Integration of voice and data in an m-commerce situation

Master Thesis
26 January, 2001

Fredrik Oljeqvist
d96-fol@nada.kth.se

Assignor	The Department of Teleinformatics at the Royal Institute of Technology
Examiner	Professor Björn Pehrson
Academic supervisor	Thomas Sjöland
Industrial supervisor	Jonas Persson
Benefactor	Across Wireless AB
Location	Årstaängsvägen 21 B, S-100 74 Stockholm

Masters thesis completed as a required part of the School of Computer Science and Technology, Telecommunication Systems program at the Royal Institute of Technology.

Abstract

The use of cellular telephones has literally exploded during the last few years. The introduction of GPRS and UMTS will provide high speed Internet connection accessible from cellular telephones. People carry their cellular telephones wherever the go and the demand for new services, such as mobile commerce, increases.

The two biggest problems with mobile commerce are presentation of information on the mobile device and security. This paper presents a solution to a secure service for mobile commerce where the customer uses his voice and a cellular telephone to purchase goods over the Internet. This solution solves the biggest problems with mobile commerce. The customer does not have to use the keypad on the cellular telephone to enter text and the purchase is secure due to the e-commerce gateway developed by Across Wireless AB.

The application consists of a demonstration online CD store called CD@Across. This application has been extended with a voice interface. The voice interface has been created using VXML and an audio browser developed by PipeBeach AB.

In addition to this an investigation of some GSM data services (USSD, SMS, GPRS and UMTS) has been done. The goal of the investigation was to analyze if the data services used today could be exchanged to some other data service in order to increase performance and usability of the application.

The final implementation functions satisfactory. There are still some things that have to be solved in order to make this application function well in all situations.

The investigation of the GSM data services showed that the data services used in the existing application are appropriate.

Table of content

1. Introduction

The telephone was originally developed for speech communication and it has been used for speech ever since Alexander Graham Bell invented it in 1875. We have all grown up with telephones and everybody, from a 2 years old kid to a very old person, can use a telephone for speech communication.

The telecommunications technologies have evolved since the birth of the telephone and during the last years it has been possible to access Internet using mobile telephones. This has led to the development of a large number of services for mobile users. Examples of these services are weather forecasts, entertainment, interactive multimedia services and various kinds of electronic commerce services.

The cellular telephone is a very good tool for electronic commerce since it contains a SIM card that is personalized and capable of carrying out security related computations. The SIM card can be used to authenticate a customer, it can be used to encrypt sensitive information and to sign messages. Many people believe that the development and standardization of PKI SIM cards will be the key to success for electronic commerce.

Even though the services developed for mobile users are tailored to be easy used with a cellular telephone many people find it difficult to browse the information using the small display on the telephone. Interactive services where the user has to type in text are troublesome to use because the keyboard on a cellular telephone is untidy to use for text input. Many people will simply not use the telephone to access these services because the technology is unfamiliar to them.

People are familiar with using the telephone for verbal communication. The research in automatic voice recognition has made it possible to create user interfaces that make it possible for the users to interact with the telephones and the services using their voice.

The combination of mobile Internet access and voice recognition makes it possible to create user-friendly services accessible for everyone through a cellular telephone. The services can be made secure by utilizing the security services provided by the SIM card.

Organization of the thesis

The goal of this master thesis is to implement a voice interface to an application called CD@Across. CD@Across is a demonstration e-commerce site developed by Across Wireless AB[1]. The master thesis also includes an investigation of new GSM data services (GPRS/UMTS) and their effect on the application.

This introduction will give the reader a brief background to the context in which the paper has been written. Chapter 2 is intended to give the reader the security background needed in the rest of the paper.

Chapter 3 is an introduction to mobile commerce. This chapter is meant to motivate the development of m-commerce services and to elucidate the security demands in mobile commerce. The facts and figures presented in this chapter should be read with some suspicion since the possibility to foresee the future is difficult.

Smart cards play an important role in secure mobile commerce. An introduction to smart cards is given in chapter 4. The GSM network is described in chapter 5. This chapter comprises information about the GSM data services, the security in GSM and the different technologies used to enhance the GSM network.

Chapter 6 gives a short description of UMTS. The implementation of UMTS still lies a few years in the future. The aim of this chapter is to give the reader a brief idea of what UMTS is and how it differs from GSM.

In chapter 7 I describe CD@Across, the demonstration application in the m-commerce system developed by Across Wireless AB. This is the application that will be enhanced by implementing a voice interface.

Chapter 8 contains an investigation of voice systems. Different systems are described with their benefits and drawbacks. In chapter 9 I explain how different GSM data services are used in the demonstration application and what alternative solutions exists. I also investigate what effects the introduction of GPRS and UMTS will have on the system.

[1] During my work Sonera Oy has bought Across Wireless AB and ID2Technology AB and formed a new company called Sonera – SmartTrust AB. The demonstration application, CD@Across, was developed by Across Wireless AB before the merge but should now be considered a Sonera SmartTrust AB product.

In chapter 10 I describe my implementation of the voice interface. Finally in chapter 11 – 13 I have drawn some conclusions about my work and the future.

2. Security background

Four classes of threats to communication over a network exist. These classes are interruption, interception, modification and fabrication [51]. The threats have to be handled in order to create a secure environment suitable for electronic commerce.

Interruption means that the flow of information is stopped. If the customer cannot get in touch with the content provider he will get annoyed and the service offered will not be used.

Interception means that the communication between customer and content provider is monitored by an unauthorized third party. This threat has to be countered, otherwise sensitive information such as credit-card numbers will be revealed.

Modification means that the information is modified or changed on the way between the communicating parties. An unauthorized part might for instance change the account-number of the content provider so that a transaction of money will benefit him instead of the content provider.

Finally fabrication can be used to forge messages of different kind. Fabrication has to be stopped in order to gain faith in a network.

The threats mentioned above can be countered by using different security services. The security services are created by using security mechanisms.

Security mechanisms

Figure 1 illustrates the idea of symmetric and asymmetric cryptography.

Symmetric Cryptography

In a symmetric cryptosystem the same key is used to encrypt and decrypt the message. This means that the content provider needs to have a unique key for each customer in order to exchange information in a secure way.

Symmetric cryptosystems have an advantage of being simple and fast. Two parties share some secret information that can be used to decrypt as well as to encrypt messages and as long as the key remains secret the system provides authentication.

Unfortunately symmetric systems have some major drawbacks. If the key is revealed the interceptor can immediately decrypt all encrypted information and he can produce fake messages stating he is the legitimate sender. Key distribution is a problem. The number of keys increases with the square of the number of users [2]. Finally the symmetric cryptosystems are fairly weak [51].

The most common symmetric cryptosystem is DES. If the DES algorithm is applied three times in a row a cryptosystem called triple DES is achieved.

Asymmetric cryptography

In an asymmetric cryptosystem both parties have a private key and a public key. Only their owners know the private keys, but the public keys are available to anyone. The sending party encrypts the message with the receiver's public key and the receiver decrypts with his own private key. This is possible due to a relation between the public- and the private keys. The keys constitute a key pair where the keys are each other's inverse. The relation is known but it is unfeasible to compute in a realistic time.

The advantages of an asymmetric cryptosystem are many. Only one key has to be kept secret. Since the other key is public the problem with key distribution does not exist. However,

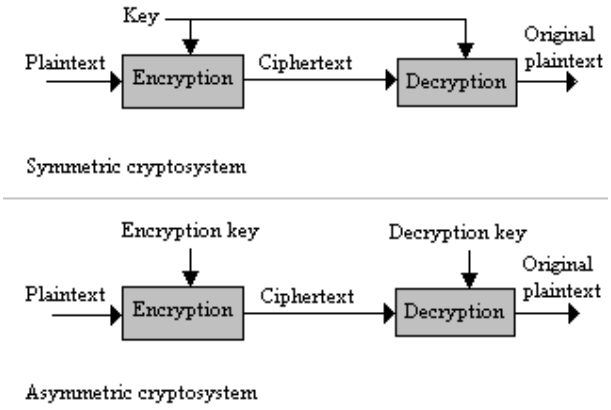

Figure 1 - Symmetric and asymmetric cryptosystems.

[2] Some symmetric cryptosystems exist where the number of keys increases linearly. An example of such a system is Kerberos [50].

the public key must be transferred to the receiver in such a way that he is absolutely sure that it is the correct public key. Public keys also provide a method for digital signing.

The drawback with asymmetric cryptosystems is that they are slower than symmetric since the use complex mathematical computing.

The most common asymmetric cryptosystem is RSA. The security in RSA is built on the difficulty to factorize large prime numbers.

Digital signing

Digital signing is a way for a party involved in a transaction over a network to put his signature on the transaction. The signature is very difficult to forge and it gives the recipient a possibility to verify that the sender is the person he claims to be.

Applying a one-way function to the message does the signing. The one-way function results in a unique digest of the message. The digest is encrypted with the private key of the person doing the signing.

When the recipient receives the message and the encrypted digest he decrypts the digest with the senders public key and retrieves the digest in clear-text. By applying the one-way function to the message and comparing the result with the received digest the sender is verified and at the same time the integrity of the message is checked.

Security services

There are three security services that are of great importance if secure transactions of information over an unsecured network are to be performed. These are Authentication, Confidentiality and Non-repudiation [2]. The services are achieved by using different security mechanisms. Authentication is achieved by using digital signing and Confidentiality by using cryptography. Non-repudiation is achieved by using digital signing and generally involves a trusted third party.

Authentication

Authentication services provide assurance of a person's identity. Authentication is used to verify that a person is the one he claims to be. In real life checking a person's identification card usually does authentication. Authentication is the most important of these three services since the others to some extent depend on it. A proper authentication fends off a potential forger's attempts to masquerade as someone he/she is not.

From an m-commerce point of view it has to be possible to verify that the customer is the person he/she claims to be.

Confidentiality

Confidentiality defends against information disclosure to unauthorized persons. It ensures that only the persons communicating can understand what the conversation is about.

In real life confidentiality is achieved by talking in private or by sending letters in envelopes. When the conversation is going on over Internet or through a radio it is more difficult to know who is listening.

In an m-commerce situation the customer does not want to reveal his/her credit-card number to anyone listening to the conversation.

Non-repudiation

The purpose of non-repudiation is to protect one legitimate user from another. Non-repudiation prevents one user from later denial of an agreement. The equivalence in real life would be that two parties sign a document and keep one copy each. The non-repudiation service itself does not prevent repudiating but it enables the other part to prove that there exists an agreement of something.

3. M-commerce

M-commerce is electronic commerce using a mobile handset such as a cellular telephone, or a wireless computer. The market for mobile commerce differs from traditional electronic commerce due to a different behavior and different expectations by the customers. The mobile telephone is used in a different way than a computer. When using a computer the customer is usually totally focused on the computer, which is quite complex to work with.

A telephone on the other hand is simple to use and is often used in a situation when the customer is occupied by another activity.

When a customer uses his mobile telephone for commerce he/she expects it to be as simple to buy things using it as it is to call from it. These expectations have made the user interfaces easy to work with.

The ease of use and the low price of the cellular telephones make the potential for mobile commerce huge [7].

Mobile commerce is estimated to grow rapidly in the near future. Some of the reasons is that mobile terminals offers flexibility, security and it is possible to create location-based services that are not found in the fixed Internet [12]. The developments of the third generation mobile communications that will provide high speed Internet access also speed up the development of services for m-commerce.

The US market for m-commerce is expected to grow 1000% in the next five years, from $90 million in 2000 to $1,2 billion in 2005 [14]. An estimation of the m-commerce market for Europe is $37,7 billions in 2004 [15].

M-commerce services includes stock trading, customers checking their bank accounts, other financial services and information services such as buying tickets to various events using the mobile terminal.

Europe has adopted m-commerce faster than the United States and is leading the evolution of m-commerce systems and services. One of the reasons for this is that Europe has one standardized system for mobile communication-GSM.

In order for the subscribers to be able to use the m-commerce services they have to be equipped with a cellular telephone capable of browsing the Internet. The market for these telephones is growing and in a near future almost all the cellular telephones sold will be equipped with a browser.

Figure 2 depicts an estimation of the number of global cellular subscribers and the numbers of cellular phones equipped with a browser. The figure also shows an estimation of the number of subscribers who will actively use the browser for m-commerce purposes [12].

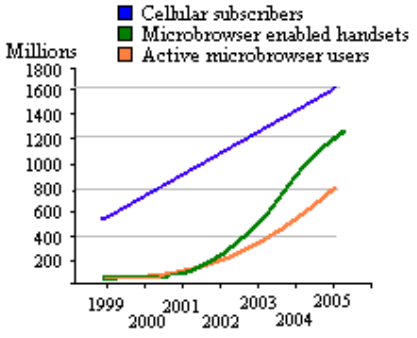

Figure 2 – This picture show the global estimation of cellular phones, the number of phones equipped with a micro-browser and numbers of users estimated to use the browser.

Security in m-commerce

If mobile commerce will grow as expected the security has to be very high. Security is one of the biggest problems on the fixed Internet. The lack of trust is one of the reasons to why e-commerce is not growing as fast as it was expected to grow on the Internet.

In order to make mobile commerce successful the customers has to have faith in the security and feel safe.

Mobile commerce using GSM cellular phones has the advantage of ordinary computers that they are personalized through the use of a SIM card. This makes the phone able to perform security-related functions such as digital signing, encryption and authentication.

Secure systems can be built in several ways. The common unit is a trusted third party that handles keys, payment etc. In the case of mobile commerce it is often the GSM operator that acts as a trusted third party. The operator verifies the subscriber and handles payment requests either through a connection to a payment provider or by charging the customer on the telephone bill.

Payment schemes

The most common payment schemes used in mobile- and electronic commerce is described below.

Mail order

Mail order is a simple and relatively safe way to do commerce over Internet. It works in the same way as traditionally mail order where the customer selects the products in a catalogue and the merchant sends the goods to the customer via mail. The customer pays for the goods when he/she fetches it at the post-office.

The only difference between the traditional mail order and the Internet based mail order is that in the latter case the customer browses the Internet instead of looking in a catalogue to find what he/she wants

This payment scheme has some drawbacks. It is troublesome for the customer to walk to the post office to fetch the goods. Another drawback is that this scheme is not suitable for merchants selling content. The content is delivered instantaneously over the Internet and thus this payment scheme does not work.

Credit cards

Some merchants use credit card numbers as the mean to pay for goods and content. The risks with giving away a credit card number has been debated for a long time. If the credit card number is sent in clear text there is a risk that a malicious person gets hold of it and uses it for his own benefit.

The merchant is also unprotected since he/she has no way to verify that the sender of the credit card number is the owner of the account. The merchant might not receive payment for what he/she delivers.

People having a positive view of using credit cards in electronic commerce states that the risk of loosing the credit card is as great when using the credit card in an ordinary shop or restaurant as it is using it on the Internet.

However, it can be concluded that using credit card numbers in electronic commerce is not a secure payment scheme unless it is used in conjunction with some other protocol that makes it secure such as Secure Electronic Transaction (SET) developed by Visa and MasterCard.

Merchant accounts

A merchant account is an account that the customer has at the merchant. The customers install some extra software in his computer that secures the link to the merchant. The link ensures that the information is not disclosed by using encryption.

Merchant accounts demand that the customer and the merchant set up some kind of agreement in prior to the actual purchase. It is also inconvenient for the customers to install new software and/or hardware for every merchant he wants to do business with.

Electronic wallets

An electronic wallet is a piece of online equipment used by the customer to protect his personal information and handles the user’s purchases over the Internet. The installation of software is done only once and then the customer can purchase goods at will as long as the merchants are using the same wallet system.

The systems using electronic wallets will not be successful unless several of the major merchants use the same wallet system.

The m-commerce demonstration system developed at Across Wireless AB uses a mobile GSM cellular telephone as an electronic wallet.

4. Smart Cards

Introduction

Smart cards play a significant role in the GSM system as well as in the m-commerce systems available today.

Since the first smart cards were developed in the early seventies the market for smart cards have literally exploded. It is estimated that more than 2,8 billions smart cards are used around the world this year [1]. A smart card is a piece of plastic the size of a credit card. Embedded in the card is a small integrated circuit chip that makes the card able to store and process data. A smart card can store more than hundred times more information than an ordinary magnetic stripe card can store.

The smart cards typically fall into two categories, memory-cards and processor-cards. In the memory-card the integrated circuit is only used to store information while the processor-card contains an operating system and a processor that can process the information in the memory.

A processor-card can perform various tasks such as advanced computing, local data processing and encryption. One of the reasons for the increasing interest in smart cards is the processor-card's ability to carry out security related functions. This is the kind of cards that offer enough flexibility to be used in an Internet-based economy.

The communication with the smart card is done through a terminal or a card reader. Some smart cards communicate through the air using an antenna. In order to ensure that equipment produced by different manufacturers of smart cards, applications and accessories can work together the hardware and the electrical interface has been standardized by ISO/IEC [32].

Smart cards are used for many different purposes such as banking, ID cards, GSM networks, telephones, buss tickets etc.

Smart cards can solve many of the problems that have been threats to mobile and electronic commerce. The reason is the combination of personalization (with cryptographic keys stored in the smart card) and processing power that the smart card offers. It is possible to uniquely identify a person and to verify that he is the person he claims to be, by using digital signing and authentication schemes. Cryptography added to this makes it feasible to exchange personal information and makes transactions over an unsecured network secure.

One of the reasons that mobile commerce is such a hot topic in Europe is that 300 million people are estimated to use the GSM network by the end of 2001 [5]. Since the GSM cell-phone contains a processor-card capable of security related tasks and since the network connections are rapidly getting faster there is a big potential for mobile commerce to evolve.

PKI-cards

PKI (Public Key Infrastructure) cards are special types of processor-cards. The difference between PKI-cards and ordinary processor-cards is that a PKI-card is equipped with a processor that can handle asymmetric encryption. This enables the card to carry out secure authentication and non-repudiation by using asymmetric cryptography.

SIM-Cards

A SIM (Subscriber Identity Module) card is a smart card with an embedded processor – a processor card – that is used in the GSM system. The SIM card is used to identify a subscriber, to store permanent and temporary network information and to store service-related data and other information needed by the GSM network in order to function.

The SIM card is described more thoroughly in the section that covers the GSM system.

5. The GSM system

Introduction

Mobile telephone systems have gone through a rapid development and are continuing to evolve. It is common to talk about three generations of mobile communication systems. The first generation was developed in the mid-eighties and comprises the analogue systems –NMT-450 (Nordic Mobile Telephony) and NMT-900 among others. It was soon obvious that these systems could not fulfill the fast growing demands of services and number of subscribers. Furthermore the different analogue systems could not cooperate.

In the beginning of 1990 the second-generation mobile communication systems were developed. This generation were digital and had greater capacity of providing services than the systems of the first generation. GSM 900 MHz, GSM 1800 MHz and GSM 1900 MHz (Global System for Mobile communication) belongs to this generation of systems. GSM was developed in Europe but spread to many different countries.

GSM is a digital network that uses Time Division Multiple Access (TDMA) and that provides speech as well as a Short Message Service (SMS) that makes it possible to send short text messages to and from GSM cellular telephones.

The increasing number of subscribers and the possibilities to offer new services has set demands of higher data speed in the communication. The subscribers also demand a global coverage of the systems. These are some reasons for the development of the third generation of mobile communications systems. International Telecommunication Union (ITU) is currently working on the standardization of this generation of systems.

The third generation of systems comprises Universal Mobile Telecommunication System (UMTS). The goal of UMTS is that the system shall be able to handle all kinds of traffic from traditional speech to video and multimedia transmissions. Roaming between different systems of the third generation is one of the most important issues.

In the beginning of the development of UMTS it was thought to be a completely new system. The large market for GSM and the huge investments in equipment done by the GSM operators has made the European Telecommunications Standards Institute (ETSI) decide to incorporate GSM into the UMTS standard.

While the third generation of mobile communications systems is being developed the second generation systems are adjusted to handle the increasing demands. These modified systems are called “generation 2.5” systems. One example of a generation 2.5 system is General Packet Radio Service (GPRS).

SIM cards

A SIM (Subscriber Identity Module) card is a smart card with an embedded processor – a processor card – that is used in the GSM system to identify and to keep track of a subscriber among other things. The SIM is the only thing that personalizes the cellular phone. This makes it possible to insert the SIM into any other cellular phone and the SIM owner will be charged for the calls.

In the GSM system two different sizes of SIM cards are standardized. These are ID-1 SIM and Plug-in SIM. While ID-1 cards have the size of a credit card the Plug-in SIM is 25x15 mm [33].

The SIM card contains different kind of information. Some of the information is stored on the card when the card is personalized and other information is stored later. Three types of subscriber related information is stored on the SIM. The first type is information that cannot be changed or read by the subscriber. Examples of this kind of data are International Mobile Subscriber Identity (IMSI) and subscriber authentication key (Ki). The IMSI is an internal subscriber identity within the GSM network and the Ki is the subscribers’ private key, used for security related operations.

The second kind of information stored on the SIM is temporary network data such as Temporary Mobile Subscriber Identity (TMSI), Location Area Identifier (LAI), cipher key (Kc) etc. This temporary information changes over time. TMSI is an internal number in the GSM network that sometimes is used instead of the IMSI. The reason for this is to prevent somebody from tracing the user by monitoring a specific IMSI. LAI identifies the geographical area that the subscriber is located in at the moment. Kc is the key that is used to cipher information to and from the cellular phone.

The third kind of information stored on the SIM is service-related data such as language preferences, advice of charge, phonebook, short messages etc.

SIM cards can be equipped with the SIM Application toolkit in order to enhance the functionality of the handset.

SIM Application Toolkit

SIM Application Toolkit is a European Telecommunications Standards Institute (ETSI) standard [34, 35] for value-added services using GSM cellular phones to do the transactions. SIM Application Toolkit makes it possible for the SIM card to interact with the handset. The interaction may be trigged by different events, such as an incoming Short Message (SM) or the subscriber pressing a button on the handset. SIM Application Toolkit gives the SIM card the possibility to control the menu system in the cellular phone and the SIM can trigger events independent of the telephone and the network. The SIM Application Toolkit is important because it makes it possible to place applications on the SIM card. These applications can perform complex tasks such as security-related computing and interaction with the user through extended menu-systems.

In order to use the value-added functionality that SIM Application Toolkit offers the subscriber needs a SIM Application Toolkit compliant cellular phone and SIM card.

Some of the advantages with SIM Application Toolkit are that it is a part of the GSM standard and most of the cellular telephones that are sold today are SIM Application Toolkit compliant [6]. Furthermore SIM Application Toolkit is used in many commercial networks for purposes such as banking-, electronic mail- and information-services. The personal information stored on the SIM card makes it possible to build secure interactive services.

The SIM Application Toolkit mechanisms that are most relevant for my work are proactive SIM and data download.

Proactive SIM

Proactive SIM gives the SIM card the ability to initiate events in the cellular phone. The SIM card can display text to the subscriber, it can send short messages (SM), and it can set up telephone calls to numbers that are stored in the SIM card.

The normal GSM communication is not disturbed by these proactive activities.

The communication between the SIM card and the handset is carried out using the T=0 protocol defined in ISO 7816-3. This protocol states that all communication is initiated by the handset. Proactive SIM use the T=0 protocol but it is enhanced with additional status-reports.

Very schematically proactive SIM works like this; the handset delivers data or polls the SIM card, the SIM card returns a status-report, which contains a special code if the SIM card wants the handset to do something. If the handset is not busy the command is executed and a response is returned to the SIM. The SIM can then, in turn, issue a new command and the communication between the handset and the SIM card can continue.

Data download

Data download is a mechanism that makes it possible for a cellular telephone to download information onto the SIM card without notifying the user. The network operators can use this feature in order to modify the information on the SIM card. The details of data download and proactive SIM can be found in [44].

System Overview

The infrastructure of the GSM system is illustrated in Figure 3. The system consists of several entities that cooperate in order to perform authentication, routing and switching.

Geographically the GSM network is built on a cell structure. A

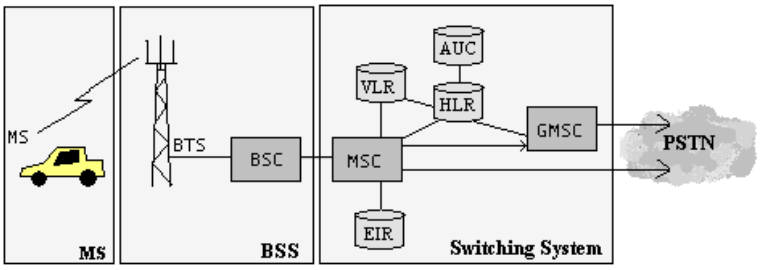

Figure 3 – Overview of the GSM system

subscriber is located in one cell, which is the geographical area that is covered by one transmitter. The size of a cell is determined by the power of the transmitter. The benefit of using the cell structure is the concept of low power transmitters, which enables reuse of frequencies.

Several cells constitute a Location Area (LA). The location area is the geographical area covered by all the transmitters controlled by one Base Station Controller (BSC.)

MS (Mobile Station)

The mobile station consists of two different entities: the SIM card and the mobile equipment. The SIM card contains information about the subscriber such as the IMSI and the cryptographic key Ki.

The mobile equipment is the hardware used to access the radio resources in the GSM network. In order to access the services that the GSM network provides both the mobile equipment and the SIM card is needed. One exception is emergency calls[3] that can be placed using only the mobile equipment.

Base Station System (BSS)

A GSM network normally consists of several Base Station Systems that connects MS in different geographical regions to the GSM Switching System. The BSS is in charge of transmitting and receiving calls and data between the mobile stations and the Switching System.

A BSS consists of one or several Base Transceiver Stations (BTS) and one Base Station Controller (BSC).

BTS

The task for the Base Transceiver Station is to handle the radio communication between the GSM network and the mobile stations subscribing to the system. The BTS consists of radio equipment and antennas and serve one cell. The BTS is usually placed in the center of a cell.

BSC

The Base Controller Station controls all the underlying BTS. The BTS handles the actual communication but all the actions taken, such as transmitting power, when and what to transmit etc. are controlled by the BSC.

Switching System

The main role for the Switching system is to manage communication between mobile users and other users. The other users can be mobile users, users on the Public Switched Telephony Network (PSTN) etc.

The Switching System contains databases with subscriber information needed for handling routing, authentication and mobility of the subscribers.

The units in the Switching System are described below.

MSC (Mobile service Switching Center)

The MSC handles the switching within the network and sets up, supervises and releases calls. It can connect telephone calls between users in a particular GSM network and it can connect calls between a subscriber in the GSM network and another network.

GMSC (Gateway Mobile services Switching Center)

GMSC is a gateway between the GSM network and some other network (such as PSTN or another GSM network). It handles routing between different kinds of networks. A telephone call from a user in the PSTN to an MS in the GSM network is routed to the GSMC that handles this particular MS. Usually each GSM operator has one or more GMSC. The National Destination Code (NDS) in the MSISDN is used for routing to the right GMSC.

HLR (Home Location Register)

Normally each GSM operator has one HLR that contains information of all the users subscribing to the network. The HLR contains information about the approximate location of all subscribers in the GSM network and what services the subscribers have access to. Among the data in the HLR, there is information about which MSC that serves the subscriber for the moment.

[3] 112 is the single emergency telephone number for the European Union. Thus, anyone travelling within the European Union can call 112 in case of emergency and get through to the emergency services in the country he/she is currently in.

VLR (Visitor Location Register)
VLR is a regional database attached to, or co-located with an MSC. It contains information about all the subscribers located in the Location Areas served by the MSC (MSC Service Area). When a subscriber enters an area served by an MSC the attached VLR will ask HLR for information about the subscriber. The VLR now has all the information needed for serving the subscriber without asking HLR each time communication is established.

AUC (Authentication Center)
The AUC is used for security purposes. It contains information and parameters used in the authentication- and encryption process. These processes are described in more detail later in this section.

EIR (Equipment Identity Register)
The EIR is a database that is used for keeping track of the mobile equipment in the network. It consists of a list of valid hardware. The mobile equipment is identified by its International Mobile Equipment Identity (IMEI). EIR can block telephone calls to equipment that is stolen or unauthorized.

Important numbers in the GSM system

There exists some important numbers in the GSM network. The numbers are used for different purposes and they are explained in detail in [52]. The most important numbers are the Mobile Station ISDN (MSISDN), International Mobile Subscriber Identity (IMSI), Temporary Mobile Subscriber Identity (TMSI) and International Mobile Equipment Identity (IMEI).

MSISDN is the telephone number to a particular subscriber. The MSISDN is divided into three fields, which are used for routing the call to the destination. IMSI is an internal number used by the GSM network to identify a subscriber. TMSI is a temporary number used instead of the IMSI in order to provide a higher level of security. IMEI finally is a number that identifies the mobile equipment (i.e. the cellular telephones) in the network.

The air interface

The radio channel between the cellular telephone and the BTS consists of two carriers, one for information from- and one for information to the telephone. The information is multiplexed onto the carriers using Time Division Multiple Access (TDMA). The air interface is described in detail in [52]. One TDMA frame consists of eight timeslots and one phone-call uses one timeslot.

Logical channels are sent on the carriers. There are many different types of logical channels but they can be divided into two main types. These are traffic channels and control channels. The traffic channels are used to transport voice and data information. The control channels are used to transport management information.

The logical channels are mapped on the physical channels (i.e. a time slot in the TDMA structure) in a certain way so that the traffic- and control channels are repeated in a cyclic way.

SMS – Short Message Service

The short message service provides the user of a cellular telephone with the ability to send short messages to and from his cellular telephone. A short message (SM) can carry 140 octets of information. This makes it possible to send 160 characters if the standard GSM alphabet is used. If another character coding scheme, such as USD2 defined in ISO/IEC10646, is used fewer characters will fit into the 140 octets.

In order to provide the subscribers with SMS the GSM network operator has to complement the network with a few units. The GSM network prepared for SMS is illustrated in Figure 4.

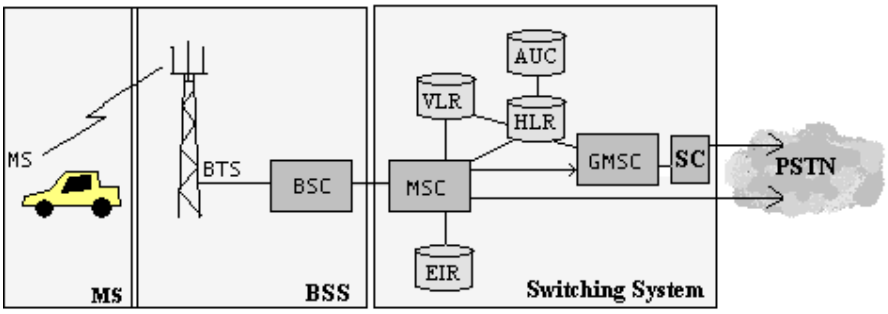

Figure 4 – The GSM network prepared for SMS.

The new entity in this network compared to an ordinary GSM network is the existence of a Service Center (SC). What is not shown is the modification that has to be done in the GMSC, the MSC and the MS. These entities have to be adjusted in order to be able to handle SM.

The SC acts as a relay station for short messages. It receives short messages, stores them and forwards them. In addition to this it reports results to the originator of the SM. The detailed functionality is of the SC is not covered by the GSM standards but is up to the operator of the SC to define. Some basic functionality is mandatory in order to supply SM. The SC should be able to submit a short massage to an MS and then wait for an acknowledgment. It should also be able to receive a SM from an MS.

If an SM is sent from the cellular phone it is called a mobile originated SM (SM MO) and if it is sent to a mobile phone it is called mobile terminated (SM MT). There is also a type of short messages used by the GSM network operator used for sending SM to all the subscribers within a certain area. This type is called cell broadcast short messages (SM CB). SM MO and SM MT are illustrated in Figure 5.

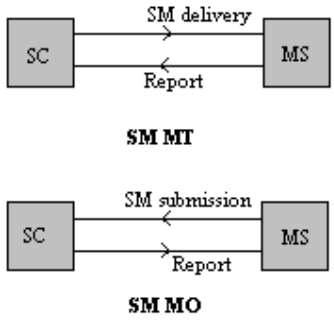

Figure 5 - SM MT and SM MO

A mobile telephone that is capable of sending and receiving short messages should be able to send and receive the short messages independently of ongoing calls. When the message has been received the telephone always sends a status report to SC.

The details of the technical realization of SMS can be found in [42].

USSD - Unstructured Supplementary Services Data

Unstructured Supplementary Services Data (USSD) is a GSM data service [46, 47] that allows interactive data communication between subscribers and applications across the GSM network. The service is optimal for communication, which structure reminds of a conversation.

The subscriber requests some information by entering a specific sequence of characters on the cellular telephone. Such a sequence might be *211#, that is used to check the balance in the Swedish telephone operator Comviq's GSM prepaid service.

The request is sent over the GSM network to a USSD server and routed to the application that handles the request. The application sends a response to the subscriber within the same signaling session. The interaction can consist of an unlimited numbers of messages sent between the subscriber and the USSD server.

Since the request and reply are sent in the same signaling session the delay between request and reply only consists of the processing time for the request and the transmission time for the messages.

USSD has some similarities with SMS in that it uses the GSM network to transmit text messages of a limited length (USSD text messages can be up to 182 characters in length). Both SIM Application Toolkit and the Wireless Application Protocol support USSD.

The major difference is that SMS is a store-and-forward service and USSD is a session-oriented service. USSD does not require a special menu choice but the messages can be entered directly (just like entering an ordinary phone number) and then sent. The drawback is that the messages are quite difficult to remember since they consist of sequences of characters that are not logical.

Security in GSM

The security aspects of GSM are defined in the GSM standards [33, 39, 40].

Introduction

Communication over the GSM network has to be protected for several reasons. One of the goals is to prevent deceivers from committing crimes such as wiretapping, identifying credentials for personal benefit and localizing individual subscribers. In order for the GSM system to be accepted by the mass market the subscribers have to have faith in the protection the system gives.

The technical solution that GSM uses, with frequency hopping, speech coding, digital modulation and the use of the TDMA architecture makes it quite difficult for an amateur to monitor the traffic on the network. These technologies do not hide the information, they only make the equipment for monitoring the traffic more expensive.

In order to get a sufficient security in the network security-services such as authentication and confidentiality have been built into the system. The design of the security mechanisms that implements the security-services is such that no sensitive information is transmitted over the network.

The security mechanisms are implemented in different elements of the GSM system.

The SIM card contains the individual subscriber authentication key (Ki), an algorithm for generating encryption keys (A8), an algorithm for authentication (A3), personal identification number (PIN) and the IMSI.

The ME contains an algorithm (A5) for ciphering data and voice transmissions and finally the GSM network contains the A3, A5 and A8 algorithms together with IMSI, TIMSI, LAI and Ki for each subscriber.

Authentication

When a cellular telephone connects to the GSM network a challenge-response mechanism authenticates the cellular phone to the network. The AUC in the GSM network sends a random number (RAND) to the ME. The ME computes a response (SRES) using the encryption algorithm A3 with the authentication key of the subscriber (Ki) and sends it to AUC.

When the AUC receives the response the identity of the subscriber is verified by repeating the computations. If the MS is successfully identified the subscriber may continue; otherwise the connection is terminated.

Confidentiality of data and voice

All data between the ME and the BSC is encrypted using the key Kc with the A5 algorithm. The SIM card computes Kc when the GSM network has authenticated the cellular phone. The computation is done by applying A8 key generation algorithm to the same RAND received by ME in the authentication process. The key for this computing is Ki.

The key Kc is changed regularly dependant of the network design and security considerations.

Confidentiality of subscriber identity

When the subscriber has been authenticated and when Kc has been computed a Temporary Mobile Subscriber Identity (TMSI) is sent to the ME. The TMSI is used to identify a subscriber in a location area. The TMSI protects the subscriber's identity because the relation between IMSI and TMSI is only known by the GSM network. Thus it is not possible for a deceiver to identify and trace a subscriber by monitoring the TMSI.

Enhancements of GSM

The development of the GSM system is divided into three separate phases. The three phases implement different features. Phase 1 contains the basic services and was completed in 1991, phase 2 was completed 1996 and contains new functions and extensions of the functions in phase 1.

Phase 2+ does not have a mandatory times schedule. This phase contains features that can be implemented successively when the technology is ready. Some of the issues covered by Phase 2+ are High Speed Circuit Switched Data (HSCSD), General Packet Radio Service (GPRS) and Enhanced Data rates for GSM Evolution (EDGE).

HSCSD and GPRS are based on new coding schemes of the data in the TDMA frames and EDGE is based on a new modulation type. The cellular telephones have to be modified in order to support these new coding techniques, modulation types and multiple slot allocation.

Universal Mobile Telecommunications System (UMTS) is a third generation system that will support both packet- and circuit switched transmission of data at rates between 144 kbps and 512 kbps, locally

even up to 2 mbps. UMTS is a new system that does not use the TDMA frame structure that is used in GSM. Therefore UMTS isn't seen as an extension of GSM but as a new system, even though it can share some parts of the existing GSM switching system (HLS, VLR etc.).

HSCSD
High Speed Circuit Switched Data makes it possible to allocate more than one timeslot in the TDMA structure for each user. By using up to four consecutive time slots and by using a new data-coding scheme, which increases the bit rate of the data in the TDMA frames, it is possible to reach a speed of 57,6 kbps. HSCSD supports the same services as today's existing GSM network but at a higher transmission rate.

GPRS
GPRS is a packet switched wireless communication service that provides connection to Internet at rates of up to 115,2 kbps. GPRS is based on the existing GSM network and will complement the circuit switched services and SMS.

The user is apparently always connected to the network and he/she is charged for the amount of data sent or received from the cellular telephone.

GPRS is a step towards EDGE and UMTS.

The high transmission rate is achieved by using up to eight consecutive timeslots in the TDMA structure combined with a new coding scheme [4]. Tunneling between mobile terminal and the Internet can give the terminal the same status as an IP host on a LAN.

EDGE
Edge gives GSM operators the ability to offer wireless multimedia, IP based services and applications at a rate of up to 384 kbps. The high rate is achieved by using new modulation (8-PSK) of the data in the air interface. By using up to eight consecutive timeslots 384 kbps is reached [4]. The change from GMSK to 8-PSK is the central change with EDGE, which prepares GSM for UMTS.

EDGE uses the same radio band, the TDMA frame structure, the same logical channels and the same carrier bandwidth as today's GSM networks.

EDGE was developed for telephone operators that will not get UMTS licenses. EDGE gives these operators a possibility to offer data services at speeds close to those in UMTS networks [17].

By using EDGE operators can offer wireless multimedia, entertainment and wireless videoconferencing [15].

GPRS – General Packet Radio Service

GPRS offers a new set of services carried by GSM. The services are described in [45]. GPRS makes it possible for the user of a mobile telephone to send and receive data using packet switched end-to-end communication at a rate of 115,2 kbps. The use of traditional GSM services is not affected by the GPRS communication. The subscribers are charged for the amount of data transferred and not for the time they are utilizing the network.

The network utilization is cost effective when transferring short frequent pieces of information and when transferring not so frequent large pieces of data [45].

The main advantages with GPRS are that it supports the Internet Protocol (IP), it uses GSM, which has a large coverage around the globe and the user seems to always be connected to the Internet.

The air interface
The radio resources are shared dynamically between GPRS and circuit switched GSM (GSM-CS). The air interface in GPRS uses the TDMA structure just like GSM-CS does. One TDMA frame consists of eight timeslots on which information is mapped. The mapping differs from the one used in GSM-CS. Some time slots are dedicated for GPRS Packet Data Channels (PDHC) and others are dedicated for GSM-CS, further some timeslots are shared between GPRS and GSM-CS and are allocated when needed. The network operator decides the number of slots reserved for different purposes.

[4] The actual implementations of GPRS and EDGE will probably not use eight consecutive timeslots. One reason is that the network operators probably will reserve some timeslots for voice calls. One consequence of this is that the practical transmission rates in GPRS and EDGE will be lower than the theoretical values.

The network architecture

The architecture of a GSM network adjusted to be able to handle GPRS is illustrated in Figure 6. The GSM network has been complemented with three units: the Serving GPRS Support Node (SGSN), the Gateway GSN and a Domain Name Server (DNS).

In addition to these three entities the Base Station System (BSC and BTS) has to be complemented with a Packet Control Unit (PCU) that handles segmentation and reassembly of data packets. It also takes care of the scheduling of signaling and data transmission over the air interface. The BSS is responsible for the allocation of radio resources. One of the responsibilities is to divide the resources between GSM-CS and GPRS.

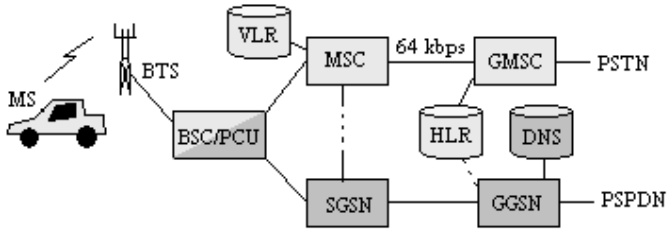

Figure 6 - The architecture of a GPRS adjuster GSM network.

BSC/PCU is connected directly to SGSN. In a traditional GSM call the information is sent from the BSC/PCU to the MSC and in a GPRS connection the packets are sent from the BSC/PCU to the SGSN.

A new type of area is introduced in GPRS. The area is called a routing area (RA) and span over a subset of the cells in a GSM-CS location area.

The SGSN:s are connected to the GGSN over the GPRS backbone. The GPRS backbone is a private IP network that is transparent to the GPRS user. The task for the network is to convey traffic between subscribers and between entities in the network.

The SGSN, which is on the same hierarchical level as the MSC in the GSM-CS network, keeps track of the subscribers as they roam around in the network. It also handles security-related functions such as authentication. In order to perform its tasks SGSN communicates with HLR and VLR. HLR and VLR contain some new – GPRS related – information. One example of this extra information is which SGSN that is currently serving a particular handset.

GGSN offers communication with external packet switched networks. GGSN offers IP routing and can be connected to IP routers. In order to be able to perform routing GGSN is connected to a DNS.

Data sent from a mobile station is tunneled through the GPRS backbone to GGSN, which forwards the data to the destination on the Internet. When a packet from an external packet switched network reaches GGSN it is routed and tunneled to the SGSN that currently serves the receiver of the packet.

Security

The security in GPRS is similar to the security in GSM. The algorithms, cryptographic keys and policies are reused in GPRS. In GPRS the encryption and authentication is handled by the SGSN while firewalls, policies and monitoring of messages are handled by the GGSN.

The mobile equipment

The mobile station differs from a mobile station used in the traditional GSM network. In order to be able to use GPRS the mobile telephone has to be able to handle multiple slot allocation and the new coding scheme for the data in the TDMA frames. A GPRS telephone can handle GSM-CS connections.

The SIM card can be a GSM-CS SIM or a GPRS SIM; the difference is that the GPRS SIM has two additional files. These files are stored in the telephone if a GPRS handset is equipped with a GSM-CS SIM.

WAP-Wireless Application Protocol

WAP is a protocol developed to meet the future demands for mobile equipment communicating with the Internet and to supply wireless Internet to the big mass of users. The standards of the protocol are developed by Wap Forum [7], an industry association consisting of several hundreds companies involved in mobile equipment, software and services. The standard makes it possible for the manufacturers to create compatible hardware, software and services by following the recommendations specified in the standards.

The protocol is optimized for handheld mobile equipment and their special characteristics, such as a small display, limited power supply, limited processing power and slow communication with the network.

In order to create a functional, flexible and easy-to-use protocol existing technologies of today's Internet are used where applicable. If the technologies are not applicable existing technologies have been modified to fit the demands that WAP put on them.

By using existing technologies existing competence's can be used in production of new services, hardware and software.

Examples of technologies reused by WAP are the use of URLs to locate files and communication using the Http protocol between the WAP-gateway and the Internet.

The WAP standard does not specify how the information is transferred through the air but states that all existing bearers used today shall be able to carry WAP. This means that WAP can be carried by for example SMS, CSD and GPRS.

Figure 7 illustrates how a WAP client communicates with the Internet. The request from the client is transferred using WAP and carried by an arbitrary bearer. The WAP gateway translates the WAP request to an Http request and sends it to the webserver. The WML page returned from the webserver is converted into WML byte-code. The byte-code is a compiled data stream, which is returned to the client and handled by the WAP browser. WML is a markup language, like HTML, optimized for low bandwidth

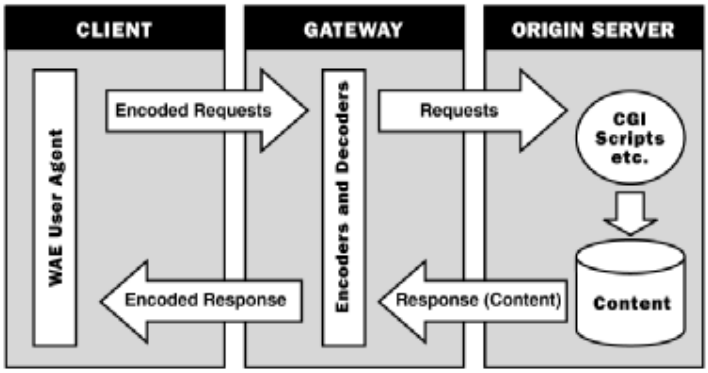

Figure 7 - The functionality of WAP [7].

6. UMTS

Introduction

UMTS is a third generation mobile system that will be the next standard for mobile services across Europe. A more detailed description of UMTS can be found in [36, 37 and 38].

UMTS will be a member of a new family of mobile telecommunications systems being developed by the International Telecommunications Union (ITU) for deployment across the world. While using different radio frequencies in different countries, every system will offer the same set of features to the users. This will allow handsets to be developed that can be carried from country to country as the user travels around the world - something already taken for granted across Europe.

Services

UMTS will support voice and data services from 384kbps up to 2 mbps and it will provide seamless telecommunications support over a wide range of heterogeneous networks. It will support speech, data, and multimedia and offers a ubiquitous service that is far more valuable to the user than current systems. UMTS will give users a consistent quality of service for voice, data, graphical, and video-based information independent of their location and access network (i.e., cellular, satellite, or fixed).

UMTS will be able to share some parts of the GSM switching system and it will support GPRS, SMS etc. Furthermore UMTS will support roaming between other third generation mobile networks; hence it will provide global personal communications to anyone anywhere. As the third generation mobile networks are being deployed they will supersede the GSM network in Europe.

Data services can be packet or circuit switched. There is some expectation that data services will be less expensive than voice, since data can be sent asynchronously and does not require a dedicated channel.

The radio interface

UMTS is based on the Wide band Code Division Multiple Access (Wide band CDMA) technology. This technology is different than the TDMA technology used in GSM. When using Wide Band CDMA all subscribers use the same frequency band at the same time. The data from each user is encoded with a pseudo orthogonal code. This code makes it possible to filter out a particular user’s data at the receiver end.

Terminals

UMTS terminals will, like other mobile terminals, be varied in their capabilities, size and sophistication. The terminals will have combined computer like features and mobile communication features. Depending on the size, processor power, memory etc. they will be able to access different type of services. ETSI have specified a set of features that they consider mandatory for all UMTS terminals. These features include the ability to identify and authenticate the user, the ability to set up and receive a connection etc.

During the deployment phase dual mode (GSM/UMTS) terminals will allow the users to access services via both GSM and UMTS radio access networks. These dual mode terminals will probably be able to provide seamless hand-over of services between the base stations.

7. Across Wireless demonstration m-commerce system

Background

Across Wireless AB has developed a platform for demonstration secure mobile commerce. The system was developed for an exhibition where Across Wireless AB wanted to show that it is possible to build functional m-commerce systems with existing technologies. Jonas Törnroth [54] developed the central part of the system – the e-commerce gateway - as a master thesis in March 2000.

This section describes the system and the different entities that the system consists of.

System overview

The m-commerce system that Across Wireless AB has developed consists of several entities. The central unit is the e-commerce gateway, the Wireless Internet Gateway (WIG) and the OTA Service Center [5].

The different units reside in different places in the GSM network and on the Internet. The Mobile station can be located anywhere in the GSM network. It lets the customer be mobile and he can always reach the merchant and the goods that the merchant offers.

The Wireless Internet gateway, the OTA Service Center and the e-commerce gateway are placed somewhere in the GSM operator's domain.

The merchant, the WEB browser and the payment provider are typically located in the Internet. Finally the bank resides in some kind of private network.

The different units in the system are illustrated in Figure 8.

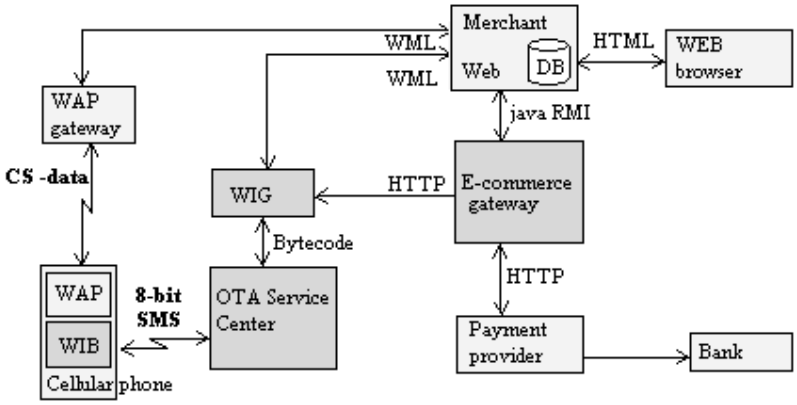

Figure 8 - An overview of Across Wireless AB's m-commerce system.

[5] The OTA Service Center has been changed to a Delivery Platform 5 (DP5) during my work. DP5 is a newer version of OTA Service Center that is more capable. The replacement has no effect on my work.

End user

It is possible for a mobile subscriber to browse the Internet using the Wireless Application Protocol (WAP) in combination with a WAP browser on the cellular phone. If the subscriber's cellular telephone does not support WAP but is SIM Application Toolkit compliant and is equipped with the Wireless Internet Browser (WIB) he can connect to the merchant using SMS instead.

The WIB is a SIM Application Toolkit application that resides on the SIM card. It makes it possible to browse the WEB using SMS. The interface between the WIB and the OTA Service Center is GSM 03.48 on top of 8-bit SMS. Each GSM 03.48 packet is transferred to the SIM and the instructions in the packet is handled by the WIB.

The WIB is developed by Across Wireless AB and has rapidly become a de facto standard.

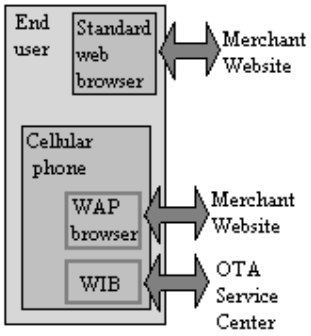

Figure 9 – Possible ways for the End user to reach the content provider.

If an operator of a GSM network would like to offer his subscribers Internet access using SMS the operator will distribute SIM cards equipped with the WIB.

It is also possible to reach the merchant's website using an ordinary web browser.

When the user or potential customer finds the preferred product he/she places an order with the merchant via their website. The possibilities for the end user to reach the content provider are illustrated in Figure 9.

GSM Operator – OTA Service Center

The Over The Air (OTA) Service Center conveys the traffic between the content provider and the end user. It consists of several different units providing different functionality.

The Wireless Internet Gateway [56] acts as a gateway between the Internet and the GSM network. The WIG makes it possible for the subscribers to access standard web applications using SMS. The content providers can supply services using standard tools and either HTML or WML, as defined by the Wap Forum [7].

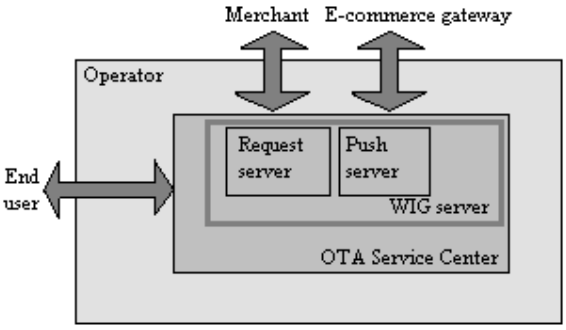

Figure 10 – The OTA Service Center.

The WIG contains a Request Server and a Push Server. The main purposes of these servers is to receive web pages from a content provider, convert them to byte code and send them to the WIB using SMS for the transport.

The Request Server waits for requests from the WIB. The requests are translated into an http request to the content provider. When the reply is returned to the WIG it is converted and sent back to the WIB.

The Push Server is similar to the Request Server. The difference is that it waits for a web page from a content provider destined for a certain receiver. The page received from the content provider is pushed to a particular end user. The OTA Service Center and its components are illustrated in Figure 10.

GSM operator - E-commerce gateway

The E-commerce gateway handles payment requests from the merchants. The logical view of the system is shown in Figure 11.

The E-commerce gateway exposes two interfaces towards the merchant. These interfaces make it possible for the merchant to utilize the services in the gateway. The interfaces are described in [56, 57 and 58].

The purpose of the notification service is to give the merchant means to send notifications to the customers.

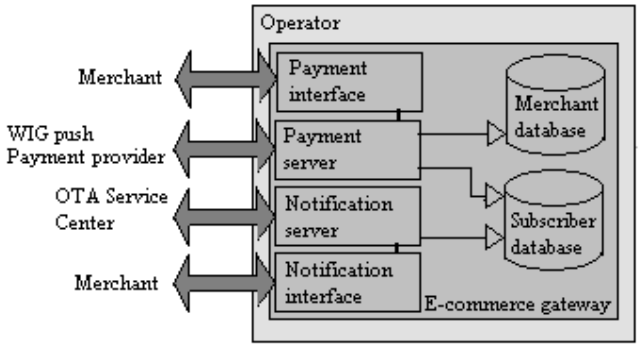

Figure 11 - Illustration of the E-commerce gateway.

The e-commerce gateway contains two databases, one for the merchants using the system and one for registered subscribers. The merchant database contains information about the payment methods supported by the merchant and the subscriber database contains the MSISDN of the subscriber and the subscribers' payment methods.

The payment service enables the merchant to submit payment requests. An incoming request is verified by checking if the MSISDN of the customer exists in the subscriber's database. If it does the payment methods of the customer and merchant are correlated. A request containing information about the purchase and the correlated payment methods are sent to the MSISDN of the customer through the WIG push module.

Merchant / Content Provider

The content providers can supply services using standard tools and either HTML or WML, as defined by the Wap Forum [7].

Requests to the merchant will come either directly from a standard web browser, from a WAP browser via a WAP gateway or from a SIM Application Toolkit browser using the OTA Service Center and the WIG.

The merchant implemented in Across Wireless AB's system consists of a simple CD-store named CD@Across. The content is very limited since the purpose of the merchant simply is to show and test the system.

Figure 12 illustrates the different ways the end user can reach the content provider.

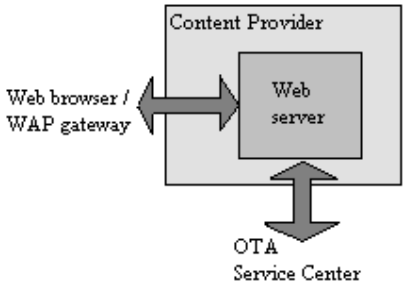

Figure 12 – Different ways for the end user to reach the Content Provider.

Payment scheme

If a customer wants to be able to use the m-commerce system the customer has to sign up with the wallet system. The reason for this is that the subscriber database in the e-commerce gateway has to contain information necessary for executing the customers' payments.

Many different merchants can connect to the e-commerce gateway. Information about what payment methods the merchants support is registered in the merchants' database.

When a customer wants to enter CD@Across he/she logs in with a username. The username is used to find the MSISDN of the customer in a user database at the merchant.

The customer browses the website using an ordinary web-browser or a cellular phone equipped with a WIB or a WAP browser. When the customer finds an album and decides to but it an order is placed with the merchant via the merchant's website.

Figure 13 illustrates the payment flow initiated by the customer requesting an album. Before the exchange starts the merchant connects to the e-commerce gateway.

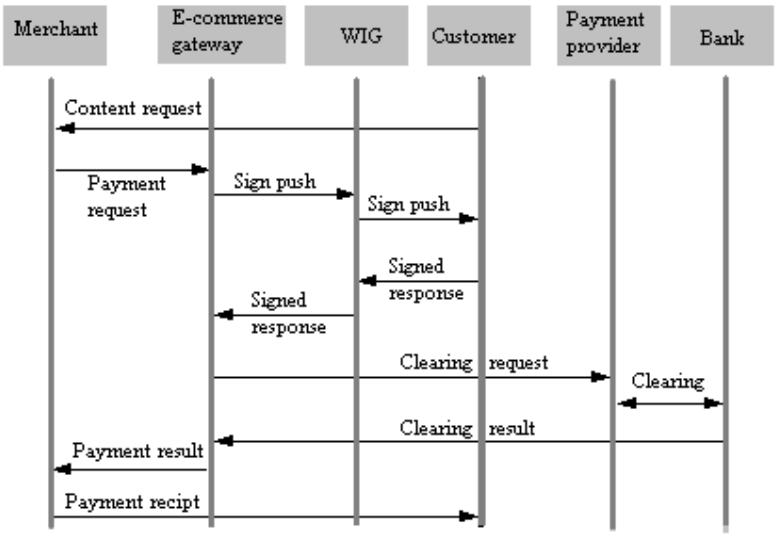

Figure 13– The payment flow during a purchase.

The customer requests an album by clicking on a button on the website. The merchant sends a payment request to the E-commerce gateway using Java RMI. The payment request contains information about the goods that the customer wants to buy. When the e-commerce gateway receives the payment request, data about the customer and the merchant is fetched from the databases. The payment methods that the customer can use are correlated with the payment methods of the merchant. A request containing information about price, goods, possible methods of payment, payment recipient etc. is sent to the customer's cellular telephone. The request is sent using the push functionality of the WIG. The WIG only forwards the information.

When the customer receives the payment request he/she can accept or deny. If the request is accepted the customer chooses a payment method and signs the request by entering a PIN code. This generates a response signed with a key stored in the SIM. The response is sent back to the E-commerce gateway through the WIG.

The e-commerce gateway sends a clearing request to the payment provider. The payment provider might handle the transaction directly or forward it to a bank, which will execute the transaction.

An answer indicating the result of the transaction is returned to the e-commerce gateway, which forwards it to the merchant. The result might also be forwarded to the customer's cellular telephone. This option might not be needed if the merchant notifies the customer through the website.

A simplified picture of this scenario is illustrated in Figure 14.

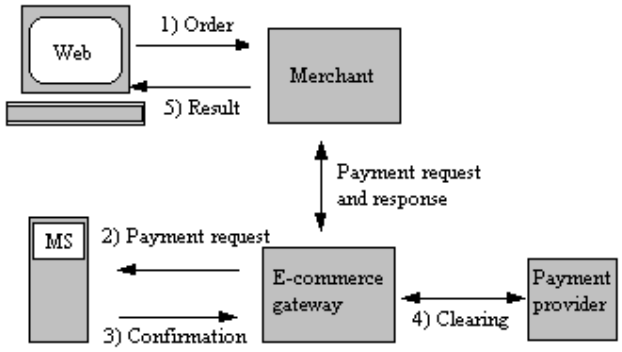

Figure 14– A user browsing for content and paying with his/her cellular phone.

Security aspects

The e-commerce platform developed by Across Wireless AB is as already explained a demonstration platform. Its purpose is to show useful situations in which the OTA Service center is involved and the possibility to create secure m-commerce applications. Therefore all security aspects are not handed but they have been analyzed and solutions to them have been given but not implemented.

A detailed description of the security in the system and what would be needed in order to make the system secure in a commercial situation can be found in [54].

8. Voice systems

Speech is the most natural way of communication for a human being. This is one of the reasons for the rapid development of voice systems i.e. systems that use the human voice as an input method for machines.

There exist various applications for voice systems of different kind. Most of the systems sold today are systems designed to give an alternative to typing for people using computers. These systems are also useful for people who are disabled in such ways that they cannot use their hands and for people who suffer from dyslexia [8].

Two types of voice systems can be identified to be useful in m-commerce situations. These are voice recognition systems and Interactive Voice Response (IVR). With these systems it is possible to automate routine call center operations such as someone wanting to buy a certain product. It is possible to reduce cost and to raise the efficiency by using these systems.

Customers accept talking to a computer rather than talking to a human being if the dialogue is designed in a good way [9]. Many customers prefer automated services. They get 24 hours service, they do not have to wait for the call to be answered and the operation can be performed simply and effective.

When it comes to shopping it is important to notice that these systems can be used by anyone without prior training or special equipment. The customers do not have to leave their homes buy what they need. Compared to shopping over Internet the voice systems are much easier to handle and they only require a telephone. Shopping using automated voice systems demands that the customer knows exactly what product he/she wants.

This section investigates IVR and voice recognition systems closer.

IVR

IVR is a system that takes a telephone call and handles it in preferred way. It can be as simple as informing the customer that all the telephone operators are busy, or it can play some nice background music to the customer while he/she is waiting to be served.

An IVR system can also be very complex and act as an operator giving 24 hours of access to the call center. If the customer knows what he/she wants the IVR can handle more orders in less time than an ordinary telephone operator can do [9]. An IVR is touch-tone driven and it can give the customer multiple choices that are selected by pressing buttons on the telephone. This makes it possible to perform effective transactions of information.

Several manufacturers are developing industrial strength systems for call centers. These systems can handle advanced telephony, computer and telephone integration (CTI) as well as touch-tone selection from a menu of options. The advanced IVR systems can also recognize a few words such as “yes/no” and single digits.

IVR is used by many companies around the world but is not yet accepted by many people.

Voice recognition

A voice recognition system tries to figure out what the person says by using large dictionaries and statistics of how often different words occur and in what order. The accuracy of voice recognition systems can be as good as 95% and if the user can correct the mistakes the system makes it is possible to reach 97% accuracy [9, 11].

The voice recognition systems can be divided into two different types. These are speaker dependent and speaker independent recognition [10]. Speaker dependent systems can recognize one particular voice after training consisting of the person that owns the voice reads a text. These systems are often used to give an alternative to typing for people using computers.

Speaker independent systems allow any person to communicate with a computer, regardless of the speakers’ gender, dialect and accent. This technology enables a computer to handle orders via telephone.

Sophisticated Automated Speech Recognition (ASR) and Natural Language (NL) systems are far more sophisticated than IVR. They operate on the level of phonemes rather than words and this makes them able to recognize very big vocabularies. This in combination with sophisticated algorithms makes it possible for customers to ask questions and engage in continuous speech. In order to work the speech has to be in a limited domain.

Discrete speech systems are voice recognition systems that use older technology then continuous speech systems. These systems require that a pause is inserted between each word. Despite this these systems can be very useful in the case of commerce using a telephone.

The discipline of speaking one word at the time address many people and the possibility the system gets of confirming each word gives great benefits [8].

These systems work well on older hardware, as the demands of processing are less than in a continuous speech system.

Continuous speech systems can recognize continuous speech. They are equipped with large dictionaries and the systems available today are speaker dependent. This means that the user has to load a personal profile before he/she can use the system.

Finally voice recognition systems can be used as authentication. This field is more related to security than to m-commerce.

IVR vs. Voice Recognition in e-commerce

IVR and voice recognition systems have different benefits. It is not the case that one type of system fits all e-commerce situations. Factors like performance, customer demand, price and core business are important factors to take into account when deciding what kind of system a company should choose.

At a first glance the voice recognition system seems to be better since they are more advanced, the technology is newer and there is a lot of research in the field. Voice recognition is a hot topic and speech recognition over the telephones is believed to be a “killer application” over a few years [10].

Even though speech recognition is such a hot topic it is important to realize that an IVR can solve the problem of automated telephone attendants just as well in certain situations. If the goods offered is limited and the dialogue is user friendly the merchant will end up with a system that is cheaper, easier to administrate, and that performs excellently.

In more complex situations where the offered goods are varied and the purchase needs a more complicated dialogue voice recognition systems will probably perform better. If a merchant decides to install a voice recognition system he has to make certain that it is a speaker independent system that can handle dialects and accents in a good way, otherwise customers will not use the system.

One example of a voice recognition system in use is the ticket ordering system at Swedish State Railways (SJ). The procedure of ordering tickets is complicated since the dialogue involves the starting place, the destination and the time of travel. In this situation an IVR would be less user-friendly.

On the other hand the telephone call to SJ is handled by an IVR that lets the customer select one out of four choices. Another example where an IVR is useful is for ordering forms from for instance the tax authorities. They supply a limited number of forms so the dialogue can be simple.

Figure 15 and Figure 16 illustrates an example of how a dialogue between a customer ordering a CD-record from CD@Across and a voice system could look like using an IVR compared to using a Voice system.

Customer	Merchant
Calls the merchant.	*Hello, and thank you for calling CD@Across. Using the telephone keypad, please spell the name of the artist you're looking for. For the letter 'q', press 7; for the letter 'z', press 8.*
Customer is interested in Madonna and presses the digits that correlate to the letters in ‘Madonna’ on the phone.	*You are looking for records with Madonna, if this is correct press 1.*
Customer presses 1.	*We have the following records with Madonna in storage: Like a prayer, Ray of Light and True Blue. Press 1 for Like a Prayer, 2 for Ray of Light and 3 for True Blue.*
Customer presses 3.	

Figure 15 - Touch-tone Driven Voice System

This example assumes that the customer knows what artist he is looking for and that he knows how the name of the artist is spelled. If this were not the case the voice system would have to supply more options in order to be user-friendly. Another weakness with this system is that the keypad of the telephone only has eight buttons associated with letters. Each button (2-9) represents three or more letters. This means that a number correlating to the letters in the artist name would represent an artist. "Madonna" would be equivalent to typing in the number "6236662" and thus the database cannot distinguish between another artist name that has the same number associated with it.

The following dialogue would occur if the touch-tone driven system were exchanged to a voice recognition system.

Customer	Merchant
Calls the merchant.	
	Hello, and thank you for calling CD@Across. What artist are you interested in?
Madonna!	
	Did you say Madonna?
Yes!	
	We have the following records by Madonna: Like a prayer, Ray of Light and True Blue. Which one do you want?
True Blue!	

Figure 16 - Voice Driven Voice System

This example assumes that the voice recognition system make a proper match on the first attempt.

It is obvious that it is more troublesome to use an IVR than a voice recognition system when using a cellular telephone. It is inconvenient for the user to bring the telephone back and forth to the ear during the session. This is not the case when using a stationary telephone where the keypad is separated from the receiver or a hands-free unit.

Conclusion

There is a lot of research in the area of voice recognition. The goal for this research is to construct robust voice recognition systems that can recognize fluent speech with a high accuracy. In order to use speaker-independent voice recognition systems the accuracy has to be at least 98% in order to be accepted by the customers. Today the speaker-dependant systems almost reach his accuracy after a few hours of training [11].

The IVR technology is old but these systems are excellent if they are used in suitable situations. IVR's will probably stay on the market for a long time.

A voice system used in e-commerce needs to be able to recognize many different voices, dialects and accent with high accuracy. Therefore discrete voice recognition systems perform better with today's technology but when the continuous speech systems get increased accuracy and will become speaker independent they will take over the market.

Unfortunately the discrete systems are getting difficult to buy since most companies are producing more advanced continuous systems [8]. But by the time the discrete systems will be extinct the continuous systems have probably reached such a high level of accuracy that they are speaker independent.

Irrespective of which system the merchant chooses it is important that the configuration is done in a careful way. The dialogue should be user friendly and the customer should not get stuck in loops or get too many choices. The opportunity to get in touch with a human operator should be available and the risk of a system crash should have been taken into account.

9. GSM data services role in the m-commerce platform

The GSM data services have been described in more or less details in the sections that cover the GSM System. This section aims to investigate if the technologies can be used and how they can cooperate in order to be deployed in the Across Wireless platform for m-commerce.

Sending and receiving data are fundamental in the e-commerce application, CD@Across, developed by Across Wireless AB. Obviously data communication is absolutely necessary when the customer accesses the merchant's website using a WAP browser or an SAT browser.

The security provided by the system partly depends on the merchant pushing a payment request to the cellular telephone of the customer. The customer accepts or denies the purchase by sending a data item back to the merchant.

When the customer uses the voice interface, which has been implemented during this master thesis, data transmissions are still used to send payment requests and replies between the merchant and the customer.

The access to CD@Across

The WAP case

When a WAP browser is used to access the CD@Across the data is sent with the WAP protocol using circuit switched data (CSD) as the bearer. The transmission speed is currently 9.6 kbps. One consequence of using CSD is that the cellular telephone is busy during the purchase. This might be inconvenient in some cases and the user has to pay for the time of the session even if a great part of the time consists of reading downloaded pages. Using CSD as the bearer should therefore be regarded as a drawback.

Another drawback with WAP carried by circuit switched data is that the time to set up the connection to the ISP is quite long. The user might have to wait for 30-60 seconds [54] before connection is established. 30 seconds feels like a long time to wait when the user is watching a small display on a cellular telephone.

Payment requests and replies can be transferred even though the cellular telephone is busy since they are carried by SMS and short messages can be transferred independently of ongoing calls.

When a customer executes a purchase from the CD Store the purchase takes about 2 minutes and 45 seconds, including set up time. When the connection has been set up the customer is active most of the time [54].

The SMS / WIB case

SMS might not look like an attractive bearer of data if bare technical facts are taken in account. In a circuit switched environment only control channels are used for transmissions, this makes the transmission rate low. The amount of data that can be transmitted in one SM is limited to 140 octets.

However, SMS has many benefits compared to the other methods. SMS has store and forward capability and it can be used in great parts of the world. SMS can be carried by GPRS and is therefore not a competitor to GPRS but rather a complement.

When the SAT browser is used to access the merchant the data is carried by SMS and sent through the WIG that translates the information in the short messages to fit the Internet. In this case there is no setup time (as in the WAP case) but since SMS is a store and forward service there is some delay in the transmission.

The time for one session consisting of eight short messages is 1 minute and 20 seconds. Only 20 seconds of this time is occupied by user action. This means that the waiting time is 1 minute [54]. The waiting time is closely related to the amount of data that is being transmitted. In order to keep the amount of data as low as possible a special set of WML pages is used by the WIB.

Another way to reduce the waiting time could be by using another SMS-C than the one used today. Some SMS service centers offers express message delivery, permitting high throughput and eliminating the latency associated with store-and-forward [21].

Alternatives

WAP over SMS

If WAP uses SMS instead of CSD as the bearer the cellular telephone will not be busy during the purchase and the set up time will be eliminated. This scheme would improve the WAP access but it would not perform better than access through the WIB. The reason is that WAP adds overhead to the

messages that are transferred. Even the simplest transaction would require several SMS messages to be sent. This might be both time consuming and expensive.

WAP over USSD
USSD might at a glance look like an alternative bearer to WAP instead of CSD that is used by the system today. It does indeed have some attractive features that could enhance the communication between the client and the merchant. USSD is session oriented just like CSD. This means that the delays between reply and response are limited to the processing time and the transmission time for the messages.

The great benefit with USSD compared to CSD is that the cellular telephone is not occupied during an USSD session. Consequently it is possible to receive ordinary phone calls during an USSD session.

The set up time is shorter than circuit switched data and the session can be released while the customer is reading the downloaded message. This prevents unnecessary costs for the user.

The drawbacks with USSD are that all GSM networks do not support USSD. In Sweden only one of the three GSM operators support USSD [20]. The GSM operator that Across Wireless cooperates with does not support USSD.

Finally GPRS is expected to be launched during 2001 or 2002. Since GPRS will be able to carry WAP in a packet switched manner at a much higher rate than USSD can it would be a waste of time and effort to change the implementation from CSD to USSD at the moment.

Using USSD instead of SMS with the WIB
USSD cannot be used by the WIB. The main reason is that although the SIM Application Toolkit supports USSD there is no mechanism that enables an incoming USSD string to be transferred to the SIM. An incoming USSD string is displayed to the user directly in the ME.

The core idea with the WIB is that incoming short messages are transferred to the SIM and the WIB behaves according to the instructions in the messages. Since this does not work with USSD, USSD cannot be used with the WIB.

The payment request & reply

The handling of the payment is separated from the actual ordering. The payment request and reply are handled the same way irrespectively of whether a WIB, a WAP browser, a Web browser or voice is used for ordering. The payment is handled in a separate session consisting of a payment request being pushed to the customer's cellular telephone and a signed reply from the customer to the e-commerce gateway, which in turn handles the transaction of money in an appropriate way.

The idea of the payment request being pushed to the customer's cellular phone is to ensure that the customer has access to his cellular telephone and knows the PIN that is used to sign the payment reply. If the cellular telephone is stolen or used by a fraud it can be assumed that the unauthorized user does not know the PIN or that the owner has blocked the phone.

The existing system uses SMS messages for the payment request. The reason is that the amount of data is small. The request and the reply fit in one short message each. As mentioned above there is no good alternative to SMS when sending the payment request and reply. USSD cannot be used since the USSD messages not can be sent directly to the SIM card. WAP is either sent over a circuit switched connection or carried by SMS. The setup time for a circuit switched connection makes it useless for these tiny messages and sending WAP over SMS seems like unnecessary work when SMS can be used directly.

Finally, SMS makes it possible to receive the payment request and to send a reply even if the customer has a circuit switched WAP session to the merchant.

The effects of GPRS & UMTS on CD@Across

Introduction
GPRS is gradually being implemented in the GSM network. The timelines for the implementation differs depending on the company that makes the timeline. One thing is certain and that is that within a few years GPRS will be functional and used by a great number of persons.

Some years further on UMTS will be implemented and it will be compatible with the existing systems. For some time three different networks will coexist, GSM-CS, GPRS and UMTS. Depending on what kind of subscription and handset the user is equipped with, he/she will be able to access different services.

The major advantages with GPRS and UMTS are that they offer higher transmission rates, they support TCP/IP and the subscriber seems to always be connected. The benefits that GPRS and UMTS will provide will have great impact on the type of services the service providers can offer their customers. Below I try to outline how these benefits will effect the CD@Across demonstration application

Coverage, access and transmission speed
GPRS will initially be an overlay network to the existing GSM network. During the deployment phase of GPRS the coverage will be limited and most subscribers will have non-GPRS cellular telephones. The initial transmission rates for GPRS is expected to be much lower than the maximum rate.

The effect of these facts is that if a merchant wants to create services that are accessible for the mass market it is better to adapt the system so that it fits technologies with a wide coverage and lower transmission rates. If an application works in the GSM network that exists today it will also work when GPRS is introduced. When GPRS is widely deployed the service can be transformed to use the features offered by GPRS. The transformation includes making a more user–friendly interface towards the customer and offering new services such as images, sound and video.

The same arguments can be given when discussing the effects that UMTS will have on the application.

GPRS/UMTS and TCP/IP
GPRS and UMTS will make it possible for the subscribers to access CD@Across using TCP/IP. A customer equipped with an HTML browser enabled GPRS/UMTS cellular telephone could access CD@Across using the ordinary HTML pages.

However TCP is a transport protocol developed and optimized for fixed connections. In a wireless network the performance of TCP is severely degraded due to the characteristics of wireless links[6]. Furthermore it is difficult to know what the different GPRS/UMTS handsets will provide.

Even though GPRS/UMTS will make it possible for the subscribers to use the TCP/IP protocol stack I find no reason for the mobile users of CD@Across to access it using TCP/IP instead of WAP.

WAP is developed for wireless networks while TCP/IP is not.

HTML vs. WML
In GPRS/UMTS the subscriber is charged for the amount of data being transferred between the telephone and the network. A web page that consists of HTML is larger than a WML page containing the same information. This means that the customer has to pay more if the information he wants is coded in HTML than if it is coded in WML.

The displays on the cellular telephones will continue to be smaller than the displays or screens of fixed equipment. WML is optimized for small displays and will therefore probably be better suited for wireless Internet browsing than HTML.

Several implementations of XML translators exist [30, 31]. These translators make it possible to convert information coded in XML to HTML or WML depending on the kind of device accessing the server. These translators might solve the problem of a content provider having to choose between HTML and WML.

Conclusion
GPRS/UMTS should be seen as complements to SMS and WAP. If a service is created with the transmission capabilities of WAP and SMS in mind the service can be used by traditional GSM users as well as by users that have GPRS/UMTS supporting cellular telephones.

SMS or WAP messages can be carried by GPRS and thus take advantage of the improved performance offered. WAP carried by GPRS reduces the set up time and SMS carried by GPRS shortens the waiting time due to a more effective signaling procedure in the network [21]. When UMTS is introduced the GPRS packets can be carried by UMTS and the transmission rate will increase even more.

Conclusion

There is not one GSM data service that is optional in all situations. They all have a niche where they are optimal, or at least perform better than the others do.

[6] On wireless links bit-errors are more frequent than on wired links with stationary hosts. TCP/IP assumes that the bit-errors is caused by congestion and acts accordingly. The transmission window size is decreased; Slow Start and Karns Algorithm [55] are initiated. These actions result in an unnecessary reduction in the bandwidth utilization.

The best suitable bearer is determined by the needs for the user in a special situation. If the customer shall choose a CD record from an online CD store WAP, SMS or WAP/SMS over GPRS suffices, but if the customer wants to listen to the record, or download it, GPRS or UMTS will be needed.

The access to CD@Across using WAP/SMS works satisfactory. Obviously it would be nice if the weaknesses of the protocols (set-up time/waiting-time) were eliminated but the advantage of reaching a big customer base are more important. As GPRS is introduces these weaknesses will disappear.

SMS is a perfect bearer of the data transmitted in the payment request. The message is short enough to fit into one short message and the time it takes for the message to reach the customer is not annoying.

When the GSM network will be upgraded with GPRS/UMTS WAP/SMS can be transmitted with GPRS and thus use the benefits of higher transmission rates and a more efficient signaling procedure.

In general it is better to develop functional, usable services that can be used by the mass market than to create trendy services that uses the latest technology but are usable for a subset of the potential customers.

10. Integration of voice & data in Across Wireless system

Introduction

What I mean with integration of voice and data is simply the combination of voice-communication and data-communication in the same session.

Across Wireless AB's e-commerce platform could be extended so that the customer places an order using his/her voice and the payment procedure is handled using data that is manipulated by security mechanisms in the SIM card. This combination of voice and data enables secure transactions and a user-friendly interface i.e. a voice interface.

As explained earlier the demonstration e-commerce platform that Across Wireless AB has developed can handle customers browsing in to the merchants website using a web browser, a WAP browser or the WIB. In order to make the platform user-friendlier demands for a voice interface to the merchant has come up.

A voice interface would satisfy the needs for people that have difficulties navigating the web through a mobile telephone with a small display and a small set of buttons.

I have chosen a method that uses a speaker independent voice recognition system. The system interacts with the customer so that the preferred product can be found by navigating the merchant website using the voice. When an order is placed the merchant communicates with the payment interface in the e-commerce gateway using the payment interface. The rest of the payment is handled as described in section 7.

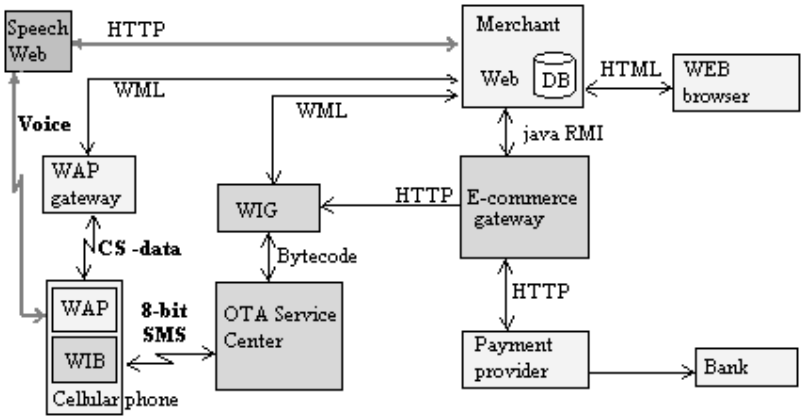

Figure 17 – The demo platform extended with a voice interface.

The new entity in the system is the SpeechWeb, which is used to create a voice interface to the merchant.

The Voice System – SpeechWeb

SpeechWeb is an application developed by PipeBeach AB, a Swedish company specialized in speech enabled mobile Internet services [18]. SpeechWeb is an audio browser server that reads the content on WebPages with a synthesized voice.

SpeechWeb can be accessed through any standard telephone. It makes the mobile telephone a convenient tool for retrieving information from the Internet. The user interacts with SpeechWeb through voice or touch-tones or both in combination.

The applications created for SpeechWeb should be written in Voice XML although SpeechWeb can read ordinary WebPages. If ordinary HTML is used instead of Voice XML the structure of the WebPages should be clean (i.e. the use of complex tables, images containing text, frames etc. should be avoided) otherwise the content might sound peculiar.

The SpeechWeb can reside anywhere on the Internet. By dialing a telephone number the user gets connected to SpeechWeb, which has a pointer, URL, to the document to read for the user. From this first document the user can navigate further on, using voice commands in interaction with SpeechWeb and forms in the documents.

Figure 18 illustrates the interaction between the user, SpeechWeb, the merchant's website and the e-commerce gateway.

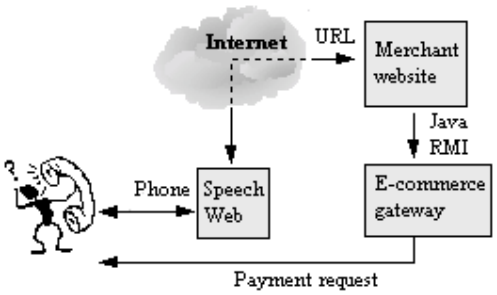

Figure 18 – Interaction between a user, the SpeechWeb, a merchant and an e-commerce gateway.

Reasons to why this system was chosen.

When I started this master thesis the intention was to use an IVR that Across Wireless AB already had but did not use.

When I started to analyze the situation of a customer buying CD records through a voice interface consisting of an IVR I realized that the interaction between the customer and the IVR would be complex and not user-friendly at all.

The first question that arose was the question of how the customer will select the artist he/she is interested in. It is too complicated to type in the letters in the artist name on the mobile handset. It takes time and it demands that the customer knows how the name is spelled. The other approach was to let the voice system reel of the names of all the artists in store and let the customer press a button when the name of the artist he is interested in is spoken. This method forces the customer to listen to lots of information and thus is not useful.

When I realized that an IVR would not suffice in order to make a user-friendly system I started to look for a company that was developing speaker independent voice recognition systems and that was willing to cooperate with me. Finally I found PipeBeach and their SpeechWeb system.

The benefit for PipeBeach is that they get an opportunity to show that their system can be used in secure m-commerce. The benefit for Across Wireless is similar.

Voice eXtensible Markup Language (VoiceXML)

Traditionally people have been using their telephones to retrieve information about various things. Today many people use the Internet to retrieve the same information. If the content providers want to provide their services to telephone users and Internet users they have to produce the same information using two different technologies.

VoiceXML, developed by The VoiceXML Forum [19], is a solution to this problem. VoiceXML is a new standard for making Internet content and information accessible via voice and phone. VoiceXML makes it is possible to create voice services using the same technology as for creating visual websites. The telephone is used for input, output and call control. It supports touch-tone input, automated speech recognition, audio recording, call transfer and much more.

With a voice markup language the content providers can provide services to their customers using the same technology that they are using to provide visual services. The advantage is obvious, many of the existing web development tools can be used, the programmers already know the technology, databases can be shared between visual- and voice services and the services can be published on existing web servers.

When large numbers of voice services are available on the Internet it will become possible to interact with several unrelated services during the same phone-call.

VoiceXML is similar to HTML. Voice services are produced by marking the content of a document with XML tags. The tags indicate how the information shall be treated by the browser. Interaction with the user is done by using links, menus and forms.

Figure 19 and Figure 20 illustrates two simple examples of VoiceXML documents.

Figure 19 illustrates how a simple “Hello World!” document can be written.

The <vxml> tag simply indicates that the document contains a dialogue.

The dialogue can be either a <form> or a <menu>. The <form> presents information and gathers input. The <menu> is used to present choices.

This example consists of a <form> and a <block>. The <form> does not expect any input, it simply synthesizes the text in the <block>.

```
<?xml version=”1.0”?>
<vxml version=”1.0”>
    <form>
        <block>Hello World!</block>
    </form>
</vxml>
```

Figure 19 - a “Hello World!” VoiceXML document.

The example in Figure 20 asks the user if he want coffee or tea. When the user says coffee or tea the application proceed with the next document.

This example contains a <field>, a <prompt> and a <grammar>. The <field> is used to catch the answer to the prompted question. The <grammar> is a set of valid answers that the user can utter.

In this case the <block> contains a link to the next document to read. Note that in this case the next file is the result of running a Java Server Page.

```
<?xml version=”1.0”?>
<vxml version=”1.0”>
    <form>
        <field name=”drink”>
            <prompt>Do you want coffee or
tea?</prompt>
            <grammar> coffee | tea </grammar>
        </field>
        <block>
            <submit next=”http://www.drink.jsp”/>
        </block>
    </form>
</vxml>
```

Figure 20 - another example of a VoiceXML document.

Normally, each document runs as an isolated application. If several documents work together to form an application there might exist links, forms, grammar, variables etc. that are common to all the documents in the application.

If this is the case one document is selected to be the application root document. When this is done the interpreter loads the application root document each time a document belonging to the application is loaded. The benefits using an application root document are that the variables in the root document can be used by all other documents in the application. Furthermore the grammar of the application root document can be set to remain active when the user is in another application document. In this way the user can always interact with common links and forms.

Grammar

The forms and links in the VoiceXML documents make it possible for the user to enter data and to interact with the document. In order to enable this the speech system has to be equipped with an Automatic Speech Recognition (ASR) module so that it can recognize what the user utters.

In VoiceXML the <grammar> tag is used to control what the user is allowed to say in order to make something happen. The user can say anything but only the utterances defined it the grammar will be recognized as valid.

The grammar can be inline grammar or grammar defined in a file. The advantage with grammar files is that the files can be used by several different VXML documents. The inline grammar is suitable in situations where the dialogue is simple.

The example in Figure 20 gives a simple example of the use of inline grammar. In this example the “submit” tag will only be executed if the user says either “coffee” or “tea”. Anything else will be unrecognized.

The complexity of the grammar depends on the application. If the user is allowed to enter entire sentences the grammar gets more complex than if the user is allowed to enter single words.

There is currently no official standard of the grammar used in speech recognition but several different grammar proposals exist. W3C is currently working on the standardization of speech recognition grammars. Right now the question is if the final standard will contain both XML syntax and (Augmented Backus-Naur Form) ABNF syntax or if it will be narrowed to one specific form [49].

An example of the difference in syntax between ABNF and XML is shown in Figure 21and Figure 22. In these examples the grammar is defined in a file and the user can choose between two artists,

Madonna and The Rembrandts. The Rembrandts will be matched by “the Rembrandts” or just “Rembrandts”.

The vertical bar in Figure 21 means OR. Thus the user can choose between Madonna or The Rembrandts

If the user wants Madonna he/she simply says “Madonna”. This matches Madonna and the variable, $artist will be tied to the word between braces.

If the user wants The Rembrandts he/she has the option to say “The”. This is indicated by the square brackets.

Even if the user say just “Rembrandts” the variable, $artist will be tied the word between braces, i.e. “the rembrandts”.

The XML grammar in Figure 22 is more complicated than the ABNF grammar in the previous figure even though they state the same rule.

The <choice> tag corresponds to the vertical bar in the ABNF syntax.

The tag keyword corresponds to the braces in the ABNF syntax. Thus the word following the “tag” keyword is the variable that will be tied to the input field if a match occurs. The token within the <count> tag is an optional token.

```
#ABNF V1.0 ISO8869-1;
language en;

public $artist = Madonna {madonna} |
[the] Rembrandts {the rembrandts};
```

Figure 21 – ABNF grammar-file syntax

```
<?xml version=”1.0”?>
<grammar xml:lang=”en”>
 <rule id=”artist” scope=”public”>
  <choice>
   <item tag=”madonna”>madonna</item>
   <item tag=”the rembrandts”>
    <count number=”optional”>The</count>
    Rembrandts
   </item>
  </choice>
 </rule>
</grammar>
```

Figure 22 –XML grammar-file syntax

The voice interface

The voice interface that I have created for CD@Across consists of several document that are dynamically created using Java Server Pages that execute and return Voice XML documents. The content of the VoiceXML pages are read by SpeechWeb and presented to the customer through voice.

I used an application root document that contains two links that the user always can follow in order to quit or to be transferred to an operator. The flowchart of the voice interface is illustrated in Figure 23.

The dashed arrow indicates that the application root document (voice_root.vxml) is loaded together with every other document in the application and that the commands “quit” and “operator” can always be issued. The words within quotation marks symbolize the words that have to be uttered in order to follow a link to sub-dialogue or to another VoiceXML document. The words within brackets symbolize that in order to follow this link the user has to match dynamic grammar specified in a database in the CD Store.

The box named “confirm” is simply a sub-dialogue, which is used to force the user to confirm the input. The reason to this is to avoid unwanted actions caused by background noise or pronunciation errors.

The note symbols means that music is played to the user and that the dialogue will continue where it was interrupted.

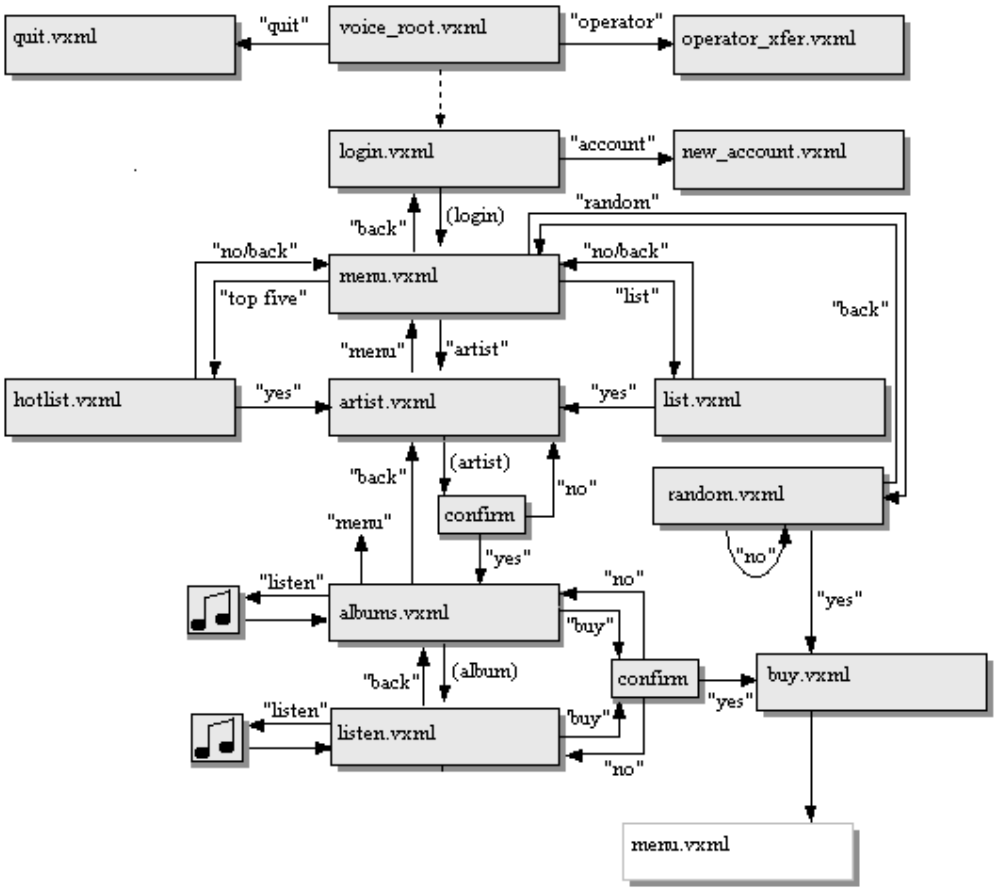

Figure 23 - a flowchart illustrating the dialogue between the merchant voice interface and the customer.

The data communication

When a customer access CD@Across using the voice interface the only data communication that occurs is when the payment request is sent from the merchant to the customer and when the customer sends a positive response to the request.

The existing platform uses SMS-PP for sending the payment request and reply. I found no reason to change this since the messages are so small. The request and the reply fit in one short message each. As I hopefully made clear in the previous section there is no good alternative to SMS in this case.

The exchange of short messages begins when the customer says, “buy”. This utterance causes CD@Across to ask the e-commerce gateway to send the payment request to the cellular telephone. The request is received after approximately 15-30 seconds. The time it takes depends on the current load of the SMS-C.

After accepting, choosing payment method and signing the customer returns the payment reply. The time it takes for the reply to be transmitted and handled is 10-15 seconds. During this time the circuit switched connection stays open and the conversation continues after the payment has been processed.

Effected parts of the system

The modifications that I have done to the existing system do not affect the functionality that already existed. I have simply added some features and made some extensions to the database in the CD Store.

The login procedure
When the user wants to access CD@Across he/she dials the telephone number and logs in. In the WAP/WIB/web browsing scenarios the user logs in by using his username. The username consists of a set of characters that sometimes are difficult to pronounce. Therefore I have given the users an additional username that is used for logging on through the voice interface. This username is simply a short unique number. The customer can log in to the system either by saying the digits in the username or by using the keys on the keypad.

The username does not play an important role from a security point of view. It is simply used to identify the customer so that the payment request is sent to the right cellular telephone. If a user, with or without purpose, uses someone else's username the payment request will be sent to the lawful owner's cellular phone. This can be annoying but as long as the lawful owner does not authorize the payment he will not be charged.

The possibility for the user to type in the username on the keypad reduces the risk that the username is revealed.

The database
The database at the merchant (CD@Across) has been extended. I have added columns in two of the tables in the database. The extensions include first and last name of the registered customers, the username intended for voice login that I mentioned above and the grammar for the artists and the albums in store.

The first and last names of the registered customers are used in order to make the interaction more personal.

The grammar for the artists and the albums is needed to recognize variants of user input and to tie utterances to distinct variables. If for example the customer wants to choose the artist "the Rembrandts" he/she might say either "the Rembrandts" or just "Rembrandts". Both these utterances should be valid and the variable "artist" should be tied to "the Rembrandts".

The grammar in the database has the ABNF form. The reason for this is that it simplifies the table entries for the administrator of the database. The ABNF grammar is converted to XML syntax when the dynamic Voice XML pages are created.

The source code
The source code of the existing system has not been modified. The changes I have done have been implemented in new Java classes that work in parallel with the existing classes. The reason for this is that I have been working with the live demo and I did not want to mess it up during my implementation.

Performance

In order to draw conclusions about the performance of the system I made a number of purchases at CD@Across. 100 of test purchases were done. In 50% of the tests I used an ordinary telephone for the communication and in the other 50% I used a cellular telephone for both the communication and the actual payment.

All the telephone calls to CD@Across were performed in the same way. I called SpeechWeb and logged in with my username. After being welcomed I selected an artist and purchased the album without listening to it.

The factors I looked at during the tests were the total time it took to order an album, the time between sending and receiving the payment request and the time from sending the reply to the result was presented to the me. Of course factors like the stability and sensitivity of the SpeechWeb was observed. The protocol and the detailed results are presented in appendix C.

It should be noted that all the test purchases were made by me. The fact that I am used to talking to the system and that I have developed it might give a misleading result.

Successful purchases
88% of the purchases were successful in the sense that I could buy the album I intended to buy. Even though some of the purchases were disturbed by background noises and/or pronunciation errors I managed to buy the preferred album.

68% of my attempts succeeded without any problems at all. This has to be considered a satisfactory result even if it is not good enough in order to work commercially. Keep in mind that the SpeechWeb system I have been using is a demonstration system. The result could probably be much better if

Across Wireless AB bought their own system and if the dialogue was altered so that it became more resistant to pronunciation errors.

Time aspects

The time it takes for a customer to buy an album depends on the actions taken by the customer. If the customer chooses to listen to the "top five" list or to the list of available artists the session will naturally take longer time than if the customer logs in, chooses an artist and an album and then buys the album without listening to it.

The shortest time for a purchase is achieved when the customer has some previous experience with the system and knows what and when to enter information. The time to buy an album for such a user is close to 2 minutes including receiving, signing and sending the payment request and reply. This can be compared to the WAP and WIB cases. In the WAP case the entire browsing/purchase takes approximately 2 minutes and 45 seconds and in the WIB case 1 minute and 20 seconds. The voice case lies somewhere in between the other cases.

Making the dialogue shorter can shorten the entire time for a purchase. In the current implementation I have tried to make the dialogue clear and easy to understand for an inexperienced customer. This might be annoying for a customer that is used to the system and wants to do a fast purchase.

The customer is active during 88% of the purchase. When I say that the customer is "active" I mean that he/she is listening, speaking or working with the cellular telephone. The inactive time consists of the time the customer is waiting for the short messages to be transferred to and from the cellular telephone.

When using a cellular telephone for both the browsing and the purchase the time for the payment request and reply is a little bit longer than when using a stationary telephone for the browsing. The reason to this is that the cellular telephone uses some of its processing power to handle the circuit switched connection and therefore it takes a little longer time to handle incoming and outgoing short messages.

Background noise

A majority of the tests have been performed in an open landscape with some background noises. Some problems with background noises occur. The tests performed outside in crowded environments also show some problems with background noises. The cellular telephone was more sensitive to background noises and other noises such as coughing etc. The reason to this is probably that the microphone of the cellular telephone is more sensitive and more exposed that the microphone of a stationary telephone.

The links located in the document root file adds some problems to my implementation when used in areas with lots of background noises. The reason is that the global links are valid throughout the session. Whenever SpeechWeb recognizes some input the input is matched with the current form grammar. If no grammar can be matched SpeechWeb will match the input with the global link that reminds most of the input. This means that a sudden noise will be matched by either some valid grammar or by "operator " or "quit".

It is very annoying when the call terminates just because a passing car honks.

Choosing long utterances for the global links can solve this problem (i.e. the risk of SpeechWeb matching a sudden noise with a combination of words is smaller than matching a short word like "quit").

PipeBeach considers the problem with global links a bug and they are working on some kind of solution to this.

Pronunciation errors

The customer's utterances might be misunderstood by SpeechWeb. If two tokens resemble each other SpeechWeb might have difficulties to separate the utterances. One example is the two tokens "buy" and "back". If the customer is careless with the pronunciation of the tokens SpeechWeb might misunderstand the customer's intention. To prevent this I have built the system so that the customer is forced to confirm certain utterances.

Stability

The SpeechWeb that I have been using during my work is a system set up for demonstration and test purpose. In the beginning of my work I had some stability problems but the system has gradually been upgraded to support the VXML standard and to be stable and robust. The upgrades have been done once or twice a week and the system has continuously become more stable and capable.

Summary
The voice interface to CD@Across works satisfactory but it has some weaknesses. If a user is aware of these weaknesses they can be handled and successful purchases can be performed. The time for a purchase lies somewhere in between the WAP and the WIB cases. In the WIB case the purchase can be performed faster than in the voice case but the voice purchase is experienced to be shorter. The reason to this probably is that in the WIB case the user is active during 25% of the purchase [54] while in the voice case the user is active during 88% of the purchase.

The stability problems I had with SpeechWeb in the beginning of my work has disappeared due to a continuous work by PipeBeach AB.

The problem with background noises disturbing the purchase is annoying. Making the purchases in a reasonable quiet place can prevent this problem.

When using a GSM cellular telephone the input (i.e. the voice) is transformed by the GSM audio codec. It seems like this transformation has a small impact on the likelihood of a successful purchase. The number of successful purchases was almost the same independently of using a stationary or a cellular telephone.

The key to a successful purchase is to talk slow, legible and to talk when asked to talk.

Access to the system

In order to be able to use the demonstration system the user has to have an account at CD@Across and the user has to be registered with the e-commerce gateway. If the user wants to perform transactions using VISA, MasterCard or some other credit card the user also has to be registered in the demonstration bank server. The transactions performed by the bank server are only demonstration transactions, no real money is transferred.

Furthermore the user has to be equipped with a SIM card with a WIB. When the user has done the proper registrations and put a compatible SIM card into his/hers cellular telephone he/she can use the system.

Note that the system can be tested with any telephone supporting DTMF tone signaling but in order to perform the payment the user has to be equipped with a cellular telephone with a WIB SIM card.

To reach CD@Across the user dials (08)-555 13 545. The user will be connected to PipeBeach gateway. In order to reach CD@Across the user has to type in the following access-code: 010453#. The user will now be able to log in using his/hers “voice username” which is a unique number that he/she got when he/she registered at CD@Across.

11. Other solutions

Commercial products that handle payments using cellular telephones exist today. Several companies have developed solutions with different functionality and security solutions. Some systems handle “real” money through transactions between different accounts, other systems use electronic cash, mail order or operator billing.

TeleVend [23] is an Israeli company that offers vending machines where the payment is done using a cellular telephone. The customer calls a telephone number that the vending machine is labeled with. The purchased item is released and the customer is charged on his phone bill. The system does not provide any security except the assumption that the customer is the owner of the phone.

SmartAxis [24] has developed a system that is based on electronic cash stored on a SIM card. It has some similarities with a pre-paid GSM subscription. The customer loads an amount of cash to the SIM card and uses it to pay for content.

Companies like Fundamo [22] and Jalda™ [25] have developed systems that use payment schemes similar to that of CD@Across. The customer identifies himself with a PIN and the payment is confirmed by sending short messages to and from the cellular telephone.

The use of voice interfaces in combination with mobile commerce is quite new and it is difficult to find information on how the transactions and security issues are handled.

One example of a solution for speech enabled mobile commerce applications is ShopTalk [28]. ShopTalk is a platform that provides access to a portal that has gathered personalized offers to the user. The voice interface is constructed using an IVR and the customer has to listen to the offers and chooses what to do by pressing the buttons on the phone. If the customer chooses to buy some goods he is connected to the company that gives the offer and the purchase is handled manually. ShopTalk has over 450.000 users.

Other systems for speech enabled mobile commerce are being developed by companies like Phone.com, Telia Mobile and by Conversay in cooperation with IIS [29]. The details of these systems are difficult to get hold of since the development is on an early stage.

The development of systems like SpeechWeb and VoiceGenie Telephony Server [26] and services like Tellme Studio [27] will probably dramatically increase the number of m-commerce sites accessible through voice interfaces. The reasons are that these systems and services make it easy to create dynamic, interactive services that can be controlled by voice.

The payment schemes used in traditional electronic commerce differ between different merchants. In the same way different methods will be used in mobile commerce until a standard that is accepted by the customers evolve. In order for the customers to use the systems they have to feel enough secure in relation to the offered service.

12. Future work

During my master thesis I have implemented a demonstration voice interface to CD@Across. If this demonstration application would be used as a full-scale commercial product several extensions would be needed or desirable.

The modifications that have to be done to the e-commerce gateway are described in [54]. The modifications include enhanced security and modification of the payment interface.

The modifications that would have to be done to the part I have implemented comprise the online registration and the transfer of the phone call to an operator. Furthermore the dialogue between the merchant's computer and the customer can be improved.

In the demonstration application the customer does not have to enter a password to verify his/hers identity. In a commercial system this might be needed since the owner of an account might be spammed with payment requests originating from a person using the owners account. The owner of the account can simply ignore the payment requests and will thus not be charged but the spamming might give the application and the company that offers it a bad reputation.

Online registration

The online registration is not yet implemented. The reason for this is that the information needed for the registration is confidential and the procedure in the demonstration environment in complicated. The user has to be registered in a database at CD@Across and in another database in the E-commerce gateway. Finally the user has to create an account in the demonstration bank.

The creation of a new account would be easy to implement for customers already using the wallet system. In this case the customer would only have be registered at CD@Across and no assets would have to be transferred over the network.

Transfer to an operator

It is not possible to be connected to an operator in my implementation. The reason to this is obvious; CD@Across is not a commercial system. Across Wireless AB does not sell any CD albums. The SpeechWeb system that PipeBeach AB has let me use does not permit transfers of telephone calls.

The dialogue

At the moment the dialogue is static in the sense that it does not care if a customer is used to the system and might want to shorten the dialogue. It might be desirable to implement some kind of feature so that a customer that is familiar to the system can navigate through it in a faster way. This can be done in several different ways. One example is to present a choice to the customer in the beginning of the conversation. If the customer says that he is experienced the dialogue is more compact than it would be for an inexperienced customer.

Another solution could be to store information about the number of visits the customer has done to CD@Across. After a certain number of successful purchases the customer would be considered experienced and the dialogue would be more compact. Information about previous purchases could be stored in a database at the merchant.

In my implementation I use inline grammar that consists of option lists. Each time a customer enters a dialogue that uses the grammar the option list has to be created through database calls. If the database would contain more artists and albums it would probably be more effective to use a grammar file that contains the grammar. A new grammar file would be created when the last album by an artist is sold or when new albums are entered into the database.

What will the future bring?

High data rates

When the transmission rates increases new services can be added to CD@Across. It is difficult to predict what the future cellular telephones will look like and how they will be used. Some people believe that all the different mobile equipment (Walkman, cellular phone, Palm Pilot etc.) will be combined in one Personal Digital Assistant (PDA). Others believe that we will continue to carry several different types of equipment but that they will communicate with each other.

The type of services that can be added due to higher transmission rates depends on how the future cellular telephones will function. A visionary view is that the cellular phone is combined with a MP3 player and that the customer downloads the music to the phone after the purchase. The customer would not have to wait for the CD album to be delivered by mail.

Voice Verification

Several companies develop equipment for voice verification over the phone. Voice verification shortens the call time since the time to say a pass phrase is shorter than the time to enter a set of characters using the keypad. A PIN can be mistyped or forgotten while the chance of someone stating the wrong pass phrase is very small.

Voice verification can make the interface towards the customer user-friendlier than it is today. The computer could, after some small talk or after a pass phrase, determine if the customer already have an account or not. If the voice is not recognized the computer can tell the new customer that an account have to be created and what the legal terms are.

Wireless Public Key Infrastructure (WPKI)

Many people believe that the solution to the security weaknesses on the Internet is a Public Key Infrastructure (PKI). WPIK is an implementation of PKI where the certificates are smaller in order to save bandwidth. The private key is stored on the SIM card in the cellular telephones and the certificates are stored in a directory server in the network. More information about PKI and WPKI can be found at [13].

The introduction of WPKI makes it possible to create an infrastructure where all customers can purchase goods from all merchants without having to trust the e-commerce gateway at the operator.

The tasks for the e-commerce gateway can then be changed so that it only relays messages between merchants and customers. The communication with the payment provider can be handled directly by the merchants.

13. Conclusion

During my work with this master thesis I have shown that it is possible to create a secure m-commerce service that combines voice and data transmissions. Voice is used in order to create a user-friendly interface, an interface that does not demand a good vision and narrow fingers. The data transmissions are used for security purposes in order to be certain that the customer is the person he claims to be.

The user does not have to enter text or to scroll WML pages using the small inconvenient keypad on the cellular telephone. Some fumbling with the keypad is unavoidable since the security is based on accepting and signing a payment request sent by the merchant via the e-commerce gateway. These actions are optimized to be simple so they hopefully won't be too difficult to perform.

My implementation is based on a voice-browsing server named SpeechWeb. This server reads voice XML documents to the caller. The use of such a system makes the implementation easy and flexible. Services can be created using standard tools and published on ordinary web servers.

The SpeechWeb system that I have been using has been a test and development system. This in combination with a non-perfect dialogue makes my application a little bit unstable. In order to make CD@Across a commercial system the dialogue has to be a bit more worked through and Across Wireless would need to buy their own SpeechWeb. Note that Across Wireless does not intend to make the application commercial; it is only used for demonstration purposes.

In addition to creating the voice interface I also investigated if and how alternative data services could be used in combination with CD@Across and how GPRS/UMTS would effect the application.

My investigation of alternative technologies for GSM data transfer showed that CD@Across works well with the technologies already used and that there is no reason to change the application.

14. List of references

World Wide Web documents

1. Choi S-Y & Whinston A, *Smart Cards, Enabling Smart Commerce in the Digital Age*, Center for Research in Electronic Comers, The University of Texas at Austin 1998, http://cism.bus.utexas.edu/works/articles/smartcardswp.html, Accessed 2000-09-15
2. *PKI smart cards*, Whitepaper, Id2Tech AB, http://www.id2tech.com/whitepapers/smartcards.asp, Accessed 2000-09-15
3. *Smart Cards and Security Overview*, http://www.smartcardbasics.com/typesofchips.html, Accessed 2000-09-15
4. Everett D B, *Smart Card Technology*, Smart Card News Ltd., Brighton, England 1997, http://www.lkv.customs.ru/sat/sat.xpress.ru/SmartCard/ISO7816-1.htm, Accessed 2000-09-15
5. *Quick GSM Statistics*, Mobile Office On-Line June 1999, http://www.mobileoffice.co.za/gsm-stats.htm , Accessed 2000-09-15
6. http://www.wirelesstoolkit.com/, Accessed 2000-09-04
7. http://www.wapforum.com, Accessed 2000-09-04
8. *Voice recognition Systems*, AbilityNet August 2000, http://www.abilitynet.co.uk/fullvis/alt_tech/voice/voct99.htm, Accessed 2000-09-19
9. Harris M, *IVR vs. the Human Touch*, Article, Call Centre Selection 1999, http://www.callcentre-selection.co.uk/ivr.html, Accessed 2000-09-19
10. Collin N, *Automated Speech Recognition*, http://www.ncollin.demon.co.uk/speechrecognition.html, Accessed 2000-09-19.
11. Alwang G, *L & H Voice Xpress Professional 4.*, Article, PC magazine July 1999, http://www.zdnet.com/pcmag/stories/firstlooks/0,6763,408623,00.html, Accessed 2000-09-20
12. *E-services – Facts and Figures*, SmartTrust 2000, http://www.smarttrust.com/e-services/services.html, Accessed 2000-09-25
13. *Public Key Infrastructure*, SmartTrust 2000, http://www.smarttrust.com/pki/index.html, Accessed 2000-11-17
14. *So far, M-commerce has flopped in US*, Newsbytes August 2000, http://www.telekomnet.com/news/8-17-00_mcommerce_flopus.asp, Accessed 2000-09-25
15. http://www.mercurycenter.com/svtech/news/breaking/reuters/docs/353855l.htm, Accessed 2000-09-25
16. *Moving towards EDGE*, Ericsson 1999, http://www.ericsson.se/wireless/products/mobsys/gsm/subpages/umts_and_3g/edge.shtml, Accessed 2000-09-27
17. *Enhanced Data rates for GSM Evolution (EDGE)*, http://www.mobilepositioning.com/edge.htm, Accessed 2000-09-27
18. http://www.pipebeach.com, Accessed 2000-10-23
19. http://www.voicexml.org, Accessed 2000-10-23
20. *GSM Info Online***,** January 1, 2001, http://www.gsmworld.com/gsminfo/gsminfo.htm, Accessed 2001-01-04
21. *Why SMS if we have GPRS?*, Logica 1999, http://www.gsmworld.com/presentations/gprs/logica.pdf, Accessed 2000-10-24
22. http://www.fundamo.com/, Accessed 2000-11-15
23. http://www.televend.com/, Accessed 2000-11-15
24. http://www.smartaxis.com/, Accessed 2000-11-15
25. httP://www.jalda.com/, Accessed 2000-10-15
26. http://www.voicegenie.com/, Accessed 2000-10-16
27. http://www.tellme.com, Accessed 2000-11-16
28. http://www.shoptalk.com/, Accessed 2000-11-16
29. http://www.nuance.com/index.htma, Accessed 2000-11-16
30. http://xml.apache.org/xalan/index.html, Accessed 2000-12-05
31. http://xml.apache.org/cocoon/, Accessed 2000-12-05

Specifications

32. *ISO/IEC 7816, Identification cards - Integrated circuit(s) cards with contacts,* ISO1998.
33. *GSM 2.17, Digital cellular telecommunications system (Phase 2+); SIM functional characteristics,* v.8.0.0, ETSI 2000.
34. *GSM 11.14, Digital cellular telecommunications system (Phase 2+); Specification of the SIM-ME interface,* v. 8.1.0, ETSI 1999.
35. *GSM 03.48, Digital cellular telecommunications system (Phase 2+); Security Mechanisms for the SIM Application Toolkit – Stage 2,* v. 8.3.0, ETSI 1999.
36. *3G TS 21.101, 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; 3rd Generation mobile system Release 1999 Specifications,* V3.1.0, 3GPP 2000.
37. *3G TS 22.038, 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; USIM/SIM Application Toolkit (USAT/SAT); Service description; Stage 1,* v. 3.2.0, 3GPP 2000.
38. *3G TS 22.100, 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; UMTS phase 1,* v.3.6.0, 3GPP 2000.
39. *GSM 02.09, Digital cellular telecommunications system (Phase 2+); Security Aspects,* v.8.0.0, ETSI 2000.
40. *GSM 03.20, Digital cellular telecommunications system (Phase 2+); Security related network functions,* v.8.0.0, ETSI 1999.
41. *GSM 01.01, Digital cellular telecommunications system (Phase 2+); General description of a GSM PLMN,* version 5.0.0, ETSI 1996.
42. *GSM 03.40, Digital cellular telecommunications system (Phase 2+); Technical realization of the SMS PP,* v.7.2.0, ETSI 1998.
43. *GSM 03.38, Digital cellular telecommunications system (Phase 2+); Alphabets and language-specific information,* version 5.6.1, ETSI 1998.
44. *GSM 11.11, Digital cellular telecommunications system (Phase 2+); Specification of the SIM – ME interface,* version 8.1.0, ETSI 1999.
45. *GSM 02.60 Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS), Service description; stage 1,* version 5.2.0, ETSI 1998.
46. *GSM 02.90, Digital cellular telecommunications system (Phase 2+); Unstructured Supplementary Service Data (USSD) - Stage 1,* v. 7.0.0, ETSI 1999.
47. *GSM 03.90, Digital cellular telecommunications system (Phase 2); USSD Stage 2,* ETSI 1996.
48. *Voice eXtensible Markup Language,* version 1.00, VoiceXML Forum 2000.
49. *Speech Recognition Grammar Specification,* working draft, W3C 2000.

Publications

50. Ford W, Computer Communications Security, Prentice Hall 1994.
51. Pfleeger C, *Security in computing,* Prentice Hall 1997.
52. *GSM System Overview,* APIS Technical Training 1998.
53. *GPRS Overview,* APIS Training & Seminars 1999.
54. Törnroth J, *Wireless Wallet,* Master Thesis, Royal Institute of Technology 2000.
55. Stephens W R, *TCP/IP Illustrated, Volume 1: The Protocols,* Addison-Wesley 1994.

Internal documents

56. Sellin A, *Product Specification –WIG,* 17390007, Across Wireless AB 1999-02-17.
57. Boëtius J, Specification – Notification Interface, Across Wireless AB 1999.
58. Törnroth J, *Specification – Payment Interface,* Across Wireless AB 1999.

15. Appendix A – List of abbreviations

Abbreviation	*Meaning*
8-PSK	8 Phase Shift Keying
ABNF	Augmented Backus-Naur Form
ASR	Automatic Speech Recognition
AUC	Authentication Center
BSS	Base Station System
CSD	Circuit Switched Data
CTI	Computer Telephone Integration
DES	Data Encryption Standard
DNS	Domain Name Server
EDGE	Enhanced Data rates for GSM Evolution
EIR	Equipment Identity Register
ETSI	European Telecommunications Standard Institute
GGSN	Gateway GPRS Support Node
GMSK	Gaussian Minimum Shift Keying
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
HSCSD	High Speed Circuit Switched Data
HTML	Hyper Text Markup Language
IMEI	International Mobile Equipment Identifier
IMSI	International Mobile Subscriber Identity
ISO	International Standards Organization
IVR	Interactive Voice Response
Kc	Cipher Key in GSM security
Ki	Internal Key in GSM security
LAI	Location Area Identifier
ME	Mobile Equipment
OTA	Over The Air
PDA	Personal Digital Assistant
PIN	Personal Identification Number
RSA	Rivest-Shamir-Adleman
SAT	SIM Application Toolkit
SC	Service Center
SGSN	Serving GPRS Support Node
SIM	Subscriber Identity Module
SM	Short Message
SMS	Short Message Service
SMS-CB	Short Message Service – Cell Broadcast
SMS-MO	Short Message Service – Mobile Originated
SMS-MT	Short Message Service – Mobile Terminated
SMS-PP	Short Message Service – Point-to-Point
STK	SIM Tool Kit
TDMA	Time Division Multiple Access
TMSI	Temporary Mobile subscriber identity
UMTS	Universal Mobile Telecommunications Systems
URL	Unified Resource Locator
USSD	Unstructured Supplementary Services Data
VXML	Voice eXtensive Markup Language
WAP	Wireless Application Protocol
WIB	Wireless Internet Browser
WIG	Wireless Internet Gateway
WPKI	Wireless Public Key Infrastructure

16. Appendix B – An example dialogue between a customer and the merchant during a purchase

Customer	Merchant
Calls the merchant.	
	Welcome to CD@Across! *Say or enter your username! Say help if you want more information.*
Help!	
	Say one digit at a time, or use the keypad on the phone, followed by hash! *If you are not welcomed by your name, you should, quit.* *You can always quit by saying, "quit", or be connected to an operator by saying, "operator". If you do not understand what I am saying, or asking, you can say, "help"!*
One, two, three!	
	I heard one, two, three! *Welcome, Fredrik Oljeqvist, to the main menu!* *Say one of: "List", "Top Five", "Artist" or "Random".* *Say "help" if you want to know more!*
Artist!	
	I heard "Artist"! *Say the name of the artist you are looking for, or say "menu"!*
Madonna!	
	I heard "Madonna". Is this correct?
Yes!	
	We have the following albums by Madonna: *American Pie, The price is $6.* *Ray of Light, The price is $4.* *Say the name of the album that you are interested in or say "menu" to get back to the main menu!*
Ray of Light!	
	Say "Listen", "Buy" or "Go Back"!
Listen!	
	Merchant plays a part of a song on the album. *Say "Listen", "Buy" or "Go Back"!*
Buy!	
	I heard "Buy!". Is this correct?
Yes!	
	A short message will be sent to your cellular telephone *Please hold on!*
Customer receives short message containing information about the purchase. The customer chooses payment method, the request is signed with a PIN and sent to the e-commerce gateway.	
	Your request has been handled by our system, The result is: Payment Cleared! *Say "menu" or "quit"!*
Quit!	
	Thank you for visiting CD at Across and welcome back! Connection terminated.

17. Appendix C – Test results

No	Success	Comment	Total time for purchase	Time to receive payment request	Time to send payment reply	Type of telephone	Date
1	Yes		2.04	24	9	Stationary	14/12-00
2	Yes		1.46	22	13	Stationary	14/12 -00
3	Yes		1.37	14	11	Stationary	14/12-00
4	Yes	Disturbed by background noise	1.44	19	11	Stationary	14/12-00
5	Yes	SW had problems understanding	2.40	28	14	Mobile	14/12-00
6	Yes	Background noises – quit during waiting for result. Got album.	1.54	30	?	Mobile	14/12-00
7	No	Background noise - quit				Mobile	14/12-00
8	Yes		2.03	27	13	Mobile	14/12-00
SpeechWeb was upgraded during the afternoon. The upgrade should eliminate some problems with understanding.							
9	Yes		2.27	26	10	Stationary	18/12-00
10	Yes		1.58	21	14	Stationary	18/12-00
11	No	Spoke outside prompt				Stationary	18/12-00
12	Yes		1.48	17	11	Stationary	18/12-00
13	Yes		1.56	26	14	Mobile	18/12-00
14	Yes		2.03	28	10	Mobile	18/12-00
15	No	Cough made call terminate				Mobile	18/12-00
16	No	Rubbed my stubble against microphone – terminated call				Mobile	18/12-00
17	Yes		2.25	21	14	Mobile	18/12-00
18	Yes	Disturbed by background noise	3.13	29	14	Mobile	18/12-00
19	Yes	Background noise - quit	2.02	17	10	Stationary	18/12-00
20	Yes	Disturbed by background noise	3.05	19	10	Stationary	18/12-00
21	No	Background noise made call terminate				Stationary	18/12-00
22	Yes		2.20	17	11	Stationary	18/12-00
23	Yes		1.57	17	10	Stationary	18/12-00
24	Yes		2.07	22	10	Mobile	18/12-00
25	Yes		3.24	1.10	14	Mobile	18/12-00
26	Yes		2.05	32	13	Mobile	18/12-00
27	Yes		2.23	19	13	Mobile	18/12-00
28	No	Background noise made call terminate				Stationary	18/12-00
Lots of background noise today. People were talking and laughing quite loud. This affected the results. There is a bug in SpeechWeb that makes the call terminate when the customer tries to follow a link. Some of the background noises were interpreted as “quit” or “operator”. This made SpeechWeb terminate call.							
29	Yes	SW had problems understanding “yes”	1.59	29	14	Mobile	19/12-00
30	Yes	SW had problems understanding “yes”	2.15	26	14	Mobile	19/12-00
31	Yes		2.07	27	13	Mobile	19/12-00
32	Yes		1.55	22	15	Mobile	19/12-00
33	No	Connection broken due to the link bug in SpeechWeb				Mobile	19/12-00
34	Yes		2.27	23	11	Stationary	19/12-00
35	Yes		2.08	18	9	Stationary	19/12-00
36	Yes		2.34	23	10	Stationary	19/12-00
37	Yes		2.08	14	11	Stationary	19/12-00
38	Yes		2.10	22	11	Stationary	19/12-00

39	Yes		2.34	21	13	Mobile	19/12-00
40	Yes		2.08	28	14	Mobile	19/12-00
41	No	Connection broken due to the link bug in SpeechWeb				Mobile	19/12-00
42	Yes	SW re-asked	2.54	28	13	Mobile	19/12-00
43	Yes		1.59	22	14	Mobile	19/12-00
44	Yes		1.54	20	11	Stationary	19/12-00
45	Yes		2.04	21	10	Stationary	19/12-00
46	Yes		2.06	20	9	Stationary	19/12-00
47	Yes	SW re-asked	2.20	15	26	Stationary	19/12-00
48	Yes		2.10	16	9	Stationary	19/12-00
No disturbing background noises today but I had the same problems with links as yesterday. The broken connections were caused by me talking unclearly. SpeechWeb interpreted my unclear utterances as “quit” or “operator”, which are global links. When SpeechWeb tried to follow these links the call was terminated. I have reported the link bug to PipeBeach AB and they will try to fix it tomorrow.							
49	Yes		2.19	16	12	Mobile	20/12-00
50	Yes		1.56	13	14	Mobile	20/12-00
51	Yes	SW re-asked	2.50	17	59	Mobile	20/12-00
52	Yes		2.23	31	23	Mobile	20/12-00
53	Yes	Background noises – quit during waiting for result. Got album.	2.00	37	?	Mobile	20/12-00
54	Yes		2.12	20	11	Stationary	20/12-00
55	Yes		2.01	21	10	Stationary	20/12-00
56	Yes		1.50	16	11	Stationary	20/12-00
57	Yes		1.55	16	11	Stationary	20/12-00
58	Yes	Background noises – quit during waiting for result. Got album.	1.42	16	?	Stationary	20/12-00
59	No	SW misunderstood - operator				Mobile	20/12-00
60	Yes	Background noises – quit during waiting for result. Got album.	2.00	17	?	Mobile	20/12-00
61	Yes		1.53	30	13	Mobile	20/12-00
62	Yes		1.54	22	15	Mobile	20/12-00
63	Yes		2.00	25	13	Mobile	20/12-00
64	Yes		2.14	23	13	Stationary	20/12-00
65	Yes		1.55	21	10	Stationary	20/12-00
66	No	Background noise - quit				Stationary	20/12-00
67	Yes		2.35	16	11	Stationary	20/12-00
68	Yes		1.49	15	12	Stationary	20/12-00
There were not so much background noises today. Some talking, coughing and laughing. The background noises did not spoil any purchases. I got what I was looking for even if the dialogue was closed two times without notifying me of the result.							
69	Yes		2.05	27	13	Mobile	21/12-00
70	Yes	SW re-asked	2.12	25	12	Mobile	21/12-00
71	Yes		2.09	29	13	Mobile	21/12-00
72	Yes		2.27	27	12	Mobile	21/12-00
73	Yes		2.15	27	14	Mobile	21/12-00
74	Yes		2.14	29	11	Stationary	21/12-00
75	Yes	Cough– quit during waiting for result. Got album.	1.22	13	11	Stationary	21/12-00
76	Yes		1.47	20	10	Stationary	21/12-00
77	Yes		1.57	22	10	Stationary	21/12-00
78	Yes		1.53	22	10	Stationary	21/12-00
79	Yes		1.55	22	13	Mobile	21/12-00
80	Yes	SW re-asked	2.10	25	14	Mobile	21/12-00
81	Yes		2.06	28	14	Mobile	21/12-00
82	No	I entered wrong payment method	2.52	28	?	Mobile	21/12-00

83	Yes		1.59	26	12	Mobile	21/12-00
84	Yes		1.56	21	12	Stationary	21/12-00
85	Yes	SW re-asked	2.12	24	11	Stationary	21/12-00
86	Yes		1.57	21	11	Stationary	21/12-00
87	Yes		1.57	20	11	Stationary	21/12-00
88	Yes		2.00	22	11	Stationary	21/12-00
Application worked fine today. Some problems with SW not understanding what I said and therefore re-asked.							
89	Yes		2.10	28	14	Mobile	22/12-00
90	Yes		2.10	26	13	Mobile	22/12-00
91	Yes		2.00	23	14	Mobile	22/12-00
92	Yes		2.08	28	14	Mobile	22/12-00
93	Yes		2.12	27	13	Mobile	22/12-00
94	Yes		2.02	24	14	Stationary	22/12-00
95	Yes	Laughed– quit during waiting for result. Got album.	1.42	25	?	Stationary	22/12-00
96	Yes		2.03	21	13	Stationary	22/12-00
97	No	Disturbed by background noise - quit				Stationary	22/12-00
98	Yes		2.02	22	10	Stationary	22/12-00
99	Yes		1.55	29	15	Mobile	22/12-00
100	Yes	Background noises– quit during waiting for result. Got album.	1.30	20	?	Stationary	22/12-00
Today there were some disturbing background noises. Christmas songs were played loudly and people were happy and joyful due to the upcoming Christmas vacation.							