

Secure M-Commerce¹

Scarlet Schwiderski-Grosche¹, Heiko Knospe²

¹Information Security Group, Royal Holloway University of London, Egham TW20 0EX, UK
Email: scarlet.schwiderski-grosche@rhul.ac.uk

²T-Systems Nova GmbH, Technologiezentrum ES21d, D-64307 Darmstadt, Germany
Email: heiko.knospe@t-systems.com

Abstract

M-commerce is a major application domain for mobile devices, enabling users to perform commercial transactions wherever they go. However, these applications require a high level of security. In this paper, we identify the special characteristics of m-commerce and reflect on some important security issues.

1 Introduction

The term *e-commerce* (electronic commerce) denotes business processes on the Internet, such as the buying and selling of goods. There is a distinction between *B2B* (business-to-business) and *B2C* (business-to-consumer) markets. In the first case, the business processes are carried out between businesses; in the latter case, they are carried out between businesses and end consumers. This general definition of e-commerce does not say anything about the kind of device that the end user employs to gain access to the Internet. The underlying technology could be wireline (e.g. using a home PC as end user device) or wireless (e.g. using a mobile phone as end user device).

The term *m-commerce* (mobile commerce) is all about wireless e-commerce, that is, where mobile devices are used to do business on the Internet, either in the B2B or B2C market. As such, m-commerce is a subset of e-commerce.

With the omnipresent availability of mobile phones (and other mobile devices), m-commerce services have a promising future, especially in the B2C market. Future applications include buying over the phone, purchase and redemption of ticket and reward schemes, travel and weather information, and writing contracts on the move. However, the success of m-commerce very much depends on the security of the underlying technologies. For example, today the chargeback rate for credit card transactions on the Internet is 15%, versus 1% for POS (Point-of-Sales) credit card transactions. Chargeback rates grow to 30% when digital products are sold. For m-commerce to take off, fraud rates have to be reduced to an acceptable level. As such, security can be regarded as an enabling factor for the success of m-commerce applications. In this paper, we discuss two main areas of m-commerce that are relevant to security, namely

- *Network technology* – in m-commerce, all data is transmitted via a mobile telecommunication network. Here, we consider existing network and service technologies for 2G (2nd Generation), 3G (3rd Generation) and other wireless systems.

¹ The work described in this paper has been supported by the European Commission through the IST Programme under Contract IST-2000-25350. The information in this document is provided as is, and no guarantee or warranty is given or implied that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.

- *M-payment* (mobile payment) – doing business on the Internet requires the payment of goods and services. M-payment systems have different requirements and characteristics than e-payment systems. Here, we give an overview of current payment technology.

In the following section, we provide background information on m-commerce. Section 3 is devoted to network technologies for m-commerce. Different approaches to m-payment will be discussed in Section 4.

2 M-commerce and its security challenges

2.1 Definition m-commerce

There are many definitions of the term m-commerce [1] [4]. Common to all definitions is that a terminal or mobile device is employed to communicate over a mobile telecommunication network. There are different views as of the purpose of this communication. Some definitions restrict m-commerce to transactions involving a monetary value, whereas other definitions generalise the term to services that involve communication, information, transaction, and entertainment. Summarising, we define m-commerce as using a mobile device for business transactions performed over a mobile telecommunication network, possibly involving the transfer of monetary values.

2.2 Mobile devices

M-commerce is not just about using mobile phones as end user devices. The following list gives an overview of different kinds of mobile devices:

- Mobile phone
- PDA (Personal Digital Assistant)
- Smart phone – the smart phone combines mobile phone and PDA technology into one device.
- Laptop
- Earpiece (as part of a Personal Area Network)

Each mobile device has certain characteristics that influence its usability, such as

- Size and colour of display
- Input device, availability of keyboard and mouse
- Memory and CPU processing power
- Network connectivity, bandwidth capacity
- Supported operating systems (e.g. PalmOS, Microsoft Pocket PC)
- Availability of internal smart card reader (e.g. for a SIM card in mobile phones)

Depending on these factors, the services that the end user can receive differ considerably. Moreover, depending on the network technology used for transmission, the bandwidth capacity varies and influences the kind of services that the end user is able to receive.

In mobile phones, there exist three solutions to internal smart cards: *single SIM*, *dual chip*, and *dual slot*. Single SIM is the solution that is most widely available today, where all confidential user information is stored on one smart card. Dual chip means that there are two smart cards in the mobile phone, one for user authentication to the network operator and one for value-added services like m-payment or digital signature. A dual slot mobile phone has a

SIM card and a card slot for a full-sized external smart card. With this solution different cards can be used one after the other. Moreover, the cards can also be used in traditional POS and ATM terminals.

2.3 Differences to e-commerce

In comparison to e-commerce, m-commerce offers both advantages and disadvantages. The following list summarises the advantages of m-commerce [1]:

- Ubiquity – the end user device is mobile, that is, the user can access m-commerce applications in real time at any place.
- Accessibility – accessibility is related to ubiquity and means that the end user is accessible anywhere at any time. Accessibility is probably the major advantage by comparison with e-commerce applications involving a wired end user device.
- Security – depending on the specific end user device, the device offers a certain level of inherent security. For example, the SIM card commonly employed in mobile phones is a smart card that stores confidential user information, such as the user's secret authentication key. As such, the mobile phone can be regarded as a smart card reader with smart card.
- Localisation – a network operator can localise registered users by using a positioning systems, such as GPS, or via GSM or UMTS network technology, and offer location-dependent services. Those services include local information services about hotels, restaurants, and amenities, travel information, emergency calls, and mobile office facilities.
- Convenience – the size and weight of mobile devices and their ubiquity and accessibility makes them an ideal tool for performing personal tasks.
- Personalisation – mobile devices are usually not shared between users. This makes it possible to adjust a mobile device to the user's needs and wishes (starting with the mobile phone housing and ringtones). On the other hand, a mobile operator can offer personalised services to its users, depending on specified user characteristics (e.g. a user may prefer Italian food) and the user's location (see above).

The following list summarises the main disadvantages of m-commerce:

- Mobile devices offer limited capabilities (such as limited display, see Section 2.2). Between mobile devices these capabilities vary so much that end user services will need to be customised accordingly.
- The heterogeneity of devices, operating systems, and network technologies is a challenge for a uniform end user platform. For this reason, standardisation bodies consisting of telecommunication companies, device manufacturers, and value-added service providers integrate their work (see Section 4.5). For example, many current mobile devices implement an IP stack to provide standard network connectivity. At the application level, the Java 2 Micro Edition (J2ME) offers a standardised application platform for heterogeneous devices.
- Mobile devices are more prone to theft and destruction. According to a government report, more than 700000 mobile phones are stolen in the UK each year [12]. Since mobile phones are highly personalised and contain confidential user information, they need to be protected according to the highest security standards.
- The communication over the air interface between mobile device and network introduces additional security threats (e.g. eavesdropping, see Section 3).

2.4 Security challenges

As mentioned earlier, m-commerce is not possible without a secure environment, especially for those transactions involving monetary value. Depending on the point of views of the different participants in an m-commerce scenario, there are different security challenges. These security challenges relate to

- the mobile device – confidential user data on the mobile device as well as the device itself should be protected from unauthorised use. The security mechanisms employed here include user authentication (e.g. PIN or password authentication), secure storage of confidential data (e.g. SIM card in mobile phones) and security of the operating system.
- the radio interface – access to a telecommunication network requires the protection of transmitted data in terms of confidentiality, integrity, and authenticity. In particular, the user's personal data should be protected from eavesdropping. Different security mechanisms for different mobile network technologies (i.e. in 2G, 3G, and other systems) will be presented in Section 3.
- the network operator infrastructure – security mechanisms for the end user often terminate in the access network. This raises questions regarding the security of the user's data within and beyond the access network. Moreover, the user receives certain services for which he/she has to pay. This often involves the network operator and he/she will want to be assured about correct charging and billing.
- the kind of m-commerce application – m-commerce applications, especially those involving payment, need to be secured to assure customers, merchants, and network operators. For example, in a payment scenario both sides will want to authenticate each other before committing to a payment. Also, the customer will want assurance about the delivery of goods or services. In addition to the authenticity, confidentiality and integrity of sent payment information, non-repudiation is important. Section 4 will review security mechanisms for m-payment.

3 Security technologies relevant for m-commerce

In this section, we give an overview of the technologies which are relevant to secure m-commerce transactions. We focus on those network and service technologies which are specific to mobile devices. The security architecture of current and potential future mobile systems has been studied in the IST SHAMAN project².

3.1 Security of Network technologies

We first discuss the security of network technologies used for mobile commerce.

GSM

GSM (Global System for Mobile Communication) is the current European standard for mobile communications. Since GSM handsets are popular and widespread, they have to be considered as the major device for mobile commerce at the moment. In the first years of GSM (beginning of the 1990s), the devices were very limited with respect to their capabilities other than telephony. Dial-in data sessions over circuit switched connections were possible but relatively slow (9.6 kBit/s) and required a separate device (computer) which reduces mobility.

² IST-2000-25350 SHAMAN, <http://www.ist-shaman.org>.

As the GSM core network was extended with more and more data service elements, the cellular phones also became more powerful. A number of data services were established:

- SMS (Short Message Service) allows the exchange of 160 character short messages over the signaling channel.
- WAP (Wireless Application Protocol) permits access to internet content and applications formatted in WML (Wireless Mark-up Language). At first, WAP was based on circuit switched connections.
- HSCSD (High Speed Circuit Switched Data) provides higher data rates by channel bundling.
- GPRS (General Packet Radio Service) extends GSM with packet oriented services. With GPRS, the mobile node can stay “always on” without blocking a connection timeslot with the base station. GPRS can also be used as a bearer service for WAP and SMS.

The basic architecture of GSM including GPRS, IN (intelligent network) and SMS components is depicted in Figure 1 below:

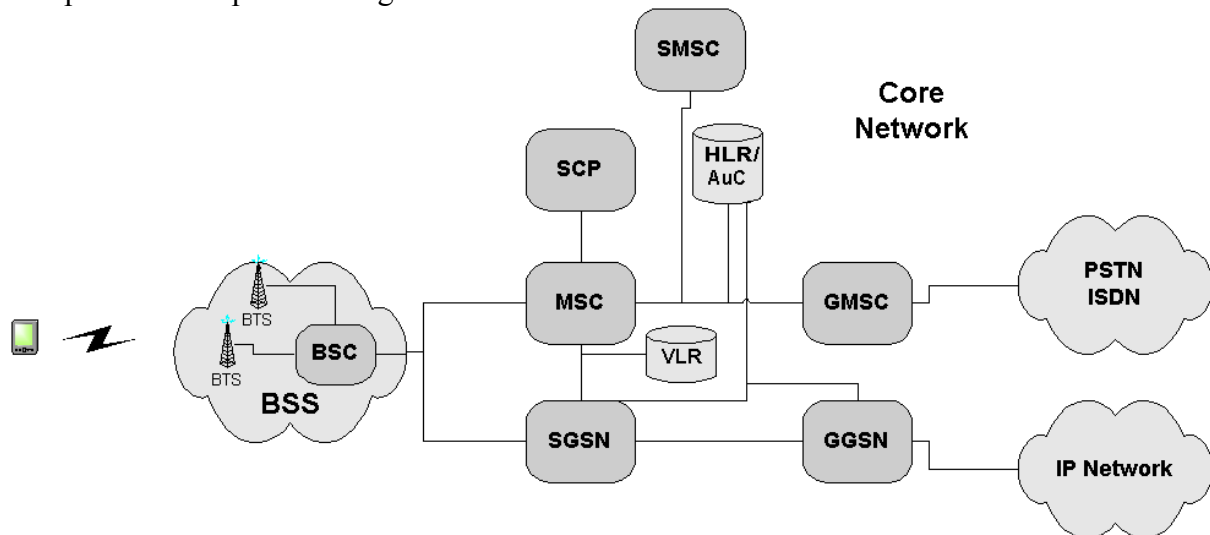


Figure 1 : GSM architecture

The mobile station communicates over the wireless interface with a base transceiver station (BTS) which is part of a base station subsystem (BSS). The base station controller (BSC) is connected with a MSC (Mobile Switching Centre) and a SGSN (Serving GPRS Support Node). The latter two are the central switching components for circuit and packet switched data.

When a customer subscribes, the GSM home network assigns the mobile station a unique identifier, the international mobile subscriber identity (IMSI), and an authentication key K_i . The IMSI and the secret authentication key K_i of the mobile station (MS) are stored in the SIM (subscriber identity module), which is assumed to be tamper proof. On the network side, the IMSI, K_i and other information are stored in the HLR (Home Location Register) and AuC (Authentication Centre).

GSM provides the following security features for the link between the mobile station and the network, as described in [5], [16]:

- IMSI confidentiality
- IMSI authentication
- User data confidentiality on physical connections
- Connectionless user data confidentiality
- Signalling information element confidentiality

GSM provides the basic security mechanisms for m-commerce transactions. In particular, the *mobile customer authenticates* towards the network with a challenge/response protocol based on the secret key K_i . Furthermore, the *wireless link between the mobile station and the BTS is encrypted* with a symmetric key which is also derived from K_i . The secret key K_i is never sent over the network. But there are weaknesses (cf. [16]): since the network is not authenticated, a false base station can perform a “man-in-the-middle” attack. The base station can suppress IMSI confidentiality and encryption and this is not even visible to the mobile station.

UMTS

UMTS (Universal Mobile Telecommunication System) is the next generation (3G) mobile telecommunication system and a further development of GSM. The major difference to GSM is the radio network (UTRAN) with its transition to the WCDMA (Wideband Code Division Multiple Access) radio technology. Two new network components, RNC (Radio Network Controller) and Node B are introduced in UTRAN. Furthermore, the security protocols have been modified and now the RNC is responsible for de-/ciphering.

The main components of the GSM/GPRS core network with MSC, SGSN etc. can be reused or evolved to UMTS. This is depicted in Figure 2 below:

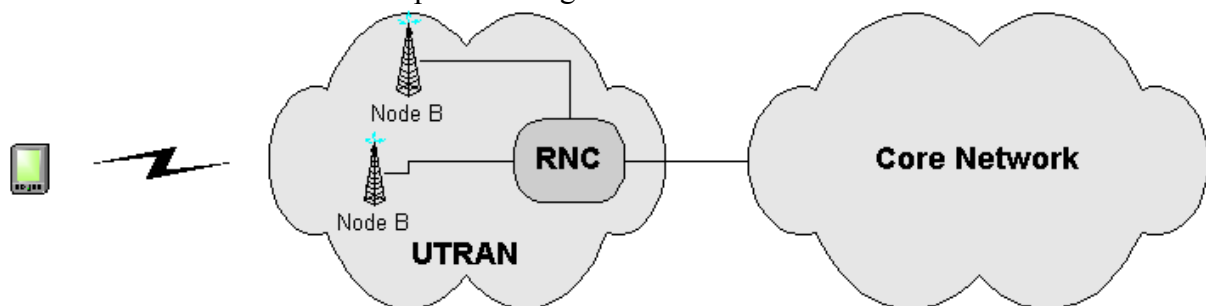


Figure 2 : UTRAN system

In general, the security architecture of UMTS is carefully designed to fix the security weaknesses of GSM (cf. [6], [16]). As is described above, the main problems of GSM originate from two facts: authentication is one way (the mobile station does not authenticate the network), and encryption is optional. In UMTS, authentication is mutual, and encryption is mandatory unless the mobile station and the network agree on an unciphered connection. In addition, integrity protection is always mandatory and protects against replay or modification of signalling messages. Sequence numbers in authentication vectors protect against reuse of authentication vectors by network impersonators. UMTS introduces new cipher algorithms and longer encryption keys. Thus, UMTS does not seem to have any obvious security holes. UMTS security is discussed in detail in article [10] in this volume.

WLAN

The IEEE standard 802.11 specifies families of Wireless Local Area Networks (WLAN) which operate in the unlicensed 2.4 GHz and 5 GHz band. The standards specify the physical layer (PHY) and the medium access control layer (MAC). For the network layer and above, WLAN employs a classical IP stack. A number of commercial products (even for PDAs) are available, and IEEE 802.11b, offering 11 Mbit/s raw bandwidth, is currently very popular. When operated in the infrastructure mode, the mobile station attaches to an Access Point (AP) which provides connectivity to fixed net IP networks (e.g. the internet) or to other mobile stations.

In the default mode, WLAN does not provide any security. This means that a mobile attacker can eavesdrop and manipulate all the wireless traffic with standard tools.

In order to provide a certain level of security, the IEEE defined WEP (Wired Equivalent Privacy). WEP was designed to provide:

- Authentication to protect the association to an AP
- Integrity protection of MAC frames
- Confidentiality of MAC frames

The protection is based on secret WEP keys of either 40 or 104 bits. Concatenated with a clear text initialisation vector, the secret key serves as input for the RC4 stream cipher. But it has been shown (cf. [17]) that *authentication and integrity protection is completely insecure and encryption at least partly insecure*. It suffices for an attacker to intercept a single successful authentication exchange between a mobile station and the AP to be able to authenticate without knowing the secret keys. Furthermore, since a CRC checksum is used for integrity protection, an attacker can modify the data and adapt the checksum accordingly. For example, if the position of commercially sensitive information (e.g. an amount) within a datagram is known, the corresponding bits can be ex-ored with any value. With a large number of intercepted frames, the WEP keys can even be recovered, breaking the encryption. Furthermore, since the WEP keys are *network keys*, preserving their secrecy is difficult for private networks and impossible for public WLAN hotspots.

In recent work of the IEEE Task group on security (TgI), the new security standard IEEE 802.1X (cf. [18]) has been adopted. 802.1X is a framework for authentication and key management which employs the Extensible Authentication Protocol (EAP, cf. [19]) for a variety of authentication mechanisms, e.g. certificate based TLS. But the weaknesses of WEP cannot be remedied by the new authentication and key management schemes in 802.1X. The IEEE is currently working towards a new standard (WEP2), and a number of proposals are in circulation.

Another approach is to employ VPN (virtual private network) technologies and in particular IPsec in order to establish *network layer* security. The IPsec protocol (or more specifically the ESP Tunnel protocol) is an internet standard (cf. [8]) for the protection of IP packets between two nodes (e.g. a mobile station and a security gateway). This architecture is depicted in Figure 3 below. Note that link layer specific information (e.g. MAC addresses) are still unprotected.

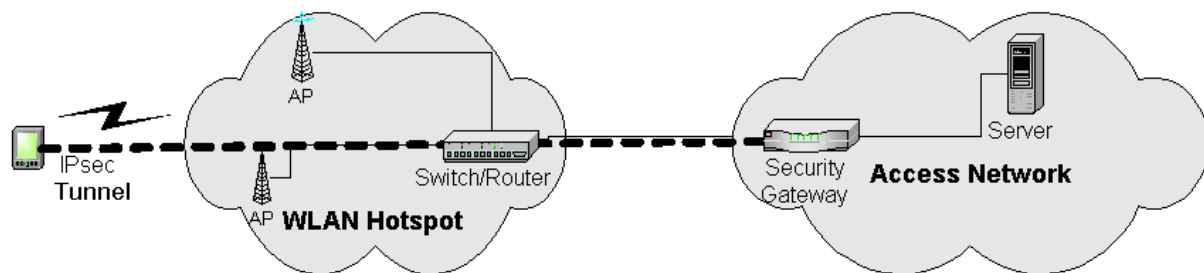


Figure 3 : WLAN security with IPsec

Bluetooth

Bluetooth is a wireless technology developed by the Bluetooth Special Interest group³ and is mainly aiming at ad hoc piconets and connections to peripheral devices. Bluetooth is also operating in the unlicensed 2.4 GHz band and can be considered as a de-facto-standard. The Bluetooth specification defines a complete OSI stack, so, unlike WLAN, it is not restricted to IP connectivity. Although raw bandwidth is limited to 1 MBit/s, the Bluetooth technology will probably often be used in the future to connect devices in the personal environment, which makes it relevant for m-commerce.

³ Bluetooth Special Interest Group: <http://www.bluetooth.com>.

Bluetooth specifies three security modes, including “no security”. Bluetooth provides link layer security with a challenge-response protocol for authentication and a stream cipher encryption of user and signalling data (cf. [20]). When the connecting devices do not share a key in advance, they have to establish an initialisation key in a pairing procedure. This is based on a PIN, which must be entered into both devices (or imported from some application). Bluetooth can currently be considered secure for small ad hoc networks, provided the pairing happens in a safe environment and the PIN is strong enough. The existing attacks are still theoretical in nature. However, privacy requirements may not be met since the Bluetooth device address (unique MAC address) allows the tracing of personal devices and hence their owner.

3.2 Transport Layer Security

The above technologies provide security for the wireless link between mobile customer and access network or access device. If the access network is considered secure and the m-commerce transaction is completely handled within the access network, this may be sufficient. But often, an m-commerce transaction involves parties outside the access network (merchant, payment service provider etc.). In this section, we discuss end-to-end security for mobile devices. This protects applications which communicate over an IP port. More information, in particular on certificates and PKI, can be found in [11].

SSL/TLS

The SSL/TLS (Internet Secure Socket Layer, cf. [9]) protocol is by far the most widely used internet security protocol. Its main application is the HTTPS protocol (HTTP over SSL), but it may also be used as a standalone protocol. SSL requires a bidirectional byte stream service (i.e. TCP). SUN has implemented a client side version of SSL for limited devices, called KSSL (Kilobyte SSL). KSSL does not offer client side authentication and only implements certain commonly used cipher suites, but it has a very small footprint and runs on small devices using the J2ME platform.

WTLS

The WAP forum has standardised a transport layer security protocol (WTLS) as part of the WAP 1 stack (cf. [21]). WTLS provides transport security between a WAP device (e.g. a mobile phone) and a WAP gateway which performs the protocol transformation to SSL/TLS. Hence, no real end-to-end security is provided and the WAP Gateway needs to be trusted. Note that the WAP Forum now proposes a WAP 2 stack (cf. [22]) which is a classical TCP/IP stack on a wireless bearer medium. This permits end-to-end SSL/TLS sessions.

3.3 Service Security

Here, we discuss the security of network services which can be used for m-commerce transactions.

Intelligent Network

With the introduction of the IN (Intelligent Network) technology to GSM networks, additional services could be realised. The IN architecture for GSM (called CAMEL, Customised Application for Mobile Enhanced network Logic, cf. [23]) was adapted from the fixed network standard ETSI Core INAP, and was originally designed for circuit switched calls (CAMEL phase 1 and 2). The IN is triggered during call handling at the MSC if the HLR entry indicates subscription to an IN service. With CAMEL phase 3 (cf. [24]), the IN services can also be applied to SMS and to packet data services. The IN component SCP (Service

Control Point) controls the call or data service via the CAMEL Application Part (CAP) protocol which runs on top of the SS7 (Signalling System Number 7) protocol.

Prominent examples of IN services are the transformation of dialled numbers (e.g. to realise Virtual Private Nets) and prepaid services. The IN platform provides some flexibility for the generation of m-commerce services. IN handling can e.g. be triggered by a specific called party, a calling party, an USSD string (requiring CAMEL phase 2), mobile originating SMS (requiring CAMEL phase 3) or mobile terminating SMS (requiring CAMEL phase 4).

The security of an IN service depends on the underlying GSM or UMTS network security (see above) and on the specific characteristics of the service application.

Parlay/OSA

Parlay/OSA (Open Service Access) is an initiative of the industry (Parlay group), ETSI and 3GPP and aims at introducing standard interfaces to network services (cf. [25]). The IN platform and their SS7 based protocols like INAP and CAP are relatively complex and generation of services is reserved to operators and manufactures. Now Parlay offers standard application programming interfaces which allows service provisioning on IT platforms using standard middleware (e.g. CORBA). The Parlay/OSA framework then provides gateway functionality between applications and Service Capability Features (SCF's) of the IN. M-Commerce applications can then access core network functionality, e.g. inquire status and location of a mobile user, send messages or place calls. Parlay/OSA applications are portable among networks which is usually not possible with IN services.

Security is an important issue, since Parlay/OSA potentially opens the core network to intruders. Parlay/OSA specifies authentication and encryption on the application layer (cf. [26]). But the security also depends on the underlying network architecture, e.g. firewalls and strict policies should protect core network components.

SMS

SMS (short message service) is a very popular data service for GSM networks. Although SMS messages are limited to 160 characters, a considerable number of m-commerce scenarios are based on this service. The sender and receiver of an SMS is identified by its IMSI which an attacker cannot forge without breaking the GSM/UMTS security mechanisms (e.g. by cloning a SIM card). Hence SMS messages can be used for authentication (at least towards the network). Furthermore, SMS data is transmitted in the GSM (UMTS) signalling plane, which ensures the confidentiality of messages. However, the protection ends in the GSM or UMTS network, there is no end-to-end security, and the network operator and its infrastructure (e.g. SMSC, Short Message Service Centre) must be trusted (when no other security mechanisms are applied to the SMS message, confer section on SIM/USIM Applications below).

USSD

The GSM Unstructured Supplementary Service Data (USSD) service allows data communication between a mobile station and either the HLR, VLR, MSC or SCP in a way transparent to the other network entities. Unlike the asynchronous SMS service, an USSD request opens a session which may induce other network operations or an USSD response before releasing the connection. Mobile originated USSD may be thought as a trigger for a network operation. USSD works with any mobile phone since the coded commands are entered in the same way as a phone number (e.g. *123#1234567890#).

With USSD, roaming can be offered for prepaid GSM customers before IN services (CAMEL) are implemented in a network. Another USSD application (requiring CAMEL phase 2) is replenishing a prepaid account by incorporating the voucher number in an USSD

string. In principle, any transaction, e.g. a payment operation, could be triggered by USSD data.

USSD possesses no separate security properties; instead it relies on the GSM/UMTS signalling plane security mechanisms.

SIM/USIM Application Toolkit

The SIM and USIM Application Toolkits (SAT and USAT respectively) allow operators and other providers to create applications which reside in the SIM/USIM. These applications can e.g. send, receive and interpret SMS or USSD strings. Currently, there exist banking applications using SAT. In [7], a format for secured packets is defined. This permits the sending application (e.g. the one residing on the SIM card) to send protected messages to the receiving application (which e.g. runs at a payment service operator). The required security mechanisms are:

- Authentication
- Message Integrity
- Replay detection and sequence integrity
- Proof of receipt and proof of execution
- Message Confidentiality
- Indication of the security mechanisms used

However, it depends on the applications whether these security mechanisms are implemented and whether their cryptographic strength is sufficient.

4 M-payment

This section gives an introduction to payment mechanisms for m-commerce. First, we will motivate the heterogeneity of payment system solutions, before categorising e-payment and m-payment systems respectively. Section 4.4 states examples of m-payment systems.

4.1 Background on payment systems

E-payment systems provide means for payment of goods or services over the Internet. In contrast to conventional payment systems, the customer sends all payment-related data to the merchant over the Internet; no further external interaction between customer and merchant is required (e.g. sending an invoice by mail or confirmation by fax). To date, there exist more than 100 different e-payment systems [15], [2], [3].

Distinctive features of payment systems

There exist a large number of payment systems for e- and m-payments. The following distinctive features of payment systems motivate this diversity.

- *Time of payment*
- *Payment amount*
- *Anonymity issues*
- *Security requirements*
- *Online or offline validation*

Time of payment denotes the relation between the initiation of a payment transaction and the actual payment. In *pre-paid payment systems*, the customer's account is debited before the payment and the amount is stored, for example, on smart cards, in specific customer accounts or as electronic cash. In *pay-now payment systems*, the customer's account is debited at the time of payment (for example, ATM card or debit card with PIN) and in *post-payment*

systems, payment can be regarded as a “payment promise” where the merchant’s account is credited before the customer’s account is debited (for example, credit card systems).

The *payment amount* has an influence on the design of electronic payment protocols. For example, payments in the order of 1 € are only viable if the incurred computational and communications overhead is kept small. Accordingly, there is a distinction between

- *micropayments* (up to about 1 €)
- *small payments* (about 1 to 10 €)
- *macropayments* (more than about 10 €)

Electronic payment systems often originate with conventional payment systems. As such, cash-like payment systems should provide *anonymity* to the customer. There are different degrees of anonymity: complete anonymity means that the customer remains anonymous to the merchant and the bank. However, in many payment systems, only partial (or no) anonymity can be provided.

The *security requirements* of electronic payment systems differ. Generally, integrity, authentication, authorisation, confidentiality, availability, and reliability issues need to be considered, depending on the specific requirements of an electronic payment system.

Offline payment validation means that no third party (e.g. a bank or credit card institution) is involved during the payment procedure, whereas *online payment validation* involves some kind of background payment server as a trusted third party. The latter causes an additional communication overhead, but reduces certain risks, e.g. double spending.

The above discussion summarises some distinctive features of payment systems. There are other issues such as

- Overhead imposed on customers and merchants (e.g. installation of software, registration)
- Performance (e.g. response times)
- Cost incurred per payment transaction
- Fulfilment of the ACID (Atomicity, Consistency, Isolation, Durability) principle for payment transactions (i.e. transactions have to be executed all or nothing, leaving the system in a consistent state, and their effect should be durable)
- National or international deployment

The above list of distinctive features gives an idea of the complexity and variety of payment systems.

4.2 Categorisation of e-payment systems

E-payment systems are typically modelled on conventional payment systems. As such, there are the following categories [3]:

- Direct cash
- Cheque
- Credit card
- Bank transfer
- Debit advice

Direct cash – in direct cash like payment systems, the customer withdraws money from the issuer, that is, the third party interacting with the customer (for example, a bank or service provider), and hands payment tokens for the payment amount to the merchant. The merchant

deposits the payment tokens with its acquirer, that is, the third party interacting with the merchant (for example, a bank or service provider). The issuer and acquirer then settle the payment. This payment scenario is sketched in Figure 4.

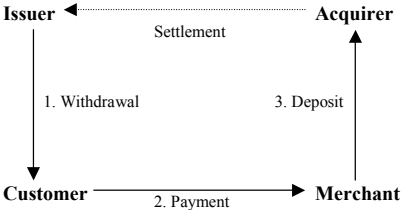


Figure 4: Direct cash like payment system

Since digital cash is trivial to copy, direct cash like payment systems involve either tamper proof hardware (i.e. smart cards) or online validation by the issuer (i.e. double spending test).

Cheque – in this scenario, the customer hands a cheque (a payment authorisation) to the merchant. The cheque is presented to the acquirer who redeems it from the issuer. Cheque-like payments are sketched in Figure 5.

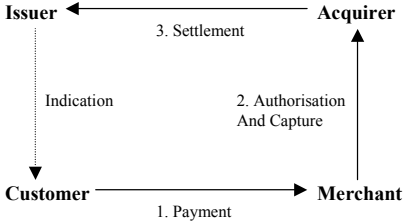


Figure 5: Cheque like payment system

Credit card – in terms of the information flow, credit card based payment systems are similar to cheque-like payment systems, with the difference that credit card based payment systems use the existing credit card infrastructure for settling the payment.

Bank transfer – the bank transfer model is sketched in Figure 6. Here, the customer instructs the issuer to transfer money to the merchant’s account at the acquirer. The merchant is notified of the incoming payment.

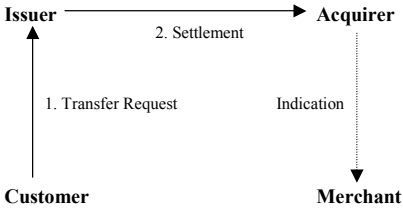


Figure 6 : Bank transfer payment system

Debit advice – this model describes the opposite case to the bank transfer model. The merchant instructs the acquirer to charge the account at the issuer. The customer is notified of the outgoing payment.

4.3 Categorisation of m-payment systems

Most e-payment systems are not suitable for use in a mobile context, that is, using a mobile device and communicating over a mobile telecommunication network. This is due to the special characteristics of mobile devices and mobile telecommunications (see Sections 2.2 and 2.3). In the following, we categorise m-payment systems according to the whereabouts of the customer’s money:

1. Software electronic coins – electronic money stored on the mobile device in file format.

2. Hardware electronic coins – electronic money stored on the mobile device on a smart card.
3. Background account – electronic money stored in a remote account at a trusted third party.

Software electronic coins – In this case, monetary value is stored on the mobile device and the customer has full control of his/her money wherever he/she goes and whatever he/she does. An electronic coin is represented as a file containing, among other information, a value, a serial number, a validity period, and the signature of the issuing bank. Since software electronic coins are easy to copy, the validity of an electronic coin depends on its uniqueness in terms of its serial number. The customer transfers electronic coins to the merchant, who forwards them to the issuing bank for the “double spending test”. In this test, it is checked whether the electronic coin has been spent beforehand. If yes, it is rejected. Otherwise, its serial number is entered into the double spending database and the money is credited to the merchant’s account⁴. The generation and storage of electronic coins is an orthogonal problem. Due to the limitations of mobile devices, electronic coins may have to be generated and stored externally, until they are downloaded onto the mobile device.

Hardware electronic coins – In this case, monetary value is stored on a secure hardware token, typically a smart card, in the mobile device. The presentation of electronic money is not important, as long as it is stored securely on the smart card. Electronic money could be represented as a simple numeric counter. In order to get to the money, the customer’s smart card and the merchant’s payment server authenticate each other and a secure channel is set up between them. Then, electronic money can be transferred from one to the other. This approach is quite attractive because smart cards provide an additional level of mobility. That means the payment smart card can also be used in POS transactions.

Background account – Here, the money is stored remotely on an account at a trusted third party. Depending on the specific payment system, the account could be a credit card account, a bank account, or an account held at the network operator. Common to all scenarios is that, on receipt of an invoice, the customer sends an authentication and authorisation message to the merchant that allows the trusted third party (that holds the account) to identify the customer and to verify the payment authorisation. The accounts can then be settled. There are numerous payment systems that fall into this category. The differences are regarding the nature of the trusted third party and the procedure to send authentication and authorisation data. For example, in some cases this data is sent in the clear (e.g. a credit card authorisation) not providing any security against eavesdropping, and in some cases this information is encrypted and digitally signed, providing anonymity to the customer (e.g. SET – Secure Electronic Transactions).

4.4 Examples of m-payment systems

In the previous section, we gave a categorisation of m-payment systems. In this section, we survey existing payment solutions for m-payment, as well as e-payment solutions that are suitable for mobile use.

Software electronic coins – There are several e-payment systems that are based on electronic coins. As for cash, one main advantage is that the customer can potentially remain completely anonymous to the merchant as well as to the bank, while staying in full control of the money. E-commerce solutions of this type include eCash, NetCash, and MilliCent. Due to the storage and processing constraints of mobile devices, an adaptation of the software is necessary. Moreover, storing electronic coins on the mobile device is problematic. One option is to run the full-fledged payment system on a home PC and download electronic coins when needed.

⁴ This description is based on the eCash electronic payment system.

Hardware electronic coins – There are various e-payment systems that implement an e-purse, that is, electronic cash on a smart card, for example GeldKarte and Mondex. In both cases, electronic money is stored on the card and can be transferred directly from the customer to the merchant. Shadow accounts are held at the bank to log transactions. Currently, these payment schemes are being adapted for m-commerce, where the GeldKarte and Mondex cards can be used in dual slot mobile phones. One smart card based system for m-commerce is already in operation; the system by BarclayCard and Cellnet uses a dual-slot Motorola mobile phone for payments of up to £50.

Background account – Depending on the type of trusted third party, there are various different approaches to a background account solution.

If the background account is held at a network operator, the charged amount is transferred to the existing billing solution and included in the customer bill. Customers pay their bills using traditional systems, such as direct debit, cheque or cash. Examples in this category are the M-Pay Bill service from Vodafone and Mobilepay by Sonera. These systems work for micro-payments only; accepting higher payments would imply that mobile operators become subject to a host of banking regulations.

If the background account is held at a credit card institution, the payment mechanism is all about secure transmission of credit card data to the credit card company. The first possibility uses a dual slot mobile phone, whereas the second possibility employs a dual chip mobile phone. The dual slot solution (for example, ItiAchat by Mastercard, Oberthur smart cards, France Télécom, Europay, and Motorola) has the advantage that it works with “normal” credit cards that can also be used in traditional POS terminals. On the other hand, dual slot mobile phones are quite bulky. This solution is favoured by credit card companies and banks, since it allows them to stay in control of the payment functionality. The dual chip solution (for example, EMPS – Electronic Mobile Payment System by MeritaNordbanken, Nokia and Visa) has the advantage that handsets can be kept small. This solution is favoured by mobile operators, because they are in control of the dual chips.

Finally, if the background account is held at a bank, the existing banking infrastructure and technology can be reused. Examples are Paybox and MobiPay by BBVA and Telefónica. Typically, the merchant receives the customer’s mobile phone number (or a pseudonym) and passes it on to the payment server, together with the payment details. The customer authorises the payment by providing a PIN.

In [13] and [14], we have studied the use of software electronic coins and hardware electronic coins for a specific application in m-commerce, namely for ad hoc payment of the access of a mobile device to an access network with which no former relationship exists (i.e. the access network is not the home network and does not have a roaming agreement with the home network). Hence, the service for which the user of the mobile device pays is the current use of this device.

4.5 Standardisation and forums

One important aspect of m-commerce is standardisation. Due to the heterogeneity of technologies for mobile devices, and the need for transmission and payment over the air interface, it is essential to find common approaches, both at a national and an international level. The following list summarises standardisation bodies and forums dealing with issues relating to m-commerce:

- PayCircle® (www.paycircle.org) – is a vendor-independent non-profit organisation. Its main focus is to accelerate the use of payment technology and develop or adopt open payment APIs (uniform Application Programming Interfaces) based on XML, SOAP, Java and other Internet languages.

- MoSign (www.mosign.de) – banks, technology partners and end-device manufacturers have joined forces in the MoSign (Mobile Signature) project to create a platform for secure, legally binding mobile transactions based on existing standards.
- Mobile Payment Forum (www.mobilepaymentforum.org) – is a global, cross-industry organisation dedicated to developing a framework for standardised, secure, and authenticated mobile commerce using payment card accounts.
- mSign (www.msign.org) – the Mobile Electronic Signature Consortium is an association of companies and organisations from the mobile phone and Internet sectors. The objective is to establish and develop a secure cross-application infrastructure for the deployment of mobile digital signatures.
- mwif (www.mwif.org) - the Mobile Wireless Internet Forum (MWIF) is an international non-profit industry association. Its mission is to drive acceptance and adoption of a single open mobile wireless and internet architecture that is independent of the access technology.
- Radicchio (www.radicchio.org) - As a non-profit organisation, Radicchio brings together market leaders to establish a common foundation for secure m-commerce by reaching a consensus on important inter-operability issues.
- Encorus (www.encorus.com) – Encorus Technologies is focused on building a flexible and open infrastructure and efficient payment processing services to drive the acceptance and usage of mobile payments worldwide.

5 Conclusions

There will be no m-commerce without security of the underlying technologies. In this paper we discussed security issues relating to network and service technologies (see Section 3) and m-payment (see Section 4). Regarding m-payment, some systems are under development or already operational. One of the main future challenges will be to unify payment solutions, providing the highest possible level of security.

6 References

- [1] T. Weitzel, W. König: "Vom E- zum M-Payment" (in German), <http://much-magic.wiwi.uni-frankfurt.de/profs/mobile/infos.htm>.
- [2] ePayment Systems Observatory, Database on E-payment Systems, <http://epso.jrc.es/>.
- [3] R. Weber: "Chablis – Market Analysis of Digital Payment Systems", Technical Report TUM-I9819, TU Munich, 1998, <http://chablis.informatik.tu-muenchen.de/Mstudy/>.
- [4] Datamonitor, Mobile Payment Systems, Whitepaper, November 2000, <http://www.sun.co.uk/wireless/resources/pdf/datamonitor.pdf>.
- [5] GSM 02.09 version 7.0.1 Release 1998. Digital cellular telecommunication system (Phase 2+); Security Aspects
- [6] 3GPP TS 33.102 3.9.0 Release 1999, 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; 3G Security; Security Architecture.
- [7] GSM 03.48 version 8.3.0 release 1999. Digital cellular telecommunication system (Phase 2+); Security Mechanisms for the SIM application toolkit.
- [8] S. Kent, R. Atkinson. Security Architecture for the Internet Protocol. RFC 2401

- [9] T. Dierks, C. Allen. The TLS protocol, Version 1.0. RFC 2246
- [10] UMTS security. This volume.
- [11] PKI in mobile systems. This volume.
- [12] BBC News, Tough penalties for mobile phone theft, 3 May 2002, [http://news.bbc.co.uk/hi/english/uk/newsid_1966000/1966247.stm].
- [13] H. Knospe and S. Schwiderski-Grosche, Future Mobile Networks: Ad-hoc Access based on Online Payment with Smartcards, 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2002), September 2002, to appear.
- [14] H. Knospe and S. Schwiderski-Grosche, Online Payment for Access to Heterogeneous Mobile Networks, IST Mobile & Wireless Telecommunications Summit 2002, Thessaloniki, Greece, June 2002.
- [15] D. O'Mahoney, M. Pierce, and H. Tewari, Electronic Payment Systems for E-Commerce, 2nd Edition, Artech House Computer Security Series.
- [16] M. Walker, T. Wright, Security Aspects. In: F. Hillebrand, GSM and UMTS : The Creation of Global Mobile Communication, John Wiley and Sons Ltd.
- [17] N. Borisov, I. Goldberg and D. Wagner, Intercepting Mobile Communications: The insecurity of 802.11. In Proceedings of MOBICOM 2001, <http://citeseer.nj.nec.com/article/borisov01intercepting.html>
- [18] IEEE Standard 802.1X-2001. Port-Based Network Access Control.
- [19] L. Blunk, J. Vollbrecht, PPP Extensible Authentication Protocol (EAP). RFC 2284.
- [20] B. Löhlein, Bluetooth - Technology, Security and Weaknesses, 6th World Multiconference on Systemics, Cybernetics, and Informatics, Orlando, Florida, USA July 2002
- [21] WAP Forum, Wireless Transport Layer Security, Version 06-Apr-2001, <http://www1.wapforum.org/tech/documents/WAP-261-WTLS-20010406-a.pdf>
- [22] WAP Forum, Wireless Application Protocol, WAP 2.0, Technical White Paper, http://www.wapforum.org/what/WAPWhite_Paper1.pdf
- [23] 3GPP TS 03.78 version 7.7.0 Release 1998. Digital cellular telecommunications system (Phase 2+); Customized Applications for Mobile Network Enhanced Logic (CAMEL) Phase 2; Stage 2
- [24] 3GPP TS 23.078 version 4.4.0 Release 4. Digital cellular telecommunications system (Phase 2+); Customized Applications for Mobile Network Enhanced Logic (CAMEL) Phase 3 - Stage 2
- [25] ETSI ES 201 915-1 V1.1.1 (2001-12) Open Service Access; Application Programming Interface, Part 1: Overview. <http://www.parlay.org/specs/index.asp>
- [26] ETSI ES 201 915-3 V1.1.1 (2001-12) Open Service Access; Application Programming Interface, Part 3: Framework. <http://www.parlay.org/specs/index.asp>