MR: Nu foreslår de i det her SA3 forslag.. De foreslår to muligheder til hvordan man kunne sikkert integrere SIM-kortet i telefonen. Og det som det er nu ikke, så har du et SIM-kort, du kan tage og fjerne. Den ene mulighed det er, at du simpelthen bare lodder SIM-kortet fast inde i telefonen. Så kan man jo så forestille sig at det skulle være et SIM-kort, som ikke bare lige var ejet af én operatør, men hvor man så kunne downloade noget kode til det, afhængigt af hvad for en operatør man vælger. Men stadig sådan et smart card, som bliver loddet fast. En anden mulighed, det er, at den kører, som du nævnte, inde i noget trusted computing agtigt. I en sikkerhedschip, som allerede er integreret i telefonen i en større sammenhæng, hvor SIM så bare er et stykke software, som bare kører blandt så meget andet, i en sikker mode. Har du nogen ideer til hvad der kan være fordele og ulemper mellem de to muligheder?

PA: Jeg skal lige forstå.. Hvordan ville man så loade data i sådan en?

MR: Det er udefineret. Det ved jeg ikke. Men det ville være en dedikeret smart card chip, som kun blev brugt til SIM-funktionaliteten. Man kunne jo forestille sig, som SIM-kort er nu, med Java kort. Hvor, som du siger, at USIM applikationen det bare er en anden Java applikation. Så kunne det jo bare være et helt almindeligt smart card, hvor man har lavet et eller andet provisioning system ved siden af... Hvor man sender en operatørs Java appletter ned i den.. og filsystem. Og så kan den passe sig selv.

PA: Altså, jeg vil sige, et software SIM er jo måske interessant i en almindelig telefon. Altså hvis vi taler M2M, så skal det helst være noget med få ben på, der ikke koster så meget. Der vil man jo skulle have, går jeg ud fra.. det er sådan lidt afhængigt af hvordan man ser på det, ikke, fordi i M2M, der har du også.. du kan købe en ret avanceret telefon som et print kort du bare stopper i, og som egentligt ikke gør noget specielt. Men jeg tror bare, at.. det er jo sådan lidt det samme, ikke. ... Det er nok sværere at certificere noget software, som man kan ændre på. Det er nok der problematikken kommer. Altså i betalingsverdenen, der ved jeg i hvert fald, at der er man mere glad for at have noget, som man kan skrue nogle skruer på og putte noget maling på, så man ved, at der ikke bliver ændret på det. Så det du tænker dig her det er sådan, at man simpelthen har et eller andet library, man linker med i sin terminal-kode, og så laver den bare SIM funktionaliteten. Hvordan skulle den så have adgang til nøgler?

[...]

PA: Det jeg tror egentligt er forskellen på de to løsninger, det er jo om.. altså, hvor let det er at certificere, ikke. Fordi lige så snart du har noget i software er det rigtig svært at holde styr på hvad det er, der foregår. Det lyder lidt som om, at det koncept, du beskriver, at det er sådan noget, som ikke alle og enhver hacker lige kan gå ind og lave om i. Men jeg tror, det et eller andet sted er det, som det står eller falder på. Det vil jo altid være nemmere og have en eller anden chip du lige knalder i, og den koster syv kroner, og den er bare sikker, det ved alle.. end det vil være at have noget der sådan er halvt sovset ind i operativsystem og sådan noget der, som måske, hvad kan man sige.. igen, det er sådan et forklaringsspørgsmål over for operatørerne, at man skal bare sørge for at det ligesom er sikkert at loade sin IMSI Ki ned i.

PAN: Dvs. ulempen ved den er, at der er en eller anden mand, der kan sætte sig til og måske lytte på benene, og sætte en eller anden device på, og måske ændre trafikken undervejs... Hvis det er en ekstern crypto-enhed eller hvad skal man sige, sikkerhedsprodukt. Hvorimod her der kan være større sikkerhed i og med at det er integreret.

PA: Ja, men det mener jeg så ikke, fordi at et eller andet sted så når du autentificerer mod GSM netværket, så modtager den en challenge og sender en response. Og der skal man bare være sikker på at der ikke sker sjove ting i mellemtiden. Og det er man måske mere sikker her end man er her. Hvorimod, at det måske - det er så afhængigt af hvor simpel den skal være ikke - men det er måske nok nemmere at lave en eller anden form for envelope-funktionalitet til at downloade i sådan en her end i sådan en. Ja, så kunne der være nogle ekstra ting i.. hvis man nu siger at udover at det er en UICC, så kan man også downloade ting sikkert til den.

MR: Så det er et spørgsmål om certificering...

PA: Det er nok der begrænsningen ligger. Fordi når ting bliver mere fleksible, bliver de også mere åbne.

MR: Men kunne man forestille sig noget, som f.eks. sådan en Trusted Platform Module, det er jo ikke et bestemt chip design, men specifikationer, som en chip skal kunne. Hvis man kunne implementere det, og få det certificeret til at det møde de specifikationer. Kunne man forestille sig at det var nok til at vise at den har de og de sikkerhedsfunktioner, og at det er grundlag nok til at beskytte en SIM-funktionalitet?

PA: Ja altså dem du skal spørge, det er jo dem der har nøglerne, faktisk.

MR: Så det er operatørerne jeg skal hen og snakke med

PA: Ja. ... Altså, jeg ved at Sonofon har jo lavet nogle tests med et softSIM med Kampstrup, som er sådan nogen, der lave målersystemer og sådan noget... hvor de har kørt nogle tests. Men hvor sikkert eller usikkert det har været, det ved jeg ikke. Men generelt set findes der jo ikke rigtig noget.. altså der er jo ikke noget i vejen for at man kunne have lavet sådan nogle system allerede, men det ser ud til at være specielt udbredt. Af hvilken grund ved jeg ikke. Altså, jeg tror det er noget med at de helst vil holde de der nøgler tæt til kroppen, for det er et eller andet sted deres egen, som de helt ikke skal slippe.

MR: Jeg har også hørt noget om at Gemalto også har lavet sådan nogle SIM-kort, som man kan lodde fast. Det er heller ikke noget de reklamerer med nogen som helt steder. Jeg ved ikke om det er fordi de vil beskytte deres forretning, eller..

PA: Nej men altså.. mange af pengene i at lave kort ligger jo i personaliseringsdelen. Og hele det at pakketere og levere nogle services. Og det er klart, at det bliver jo lidt mere komplekst ikke. Det er jo ikke sådan at det ikke kunne lade sig gøre. Jeg tror i virkeligheden at der er mange der er villige til at betale en hel del mere for sådan et SIM-kort, hvis bare man kunne lave det lidt mere robust. Men det er jo nogle standard chip typer, der bliver brugt i dag til SIM-kort, og der er ikke noget i vejen for at man kunne få det i en standard-packaging. ... Så det var et rigtig godt mudret svar jeg gav på det her.

MR: Jeg har kun et skema tilbage. Og det hænger sammen med muligheder og ulemper - eller opportunities og threats - ved alt det her med at flytte SIM funktionaliteten fra SIM kortet til telefonen. Jeg ved ikke om du har nogle forslag til hvad der eventuelt kan stå, og hvor vigtige de ting er, og hvad sandsynligheden for dem er. Nu har jeg nogle opportunities her jeg har lavet.

PA: Hmm.. "new business opportunities [in the M2M sector]" ... Altså, det er det mest oplagt ikke. Og det er skide besværligt at bruge et SIM kort i et jordspyd der sidder ude i en eller anden mark, der skal måle luftfugtighed eller et eller andet. Så derfor så.. hvis man skal have noget, der virkelig skal drive det her, ikke, så synes jeg det er et godt sted at starte. Og jeg ved også, der er en efterspørgsel efter det.

MR: Hvad er sandsynligheden for at det kan lade sig gøre at få gennemført det?

PA: Ja.. altså, hvis man kigger på ideen og efterspørgslen alene, så vil jeg sige, at den er der allerede.

PA: "increase competition in the M2M market" [...]

MR: Når du skriver 5 her, tager du så også hensyn til om operatørerne er villige til at...

PA: Nej, altså det er det jeg siger, at.. hvis man kigger på muligheden alene og den efterspørgsel der er, og glemmer alt om de forbehold, operatørerne tager.

PAN: Så det er ren og skær set ud fra markedet, og ikke for dem, der er de aktører...

PA: Ja, fordi jeg mener jo at.. probability den er certain, fordi efterspørgslen er der i dag, det er ligesom derfor du er startet på det her. ... "increased competition".. hvilken form for competition menes der?

MR: Omkring operatørerne.. at der bliver større konkurrence mellem dem på M2M markedet.

PA: Ja, men det ved jeg ikke om jeg... Ja der bliver selvfølgelig større konkurrence for dem der indfører det først ikke. I og med det er noget der er efterspurgt. Men når de alle sammen har det, så vil det igen være svært at...

MR: Du synes ikke det er særligt relevant?

PA: Nå, men man kan sige, at hvis man ikke har det...

MR: Ideen med spørgsmålet er, at hvis man har et virtuelt SIM, så vil det være nemmere at skabe konkurrence mellem operatørerne fordi man ikke skulle hen og skulle skift SIM kort i enhederne, og at man måske ville kunne presse operatørerne lidt på prisen. Mere end man ville kunne end hvis det var besværligt.

PA: Ja... altså det er da klart, at du vil da få en højere churn. Men jeg tror bare at operatørerne også samtidigt vil bibeholde de mekanismer de har i dag med at låse ting fast, og sørge for at man kontraktligt har en forpligtelse. Og hvis ikke man kan lave en SIM-lock funktionalitet, så vil man muligvis bare sige, jamen du skal betale uanset om du skifter operatør.

MR: Så det mener du ikke er særligt vigtigt..?

PA: Jo, men man kan sige, jeg tror da.. jeg tror du har ret i at det selvfølgelig vil stramme eller skærpe konkurrencen. Men jeg tror også, at de vil gøre deres til at man ikke bare sådan kan flakse rundt mellem alle mulige. Jeg vil sige at den er i hvert fald 3. Og hvad kan man sige..

competition generelt er en vigtig parameter når man laver ting.. fordi det kigger de på hele tiden.

PA: "eliminate cost of the SIM card for operators".. Jeg tror ikke så meget at det er udgiften ved det, fordi den er minimal, og det er noget, som de sagtens kan dække ind.

MR: Også logistisk?

PA: Nej. Og det er det jeg tror.. Jeg tror det er mere de praktiske problemer.

MR: Jeg hørte noget om i slutningen af 90'erne, hvor der var en stor efterspørgsel på SIM kort lige pludselig, fordi der kom mange flere kunder til. Og samtidig var der problemer med at levere silicium til det..

PA: Ja, der var lige lidt jordskælv over i Thailand

MR: ...så der var nogen operatører, som stod uden SIM kort. Og det er en af årsagerne til at man nu har flere SIM-kort leverandører, som regel. Man kunne forestille sig, at de ville være glade for at skulle slippe for det problem.

PA: Ja. Helt sikkert. Jeg mener også, at den er rigtig høj, for det er noget de fokuserer meget på. Og de skal hele tiden sørge for, at de har nok stock af SIM-kort. ... "probability".. den er høj. .. "importance"..

PA: "introduction of entry-level handsets without SIM card" ... Altså du tænker på at man.. uden for M2M.. Altså, jeg mener at det er en oplagt ting også at flytte hele den her meget dynamiske ID over i almindelige handsets.

PAN: Ved du hvad entry-level handsets er for noget?

PA: Ja, det er sådan en skod-telefon, går jeg ud fra. Men i virkeligheden kan det overføres på alle typer handsets. Jeg mener også at det er "very likely" at man går over til sådan noget, hvis man allerede har infrastrukturen til at gøre de her ting.

PA: "smaller handsets. more flexibility" .. Jeg tror sgu ikke de kan blive mindre. ... Det er jo vigtigt i M2M markedet, at det er kompakt. Men jeg tror ikke at det er noget.. jeg mener ikke at det er vigtigt.

PA: Så kan man sige, "Threats". ... "Too much churn"

MR: Er det operatørernes frygt?

PA: Ja.

MR: .. sandsynligheden for at det bliver en showstopper ..

PA: Ja.

PAN: Ville Ane vide noget om det? .. Hun ved i hvert fald noget om churn.

PA: Hun ved jo noget om taletidskort. Hun ved jo noget om de hersens abonnementer, som folk betaler mange penge for fordi de skal lave nogle suspekte ting. ... Men altså, det er klart, at hele det her med at shoppe rundt på.. altså man kan sige.. hvis man nu tager på ferie på Tenerife, vil man så også lige snuppe et abonnement dernede fordi det var billigere. Det er selvfølgelig en oplagt ting, ikke altså. Og det er alle de der mekanismer man ligesom skal have et overblik over.. hvad kan folk finde på, og hvad bør man kunne gardere sig imod.. hvad vil operatørerne stille som krav for at det her skal fungere. Og det tror jeg nok jeg vil spørge operatøren om. Men det er helt sikkert, at de fokuserer på det.

PAN: Kunne Thomas [Thomas Hensing, TDC] være en mulighed at spørge om det?

PA: Nej, fordi han har ikke noget produktansvar. Det er nok Anes kollega, der sidder med ansvaret for post-paid subscriptions i TDC, der måske ville kunne sige noget om det.

PAN: Så det kan vi spørge Ane om, eventuelt, hvad navnet er på vedkommende.

PA: Så kan man sige.. jeg vil også sige, at det her med, at man kræver en infrastruktur.. "need for new signalling infrastructure"..

MR: Så er det et spørgsmål om, hvad man finder på der, og hvor kompliceret det er.

PA: Ja. Altså, jeg tror, at det er nok farligt, at udtænke noget, som kræver det. Altså burde man ikke kunne lave en løsning som kunne eksistere i dag og gøre det samme.

MR: Eller maksimum kræver en firmware update, eller sådan et eller andet.

PA: Ja.. Ja, men altså jeg tror ikke på, at man.. Altså selvfølgelig skal du have en ny terminal, eller noget nyt hardware af en eller anden art.

MR: Hvis det kunne være centralt hos operatørerne, så er sandsynligheden for, at...

PA: Ja, men jeg ville nok sige, at så skulle Nokia måske bare lave - eller hver producent - bare lave en roaming agreement med en masse operatører, således hvor i verden man tænder så ville man i hvert fald have de der 5 kroner det ville kræve for ligesom at få de 3 SMS'er.

MR: Ja det ville så være i stedet for den her "Registration Service", så kunne ham her også lave roaming agreements med folk.

PA: Altså, jeg tror stadig du skal have en "registration service", men jeg kan bare ikke lide tanken om, at hver operatør skal være med til at man skal bruge deres netværk til det her, på en eller anden måde. For det vil et eller andet sted kræve, at...

MR: Ellers kunne man forestille sig nogle begrænsninger i telefonen med at.. hvis Nokia, eller hvem det nu var, de gav telefonen et abonnement fra start af, som var fuldt funktionelt, ud over at man indbygget i telefonen kunne lave noget call control, som gjorde at man kun kunne ringe til Nokia og så et SIM-kort eller et eller andet. ... så i ham hers [VPLMN] øjne, så var Nokia - eller en eller anden - bare en helt almindelig operatør.

PA: Ja, det tror jeg ville være det rigtig, for så har du klaret hele infrastrukturen med det her med, så behøver du ikke... lad mig se. ... Du sagde noget med at man skulle have en begrænset ret til nogle ting. Spørgsmålet er om man kan implementere det i dag, fordi du kører jo allerede.. i forbindelse med at du attacher til et netværk i udlandet kører du noget roaming, hvor du ryger helt tilbage til ding egen operatør, og hvor den udsteder nogle midlertidige IMSI'er og sådan noget, som den downloader til den lokale operatør.

MR: Ja, det er i sammenhæng med den der autentification vektor.

PA: Ja.. altså.. hvis jeg skulle lave sådan et system ville jeg sørge for at det ville virke i dag. Med den infrastruktur vi har i dag. Og så kan man sige, at så kunne det godt være, at du roamede tilbage til den registration service, eller hvem det nu var.. og så kunne den kun bruges til det.. og så når den ligesom var oppe at køre, jamen så kunne den så logge på med den rigtige identitet.

MR: Det vil sige, at hvis ikke man gør det, så er sandsynligheden for at det her det ikke fungerer, den er ret høj.

PA: Ja.

[...]

MR: Lige en anden trussel, som du gerne må overveje... Det er, at hvis operatørerne skal sige god for device'en inden de sender deres SIM-kort.. er der så en risiko for at de kan lægge pres på device manufacturer'en eller sådan noget, ved at banne deres devices, simpelthen, fra deres netværk, og ikke lige sælge SIM kort til dem..

PA: Jamen.. det kan de jo et eller andet sted allerede i dag, fordi.. operatørerne lave nogle aftaler.. de går ind og tester alle terminaler i dag, og så anbefaler de terminalerne..

MR: Er det type approval du snakker om?

PA: Jaa.. altså, jeg ved ikke rigtig hvad der sker, hvis man ikke har type approved sin terminal.

MR: Der sker ikke noget. Det har jeg nemlig spurgt om, og de prototyper, som Nokia går og bruger internt, de er jo ikke type approved, og de virker udmærket - også uden for huset. Og jeg læste specifikationerne, og der står, at operatørerne har mulighed for at check IMEI'en, og checke om den er type approved og om den er ok, men det bliver ikke brugt. Og jeg tror det er outdated.

PA: Jeg ved ikke engang om de lave tyveri-check på IMEI'en mere..

MR: Og IMEI'en er jo.. altså, i teorien så kan man jo ændre den. I Nokias moderne telefoner bruger de det der sikkerhedssystem bl.a. til at sørge for at IMEI'en ikke bliver faked, men i de fleste telefoner, kan man lave den om.

PA: Jamen altså, det er da klart, at der vil de have en pressionsmulighed over for leverandøren. Men altså.. jeg tror ikke det er nogen stor trussel. Fordi der e jo i dag allerede et samarbejde.. altså man kan sige, at hvis leverandørerne ikke allerede samarbejdede med operatørerne, så ville, altså.. operatørerne går jo tit ind og køber en masse telefoner, som skal subsidieres, ikke. Og derfor, hvis ikke man ligesom har den dialog, og det tillidsforhold, jamen så... Jeg tror ikke det er nogen stor trussel, det må jeg sige.

MR: Hvad med sådan noget som liability. Hvem ville være ansvarlig, hvis nu der var et eller andet der gik galt med et her... hvis nu operatørernes netværk blev misbrugt eller sådan noget.

PA: Ja.. det er jo faktisk et godt spørgsmål, fordi.. I dag, der er jo.. hvis du kigger på analogien til betalingsverdenen, der er det jo Visa og MasterCard, der kræver at man skal have de her PCI certificering. Og derved er det jo også dem, der står inde og betaler, hvis der er svindel i systemet. Hvordan det vil være her, det er jo et rigtigt godt spørgsmål, for der er jo ikke nogen global organisation, der kan gå ind og sige, nå men det dækker vi. Som jeg ser det, så må operatørerne ligesom selv sørge for, at det er sikkert nok. Men hvis TNO eller T-Systems så går ind og laver en fejl i deres certificering, så er det operatøren der har et problem.

MR: Kunne man forestille sig, at det skulle være Nokia, for eksempel, der stod med ansvaret?

PA: Altså, det er jo selvfølgelig Nokia der står med ansvaret hvis der er for mange mikrobølger i telefonen, og man får ristet hjernen, når man bruger den. Så jeg ved ikke om man kan bruge analogien over til det.. det er måske ikke relevant. Det er et interessant.. Det er i hvert fald noget, som man bør overveje.

MR: Tak. Jeg har ikke flere spørgsmål.

[...]

PAN: Er der nogen ting, som du mener.. [...] er der nogen ting der mangler.. som du vil tilføje ud over det i har talk om indtil videre?

PA: Nej.. altså nu har jeg måske ikke overveje så meget de her lidt mere bløde spørgsmål. Altså det jeg sådan først kommer i tanker om når jeg hører virtuelt SIM, det er alle de her tekniske ting, man skal lave.. og hvordan man skal abstrahere og kunne forudse, hvordan verden ser ud i fremtiden. ... Nej, altså jeg tror vi har været omkring de væsentlige ting. Og det er jo det, at sikre, at det er sikker, og at man kan få det valideret, og at man kan bevise over for operatørerne at det her er lige så sikkert som, at de putter deres nøgler i SIM kortet.

A.3 Jens Benner, 16-10-2007

Subject:	Virtual SIM		
Date:	16-10-2007	Location:	Nokia, Copenhagen
Type:	Interview	Output:	Audio recording
Participants:	Jens Benner, Business Development Manager, Nokia Marc Richarme, Student, DTU		

Interview has been recorded, but not transcribed.

A.4 Valtteri Niemi et al., 19-10-2007

Subject:	Software SIM and Onboard Credentials		
Date:	19-10-2007	Location:	Nokia, Helsinki
Type:	Meeting	Output:	Audio recording
Participants:	Valtteri Niemi, Research Fellow, Distinguished RL, Nokia Jan-Erik Ekberg, Principal Member of Engineering Staff, Nokia Silke Holtmanns, Principal Member of Research Staff, Nokia Aarne Rantala, External Consultant, Nokia Peter Vestergaard, Manager, Smart Cards, Nokia Marc Richarme, Student, DTU		

Interview has been recorded, but not transcribed.

A.5 Lauri Paatero, 19-10-2007

Subject:	Interview regarding Virtual SIM project		
Date:	19-10-2007	Location:	Nokia, Helsinki
Type:	Interview	Output:	Audio recording
Participants:	Lauri Paatero, Senior Specialist, Security, Nokia Peter Vestergaard, Manager, Smart Cards, Nokia Marc Richarme, Student, DTU		

Interview has been recorded, but not transcribed.

A.6 Tjaard Meier, 07-12-2007

Subject:	Simless handset		
Date:	07-12-2007	Location:	N/A
Type:	E-mail	Output:	E-mail conversation (2 mails)
Participants:	Tjaard Meier, Senior PPM Concepting, Nokia Marc Richarme, Student, DTU		

From:	Richarme Marc (EXT-Adecco/Copenhagen)
Sent:	Thursday, December 06, 2007 9:28 PM
To:	Meier Tjaard (Nokia-MP/Beijing)
Cc:	Benner Jens (Nokia-MP/Copenhagen)
Subject:	Simless handset

Hi Tjaard

I'm currently involved in a project studying the possibility of removing the SIM card from handsets. On this matter, Jens Benner advised me contact you regarding the benefits for Nokia.

Specifically, I would like to know the cost of the SIM card connector in mass produced phones, and your oppinion on the value of the saved space if the SIM disappears: I presume that the PCB floorspace could be used for other purposes or the phone made smaller, and additionally I presume that the placement of the SIM connector poses some design constraints. But I'm not sure how important those factors are? Would removing the SIM card make it possible to make phones with more features than is possible with the current design constraints?

I'm looking forward to hearing your oppinion on this matter.

Best regards
Marc Richarme

From:	Meier Tjaard (Nokia-MP/Beijing)
Sent:	7. december 2007 05:33
To:	Richarme Marc (EXT-Adecco/Copenhagen)
Cc:	Benner Jens (Nokia-MP/Copenhagen)
Subject:	RE: Simless handset

Hi Marc,

The cost is only relevant for products on extreme low end of the portfolio. It consists of the connector, some holding and shielding features in the mechanics and some electrical components related to the interface and protection from for example Electro Stratic Discharge. I have no detailed breakdown at this point but I would estimate this to be only 0.10 to 0.12 EUR.

The main benefit would come in terms of size. The SIM card is one of the main blocks to place when creating a product architecture. To make a phone thin the trick is to place things next to eachother. The space next to the battery is available for the chipsets (processors, RF chips, memory, FM radio, BT, GPS, ...), a camera, audio components (microphone, speakerr, vibrator) and also the SIM card. Stacking of these elements is an option (up to the battery thickness) but poses a problem for components that need an electrical connection to the board which all of these components do need. Our Barracuda product (Jens can show you one) is a good example on how we manage to stack the SIM but keeping it next to the battery. A technical challenge integrating a SIM is that the lines are sensitive to coupling from the antenna and need to be shielded. This is why it is uncommon to place the SIM card far away from the chipset (the connection lines would act as receiving antenna's). One additional challenge with the intergration of SIM cards is that they have to be user changeable. So ease of use is another complicating factor.

Hope this helps,

BR...Tjaard

A.7 Claus Rasmussen, 29-12-2007

Subject:	Meeting regarding Virtual SIM concept		
Date:	29-12-2007	Location:	Copenhagen
Type:	Discussion	Output:	Audio recording
Participants:	Claus Rasmussen, Field Application Security Engineer, Cloakware Marc Richarme, Student, DTU		

Interview has been recorded, but not transcribed.

A.8 Ross Campbell, 07-01-2008

Subject:	SIM-less phone discussion		
Date:	07-01-2008	Location:	Telia Danmark, Copenhagen
Type:	Meeting	Output:	Notes
Participants:	Ross Campbell, System Manager, TeliaSonera Marc Richarme, Student, DTU		

Key numbers:

Price of SIM card:

Native OS: 0.6 EUR	JavaCard 64k: 1.4 EUR

Packaging & logistics costs per issued card:

High-end: 1.5 EUR	Low-end: 0.6 EUR

Number of different SIM card suppliers for TeliaSonera:

Up till now: 6 From now on: 3

Number of SIM cards ordered per year: 800.000

(for 1.300.000 subscribers in Denmark – this number is a bit higher for Telia than other operators, maybe because of higher number of pre-paid subscriptions, and because of the Danish market, where users throw their phone away after 6 month and then churn or buy a new subsidized phone, including new pre-paid card)

VAS:

SmartTrust OTA platform can manage cards from all vendors. OTA SMS format is the same, but with proprietary extensions

Telia was originally interested in Smart Card Web Server (SCWS), but lost interest and is not planning on using it. SCWS needs SIM OTA updates in order to be updated, whereas the operator portal can provide up-to-date VAS content in the browser interface. Telia is pushing its customers to adopt GPRS/internet so that more will use the portal / VAS services.

Telia is using the STK operator menu for VAS, but most operators agree that the STK menu will soon reach its end of life, partly because it is usually buried in the phone menus.

Main focus for VAS is on the WAP portal.

SIM Toolkit:

Used in TeliaSonera:

- Operator Menu: Not so important (see above)
- Smart Roaming: reorganizes PLMN list (e.g. some Nokia phones pick a “random” network if none of the top eight entries in the list is found), uses network measurements, timers and refresh command to get the phone to use the best available network.
- IMEI Tracker: Sends SMS to operator when SIM is inserted in new handset, such that operator can update device settings.
- Dual IMSI: Complex application allowing to change between home / work subscription by

rebooting phone and entering different PIN code.

Others:

- IMEI Lock: Used by e.g. delivery companies, M2M, to lock the SIM card to a single terminal, so that it cannot be abused if stolen.

OTA Updates:

- File management:
  - PLMN list update
  - Rebranding (operator name)
  - Phonebook updates (FDN, ADN) e.g. new customer service number
- Updates to operator menu
- Card applet management: rarely used – typical applet is 3kb requiring 26 SMS messages; big applets can use up to 50 SMS messages per subscriber

90% of updates fit inside single SMS – this might partially be due to the complexity of larger updates.

Applet of 3kb size (26 SMS) takes approx. 8–10 min. per subscriber. Update of PLMN list for all subscribers takes 1–3 month.

SIM Logistics:

When SIM card stocks run low, a batch of 50,000–100,000 cards is ordered to the card vendor. Operator and card vendor have previously agreed on card profiles defining the content of the card (features, file system structure, applications, etc.), the graphical design, etc. A request file is generated and sent electronically to the card vendor containing IMSI/ICCID range and card profile. Card vendor have a stock of card branded for the operator, which is then personalized upon reception of this request. Card vendor generates all card-specific numbers and personalizes the card electronically (IMEI, IMSI, Ki & OTA keys, MSISDN – only for prepaid card, as postpaid MSISDNs are never present on the card) and physically (IMSI and MSISDN (for postpaid) etched to the card). Typical SIM card (64k JavaCard) costs 1.4 EUR.

The entire batch of cards is sent to a logistics packing company (3rd party), which makes the physical package, including user agreement document (which must be enclosed as a physical document by law in Denmark – in some other countries it is enough to refer to a web page,) and adds stickers with IMSI/MSISDN and bar-code. This company then sends out smaller batches of packaged cards directly to retail stores (Telia shops, Merlin, ElGiganten, super markets, kiosks). Services performed by the packaging company costs between 0.6 and 1.5 EUR per card.

The card vendor also sends three files (electronically, encrypted using pre-shared key) back to the operator:

- One file that is used for the billing system (BSS) and HLR/AuC containing Ki, ICCID, MSISDN.
- Two files used for the OTA update system, one containing the ICCID, MSISDN and IMSI, and the other containing OTA update keys: ICCID, KIC, KID, KIK

Pre-paid card are activated immediately. Post-paid are activated when sold in the shop. Activation happens on the network-side, not in the card. When a post-paid card is sold, user can typically choose MSISDN, which is store on the network side, and never written to the card.

When a card is activated in the store, a request for device customization is sent to the OTA system, which is typically processed 24 hrs later. If phone is turned of this simply fails. Other operators use network equipment which detects when new handset is attached to the network for the first time and then does device provisioning.

Standardization:

Motorola's proposal was rejected; TeliaSonera voted against. Will try to find out the motivation for this. It probably is a don't-care-so-don't-annoy-people-who-do vote.

Obstacles for getting operators approval:

- Fear of losing control over device / of giving handset manufacturers power
- Fear of change / of the unknown
- Card vendor's lobbyism
- Prejudice

Pros & Cons:

Cost savings of the SIM card, and especially the SIM logistics are a big deal to operators. New handsets have good SIM lock mechanism (due to trusted hw platform): VSIM could mean that this protection level would be standard in all phones.

A VAS platform that is easier to develop for than the SIM could open up for more 3rd party developers/content providers, which could result in new services

"Perception of a less secure environment" is likely, so some work is needed here.

Algorithm:

Most (if not all) operators use MILENAGE algorithm. It shouldn't be a problem if using MILENAGE for VSIM was a mandatory. Network equipment manufacturers might not even support other algorithms.

Provisioning methods:

– "Over an insecure out-of-band channel" could be very difficult to manage in stores and kiosks.

– "Over mobile GSM/UMTS network" seems to preserve current logistics, by allowing subscriptions to be sold by "paper slip". This could also be very relevant for developing markets. This looks like the most promising solution, but needs changes to network elements, which could be challenging.

– "Using a secure token" has no real advantages over using a SIM card.

Miscellaneous:

With SIM, if user's phone breaks, he can move the SIM card to another phone, what will happen with VSIM.

Operators are testing VAS services for different SIM card configurations and with different handsets, this is a lot of combinations. Interoperability issues are common. Typically, only handset specs can be checked against VAS requirements before handset is put to market, and with high-appeal handsets (e.g. N95), handset is put to market regardless of testing/conformance. 3rd party conformance tests before handset is put to market by manufacturer would alleviate this problem.

There is virtually no fraud related to the SIM card.

SIM Lock is crucial.

A.9 Ross Campbell, 08-01-2008

Subject:	Follow-up to SIM-less phone discussion		
Date:	07-01-2008	Location:	Telia Danmark, Copenhagen
Type:	E-mail	Output:	E-mail conversation (3 mails)
Participants:	Ross Campbell, System Manager, TeliaSonera Marc Richarme, Student, DTU		

From:	Ross.Campbell@teliasonera.com
Sent:	7. januar 2008 13:21
To:	Richarme Marc (EXT-Adecco/Copenhagen)
Subject:	SIM card, OTA and provisioning

Hi Marc,

Here are the slides I used from today.

Let me know any questions you have and I'll try and get them answered, and hopefully I can arrange a meeting with someone from Fraud for next time.

Kind Regards,
/Ross.
Ross Campbell
System Manager - Common Development

Telia Danmark
Holmbladsgade 139 Mobil: +45 2610 0529
2300 København S Telefon: +45 8233 7000
www.telia.dk

From:	ext-marc.richarme@nokia.com
Sent:	8. januar 2008 13:06
To:	CAMPBELL, Ross
Subject:	Followup

Hi Ross,

Below are the notes I've taken from our meeting, in case you are interested. If you have any corrections, please let me know.

I've contacted someone from Nokia Siemens Networks regarding the feasability of the OTA provisioning system we talked about, and I'm waiting for the reply.

Another concept I've thought about: Imagine you can download a VSIM to a phone using a "paper slip" code... You could potentially download the VSIM to more than one phone simultaneously, and either only have one active phone, or have both phones active at the same time (and have both ring when a call is received.) This way users could take a different phone with them in the morning, depending on their mood or on the style of their handbag or whatever.

For operators, this could generate revenue by charging users a fee for additional devices, or by data synchronisation services & traffic.

Handset manufacturers might sell more devices, if people are using more than one at the same time.
Furthermore, this could be an additional driver for the introduction of VIM (and the consequent cost savings), and users might also see this as a justification for the paradigm change.

Any comments on the idea? How would operators percieve such a use case?
As I understand, the current barier for such a system is that operators are charged for the HLR on a per-IMSI basis, is this true?

WRT fraud, I'm still working on formulating the questions, but I hope you will be able to set up a meeting.

Kind regards,
Marc

[original message included the notes shown in section A.8]

From:	Ross.Campbell@teliasonera.com
Sent:	8. januar 2008 15:47
To:	Richarme Marc (EXT-Adecco/Copenhagen)
Subject:	RE: Followup

Hi Marc,

Your idea is good. I think similar functionality is already available with 'cloned' SIM's. We don't have this in Telia Denmark, but I understand that other countries and also compainies within denmark offer this product to customers. I'm sure the NSN guys will know more about the service that exists today, although I can see benefits of this being software based instead of on physical SIM cards.

Operators pay differently for the HLR, but you are right that most of them will be at least partly based on a licence fee for the number of IMSI's. It depends on the contract they negociate with e.g. Ericsson.

Your memory is very good. One small clarification for number of SIM card vendors: TS has about 6 just now, but has agreed only to use 3 from now on. (basically just the Sourcing department helping to minimise costs and make things simpler for the business)

I found out why TeliaSonera voted against the suggestion of Motorola, but I can't comment on it unfortunately. One of those confidential things I'm afraid.

Personally I think many operators have a focus on keeping 'ownership' of the customer and the services offered for the customer using the operators network/subscription. It's also related to the ability to guarantee a good quality of experience if the SIM unit is tested and approved by the operator so they are sure of what the customer has in hand.

Let me know when you have some fraud questions ready.
/Ross.

A.10 Stefan Kaliner, 10-01-2008

Subject:	Future of the SIM card		
Date:	10-01-2008	Location:	Nokia, Copenhagen
Type:	Meeting	Output:	Notes (only concerning SIM future)
Participants:	Stefan Kaliner, Head of UICC Development, T-Mobile International Martin Froels, Smartcard and Systems Development, T-Mobile Int. Jais Agertoft, Technology Manager, SW, Nokia Jens-Ole Madsen, Specialist, SW, Nokia Keld Stougaard, SW Design Engineer, Nokia Peter Vestergaard, Manager, Smart Cards, Nokia Marc Richarme, Student, DTU		

The opinions expressed below are the personal opinions of Stefan Kaliner on the future of the SIM card.

"In the future, we will still have mobile communication, we will still need to identify the customer, but the rest is completely open."

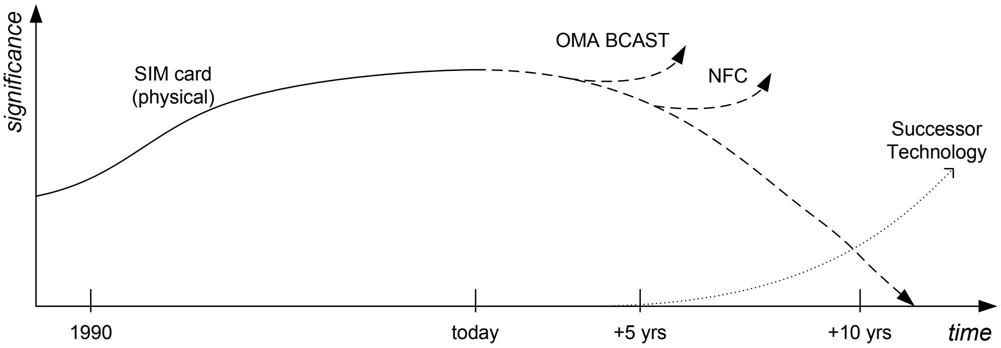


Figure: Whiteboard drawing made by Stefan Kaliner during the meeting (copied from memory).

In the early nineties, SIM was taken very seriously, and it quickly evolved. Then STK was introduced and JavaCards were introduces, which extended its role and its significance. Today the significance-curve is flattening out, and with the features introduced in the SIM it is more and more difficult to imagine how you can make a business case out of it: "USB – what is it good for?". There is a chance/risk that the significance of the SIM will fall, and that it will eventually disappear (in its current from). This could be delayed by new SIM-cased features such as OMA BCAST or NFC, which could make the significance rise again.

"In a nutshell, at some point, the SIM will be completely gone." [in the sense of a removable smart card]

It will take at least 5 years until a SIM alternative reaches the market and at least 10 before it becomes a widely-used alternative to the SIM. It is probable that a SIM alternative would first be adopted by smaller operators or in niche markets, which would then subsequently drive mass-adoption.

On the possibility of removing the physical SIM card:

A software-only SIM (without trusted hardware) might not be accepted by operators, but security is their main concern regarding a "software SIM". Assuming the security issues were resolved, their main concern would be control:

For operators, the SIM card is an asset. Its functionality and design is entirely under the control of the operator. He can potentially control every bit it contains and every service it provides.

This is why operators favour the OMA BCAST standard, which uses the SIM: The operator sits in the middle, and is in control. Of course, operators would like to keep it this way in the future.

For operators, the optimum solution would be a SIM like hardware module, integrated in the phone by manufacturers, but under control of the operator buying the handsets. This way, operators would still be in control, but avoid the disadvantages of having a removable SIM.

“The SIM has a significant logistics impact. Getting rid of it would be a nice option, but not at any price.”

Operators wouldn’t mind if users were not able to change subscription on their phone, but regulators and users would probably not accept this. “I don’t think there will be a way around operator swap.”

But they do not want a situation where users change operators every week, when there is a promotion.

Jais Agertoft points out that in some cases, this is almost the situation today in Denmark, where users can buy a subscription online (SIM arrives in 1–2 days by post) or prepaid in gas stations, etc. The old SIM is simply thrown out after a short period. Some promotions give more air-time than the initial acquisition cost, which can be a motivation for such behaviour.

On the M2M study item in SA3:

TMO are considering it, and following the discussions. There are some requirements from the automobile industry for new SIMs with better mechanical capabilities. They are not questioning the SIM itself, but just the form factor. Probably, this will result in a new SIM form factor resembling a “normal” integrated circuit. It is possible that it will be allowed to be fixed (e.g. soldered to a PCB), which is currently not allowed in the specifications. This will, however, only be applicable for M2M purposes. Some car manufacturers are currently burying GSM terminals containing a SIM deep down in the motor compartment, so that the SIM is effectively non-removable.

Impact on the eco-system:

This will have a serious impact on the eco-system. “What would we need SIM vendors for.” Gemalto is thinking about these scenarios, which can be seen by looking at their web page. They like to think of (and sell) themselves as digital security providers, and they to not mention smart cards on their front page. But in the end, they are in the business of smart cards, and a SIM card alternative would be extremely disruptive to their business.

On the possibility of updating the network infrastructure for VSIM:

“It is possible”... “It cannot be excluded that some characteristics of the network could be changed”. The advantages must be balanced.

Algorithm:

T-Mobile uses custom algorithms, and has done so both for GSM and UMTS. Different custom variations of the MILENAGE algorithm are used in different countries. The motivation for this is an assumption of extra security by not using the open standard.

When COMP128 was broken, Vodafone was affected and not T-Mobile, but the impact was very limited, so the value of having a custom algorithm could be discussed.

Many other operators have a custom algorithm as well, and operators would like to keep it this way in the future. However, it is probably negotiable, as this might be a serious obstacle for a software-SIM approach.

OTA Personalization:

Operators are very interested in being able to personalise the SIM card OTA. That is, the IMSI, etc. are only assigned to the SIM once it is in the hands of the customer. Currently these numbers are allocated when SIMs are ordered from the vendor [note: in lean-terminology, these db-entries are wasted inventory]. This is because production personalisation is expensive and because of the cost of having entries in the subscriber databases, which are unused between the time where a SIM card is ordered from the card vendor and the time where it is sold to a user.

Role of smart card vendors:

Making the SIM software is a big job. The plastic is nothing. Then there is personalization, packaging and distribution, and dealing with operators.

In some cases they also sell VAS software and other solutions to operators. TMO usually don't use this. TMO even has a self-made OTA platform.

In the case of TMO, the SIM vendor is responsible for personalisation and shipment to the stores. Some operators split up this job such that shipment is handled by another company.

When SIM vendors are developing a SIM platform for T-Mobile, the contract forces them to sell the intellectual property rights to the competitors. This way all suppliers can produce the same chip, both in terms of hardware and software. This removes interoperability issues between suppliers, and there is no dependence on a single supplier. However, if there is a problem, it is on all cards.

SIM vendors also have electronic interfaces to operators, both for ordering and for sending the resulting SIM data & keys.

Liability:

For the SIM card, the liability is distributed. Every party is liable for the production/development it does. E.g. a SIM vendor is not responsible for software he has licensed from someone else. If a problem is in the hardware, SIM vendors would hold the chip supplier liable.

For a software SIM model, this would be one of the interesting questions.

Concluding remark:

"I'm absolutely convinced that the SIM won't be around forever, there will be some successor technologies, and it is up to us to design them"

A.11 Palle Staffeldt, 10-01-2008

Subject:	SIM card idle current consumption measurements		
Date:	01-02-2008	Location:	Nokia, Copenhagen
Type:	Discussion	Output:	Notes
Participants:	Palle Staffeldt, Senior HW Design Engineer, Nokia Marc Richarme, Student, DTU		

The SIM card is mostly trouble. It was once a modern part of the phone, but is now outdated.

There are at least three advantages to removing the SIM:

1. Power consumption. The SIM card is a big consumer of power when the phone is in idle mode, and thereby reducing standby time. Furthermore, the power consumption of the SIM varies between card types, making the standby time experienced by the end-user inconsistent and to some degree out of Nokia's control.
2. Board space. The SIM is a major constraint in designing the phone, and takes up an unreasonable amount of board space. Furthermore, it is very difficult to handle with regard to EMI.
3. A very high percentage of faults are related to the SIM card. Mostly due to its removable nature.

Appendix B

Gathered data

B.1 SIM use case comparison

This section describes various usage scenarios related to subscription management and the SIM, and delineates how these scenarios would work with the different provisioning methods described in section 4.6.1 (except for provisioning in factory):

- Smart card: This is the status quo with the SIM card.
- Local connection: The VSIM is provisioned by means of a 'local connection' to the terminal. This could be an USB cable, Bluetooth, IrDA, or similar connection.
- OTA: In this case, the subscription is downloaded from the desired operator over the mobile network.

The set of scenarios discussed below has been gathered from the information collected during the conducted interviews, combined with ideas obtained during a brainstorming session.

7.1.1.1 Borrowed phone

Today, users can borrow any GSM/UMTS handset, and use it with their own subscription. This could be motivated by several factors, for example: the user runs out of battery or the user's own phone is broken and is being repaired. (This scenario assumes that the temporary phone is not SIM locked.)

- Smart card: The user places his own SIM card in the borrowed phone. He is now able to turn on the phone with his own PIN code and use it with his subscription. When done, the original SIM card is put back into the phone.
- Local connection: Not practically feasible.
- OTA: The "Change subscription" menu is activated on the borrowed phone, allowing the user to choose his operator from a list. The user is then taken to a WAP page belonging to that operator, allowing him to download a VSIM to the phone, after entering his credentials (e.g. phone number and password). The user can choose between a temporary or a permanent VSIM, where the temporary one only works until the next time the phone is turned off (or the VSIM is removed manually.) Once the

VSIM is removed (when the phone is returned to its owner), the previous VSIM is re-activated, bringing back the phone to its original state.

7.1.1.2 Borrowed subscription

Today, users can borrow a friend's SIM card and use the friend's subscription in their own phone (assuming it isn't SIM locked). The motivation could be that the user runs out of credit on his prepaid card, or he is roaming abroad but does not have a subscription allowing this.

This scenario essentially works the same way as the previous one (borrowed phone) for all cases.

7.1.1.3 Subscription bought over the internet

Some operators allow users to create a subscription on-line by entering their personal and billing information on a secure web page. Once the user has paid, the SIM must be placed in the user's terminal.

- Smart card: The SIM card is sent to the user by post.
- Local connection: The user enters his phone's EMEI number on the web page, and can then download the VSIM card to this personal computer. He then uses the phone's connectivity software to install the SIM, using either a cable connection or Bluetooth.
- OTA: Several options, for example:
    2) The user is presented with a numeric code that can be entered in the "Change subscription" menu of the phone to initiate a VSIM download.
    3) The user enters the "Change subscription" menu of the phone, chooses his operator, enters his credentials (e.g. phone number and password), and can the download a VSIM (same procedure as in the "borrowed phone" scenario).

7.1.1.4 Subscription bought in an operator store

This is a very common way of selling subscriptions – both pre- and post-paid ones, and often with subsidised phones.

- Smart card: The SIM card is handed to the user in a package also containing his phone number and his PIN and PUK codes. The user must either put the SIM card in the handset himself, or the staff can help with this (often tedious) procedure.
- Local connection: Either the staff installs the VSIM at the store using a cable or Bluetooth connection or the user is instructed to follow the instruction on a web page, where the phone's EMEI will need to be entered (same procedure as in the "subscription bought over the internet" scenario.)
- OTA: The user is provided with a package containing the phone number, his PIN and PUK codes, and either a password or a numeric subscription activation code, which is used for one of the following provisioning options (which correspond to those in the "subscription bought over the internet" scenario):
    1) The subscription activation code is entered in the "Change subscription" menu of the phone to initiate a VSIM download; or
    2) The user enters the "Change subscription" menu of the phone, chooses his operator, enters his credentials (e.g. phone number and password), and can then download a VSIM.

7.1.1.5 Subscription bought in 3rd party store

This is also a very common way of selling subscriptions (e.g. in a gas station).

- Smart card: Same procedure as the previous scenario (subscription bought in an operator store).
- Local connection: Same as previous scenario, except that either the user is provided with a CD containing the VSIM, which must then be installed using a personal computer, or the staff can install the VSIM at the store using a cable or Bluetooth connection. (Same procedure as in the "subscription bought over the internet" scenario.)
- OTA: Same procedure as previous scenario.

7.1.1.6 Transfer phonebook contacts or messages between terminals

This scenario is relevant in two cases. The first one is when a user gets a new terminal and wants to transfer his personal data. The other is in the case where a user regularly uses more than one terminal (multiple device ownership), where he would like to use the same subscription of each terminal, and have the same contact data, etc., available.

- Smart card: User stores his contacts and text messages on the SIM card, so that they are always available in the terminal holding the card.
- VSIM: No physical token is exchanged if a user decides to change terminal, so this data must be synchronized by other means. Any stakeholder (operators, terminal vendors or third parties) could provide a service to synchronize this data over-the-air. Nokia currently provides a PC-based utility to transfer user data when changing phone, as well as a PC-based synchronization tool for daily use. Another example is ZYB, which is a web-based service allowing users to back up their contact data and synchronize it between multiple devices.

7.1.1.7 Developing countries: subscription bought from street merchant

Developing countries have primitive sales channels, and there is no common access to the Internet or computers. Often subscriptions and air-time is bought from street merchants, at markets, or in primitive kiosks.

- Smart card: Client is handed a package containing a SIM card, phone number and PIN/PUK codes.
- Local connection: Not practically feasible in developing countries.
- OTA: Client is handed a paper slip containing a subscription activation code (and possibly a phone number). The user goes to the "Change subscription" menu and enters the code, which triggers the download of a VSIM. Soon after, the user receives an SMS containing his phone number and the amount of credits on the subscription. The user can perform this procedure in front of the merchant, to ensure the validity of the code he has bought.

7.1.1.8 Developing countries: village phone

In some areas, people cannot afford to have their personal phone, so it is common that a "village phone" is shared by members of a community. In this case, each user uses his personal subscription, but the common handset.

- Smart card: Either users must change SIM cards before every use, or some complex application in a SIM toolkit application and/or the network can be used to switch between multiple subscribers sharing a single SIM card.
- Local connection: Users can potentially switch between several installed VSIMs (see below), but provisioning using a local connection is not feasible in developing countries.
- OTA: Several VSIMs can potentially be installed on the handset, and a menu can be used to switch between them. This can be implemented either by a single operator, or in a more fundamental way, such that several subscriptions form different operators are allowed on the same handset.

7.1.1.9 Home/work subscription

Many people have a “work” subscription provided by their employer, but wish to use their personal subscription when not working.

- Smart card: Users mush change SIM cards when going to/from work. A few operators use a “Dual IMSI” SIM toolkit application allowing two subscriptions (from the same operator) to be stored on the same SIM card. The user must restart phone and enter a different PIN depending on which subscription he wishes to activate.
- VSIM: Several VSIMs can potentially be installed on the handset, and a menu can be used to switch between them. This can be implemented either by a single operator, or in a more fundamental way, such that several subscriptions form different operators are allowed on the same handset (basically the same concept as for the “village phone”.)

7.1.1.10 SIM lock

Today, phones sold along with a subscription are cheaper than their market price (subsidized), but the phone is locked to the original operator for a period of time.

- Smart card: Phone software is responsible for ensuring that a different operator’s SIM cannot be used until the original operator allows the phone to be unlocked. Traditionally, this protection was easy to circumvent by crackers, and this is a very lucrative business. Newer high-end phones have a SIM lock protection that is very difficult to bypass.
- VSIM: Phones supporting VSIMs have secure hardware protection mechanisms ensuring the security of installed VSIMs. This mechanism will provide a very high level of SIM lock protection for all handsets.

B.2 SIM faults

B.2.1 Summary

The data presented in the following table show the number and cost of warranty-repairs related to the SIM card for all Nokia phones during the year 2007. The table includes part costs and service costs.

This information was gathered with the help of Terhi Pere, who added the following information in an e-mail:

> *Hi Marc,*
>
> *Yes for the part for part costs and labor costs. In addition to that there are overheads and transportation costs. I would take this figure as a very indicative saying minimum rough estimation is 2.5MEuros. Then you should note that China volumes are not covered by the cubes in 2007 except for the last months, therefore I would use this 2.5 MEuros as a rough minimum estimation for the costs.*
>
> *BR*
>
> *Terhi*

As Terhi's mail also state, there are some overhead costs involved in addition to the costs shown in the table, and data for China is only included starting October 2007. Furthermore, these numbers only cover warranty repairs, and not actual faults.

The total cost for the costs included in the table is 2.3 million euros, but the actual number is estimated to at least 2.5 million euros.

B.2.2 Explanation of columns

Nmbr Of Accpt WOs	Total number of work orders with validation status "accepted", including accepted work orders and accepted swaps.
Nmbr Of Accpt Key Reprs	Total number of Key Repairs from accepted work orders that were repaired.
Nmbr Of Accpt Repr Detls	Total number of repair details from accepted repairs. One repair may contain more than one repair detail.
Nmbr of WOs With Srvc Costs	Total number of work orders with validation status "accepted" and with service costs specified, including accepted work orders and accepted swaps.
Nmbr of WOs With Part Costs	Total number of work orders with validation status "accepted" and with part costs specified, including accepted work orders and accepted swaps.
Part Costs From Key Reprs	Cost of replaced parts (in euros) related to key repairs, but doesn't include part costs for Accessory Replacements, for which only Service Costs are included.
Part Costs	Cost of replaced parts (in euros), but doesn't include part costs for Accessory Replacements, for which only Service Costs are included.
Srvc Costs	Service costs (in euros)
Wrty Cost	Part costs + service costs (euros)

B.2.3 Warranty repair metrics

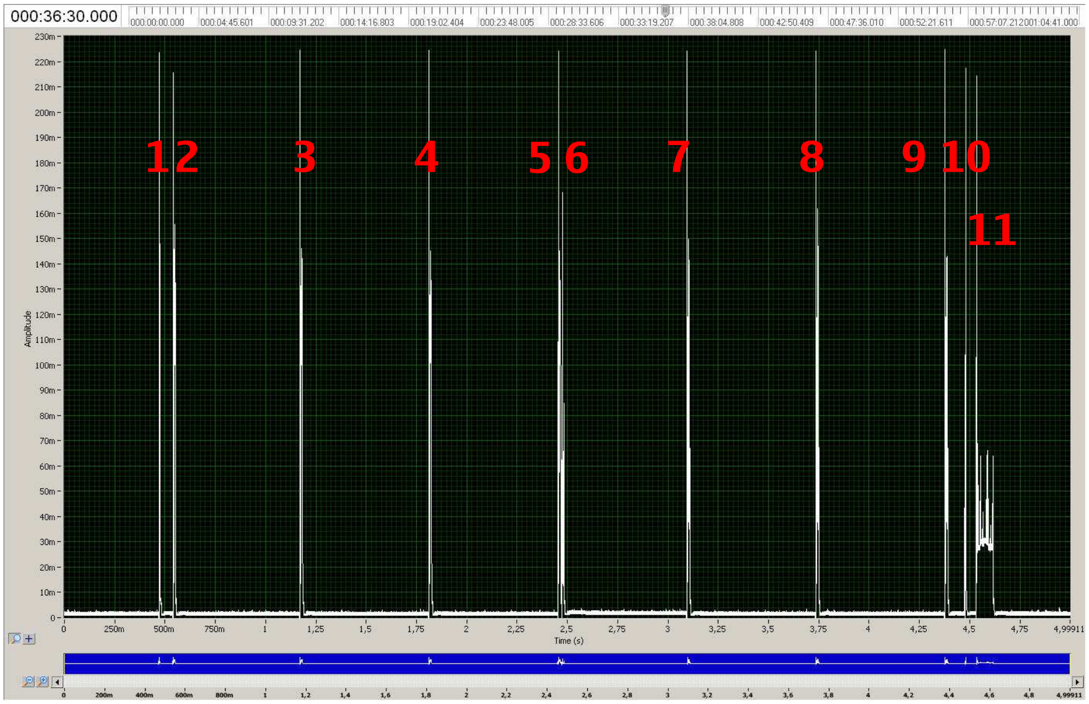
Component	Nmbr Of Accpt Wos	Nmbr Of Accpt Key Reprs	Nmbr Of Accpt Repr Detls	Nmbr Of Wos With Srvc Costs	Nmbr Of Wos With Part Costs	Part Costs from Key repairs	Part Costs	Srvc Costs	Wrty Cost
0039546 - MYLAR SIM POLYESTER BLK 57P31	97	97	1740	97	97	13,23	4270,57	394,92	4665
0202755 - 1RC SDSIM PWB MODULE	2	2	2	2	0	0	0	30,39	30
0202755 - 1RC SDSIM PWB MODULE	64	64	214	64	61	232,8	917,02	941,96	1859
0263490 - SIM RETAINER ASSEMBLY DMC07675	3	3	4	3	0	0	0	61,39	61
0263490 - SIM RETAINER ASSEMBLY DMC07675	4	4	4	4	4	13,02	14,07	79,75	94
0263990 - SIM SUPPORT ASSEMBLY	21	21	120	21	21	3,49	193	446,23	639
0263990 - SIM SUPPORT ASSEMBLY	4	4	18	4	3	0	26	82,98	109
0264324 - SIM READER ASSY 040-021073 P2730	1	1	1	1	0	0	0	11,3	11
0264324 - SIM READER ASSY 040-021073 P2730	1	1	1	1	0	0	0	89,86	90
0264380 - SIM READER LID ASSY 040-024405 P2730	48	48	60	48	1	0	15	965,64	981
0264380 - SIM READER LID ASSY 040-024405 P2730	58	58	90	58	58	13,83	55,37	1172,1	1227
0269084 - SIM RETAINER ASSEMBLY BLACK	2	2	2	2	0	0	0	30,2	30
0269084 - SIM RETAINER ASSEMBLY BLACK	15	15	22	15	15	37,12	85,88	238,18	324
0269393 - SIM SUPPORT ASSEMBLY P2913	7	7	40	7	7	1,22	49,88	187,48	237
4120071 - ASIP SIM FILTER BGA8	48	48	48	48	0	0	0	731,26	731
4120071 - ASIP SIM FILTER BGA8	1936	1936	2671	1935	1924	300,7	3771,26	28923	32694
4129071 - ASIP SIM INTERFACE ** PB-FREE **	434	438	727	434	134	0	2057,29	6636	8693
4129071 - ASIP SIM INTERFACE ** PB-FREE **	18379	18412	38384	18359	18089	2628	46041,4	281876	327918
4129257 - ASIP SIM INTERFACE **low cap** BGA8	2765	2770	5640	2765	1293	2,37	36941	40234	77175
4129257 - ASIP SIM INTERFACE **low cap** BGA8	51528	51727	115445	51275	48335	4759	141888	1E+06	1177527
4129281 - ASIP SIM ESD/EMI FILT 400UM BGA8	506	509	920	506	176	0	1427,05	8082,7	9510
4129281 - ASIP SIM ESD/EMI FILT 400UM BGA8	12822	12864	28354	12822	12791	1565	30620,9	242319	272940
4900224 - CONN SIM CARD 6POL 2.54MM SMD ST	4	4	4	4	2	0,48	0	69,24	69
5400085 - SIM CARD READER 2X3POL P2.54 SM	4	4	4	4	4	0,77	0	42,24	42
5400085 - SIM CARD READER 2X3POL P2.54 SM	1	1	1	1	0	0	0	14,39	14
5400169 - SIM CONN 2X3POL P2.54 SPR 15V 1A	1	1	1	1	1	0,11	0	20,56	21
5400313 - SM SIM CONNECTOR 6POL P2.54	85	85	346	85	85	8,26	318,74	1708,6	2027
5400313 - SM SIM CONNECTOR 6POL P2.54	2	2	13	2	2	0	84	29,72	114
5400329 - SPRING SWITCH SIM CONN 2X2POL	49	49	97	49	49	20,58	86	481,22	567
5402001 - SMD SIM CONN 2X3POL P2.54 H 1.6MM	1	1	6	1	1	0,27	33,92	0	34
5407051 - SM SIM CONN 6POL P2.54 H1.5	858	865	2175	848	699	97,6	5585,58	14434	20019
5407051 - SM SIM CONN 6POL P2.54 H1.5	298	298	341	298	8	0	716,36	4201,4	4918
5407091 - SM SIM CONN 6POL P2.54	859	860	1736	859	837	111,6	1492,22	16719	18211
5407091 - SM SIM CONN 6POL P2.54	178	179	222	178	23	3,6	353,73	3473	3827
5407105 - SM SIM CONN 6POL P2.54 H1.8	2912	2919	6561	2890	2506	447,2	6899,46	36117	43016

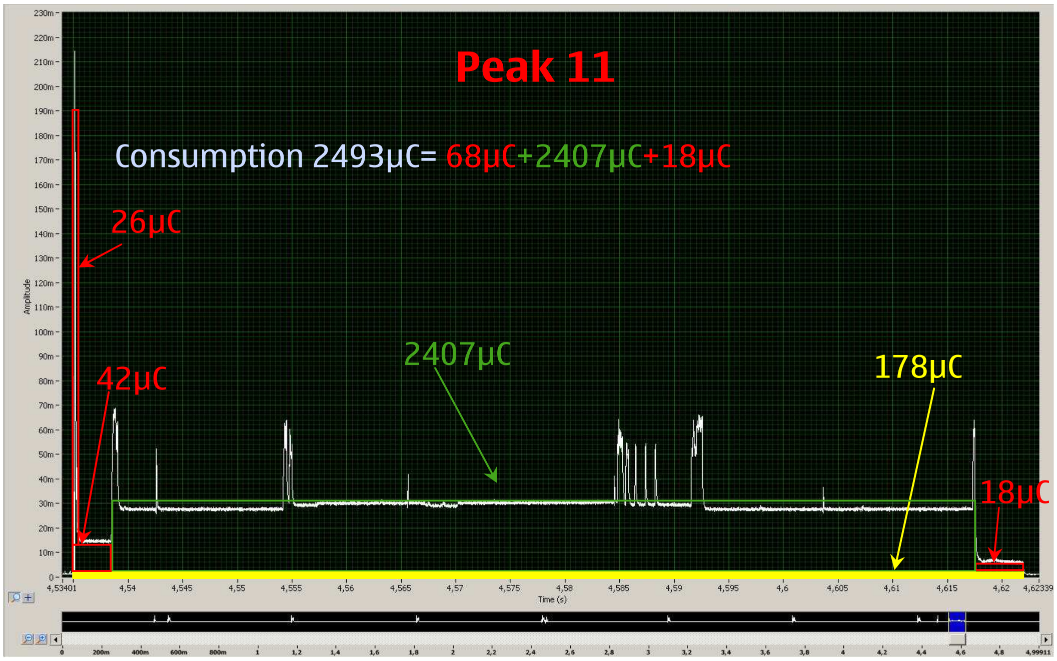
5407105 - SM SIM CONN 6POL P2.54 H1.8	751	751	933	751	90	2,45	5630,03	9236,4	14866
5407225 - SM SIM CONN 6POL P2.54 H1.3	1	1	1	1	0	0	0	19,02	19
5407225 - SM SIM CONN 6POL P2.54 H1.3	1	1	1	1	1	0,12	0	20,09	20
5407376 - CONN SIM 2X3POL P2.54 15V	1	1	1	1	0	0	0	9,5	10
5407376 - CONN SIM 2X3POL P2.54 15V	6	6	12	6	5	1,03	14,06	51,51	66
5409033 - SIM CARD READER CCM04-5004 2X3SMD	2	2	2	2	0	0	0	25,86	26
5409065 - SM SIM CARD CONN 2X3POL P2.54	1	1	1	1	1	0,23	0	19,06	19
5409117 - SM SIM CONN 2X3POL SPR P2.54 0.5A	1	1	10	1	1	0,41	4,07	0	4
5409145 - SM SIM CONN 2X3POL P2.54 H 1.95MM	1	1	1	1	0	0	0	13	13
5409215 - SM SIM CONN 2X3POL P2.54 SPR	1	1	2	1	1	0,39	1,12	7,86	9
5409219 - SM SIM CONNECTOR 6POL P2.54	1598	1603	3783	1592	1512	144,9	4194,88	23570	27764
5409219 - SM SIM CONNECTOR 6POL P2.54	400	402	426	400	8	8,6	40,43	4719,8	4760
5409273 - SIM CONN 2X3POL H 2.20MM	8	8	10	8	2	0	1,54	100,75	102
5409273 - SIM CONN 2X3POL H 2.20MM	33	33	117	32	25	2,58	195,73	493,69	689
5409317 - SM SIM CONN 2X3POL P2.54MM	1	1	1	1	1	0,2	0	16,49	16
5434003 - SM SIM CONN 2X3POL P2.54 H2.2	137	137	330	136	133	25,82	660,67	1580,4	2241
5434003 - SM SIM CONN 2X3POL P2.54 H2.2	32	32	39	32	4	0	857,49	381,26	1239
5469046 - CONN SIM Scalable Block 0.5-1.2mm H0.7mm	6	6	19	6	3	0	354,64	161,43	516
5469046 - CONN SIM Scalable Block 0.5-1.2mm H0.7mm	96	96	410	96	96	10,98	1190,52	3236,6	4427
5469195 - CONN SIM 2X3POL P2.54 15V 1A	28	28	56	28	28	3,49	471,77	511,64	983
5469195 - CONN SIM 2X3POL P2.54 15V 1A	5	5	8	5	3	0	1041	60,39	1101
5469196 - SIM TF READER P2910	78	78	95	78	3	0	29,25	1924	1953
5469196 - SIM TF READER P2910	490	491	1170	490	489	420,5	3504,83	10426	13931
5469226 - CONN SIM 2X3POL P2.54 15V 1A H3.4mm	421	421	1266	421	421	42,01	1495,26	10792	12288
5469226 - CONN SIM 2X3POL P2.54 15V 1A H3.4mm	22	22	30	22	7	0	1635,41	381,36	2017
5469283 - CONN SIM 2X3POL P2.54 15V 0.5A	196	196	245	187	82	18,66	102,95	4091,2	4194
5469283 - CONN SIM 2X3POL P2.54 15V 0.5A	61	61	62	61	0	0	0	604,56	605
5469315 - SM SIM CONN 2X3POL P2.54 50V 0.5A	189	189	683	189	188	46,08	982,62	4813,4	5796
5469315 - SM SIM CONN 2X3POL P2.54 50V 0.5A	11	11	15	11	1	0	211,16	154,07	365
5469415 - SM SIM CONN 11POL H2.0	103	103	108	103	7	0,54	468,55	1306,8	1775
5469415 - SM SIM CONN 11POL H2.0	1054	1054	5321	1049	1003	304,5	6842,38	24312	31155
5469419 - SM SIM CONN 2X3POL P2.54 H4.6	2	2	3	2	1	0	184	29,86	214
5469419 - SM SIM CONN 2X3POL P2.54 H4.6	29	30	56	29	28	3,9	250,91	465,07	716
5469459 - CONN SIM READER 8POL 0.01V 0.2A	1	1	1	1	1	0,94	1,12	7,86	9
5469487 - SM SIM CONN 6POL P2.54	3	3	12	3	2	0	96,59	49,47	146
5469487 - SM SIM CONN 6POL P2.54	48	49	95	48	48	11,98	94,55	199,03	294
5469505 - CONN SIM 2X3POL P2.54	1	1	1	1	1	0,42	0	19,2	19
5469729 - CONN SIM SM 6POL P2.54 H1.05	2259	2271	8401	2245	2118	283,7	8845,58	46604	55449
5469729 - CONN SIM SM 6POL	797	797	984	797	59	14,25	529,25	12256	12785

P2.54 H1.05									
5469765 - CONN SM SIM 6POL P2.54 H5.0	3	3	5	3	2	0	0	73,54	74
5469765 - CONN SM SIM 6POL P2.54 H5.0	22	22	32	22	22	4,27	64,14	224,23	288
5469791 - SIM TF READER	98	98	112	98	5	0	1106,39	1087,3	2194
5469791 - SIM TF READER	256	256	441	256	256	233,4	2913,62	5040,1	7954
5469809 - SM SIM CONN 2X3POL P2.54	925	925	2033	923	883	143,8	5120,99	14874	19995
5469809 - SM SIM CONN 2X3POL P2.54	261	261	339	261	20	0	3114,82	3343,8	6459
5469853 - CONN SIM SM 6POL h=0.7	75	75	124	75	13	0	848,98	1676,1	2525
5469853 - CONN SIM SM 6POL h=0.7	1253	1254	3924	1252	1241	428,8	10426,7	30290	40716
5469921 - SIM READER COVER	18	18	22	18	18	10,27	15,57	393,65	409
5469921 - SIM READER COVER	3	3	3	3	0	0	0	55,95	56
5469927 - SM SIM Connector	7	7	46	7	7	0,73	114,19	232,94	347
5469927 - SM SIM Connector	13	13	14	13	1	0	0	184,18	184
6442483 - SIM FLAP	26	26	28	26	26	4,68	1,12	438,1	439
6442565 - SIM FLAP	12	12	12	12	12	3,24	0	202,2	202
6442881 - SDSIM SUPPORT FRAME STAINLESS 040-014365	1	1	1	1	0	0	0	0	0
6443334 - SIM BRACE P3358	27	27	99	27	27	6,59	152,41	678,08	830
9402341 - SIM INSULATOR PRINTED 040-032003	12	12	361	12	12	1,68	1898,32	44,77	1943
9402341 - SIM INSULATOR PRINTED 040-032003	1	1	1	1	0	0	0	10,86	11
9403162 - SIM CARD LABEL P2675	1	1	1	1	0	0	0	9,5	10
9403162 - SIM CARD LABEL P2675	1	1	1	1	1	0,06	0	9,44	9
9450588 - SIM GUIDANCE	10	10	10	10	0	0	0	192	192
9450588 - SIM GUIDANCE	3	3	3	3	3	1,08	0	57,6	58
9457740 - A-COVER ASSY DMC00711 SIMPLEX	22	22	24	22	22	24,7	31,14	245,57	277
9460597 - SIM-COVER DMD11555	75	75	114	75	75	6,78	12,06	1837	1849
9460597 - SIM-COVER DMD11555	6	6	7	6	0	0	0	108,8	109
9560306 - SIM-GRIP DMD11281 HDJ12	480	480	612	478	399	74,14	97,44	6417,8	6515
9560306 - SIM-GRIP DMD11281 HDJ12	127	127	166	127	1	0	5	1678,6	1684
9590845 - SIM SUPPORT SHIELD CuNi 18Zn20 F610	2	2	13	2	2	0,43	13,18	66,17	79
9590876 - SIM LID SUS P2636	47	47	53	47	1	0	11,23	688,05	699
9590876 - SIM LID SUS P2636	112	112	205	112	112	23,7	128,23	1945,9	2074
9590887 - SIM COVER P2130	3	3	4	3	0	0	0	62,4	62
9590887 - SIM COVER P2130	120	120	266	120	120	17,33	143,65	2551,1	2695
9591220 - SIM SHIELD COVER 040-023264	4	4	20	4	4	0,28	74	72,49	146
9591224 - SDSIM FLEX GASKET 040-024317	1	1	1	1	0	0	0	12,1	12
9591224 - SDSIM FLEX GASKET 040-024317	2	2	2	2	2	0,16	0	23,88	24
9901602 - SIM SLEDGE	4	4	5	4	4	2,12	0	36,72	37
									2314182

B.3 Power measurements

The SIM card is usually polled approximately every 25 seconds for SIM toolkit commands. When the terminal is idle, it goes into a low power mode, but polling wakes up the processor every 25 seconds by default. The first figure shows the current consumption of an idle Nokia 6500 slide terminal. The peak labelled ‘11’ has been identified as a SIM polling event, and the pattern was indeed found to occur every 25 seconds.





The second figure shows a detailed view of the power peek caused by SIM polling. The yellow area is the power normally consumed when the terminal is in ‘sleep-mode’; the red areas are wasted power due to starting and stopping the terminal’s processor, and the green areas are power consumed by the processor doing ‘useful’ work communicating with the SIM card. As can be seen from the figure, the consumption for this one peak is 2.493mC after the normal sleep current has been subtracted.

The 6500 has a 900mAh battery and a standby time of 320 hours (source: nokia.com). Thus, the average consumed charge over a 25 second period is:

$$25s \times \frac{900mAh}{320h} = 70{,}31mC$$

Subtracting from this the 2.493mC incurred by the current peak gives a charge of 67.82mC that would have been what was consumed over 25 seconds had it not been for the peak. Calculating back into standby time gives us:

$$25s \times \frac{900mAh}{67.82mC} = 332h$$

Thus, without the SIM polling, standby time for this phone would increased by approximately 3,75%.

Both figures are courtesy of Palle Staffeldt.

B.4 SIM cost estimations

B.4.1 SIM cost for operators

B.4.1.1 Information from Telia

Information from Telia regarding the cost of SIM cards obtained during the interview transcribed in Appendix A.8:

- Price of SIM card:
  - Native OS: 0.6 EUR
  - JavaCard 64k: 1.4 EUR
- Packaging & logistics costs per issued card:
  - High-end: 1.5 EUR
  - Low-end: 0.6 EUR
- Number of SIM cards ordered per year:
  - 800.000 (for 1.300.000 subscribers – this number is a bit higher for Telia than other operators due to the high number of pre-paid subscribers)

B.4.1.2 Information from Gemalto

The following table provides an estimation of the total revenue of the SIM card market, based on key-numbers for Gemalto from 2006. Source: Gemalto (not the estimated numbers).

Gemalto revenue in telecom sector in 2006	994.000.000€
Gemalto SIM cards sold in 2006	~1.000.000.000
Gemalto market share in 2006	48%
Estimated total SIM card revenue in 2006	100/48 * 994,000,000€ = 2,070,800,000€
Estimated SIM cards sold in 2006	100/48 * 1,000,000,000 = 2,083,000,000
Estimated average price per SIM card	0.99€

B.4.2 Terminal bill of materials

The following table shows a breakdown of the Bill Of Materials for various categories of mobile devices – source: Informa Telecoms & Media [115]. In addition, the estimated contribution of the SIM to the total BOM, "SIM % of BOM", and the estimated increase in device vendor margin if this cost was removed from the BOM have been calculated for each category. The costs of the SIM-related components used in these calculations are shown in the third row, "SIM-related BOM". These numbers have been calculated based on an estimated cost of 0.15€ (0.24$) for 2008, and extrapolated to the other years based on the cost evolution of the "Other variable" row for basic phones.

Note that this calculation does not incorporate additional costs due to the use of the VSIM. Once such cost could be per-device key-generation during production.

US$	2006	2007	2008	2009	2010	2011	2012
SIM-related BOM (x)							
	0.36	0.29	0.24	0.21	0.20	0.20	0.20
Basic phones							
Modem & multimedia chipset*, **	8.97	6.50	4.88	3.90	3.54	3.44	3.39
Display & camera modules ***	4.55	4.22	3.97	3.87	3.87	3.91	3.93
Battery	2.65	2.31	2.23	2.24	2.28	2.31	2.34
Terminal software BOM (++)	1.44	1.50	1.36	1.28	1.32	1.70	1.98
Wireless IPRs	4.32	3.49	2.53	1.92	1.61	1.70	1.62
Mechanical/Plastic (‡ ‡)	2.69	2.18	1.81	1.55	1.50	1.50	1.50
Added value features	2.84	2.65	2.37	2.18	2.10	2.00	2.00
Other variable†	9.12	6.72	5.42	4.64	4.38	4.47	4.53
Total BOM	36.58	29.57	24.56	21.58	20.60	21.04	21.29
Device vendor Margin	7.32	5.50	4.30	3.45	3.37	3.48	3.52
Average sale price	52.23	41.73	34.34	29.79	28.53	29.18	29.53
SIM % of BOM	0.98%	0.98%	0.98%	0.95%	0.97%	0.95%	0.93%
VSIM increase in Margin	4.87%	5.26%	5.58%	5.96%	5.90%	5.72%	5.65%
Low feature phones							
Modem & multimedia chipset*, **	18.64	13.55	9.85	8.41	7.78	7.48	7.56
Display & camera modules***	6.84	6.06	5.20	5.40	5.69	6.03	6.24
Battery	4.77	4.38	3.99	3.66	3.50	3.42	3.40
Terminal software BOM (++)	3.21	3.22	3.09	3.03	3.02	3.07	4.19
Wireless IPRs	8.67	6.46	4.84	3.93	3.48	2.56	2.26
Mechanical/Plastic (‡ ‡)	3.84	3.32	2.77	2.35	2.26	2.26	2.29
Additional features‡	4.33	4.16	3.92	3.68	3.57	3.52	3.55
Other variable†	15.74	10.37	7.57	6.70	6.30	6.09	6.49
Total BOM	66.04	51.52	41.23	37.17	35.58	34.42	35.97
Device vendor Margin	15.38	10.50	7.42	6.13	5.73	5.58	5.83
Average sale price	96.89	73.80	57.90	51.53	49.16	47.59	49.74
SIM % of BOM	0.54%	0.56%	0.58%	0.55%	0.56%	0.58%	0.55%
VSIM increase in Margin	2.32%	2.75%	3.23%	3.35%	3.47%	3.56%	3.41%
Feature rich phones (+)							
Modem & multimedia chipset*, **	35.20	28.07	24.18	22.16	20.97	19.36	19.44
Display & camera modules ***	42.88	39.20	35.17	35.67	37.56	38.25	36.98
Battery	6.22	5.82	5.41	5.18	5.15	5.19	5.31
Terminal software BOM (++)	12.29	7.84	6.47	5.74	5.12	4.16	3.88
Wireless IPRs	18.43	15.91	13.13	11.66	10.40	7.40	6.32
Mechanical/Plastic (‡ ‡)	5.46	4.93	4.60	4.29	4.15	4.00	4.20
Additional feature‡	8.48	8.00	7.44	7.00	6.52	6.40	6.60
Other variable†	36.75	27.11	21.69	20.18	19.23	17.80	17.62
Total BOM	165.71	136.89	118.09	111.88	109.11	102.57	100.36
Device vendor Margin	44.74	33.81	24.80	20.25	18.77	17.44	15.56
Average sale price	250.44	203.14	170.04	157.24	152.18	142.80	137.94
SIM % of BOM	0.22%	0.21%	0.20%	0.18%	0.18%	0.19%	0.20%
VSIM increase in Margin	0.80%	0.85%	0.97%	1.01%	1.06%	1.14%	1.28%

Low-end smartphones							
Modem & multimedia chipset*, **	44.25	38.63	35.50	31.48	26.85	21.49	19.96
Display & camera modules***	54.86	49.94	44.83	46.60	50.32	52.02	50.36
Battery	6.72	6.29	5.92	5.85	5.89	5.60	5.47
Terminal software BOM (++)	26.51	25.08	20.09	14.66	12.80	9.34	8.24
Wireless IPRs	24.47	20.52	16.43	14.66	12.80	9.34	7.30
Mechanical/Plastic (‡ ‡)	5.46	4.93	4.60	4.29	4.15	4.15	4.20
Additional feature‡	9.20	8.63	8.20	7.83	7.32	7.20	7.25
Other variable†	48.87	38.04	30.50	27.58	25.71	21.28	20.55
Total BOM	220.35	192.07	166.06	152.95	145.85	130.43	123.33
Device vendor Margin	72.72	58.77	44.01	31.20	27.71	23.74	22.32
Average sale price	348.76	298.51	249.98	219.14	206.53	183.45	173.32
SIM % of BOM	0.16%	0.15%	0.14%	0.13%	0.14%	0.15%	0.16%
VSIM increase in Margin	0.49%	0.49%	0.55%	0.66%	0.72%	0.84%	0.89%
High-end smartphones							
Modem & multimedia chipset*, **	54.64	50.89	44.92	38.72	30.62	27.13	25.61
Display & camera modules ***	83.79	73.28	61.57	58.28	56.55	54.06	50.37
Battery	7.04	6.58	6.28	6.31	6.36	6.05	6.17
Terminal software BOM (++)	44.07	37.29	24.93	19.82	16.09	11.02	9.46
Wireless IPRs	19.80	16.75	13.43	13.21	13.17	9.78	8.74
Mechanical/Plastic (‡ ‡)	6.60	6.00	5.43	4.80	4.45	4.32	4.40
Additional feature†	13.66	12.49	11.32	10.54	10.00	9.60	10.00
Other variable‡	63.14	50.21	37.77	33.37	29.37	24.76	23.52
Total BOM	292.75	253.49	205.65	185.06	166.61	146.71	138.27
Device vendor Margin	102.46	83.65	63.75	46.08	36.65	29.34	26.96
Average sale price	470.30	401.20	320.59	275.05	241.88	209.51	196.63
SIM % of BOM	0.12%	0.11%	0.12%	0.11%	0.12%	0.14%	0.14%
VSIM increase in Margin	0.35%	0.35%	0.38%	0.45%	0.54%	0.68%	0.74%
Average across total handset market							
Modem & multimedia chipset*, **	23.52	19.49	17.51	16.78	16.27	15.37	15.78
Display & camera modules ***	20.97	21.09	20.96	23.39	26.85	29.73	30.95
Battery	4.89	4.60	4.42	4.36	4.47	4.55	4.71
Terminal software BOM (++)	7.75	6.88	6.30	5.90	5.83	5.06	5.19
Wireless IPRs	11.55	10.05	8.64	8.11	7.75	6.04	5.34
Mechanical/Plastic (‡ ‡)	4.21	3.78	3.49	3.26	3.26	3.32	3.54
Additional feature†	5.68	5.61	5.51	5.44	5.37	5.51	5.85
Other variable‡	23.20	17.96	15.24	14.94	15.07	14.36	14.94
Total BOM	101.75	89.45	82.06	82.18	84.88	83.96	86.32
Device vendor Margin	27.26	22.32	18.25	15.60	15.29	14.74	14.54
Average sale price	153.52	133.00	119.37	116.35	119.20	117.45	120.02
SIM % of BOM	0.35%	0.32%	0.29%	0.25%	0.23%	0.24%	0.23%
VSIM increase in Margin	1.31%	1.30%	1.32%	1.32%	1.30%	1.35%	1.37%

Table notes: (x) Assumed to be 0.24$ in 2008, adjusted for each year according to the "Other variable" cost for "Basic phones", (+) excludes Smartphones,* include Antenna, power amplifiers and power man, ** including applications processors, hardware accelerators, and memory *** include related chipsets, † features include Bluetooth, GPS, WLAN, music player, stereo speaker and others, ‡ other discrete components, packaging, manufacturing, assembly, IOT test & validation, product design, customisation costs, etc, ? based on NiMH or large cells Li-Ion technology, (++) include royalties, software integration and testing costs, ‡ ‡ Keypad/Speaker/Microphone/PCB/Mould/charger. Source: original numbers: Informa Telecoms & Media; SIM-related calculations: own work.

Appendix C

Miscellaneous documents

C.1 Motorola's SoftSIM proposal

Technical Specification Group Services and System Aspects ***TSGS#38(07)0768***
Meeting #38, 03 - 06 December 2007, Cancun, MEXICO

Title: Introduction of SoftSIM into 3GPP
Document for: Discussion
Agenda Item: 12
Source: Motorola

Summary

3GPP SA are requested to consider whether the time is right to start work on replacing the current hardware SIM card with a secure downloadable version that is stored and runs directly in a secure environment on the UE.

Introduction and Discussion

At the last SA plenary (#37) the study item on "Remote management of USIM application on M2M Equipment" was approved, and the work is under way in SA3. The obvious question that some of our customers are starting to ask is whether this will affect the "normal" phone, in our opinion it will not. However the follow up question tends to be is the time right to start the standards process in 3GPP.

The SIM card is ubiquitous within present day GSM systems. As such, 3GPP has continued this with the inclusion of the USIM application on the UICC for next generation systems. Additionally, the Long Term Evolution (LTE) initiative within 3GPP, which focuses on defining a new air interface and access network specification, has specified (technically noted in a meeting report) a USIM on UICC as the standard for device access to the LTE air interface. There are fair justifications for the inclusion of a smart card based token within 3GPP. However, there are also equally fair arguments for examining alternative solutions. Often the arguments for and/or against the use of a hardware-token SIM are comprised more of dogma than anything else. Some carriers will argue the impossibility of a system without the SIM. Others will point to

3GPP2 or WiMax as systems that have no such SIM requirement. Thus, there is no absolute answer, and the determination of whether a softSIM is a viable approach depends both on the technical merits of the approach as well as the perceived business value to involved parties.

Pros and Cons

- Manufacturers
    - Pros
        - To allow even more stylish and slim phones that are already very popular in the market place, removing the SIM will allow manufacturers to go even further in creative designs.
        - Not only claiming the additional real estate in the terminal is a key pro for implementing the softSIM, but the additional cost savings in connectors and additional parts will benefit the industry by allowing for lower cost phones.
        - Smartcard manufacturers, giving them a chance for them to share their expertise and software solutions with terminal manufacturers
    - Cons
        - The need for secure execution and storage, and provisioning of some form of secure identification/certificate.
- Operators
    - Pros
        - Lower cost, especially in the ultra low tier market such as those defined by the GSM association.
        - The ability to expand the user base beyond conventional phone to the application space, e.g cameras, MP3 players, games, toys...
        - To the operator, having the ease, flexibility and simplicity of having a softSIM in the terminal will be a major pro. Especially when it comes to switching between different types of networks other than GSM or 3G, like WiFi. With the rapidly changing technology, this will also provide benefit for future applications.
    - Cons
        - The perception of a less secure environment
        - Need to change provisioning systems.

Motorola believe the pros outweigh the cons. Motorola has been developing security architectures in our handsets that can be built upon to provide the additional security you find in smartcards. The softSIM will benefit not only manufactures and operators, but also the customers that use our products. The flexibility the softSIM gives the end user will allow them not to worry about accessing and swapping the little card that will allow them to use another phone, have multiple numbers and operators on the same phone, and other rich features.

Proposal

SA plenary have a brief discussion, if enough support is shown both within and outside the meeting Motorola will start to generate a study item to be elaborated in the working groups and discussed at the next SA plenary meeting.

C.2 Implementation of the MILENAGE algorithm using the OnBoard Credentials platform

Source: This appendix is taken from [98], appendix D.3.

As an example of a real-world algorithm that has been implemented for the ObC architecture, here is presented the core part of the Milenage/3G algorithm - the UMTS/3G authetication function. The code is Lua, and the native compiler is used after a pre-processing step with the MPP pre-processor. The code is commented, but note the env in(x) functions for inputting data, env out(x) for presenting the result, and the macro functions (resolving into external function calls) identifiable by the '#'-prefix. The 'unseal' and 'seal' operations unwrap (and re-wrap) encrypted data for this script, and e.g. #aes enc is a typical example of the invocation of an external cryptographic function.

On a practical level, this function constitutes the security core of a UICC - the 3G SIM card. Thus, in principle, this code, suitably provisioned with keys and operator constants, could be used to authenticate to a mobile phone operator.

```
-- -----------------------------------
-- Milenage 3G security kernel
-- (c) Nokia Research Center 2007
-- -----------------------------------

-- -----------------------------------
-- Input data (key) and unseal
-- Input is expected to be 8 shorts (16 bytes).
-- -----------------------------------
#env_in(ii)
#unseal(n,ii,kk)

-- -----------------------------------
-- Challenge input for Milenage kernel,
-- another 8 shorts.
-- -----------------------------------
#env_in(rn)

-- -----------------------------------
-- Operator Variant Algorithm Configuration
-- Field, another 8 shorts.
-- -----------------------------------
#env_in(ii)

-- -----------------------------------
-- Function number, scalar.
-- -----------------------------------
#env_in(fn)

-- ------------------------------
-- Run the Kernel itself
-- ------------------------------

i = 0

while i < 8 do
  rn[i] = rn[i] ^ ii[i]
  i = i + 1
end

#aes_enc(n, kk, rn, ww)
```

```
while 0 < i do
  i = i - 1
  ww[i] = ww[i] ^ ii[i]
end

if fn == 2 then
  rf = 0
  cf = 1
elseif fn == 3 then
  rf = 2
  cf = 2
elseif fn == 4 then
  rf = 4
  cf = 4
else
  rf = 6
  cf = 8
end

while i < 8 do
  wz[i] = ww[(i+rf)%8]
  i = i + 1
end

wz[7] = wz[7] ^ cf

#aes_enc(n, kk, wz, ww)

while 0 < i do
  i = i - 1
  wz[i] = ww[i] ^ ii[i]
end

-- ------------------------------
-- Return f(n)
-- ------------------------------
#env_out(wz)

end
```