On the Potential of High Density Smart Cards

Keith E. Mayes, Konstantinos Markantonakis

Smart Card Centre, Information Security Group,
Royal Holloway, University of London,
Egham, Surrey,
TW20 OEX, UK
{K.Markantonakis, Keith.Mayes}@rhul.ac.uk

1. Introduction

It is generally accepted that we live in a time of rapid change. Many advances are due to the introduction of new technologies that not only bring new opportunities but also can have a profound effect on human behaviours. Technology is always advancing steadily, but every so often there is evidence of a momentous leap forward. In fact this report is produced courtesy of a laptop, mobile phone and the Internet i.e. powerful technologies that we regard as essential to modern life yet would have been considered science fiction only 30 years earlier. This "information technology" has empowered individuals like never before however it is indiscriminate and works just as well for those with good or bad intentions. Therefore to safeguard the good from the bad the field of Information Security was born. Information Security has been steadily evolving and has developed special technologies of its own to use in the fight to keep systems, information and individuals secure. One of these is the Smart Card (SC) that has proven itself to be tamper-resistant security token that can be deployed to end-users to secure systems and protocols. The cards have been very successful and have evolved steadily to a surprisingly level of sophistication, however we may be just at the point when the smart card makes a momentous and assumption-shaking leap forward. The reason for this view arises because Smart Card vendors now have pre-production prototypes of High Density Smart Cards (that we will refer to as HDSCs) aimed initially at the mobile communications market, which could change our perception of security tokens and the associated system architectures and processes. This paper explores this possibility by first reviewing the current roles and limitations of conventional smart cards and then comparing with the HDSC. Finally some thoughts will be presented on the business reality of introducing HDSCs.

2. Conventional Smart Cards

Smart cards in various forms have been around for over 30 years, although the pace of technical development was fuelled by the introduction of GSM[1] SIM [1] cards in the earlier nineties. There are two main forms in use today i.e. cards with contacts and cards that are contact-less [3]. This paper will focus on the contact variety. The smart card may be regarded as a small tamper-resistant computing platform. As can be seen in Figure 1, at the heart of the platform is a micro-controller chip equipped with CPU, logic and various types of memories.

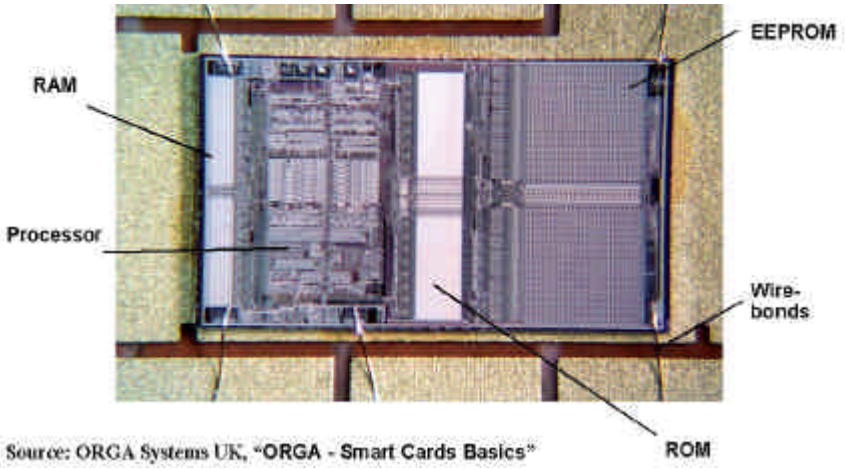

Figure 1 Exposed View of Smart Card Chip

2.1. Memory Size Limits

It has been accepted for many years that smart card memory is relatively small and is therefore not suited for any bulk storage or memory hungry applications. There are several reasons for this but they mainly relate to the physical area of the chip. Generally the bigger the chip area the higher the cost of the device and as smart cards are normally purchased in large volumes there is great pressure to minimise the cost. However standardisation also places a bound on the maximum size of the chip. ISO10373 [4] specifies that a card must survive certain bending and twisting stresses. If the chip is made too large it would not survive these tests as the stresses would break the fine wires between the chip and the card contacts. Typically a smart card chip area would not exceed about 16mm2. With restricted chip area there is a natural tendency to maximise the use of area-efficient memory types. The packing density for the typical memories can vary with production process but 1:4:16 (ROM: EEPROM: RAM) can be used by means of example. So in a given area you could have 256k ROM or 64k EEROM or 16k RAM. The ROM is fine for the operating system and some well proven common applications but it cannot be changed which is a serious drawback for the flexibility and lifetime of the card. EEPROM does not suffer from this but is a limited space for data and application storage. The small about of RAM limits the run-time behaviour of the device and is a chakkenge for the developers. However the support for custom application hosting is surprisingly advanced. Modern cards often incorporate a JVM [5] and Global Platform [6] to support applet loading and security domain management. It is therefore rather a shame that the available memory is so constrained. Taking the conventional approach to smart card technology provides a card today with about 64-128k EEPROM and Moores Law will probably double this every 18 month. It is not surprising that the Smart card is usually discounted as a mass storage device.

2.2. Interface Speed Limits

The interfaces between the smart card and the reader device is well standardised and makes use of a single bi-directional I/O line in addition to power, clock and reset connections as shown in Fig 2.

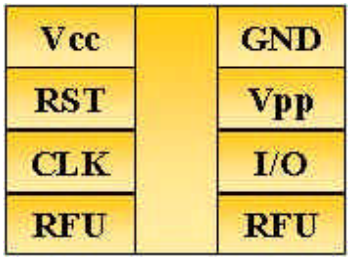

Figure 2 Smart Card Contacts

This permits half-duplex communication with the reader in control as the "Master" with the smart card acting as the responder or "Slave". In some implementations it is logically it is possible for

the smart card to be in temporary control of the dialog such as in the Proactive SIM Toolkit feature [7] used in mobile communication. A proactive SIM can tell the handset by means of a special format response message that it wishes to execute its own functionality, such as place a call, send a message or set-up a new SIM Toolkit phone menu. Whilst these measures permit a useful level of bi-directional communications at a logical level, the resulting performance is really suited to small infrequent message transfer. The first limit to appreciate is the transfer rate between the smart card and reader devices. The standards were produced around the time that a 9.6k baud modem or I/O port was considered fast and so the default common interface speed for the smart card was set to approx. 13.4k baud. As the smart card was usually considered a restricted platform there was no great pressure to dramatically increase this speed. There are some speeded up modes available which can take the raw bit rate to 78k baud. However, this is only possible if both the card and host device/reader are compliant. There are 2 low level transport protocols used within smart cards usually referred to as T=0 and T=1. The former is a byte oriented protocol and is typically the default used in (U)SIM cards whereas the latter is a block protocol commonly used in banking applications such as EMV[8] cards.
The message protocol sits on top of the transport protocol and consists of Application Protocol Data Units (APDUs) [9] plus associated responses. The protocol is well standardised but was never expected to cope with bulk transfer of information between card and readers. The APDU data field is limited to 255 bytes (plus header information) and so a larger amount of data has to be segmented and sent a message at a time. Even if the transfer rate could be speeded up for a large transfer to the card e.g. some rapid auto acknowledge of messages this may not help the situation as there is another limit due to the I/O buffer size. These buffers are typically very small e.g. enough for a message. Technically it would be possible to increase the size of the buffers during chip design but we recall that RAM is an expensive memory and its use tends to be minimised. The general observation from this section is that the interface speeds are slow because a variety of interacting factors and to make a real difference would require a radical overhaul of the interface

2.3. Processing Speed

Early smart cards had very simple micro-controllers and for a while an 8 bit CPU was considered quite exotic. Small CPUs are still in use however the popularity of Java card has lead to improvements in smart card CPU capabilities. Essentially Java is designed to work best with 32-bit CPUs and such devices are becoming increasingly common particularly for (U)SIMs. The overall processing rate of the CPU is strongly influenced by the clock speed. Typically smart cards do not have their own internal clocks and so rely on the clock signal provided by the host device or reader. The supplied clock speed is usually 5MHz. Whilst this was once considered fast, PC CPUs now run almost 1000 times faster. To compensate for the slow clock, modern smart cards incorporate frequency multipliers to reach more respectable speeds and they also employ specialist crypto co-processors for the rapid execution of security functionality such as digital signatures. The CPU's capabilities are fundamentally limited by the factors that affected memory size i.e. chip area and cost, however there is another restriction to appreciate. The effect is that Smart Card CPUs sometimes have the capability to run much faster than they do in practice. This situation arises because the smart card is dependent on the host device for the supply of power. In GSM mobile communications for example, standardisation has restricted the peak supply current to 10mA [7] to the smart card and at the same time there is pressure to reduce the operating voltages from 5 to 3 to 1.8v. Constraining power in this manner may prevent the smart card from running at the full rate and re-enforces the assumption that a smart card device is not used for high speed processing. The CPU speed is also inter-linked with the memory size and interface rates as a slow processor cannot manipulate bulk date or perhaps service interface buffers quickly enough.

2.4. Programming and run-time environment

Once removed from the very low-level limitations of the smart card, its role as an applications platform is surprisingly sophisticated. There are various operating systems, virtual machines and runtime environments [5,10,11] and to describe and compare them all would beyond the scope of this work. Therefore the arguments here will be restricted to Java card which is becoming the de-facto standard for high-end smart cards. It is a common misconception to consider Java as an operating system whereas it is in fact a virtual machine and run time environment layered upon a vendor's own operating system (OS) as shown in Fig 3.

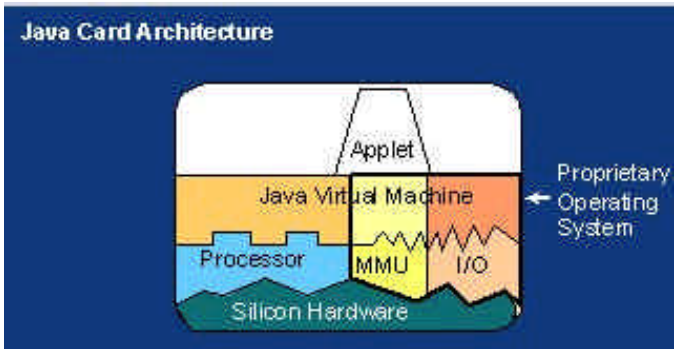

Figure 3 Java Card Architecture (source Sagem-Orga)

Whilst the standard defines a restricted sub-set of Java it is still useful for rapid application development and chip platform independence. Java card and global platform are usually strongly linked allowing support for application loading and security domain management. The security and management aspects have the flexibility to support a range of security, operational and business models. It would therefore appear that the Java card approach does not have strongly limiting factors and if more memory and processing resources were available they would be used appropriately. However there is still the reliance on the Vendor OS and if this was designed around small memories and files it may throttle back application performance.

3. The High Density Smart Card (HDSC)

The HDSC is a response by the industry to address some of the limitations that have applied to conventional smart cards and thereby challenge assumptions about how they can and cannot be used. The companies driving the development of HDSC devices include Renesas, Samsung, M-Systems, Atmel and ST with the major Smart Card Vendors such as Giesecke & Devrient, Gemplus, Sagem-Orga and Axalto etc, trying to drive the business opportunity. Although the memory size is the headline feature, the implications of this go much further and there is potential for smart cards to take a "momentous leap" forwards.

3.1. Memory

We have mentioned that technology makes a steady progress fuelled by Moores law. Based on this we would expect smart card memory to double roughly every 18 months. This would not make the smart card relatively more attractive as a storage location as all other memories in use would also increase based on the same law. A momentous leap needs to make a radical shift in card storage today that changes the perception of smart card storage. A ten-fold increase in memory might be considered radical and so the thousand-fold+ increase of the HDSC justifies the momentous category. Figure 4 gives an example of a HDSC from Samsung in a plug-in SIM format, which in theory could come to replace a conventional SIM

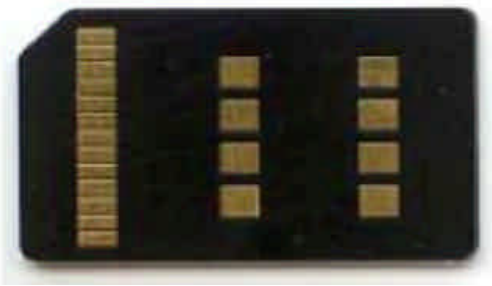

Figure 4 Samsung HDSC

Instead of designing for a 128kb card you could consider a 128Mb card and in fact 1Gb prototypes already exist. The storage is now equivalent to USB flash memory dongles but retaining the managed security token capabilities of the smart card. Only a few years ago this size of memory

was restricted to hard-drives but is now a tamper-resistant security token. The reason this is possible is due to a combination of smart card and flash memory technologies but also a willingness to change or break-free of some of the older ISO standards elements. Bending & twisting tests are fine for a physical card that sits in the users pocket or wallet but unnecessary as a SIM card or a PC peripheral.

3.2. Interface Speed

As mentioned earlier, a large memory cannot really be exploited without much faster interface speeds (and source clock) otherwise it will be too time-consuming to read/write the data. This is being addressed within the standards [12] where a corresponding leap in interface speed is proposed. There are in fact 2 candidate approaches that are being hotly debated. The first is based on USB technology whilst the second is based on a memory card interface (MMC). The weight of support appears to be behind the USB solution whereas the various company interests and IPR seem to prevent a final decision from being made. Both would probably do the job and so the industry risks wasting an opportunity to forge ahead with a standardised solution. Supporters of USB like the potential for PC compatibility, whereas handset vendors already have MMC interfaces in some models and claim that this interface would lead to a faster deployment of compatible handsets. Whatever the final choice a significant speed increase over conventional SCs is expected and some approximate rates are given in Table 2.

Table 1 SC and HDSC Interface Speed comparison

	I/O Speed	
	Typical	Maximum
SC	9.6kbits/s	78kbits/s
HDSC (USB)		8Mb/s

3.3. CPU Speed

A larger chip opens the way to more sophisticated and faster processors and increased availability of RAM would help with buffering and caching. CPU clock speeds in conventional SCs can reach up to about 66MHz using internal frequency multipliers and extending far beyond this still faces the issue of the power supply limitation. However in the current mood there seems to be a readiness to radically overall the existing standards and this might also extend to specifying a more generous power supply from the host device.

4. Implications

Purchasers of conventional smart cards might regard the HDSC as just a smart card with a lot of memory, but then they would have lived with the normal smart card role and limitations for so long that they are not on the look-out for a momentous-leap. Indeed the leap may not be of most benefit to the smart card's conventional role but in new applications and business opportunities. The first question is whether there really is a role for a large secured memory and an embedded tamper-resistant token.

4.1. Large Memory

The storage of sensitive data is a big issue whether it relates to personal privacy information, health records or the latest downloaded music or video clip. Where you decide to store this information shapes the architecture of the overall solution, its security, usability, portability and cost. There are similar issues associated with application hosting and portable memories are now so large that there is a real alternative to installing programs on PC hard drives. Table 2 attempts to summarise some characteristics of the various storage options and suggests that the HDSC might be good choice when high security and management is required.

Table 2 Memory Characteristics

	Mobile handset memory	Flash-Dongle	Flash-Dongle with password access	Conventional Smart Card	HDSC
Security	0	0	1	2	2
Capacity	2	2	2	0	2
Management	0	0	1	2	2

0 = Low, 1=Medium, 2= High

4.1.1. Data Rights Management

For Data Rights Management solutions the industry requires a large amount of data storage, coupled with a robust security solution and system for the management of portability of user rights. The HDSC could certainly be considered as a candidate. Not only does it have a large secured storage capacity but it is also portable between devices. Whilst some flash memory sticks have similar features they lack the centralised issuance and management functionality that the HDSC inherits from its SC predecessor. Data Rights could also move in a slightly different direction by relating to software program licences and execution. It is already possible to run programs direct from USB flash memories as an alternative to installing a program on a PC drive and the HDSC could provide formalised security for software

licensing. Indeed the users personal data and computing environment could be completely portable yet secure and any available computers could be used as temporary hosts.

4.1.2. Personal Data

For personal data and perhaps health record information there is a lot of discussion about conventional smart cards, however the most emotive debates are usually reserved for the databases that will hold the information. The HDSC would offer safe storage for personal information without exposing it to the vulnerabilities of a large centralised database. A distributed database could then exist as an alternative to the centralised approach or to complement it. Users could be in complete control of how their local information is dynamically released and used. Clearly a fallback solution would be needed to restore faulty, lost or stolen tokens but this need not reveal unencrypted information to unauthorised third parties.

Biometrics are a particular class of personal data and there is often debate about where this is best stored but also where verification or matching should take place. Ideally matching is carried out within the same tamper-resistant token used to store the reference information as that means that the reference need never level the protected device. Clearly the memory of the HDSC is sufficient for not only storing large amounts of biometrics data but also the algorithms used for matching and any associated certificates hashes or signatures.

4.1.3. Communication and Service Settings

Communications and applications access is becoming increasingly complex. This complexity translates to more configuration options and settings that must be tailored to the individual customers account and hardware whether that is a computer, PDA or mobile phone. If a customer accesses many services then the number of settings multiplies and different sets may be appropriate for geographic regions or during the lifetime of the systems and services. When a communication link is available then it is possible for the centralised system to dynamically download settings information but as the number of customers and services increases this can become difficult especially when customer participation is required. More convenient for the customers would be if all settings for all foreseeable devices and services were securely pre-stored with the HDSC and could be used automatically even when a communication link to the network is not available.

4.2. High Speed Interface and Processing

As mentioned previously, to exploit a large memory requires the fast interface speed that the HDSC provides. However a high-speed interface to a security token also provides opportunity to radically alter security architecture and protocols. The algorithms and security processes currently in use are a compromise based on security requirement, capabilities of the token and the maximum allowable time to execute. If the CPU in the HDSC runs faster than a SC, then more complex algorithms may be used and as more memory and I/O speed is available there is no need to use special compact formats for security data storage. It is also possible to challenge the long held assumption that the SC is too slow to carry out high rate encryption and decryption of the kind necessary to secure audio and video streams. For high rate media transfers such as in satellite TV [13] the solution was to use SCs to deliver session keys into untrusted hardware that could cope with the required rates. This had lead to attacks on security such as session key redistribution and card-sharing methods [14] used in open-satellite receivers. With a HDSC it may be possible to encrypt/decrypt directly within the trusted device and so the session keys need never pass into an untrusted environment. Clearly the feasibility of on-card ciphering/deciphering is also directly dependent on the speed of the CPU plus associated crypto-coprocessors and indirectly on the available power. There would also need to be changes to the messaging protocols as short APDU messages sent over a half-duplex link are unlikely to be the best way to maintain the necessary data flows.

4.3. Application Environment

The HDSC could initially retain the SC Java card platform/environment, with Global Platform for application and security domain management. However it is questionable whether this is how things would remain, as Java card was defined for a limited processing device and because of the "momentous leap" we now have a far more powerful platform. Potentially more capable versions of Java could be considered comparable with those intended for mobile handsets or normal PCs.
This may not be an easy transition, as whilst it could be relatively easy to redefine the Java Virtual Machine, the underlying OS may still be of the design used for conventional SCs. If the (U)SIM application is considered as an example we see a structure of relatively small files arranged in a hierarchical fashion. An OS that is designed and optimised for this scenario is unlikely to be suited for high-speed data processing on much larger data files. With such an increase in capability it is also reasonable to expect that general PC development practices to move into the HDSC space. In fact it is seriously proposed that the HDSC becomes an IP server platform which could lead to countless possibilities. Figure 4 gives an architectural view how new versions of Java card might accommodate both classic applets and server functionality.

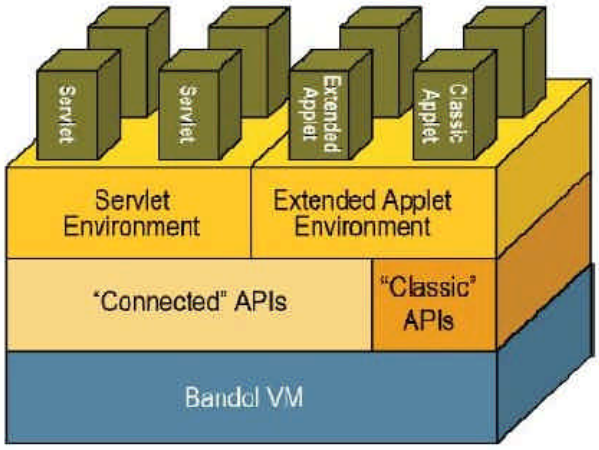

Figure 5 New Java Card Architecture - Bandol (source sagem-orga)

A user could for example browse through megabytes of automated customer care information and cached data or the IP communications could simply be used as the normal communications method for the HDSC rather than the master-slave APDU exchanges used in SCs. Of course with the good comes the bad and too much freedom on the HDSC platform will import the problems from the IP space and we may find it necessary to pat more attention to firewalls, trusted application lists, intrusion detection and virus checkers. However the big advantage is that this all takes place on a security managed and tamper-resistant platform.

5. Industry Impact

Having a new technology does not guarantee that anyone will buy it and in fact people rarely buy technology but rather the service or application that it enables. HDSCs are initially being touted to mobile network operators and as Intel and the GSM Association jointly announced plans [15] to fit a (U)SIM card in every laptop then HDSCs are also likely to be candidates for the PC market. If we take the SIM replacement as an example then we can soon discover that the introduction of this technology is not trivial and will not be achieved overnight. Firstly a (U)SIM is a commodity device and recent years have seen a downward trend in prices due to pressure from large purchasing groups, yet the HDSCs will be more expensive than SCs. A large purchasing group might order say 20 million cards per year and if the device suddenly becomes 5 Euros more expensive then the group has to sign an additional cheque for 100Million Euros. Very few companies would be prepared to do this and only then if they had a guaranteed scheme to leverage 100million+ from the introduction of HDSCs. Therefore the early mass-market introduction of HDSCs as SIM alternatives seems unlikely at least with respect to the existing purchasing model. The cards could be used for specialist applications such

as with high margin corporate customers as part of a business operation solution. There is also the possibility that end-users would wish to buy the cards as accessories. The cost of a HDSC could be similar to the cost of a normal flash memory card which customers are used to buying. In fact an interesting possibility is for the customer to buy a SIM-replacement HDSC based on some pre-installed media and download service. The linkage to media fits nicely with the capabilities of the device, supports portability, churn reduction and opens a new revenue stream without increasing operator (U)SIM purchasing costs. On the downside, the operator's claim to ownership and control of the SC/HDSC might be considerably weakened by the fact that it is purchased by the customer. Assuming that hurdle could be overcome then once a user has a HDSC then it is also a general-purpose application platform that could be used to offer a range of new revenue generating services. Other areas where operators could make or save money include existing service activation and customer care. Handset devices and advanced services may require complex configuration, to the extent where a proportion of users either are unable to activate a service or find it too much trouble. If the HDSC could have stored knowledge of common handset and service activation settings then a higher proportion of users would be enabled for the value added services that the operator is trying to market. Customer care is relevant as help-line calls are expensive to deal with and the process can cause increased customer dissatisfaction. If the HDSC can provide automated off-line help and intelligent access to network help services then the operator costs may be reduced and customer experience improved. This is of course ignoring the most important topic of handset compatibility. There is a significant lag between a standard being published and compatible handsets being manufactured. There is then an even longer delay until a significant proportion of the handsets in the market has the required capability. As the standardisation of the HDSC and in particular its physical interface is not yet standardised and the international experts continue to squabble, the mass-market commercial deployment of HDSCs seems some way off.

6. Concluding Remarks

The introduction of the HDSC appears to be a momentous leap forward in SC technology. Notably a 1000 fold increase in memory, a 100 times faster interface plus the promise of swifter processors. Such a change rips through many of the historical limitations of the SC and challenges our assumptions about the future role of smart cards. If the smart card is no longer a limited platform then it may be drawn into mainstream development methods common with the PC industry and even the prospect of an IP server based on a HDSC is no longer an idea to be ridiculed. Security architectures are also challenged, as a token that can also cipher/decipher high rate data can eliminate security weaknesses from other designs. However the introduction of the HDSC is by no means certain and for example it is unlikely that mobile operators would mass purchase more costly alternatives to the (U)SIM unless there was a very compelling application such as a DRM solution. More likely there would be a gradual introduction for specialist corporate customers or devices purchased by end-users. Whilst the latter solves the operator cost problem it touches on an interesting area as operators control SIMs by virtue of ownership, but would the customer own the HDSC? The biggest obstacle to the early introduction of HDSCs, seems to be the standardisation of the high-speed interface and the time it will take for compatible host devices to significantly penetrate the market. Perhaps the most important and unanswered question is whether the main buyers of smart cards actually want a momentous leap forward in capability or whether a low cost and expendable security token is still the main requirement

[1] M. Mouly, M-B Pautet, The GSM System for Mobile Communications, Cell & Sys. Correspondence 1992
[2] ETSI. GSM 11:11 - Digital cellular telecommunications system (phase 2+); Specification of the Subscriber Identity Module - Mobile Equipment (SIM-ME) interface, Version 8.3.0 1999
[3] International Standard Organisation, " ISO/IEC 14443-X Identification cards - Contact-less integrated circuit(s) cards - Proximity cards" 2000.
[4] International Standard Organisation “ISO/IEC 10373, Identification cards - Test Methods - Part 1 General characteristics” 2003.
[5] Java Card, “Java Card API 2.2.1 Reference Implementation”, 2002, URL: http://www.Javasoft.com/products/Javacard/
[6] GlobalPlatform, “GlobalPlatform Card Specification”, Ver. 2.1, 2001, www.globalplatform.org
[7] ETSI. GSM 11:14 - Digital cellular telecommunications system (phase 2+); Specification for the SIM Application Toolkit for the Subscriber Identity Module - Mobile Equipment (SIM-ME) interface, Version 8.3.0 1999
[8] EMV1, “EMV’96 Integrated Circuit Card Specification for Payment Systems”, Version 3.0, http://www.europay.com/Pdf/EMV_card.pdf.es 1996
[9] International Standard Organisation, “ISO/IEC 7816, Information technology - Identification cards - Integrated circuit(s) cards with contacts- Part 4 Interindustry commands for interchange” (1995)
[10] Multos, 2002, www.multos.org
[11] “Windows Powered Smart Cards”, 1999, URL: http://www.microsoft.com
[12] ETSI SCP TS 102.412 release 7.
[13] ETSI, Digital Video Broadcasting (DVB); Support for use of Scrambling and Conditional Access (CA) within Digital Broadcasting Systems, European Telecommunications Standards Institute (ETSI), Sophia Antipolis, France, ETSI Technical Report ETR 289, Oct. 1996.
[14] L. Francis, W. Sirett, K. Markantonakis, K. Mayes, "Countermeasures for Attacks on Satellite TV Cards using Open Receivers", Third Australasian Information Security Workshop (AISW2005): Digital Rights Management.
[15] GSM Association and Intel Press Release " GSMA and Intel to enable laptops to connect automatically to high-speed 3G broadband data and Wi-Fi networks via SIM" Feb 14th 2005 Barcelona www.gsmworld.com/news/press_2006/press06_10.shtml