Technical Report

UCAM-CL-TR-372
ISSN 1476-2986

Number 372

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

A brief history of mobile telephony

Stefan G. Hild

January 1995

15 JJ Thomson Avenue
Cambridge CB3 0FD
United Kingdom
phone +44 1223 763500

http://www.cl.cam.ac.uk/

Technical reports published by the University of Cambridge Computer Laboratory are freely available via the Internet:

http://www.cl.cam.ac.uk/techreports/

ISSN 1476-2986

A Brief History of Mobile Telephony

Stefan G. Hild
University of Cambridge Computer Laboratory
New Museums Site
Pembroke Street
Cambridge CB2 3QG
Stefan.Hild@cl.cam.ac.uk

January 1995

Abstract

Mobile telephony has gone through a decade of tremendous change and progress. Today, mobile phones are an indispensable tool for many professionals, and have great potential to become vital components in mobile data communication applications. In this survey we will attempt to present some of the milestones from the route which mobile telephony has taken over the past decades while developing from an experimental system with limited capabilities to a mature technology (Section 1), followed by a more detailed introduction into the modern pan-European GSM standard (Section 2). Section 3 is devoted to the data communication services, covering two packet-oriented data only networks as well as the data services planned for the GSM system. Section 4 covers some security issues, and section 5 gives an insight into the realities today with details of some networks available in the UK. Finally, section 6 concludes this overview with a brief look into the future.

1 Humble beginnings

Wireless communication had a modest start on October 23, 1915, when AT&T engineers managed to transmit radio-telephone signals from Arlington, Virginia to both Paris in France and Pearl Harbour, Hawaii. The technology went into commercial service shortly after World War I, relaying 30 calls per day at a cost of $75 for a three minute connection [DDF92].

The first significant use of radio-based mobile communication system can be traced back as far as 1921, when the Detroit police department installed 2MHz-radio transmitters in their police cars [PG89]. Even though this early system had nothing in common with telephony as we know it today, the benefits of mobile communications were quickly discovered and were only hampered by the lack of channels available in this low-frequency band. Progressively higher frequencies were introduced in subsequent years, opening up more channels.

The lack of channels posed the most serious problem, prohibiting more user-friendly full-duplex systems rather than the "push-to-talk" approach. In addition, it was not possible to make calls into the public telephone network.

Shortly after World War II Bell Systems proposed a *Public Correspondence System* which would remedy at least the last restriction. It eventually went into service in St. Louis in 1946 [You79], operated at 150MHz and provided three speech channels spaced at 120kHz. Calls could be made into the public telephone network but had to be set-up manually. Sill, half-duplex connections where used.

Extending conventional radio technology to allow for a large user community and full-duplex connection would have required far too much bandwidth: for full-duplex the bandwidth requirement doubles; each call requires its own channel allocation, thus n calls will requires n times that bandwidth. It is clear that even a small user community would use up all the bandwidth available very quickly.

The problem was solved by splitting up the area into small cells (see figure 1). Each cell is served by a base-station and allows a fixed maximum number of connections at any one time. The frequencies used in any cell can be re-used in other cells at some safe distance. Although there is still a maximum

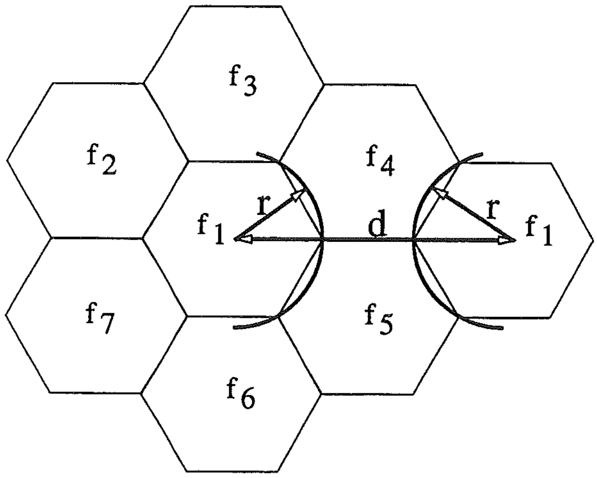

Figure 1: Typical cellular schema. 7 sets of frequencies are sufficient to cover an arbitrarily large area, providing that the repeat-distance d is larger that twice the maximum radius r covered by each individual transmitter.

number of connections each basestation can support in one cell, cells can be made very small in areas of high demand. This largely solved the problem of bandwidth, but created a system that is highly complicated to manage and operate. To name only two of the most pressing problems: Firstly, mobile users may move out of one cell and into another and the system has to cope with this (through a process called *hand-off* and *hand-over*), secondly, callers can not be expected to know the exact location of the mobile phone, so the system has to be able to keep track of the mobile users (through a process called *locating*). The design of cells itself proved to be immensely complex: natural landscape and tall buildings influence radio propagation and the selection of suitable cell-sites requires much planning as well as a fair amount of practical tests [JL90]. Although the concept wasn't really put into practise before 1978 in the *Advance Mobile Phone System system*, the cellular concept itself was conceived as early as 1947 by D. H. Ring of Bell Laboratories in an unpublished work [You79].

1964 and 1969 saw the arrival of two new systems in the USA, called *MJ* and *MK*, respectively. The first operated in the 150MHz range, the latter close to 450MHz. Both systems employed a set of transmitters operating at slightly different frequencies as in the cellular-setup detailed above; they did not, however, make provision for hand-offs and were not able to locate mobile units. Thus, even though calls could be set up automatically, the locations of the mobile unit had to be know in advance [You79].

At the same time similar system were developed in Germany, which were largely hampered by the same restrictions. With the sort of imagination and creativity each German appears to be born with, the German systems were called the *A*-system and the *B*-system (of which two varieties were in operation, duly called *B1* and *B2*).

1.1 First real mobile telephony

First proposals for a 'proper' cellular telephone system which employed the cellular concept including *locating* and *handoffs* where made in the USA by the Bell Laboratories in the late 1960s. Two systems subsequently went into field trials: AMPS (Advanced Mobile Telephone System) in Chicago and the ARTS (American Radio Telephone Service) in Washington DC. The latter was designed to support handheld equipment as well as car telephones. However, it was not before 1983 that those systems went into commercial service [PG89].

The cradle of modern mobile telephony stood far away from the traditionally strong high-tech countries - in northern Europe. The Scandinavian countries where among the first to recognise the potential of radio based mobile communication: this was the only technique that made if feasible to establish a communication infrastructure in its thinly populated areas. The first small mobile telephone system was in public operation in Sweden as early as 1978 [Lam84]. The system really took off three years later, when the *Nordic Mobile Telephony System (NMT)* went into public service in 1981. Until the mid-1990s the original NMT and its successor system were the most widespread systems, being in operation not only in Sweden, Norway, Finland and Denmark, but also in Spain, Tunisia, Saudi Arabia, the Netherlands, Austria and Ireland. Needless to say, it also had, by far, the biggest user community.

Standards started to evolve in the early 1980s. However, the markets where exclusively national; in 1978, W. Young of Bell Laboratories remarked on the issue of widespread availability: *"Neither the characteristic nor nationwide compatibility necessarily implies universal coverage."*

As a consequence, many incompatible standards evolved in different countries, and by and large the phones sold in one country could not be operated in another country. Where several systems were operational in the same country, it was even possible that some phones worked in one city but did not work in another city. Some of the more interesting systems are introduced in some detail in the following subsections, followed by a table giving a summary of key technical details (figure 2).

1.1.1 NMT450 and NMT900

As mentioned above, the Nordic Mobile Telephony System NMT started public service in 1981 and therefore holds the record of being the longest established mobile phone system. It was developed by Ericsson Radio Systems which is still a very big player in mobile communication systems today.

The first system, known as NMT450, operated in the 450 to 470MHz band. The original channel-allocation was initially thought to be very generous, but the bandwidth available filled up more rapidly than expected and the successor system, NMT900, was assigned an even greater bandwidth in the 900MHz-band. Each cell is serviced by a base station and a number of those is serviced by a mobile switching centre which provides access to the public switched telephone network [Lam84].

The NMT system was hugely successful. Not only did it achieve high market penetration (in 1991, 6% of all Swedes had a mobile phone and the annual growth rate was put at 1.5% [HL91]), but it also offered wide coverage and even international roaming in the Nordic countries. 10 years after its introduction, the inventors triumphantly remarked [HL91]: *"Even in the USA, roaming and call-delivery features are implemented only at a local level."*

1.1.2 ARTS, AMPS and TACS

The technologies of both the ARTS and the AMPS system where eventually merged by the FCC into one standard. The resulting system became available to the public in 1983. The system was less successful than its counterpart in Sweden and suffered under larger cells with greater overlap. Heavy call charges and equipment prices between $2400 and $2700 in 1984 did not help the situation.

Despite the problems with AMPS the British Government decided to ignore the more successful and widespread NMT effort and developed a derivative of AMPS for the British environment. The system was called the *Total Access Communications System* or *TACS* for short, and was licenced to Racal Millicom (*Vodafone*) and Telecom-Securior Cellular Radio (*Cellnet*). Unlike AMPS, TACS was to operate in the 860-960MHz band and to use a channel spacing of 25kHz. TACS was later extended to make use of additionally released frequencies (*ETACS*) [Lam84].

1.1.3 C-900

The German *C*-system (a continuation of the naming tradition started with the A- and B-systems) was developed in the early 1980s and was the first system worldwide to use digital signalling. The speech channel still operated in analogue mode, but was encrypted. At its time the C-900 system was the technologically most advanced and most complex system available. It was the sole effort of Siemens and, at the time it went into service in 1984, the system covered 70% of West-Germany through 95 base stations. Like the other systems, it was later expanded to operate in the 900MHz-band, but at first was restricted to the 450MHz-band. One of the most advance features was its ability to manage cell boundaries dynamically by varying the output-power of adjacent cells. By doing so cells could adjust to varying demand and traffic load. The output power of the mobile unit could be reduced to the minimum required for a clear connection, thus reducing interference elsewhere in the network.

Channel allocation was extremely efficient. On the down side, the extremely complex system architecture resulted in Siemens being the sole supplier of C-phones and kept unit prices and call charges high. Nevertheless, the system replaced both the early A-system and the B1-system, and in December 1994 the B2-system was finally taken out of operation. In early 1995, the C-system came close to 1 million subscribers, reaching the limit of the system's capabilities [Kai94], [Lam84].

2 Move towards a world standard - GSM

Even though different systems were introduced in different countries, the nasty habit of radio waves to cross country borders without any respect necessitated some degree of collaboration fairly early on. The frequencies used for mobile telephony in Europe where agreed upon at the *World Administrative Radio Conference* in 1979. At the same conference, it was decided to set up a research group that would define a common standard for the future. Such a standard, it was thought, would have many advantages for both the network operator and the

System	Intro.	Frequency (MHz)	Channel spacing (kHz)	Duplex spacing (MHz)	Signalling channel	Scrambling	Coverage
NMT450, NMT900	1981	450-470, 860-960	25	10, 45	combined	no	Scandinavia, ...
AMPS, ARTS (FCC)	1983	825-890	30	45	separate	no	USA
C	1984	450, 900			both	yes	Germany
TACS, ETACS	1985	860-960	25	45	separate	no	UK

Figure 2: Some cellular standards

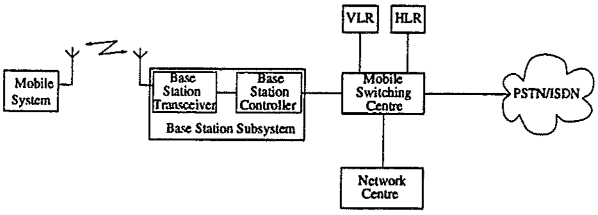

Figure 3: GSM network overview

customer: firstly, standardisation throughout large economic entities means that economies of scale can be reached that will lead to a sharp decrease in unit-prices which will in turn lead to more demand and a higher profitability of the networks. The end-user does not only benefit from reduced equipment costs, but will also be able to use the equipment wherever the standard has been adopted.

The first practical steps to this goal were done in 1982 under the watchful eyes of the *Council of European Posts and Telecommunications (CEPT)* which set up the *Group Speciale Mobile*, or *GSM* for short. In order to please the English speaker the initials GSM where later re-assigned to stand for *Global System of Mobile communications* and the system developed is now colloquially known as the GSM-system.

2.1 Network Architecture

The design group undertook the difficult task of specifying the interfaces and technologies well enough to enable equipment supplied by different manufacturers to work together, while leaving at the same time enough scope for design options. The basic elements defined by GSM are the *mobile stations (MS)* comprising the actual mobile phone, the *base stations (BS)* which in turn consist of one or more actual *base station transceivers (BTS)*, a *base station controller (BSC)* which logically administers the BTSes, and the *mobile switching centres (MSC)*. All the elements can be seen in figure 3. GSM describes the functions of each element and the protocol to be used between them. Of particular interest is the radio-interface used between the mobile station (i.e. the mobile phone) and the base station. GSM also goes into great depth in defining exact standards for protocols between the base station and the mobile switching centre and for the protocol between the different mobile switching centres in the network.

Each mobile station is recognised by the network by its unique *International Mobile Station Identity (IMSI)*. *Subscriber Identity Modules (SIM)* are used to personalise the phone; typically, SIMs have the same size and shape as conventional credit cards but have an in-build chip which stores the owner's identity encoded in the *SIM-ID*-number, the mobile phone number and the list of services the user is subscribed to.

When calls are made from the mobile (so-called *Mobile Originated Calls MOCs*) the phone transmits its IMSI, the number dialled, and the SIM-ID. The base station authenticates first the actual phone, then the subscriber and finally checks if the subscriber is entitled to receive the requested service. In this process the system assigns encryption keys for the duration of the connection and the connection is then established.

Obviously, the process of connecting a call to a mobile unit (a so-called *Mobile Terminated Call MTC*) is somewhat more complicated. To keep

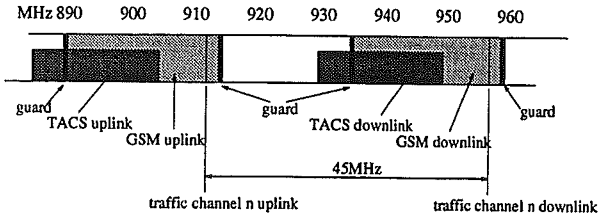

Figure 4: Frequency band allocated to GSM in the UK.

track of moving mobile stations the network maintains a database containing the *Home Location Register (HLR)* and the *Visited Location Register (VLR)* for each SIM-ID. The HLR is static and it records the home-address for each SIM. Each home-location is responsible for the network administration and billing for its subscribers. This also allows calls to be billed to the national operator of foreign phones. On the contrary, the VLR stores the current location of all SIMs. The VLR is updated regularly by the base stations, which interrogate all the mobiles in their cells at regular intervals. Mobile Terminated Calls are routed by the network by firstly identifying the SIM-ID from the requested telephone number, and secondly examining the VLR for that ID. The same checks as for mobile originated calls are then made and the call is finally connected.

2.2 The Radio Interface

The frequency band available to GSM is split into two halves: the band between 890-915MHz is used for the uplink, ie. for the direction mobile station to base station; the band between 935-960MHz is used for the downlink in the opposite direction. However, this overlaps with the frequency band used for other mobile systems, so only a part of the band is currently in use. In the UK, the TACS systems in operation occupy the lower 15MHz of both bands and therefore only the upper 10MHz are currently used for GSM (see figure 4).

A guard band of 200kHz protects the frequency bands on either end. With the 200kHz carrier spacing used in GSM, this leaves 124 possible channels. Up- and downlink are always separated by 45MHz.

GSM uses *Gaussian minimum shift key modulation (GMSK)*, which provides a spectral efficiency of 1.35 bps/Hz [CMS93]. This amounts to a bit rate of 270.838kbits/s. Time division multiplexing is used to assign 8 logical channels onto each physical carrier channel. The physical channel is split into time slots of roughly 577μs duration, which at the bit rate used corresponds to 156.25 bits. These time slots are also referred to as bursts. A logical channel is therefore defined by its frequency and the TDMA frame time slot number.

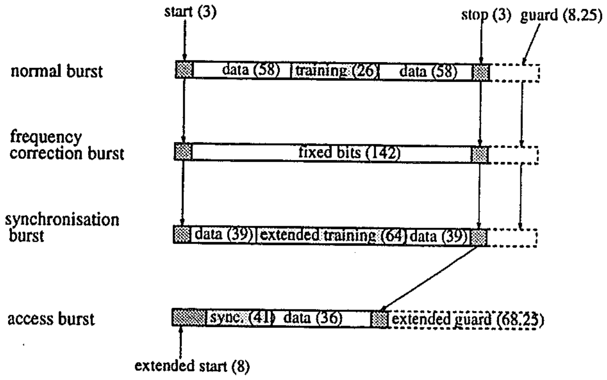

Figure 5: Burst types for GSM.

Different burst-types have been defined for GSM [Hod90]. All bursts have a sequence of start bits (mostly 3), followed by a variable number of data bits and finally a guard-period which is primarily used for synchronisation purposes. Since the different slots in the time-division schema are filled by different mobile stations, synchronisation can only be guaranteed by introducing this idle-period: Each mobile station injects a single burst into the TDMA-structure, a process that has to be both accurate and timely otherwise the integrity of the TDMA structures is at risk. The mobile phones have to perform this largely blindly, since output power is reduced to the minimum required to maintain connection to the base station and therefore more remote mobile phones are, in general, unable to 'listen' to the transmissions of other mobile phones and to time their transmissions accordingly. The base station monitors timeliness for each mobile and sends frequency correction burst and synchronisation burst to mobiles, if required.

The exact specification for the different burst-types can be seen in figure 5.

One TDMA frame is made up of eight bursts or 4.616ms. The frames are grouped together into multiframes of either 120ms duration or 235.4ms duration. The first comprises 26 TDMA frames, the latter 51. Handoffs always coincide with the end of a multiframe. Superframes are then comprised of either 51 or 26 multiframes, given a duration of 6.12s. The mobile uses the duration of a superframe to form an average of the detected signal strengths.

A further level is added in the form of hyperframes, consisting of 2048 superframe structures. A hyperframe has a duration of 3 hours 28min 53s 760ms. This long frame is used to support the encryption algorithm used for the encrypted data in the time slots. Figure 6 gives an outline of the framing structures.

8-slot TDMA together with the 248 physical half-duplex channels currently available in the UK give a total of 1984 logical half-duplex channels, or roughly 283 per cell. However, only a part of those are used as traffic-channels to convey speech. In addition GSM defines a great number of control channels which allow the system to engage in a highly active exchange of data.

Two forms have been specified for the traffic channel: the *full rate traffic channel*, which carries user data and speech at a gross rate of 22kbits/s, and the *half rate traffic channel* which will eventually take over speech transmission if suitable encoding algorithms become available.

Three different varieties of control channels have been defined. *Broadcast channels* are downlink only-channels and are used by the base stations to inform the mobile stations, in particular about TDMA synchronisation via the *Frequency Correction Channel (FCCH)* and the *Synchronisation Channel (SCH)*. The *Broadcast Control Channel (BCCH)* is used by the base stations to transmit their identifications to allow mobile stations to detect signal strength for a possible hand-over.

The *Common control channel (CCCH)* is used for classical signalling purposes such as establishing a connection, allocating traffic channels, etc.

Finally, three types of the *Dedicated Control Channels* have been defined, at least one of which is set up for each traffic channel and is used to maintain data communication throughout the conversation.

Since GSM not only supports voice traffic but also data transmission, it is necessary to inform the remote end as well as the intermediate systems which sort of traffic is going to be exchanged on that particular channel. This is done using a *Bearer Capability Information Element (BCIE)*, which contains information on the type of traffic (speech, fax, data, etc), the data rate (300 bits/sec, 1.2, 2.4, 4.8 or 9.6kbits/sec), word length, stop bits and other information.

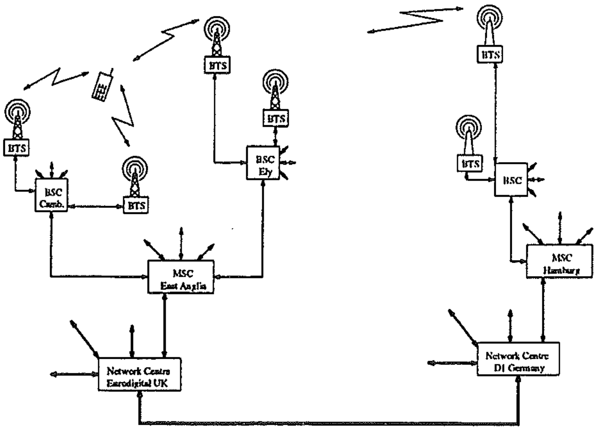

Figure 7: GSM handover procedure

2.3 Handovers

The signal strength and quality is assessed throughout the connection for the up- and downlink by the mobile unit and the base station transceiver servicing the call. If the assessed quality drops below a pre-defined threshold, the base station transceiver attempts to allocate a different channel within the same cell. If the mobile unit has moved outside the cell-boundaries of that base station transceiver this will fail and a hand-over to another cell is required.

The mobile unit is only active during 2 out of 8 time slots in each TDMA-frame (one for the uplink and one for the downlink), and can use some of the other time slots to assess signal strengths from other base stations. This is done by tuning in to the Broadcast Control Channel (BCCH) of the surrounding cells and averaging a number of signal strength measurements over the period of a superframe (6.12 seconds). Messages on the BCCH carry a unique *base station identification code (BSIC)* which allows the mobile station to determine which base station it is listening to. Once the BSICs and the signal strengths of the six strongest surrounding base stations have been determined, this information is transmitted to the current base station via the slow-associated control channel (SACCH). The base station then evaluates this information by checking if one of the base-stations listed is willing to accept the call, and the call is then handed over on either base station controller-, mobile switching centre- or network centre-level.

Thus, handovers are initiated by the mobile station, but they may be executed on any level between the base station controller and the network centre. Due to the different complexities involved,

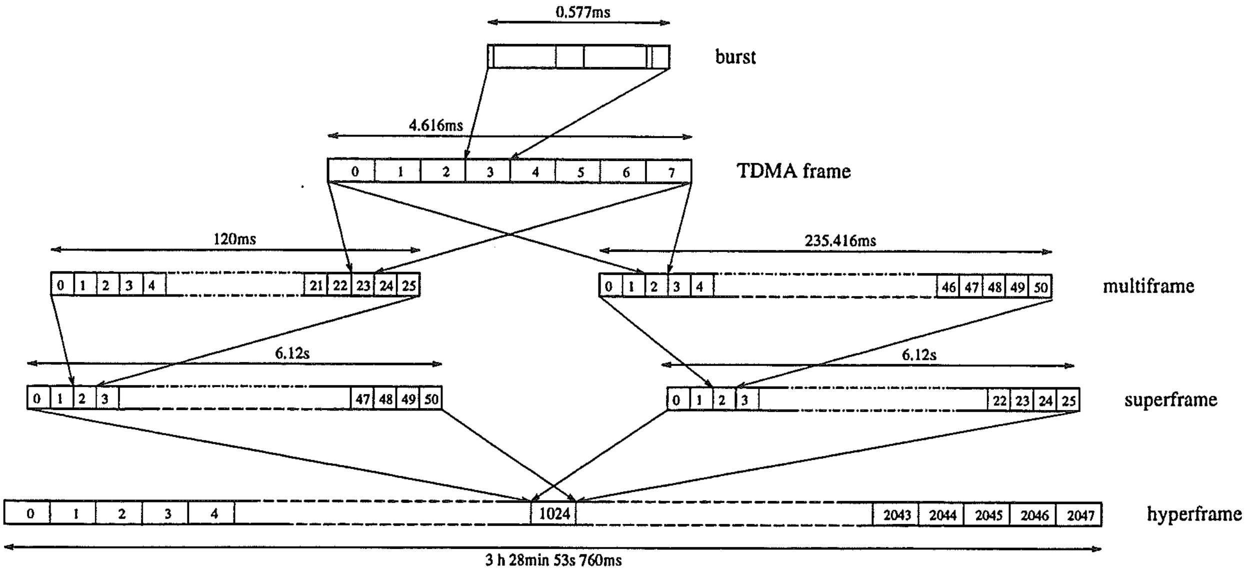

Figure 6: GSM frame structure

a handover may take between 190ms and 320ms in set-up time. The actual switch, however, occurs much more rapidly at the end of the current multiframe and is agreed upon in advance over one of the control channels. The process and the different levels within the GSM infrastructure are illustrated in figure 7.

2.4 GSM speech coding

GSM employs the most complex speech coding and channel coding currently in use by any commercial mobile telephone system. The coding itself is done using a *regular pulse excited linear predictive coder with long-term pitch* or *RPE-LTP* for short. The coder is block-based and operates on speech samples of 20ms duration. The output for each block is 260 bits.

Due to the high error rates incurred by the radio-interface forward error correction is a must. However, these techniques are very expensive in terms of bandwidth and have to be applied very carefully for maximum effect. Subjective testing has shown that 182 bits of the 260 produced by the speech coder are extremely sensitive to transmission errors and cause a significant decrease in speech quality if in error. Those bits were termed *Class 1-bits*, whereas the remaining 78 bits are *Class 2-bits.*

The class 1-bits were further analysed and subdivided into two groups: 50 *Class 1a* bits and 132 *Class 1b*-bits with the first group showing higher sensitivity to transmission errors than the latter [BGMF94].

A 3-bit cyclic redundancy check is used on the class 1a bits. If, on the other end, the CRC-bits reveal that the class 1a-bits have been corrupted, it was found to be better for the perceived speech quality just to ignore the entire burst. For transmission, the bits are re-ordered in such a way that first all class 1a-bits are transmitted, then the three CRC-bits, and finally the class 1b-bits. A trail of 4 zeros is added to this group, to enable the decoder to verify the integrity of the block. A convolution coder is applied to those 189 bits, resulting in a highly protected 378 bit-block. To this the remaining 78 class-2 bits are appended. Thus, each 20ms-block of speech results in a 456 bits, or a net transmission speed of 22.8kbits/s (see figure 8).

In mobile environments, transmission errors tend to occur in bursts, sometimes longer than the transmission bursts used in GSM, which may lead to the loss of an entire transmission burst. If data frames are re-distributed over several bursts it may be possible to restrict the number of transmission errors falling within the same 20ms-speech sample. The 456 bits of each data frame are therefore re-ordered and interleaved over 8 TDMA frames for final transmission [Kai94].

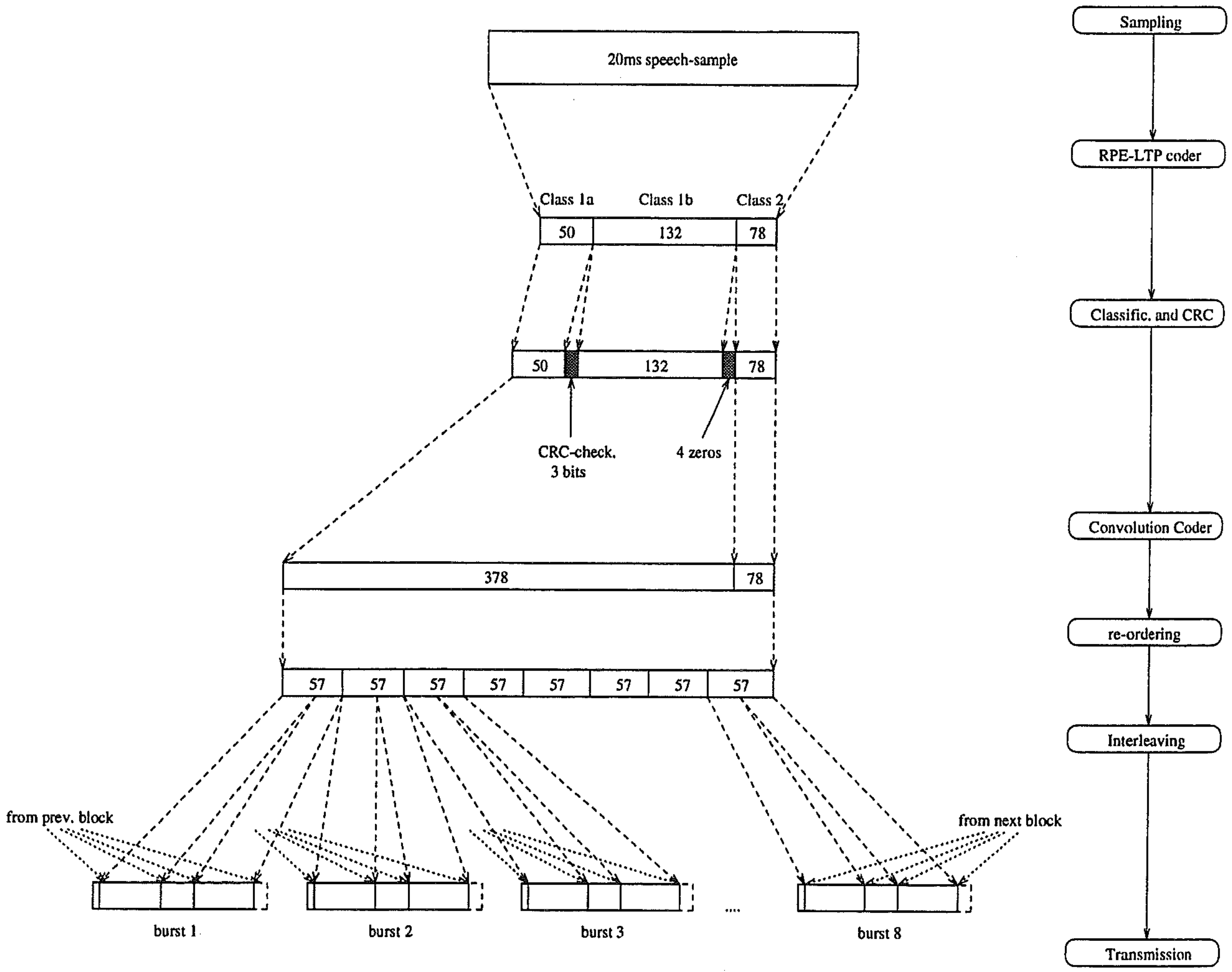

Figure 8: GSM speech coding schema, including interleaving onto transmission bursts

2.5 DCS1800, PCN

DCS1800 is specified by the European Telecommunications Standards Institute (ETSI) as a closed derivate of GSM. Whilst the basic technologies are very similar, DCS1800 differs in three points: Firstly, it operates in the newly freed 1800MHz-range, with the uplink and downlink-bands at 1710-1785MHz and 1805-1880MHz, respectively. This provides a total theoretical capacity of 375 radio carriers. TDMA is used, as in GSM, and full- and half-rate traffic channels are supported.

Secondly, output power for DCS1800 phones is restricted to 1000mW (peak) for car-phones and 250mW (peak) for hand-held units.

Thirdly and finally, the standard allows roaming between different national operators, where networks overlap. Recently, roaming agreements have even been signed between network operators in different countries, in particular between operators in London, UK, and in Leipzig, Germany [Goo94]. However, whereas the GSM-standard commits its operators to such agreements, it is left to the individual PCN-operators to decide on this issue.

The higher frequency band necessitates some changes to the network topology itself: maximum propagation is reduced, so cells are smaller. Typically cells in urban environments have radii from 0.4km, rural areas are covered with 5km-cells. This is in contrast to the GSM system, where cells can cover a radius of up to 30km. In urban areas however, GSM cells also tend to be small in order to provide high level of service.

In this setting, micro-cells can be established to cover small areas in railway stations, airport terminals or shopping malls. It is even possible to

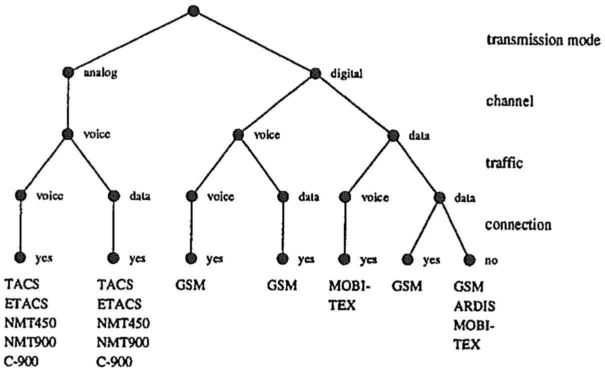

Figure 9: Classification of radio communication networks

integrate pico-cells for indoor-coverage. This availability of the mobile system which clearly is beyond the scope of the old style car-phone prompted the classification as a *personal communications network* or *PCN*.

As for the non-RF part, PCN-networks are identical to GSM from the technical point of view. [Pot92]

3 Mobile data communication

In the early days of computing, networks where used to connect terminals to mainframes. With the advent of cheap personal computers, the simple terminal networks were developed into data exchange networks allowing data to be down-loaded from remote file servers across dedicated lines as well as conventional telephone connections. Modems on either end converted the digital stream of data into a corresponding analogue signal which could be conveyed using the normal PSTN network.

The introduction of small, mobile computing equipment in the mid 80s also sparked off enormous interest in mobile data communication networks. The success of mobile telephone systems in the 90s helped to convince network operators that there may also be a market for data networks and a number of solutions where developed and introduced. Figure 9 gives a rudimentary classification based on transmission mode (analogue or digital) and connection mode (connection oriented or packet oriented).

3.1 Connection Oriented Data

As outlined above, by using a modem on either end a conventional telephone connection can be set up to a remote machine and data can be transmitted. From the network's point of view, this is equivalent to a voice connection and consequently treated in exactly the same manner: call charges are based on connection time and the quality of the connection is the same as for voice calls.

3.1.1 Modems and mobile telephony

The use of modems with mobile phones is, unfortunately, not as straightforward as one might expect. Connections from mobile phones are significantly more prone to noise and will thus introduce more transmission errors than wired networks, but more significantly they will provide less bandwidth than wired network. Thus, modems that make full use of the bandwidth provided by conventional PSTN connections will fail to work on analogue phones.

For analogue phones, special modems are available that take into account the characteristics of mobile connections and can be used in conjunction with some phones on the NMT, AMPS and ETACS systems. The bit rates are around 2.4kbits/sec and are therefore significantly below bit rates achievable on wired networks.

As for digital networks, due to the complex speech encoding simple modems cannot be used on speech channels. For GSM, a special data channel has been defined instead.

3.1.2 GSM data channels

Being a digital system from the outset, GSM lends itself perfectly to data transmission. In this chapter, 'GSM' may also be read as 'DCS1800' or 'PCN', for those standards offer the same technology with respect to data transmission as GSM does.

Each traffic channel may, via a BCIE, be configured as a data channel at any time. Figure 10 gives an overview of the network connections available. In general, mobile users are offered a 1B+D ISDN configuration of traffic and signalling channels. Obviously, the full bandwidth of 64kbits/sec is not available on the ISDN B-channel. The conversion from the bit-rate used on the mobile side to the full B-channel bandwidth is done by the mobile switching centre using the standard V.110 protocol.

The bit rates offered for data transmission on GSMs traffic-channels that can be configured

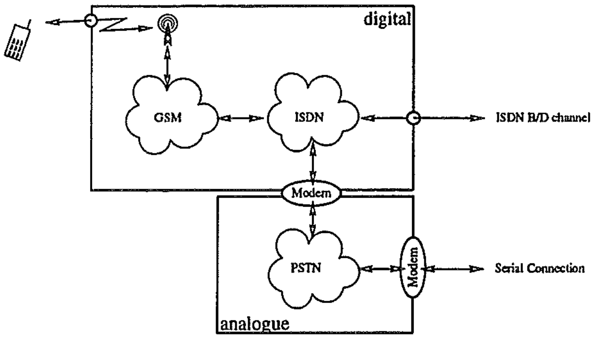

Figure 10: GSM data channel overview

through BCIEs are 300 bits/sec, 1.2, 2.4, 4.8 and 9.6 kbits/sec. The ISDN D-channel is mirrored in the GSM network through the Common Control Channel (CCC) and it can be used for various add-on services such as *Short Messages Service*, paging, etc.

In general, traffic channels that have been configured as data channels through some BCIE are treated differently from traffic channels that carry voice traffic. Though it would be possible to treat both in the same manner, it has been shown that it is advantageous for the network operator to give priority to voice traffic. Most networks therefore employ a *frequency hopping* schema that immediately interrupts a data connection should a frequency be required for a waiting voice connection. Though the network attempts to identify a new channel for the interrupted data connection, this can not always be guaranteed. Effectively, this means that data channels are slightly more prone to break-ups than voice channels, and in heavily loaded areas it may be difficult to establish a data connection at all. Some network providers set aside one particular logical channel as a data-only channel to offer some minimum service to data-customers. The throughput of the system does suffer slightly under this schema because of the time required by the system to switch channels.

The GSM specification defines two data services - the *transparent mode* and the *non-transparent mode.*

In transparent mode, no error correction beyond FEC is employed. The end-user thus experiences a fixed throughput and fixed delay but also has to deal with variable error rates. The delays experienced are dependent on the load in the network and the exact network implementation, but in the UK delays in the oder of 200ms are the norm (and are therefore roughly the same as for a single satellite hop). The error rate depends on the reception quality and the amount of FEC employed, which in turn depends on the bit rate chosen by the user. Table in figure 3.1.2 gives some projected performance characteristics (since no data service is currently operational, actual performance may differ).

No data is currently available on typical bit error rates. It has been reported, though, that the transmission itself is extremely prone to transmission errors and without forward error correction bit error rates are in the region of 10^{-3}.

Due to the strong channel coding and the forward error correction schemes, data that has been corrupted by error bursts or other transmission errors is not being passed on until the channel is lost entirely. The FEC then goes out of control and the user sees random data until the network detects the fault and clears the call after roughly 12 seconds. Handovers usually require between 190ms and 320ms. During this time data may be affected and the error rate may peak beyond the capabilities of the built-in FEC. Again, due to the non-availability of data services more detailed performance data is not yet available. Throughput is constant, so flow control is simple.

Non-transparent mode offers a 'virtual circuit'-quality connection. In addition to user data, control information and forward error correction, an additional automatic repeat request-protocol is being employed that allows to improve the quality of service dramatically. Error rate is now fixed at an extremely low level, but this has to be payed for by variable throughput (below that of the transparent service) and variable delay. The minimum delay is the same as for the transparent mode, but may grow due to the re-transmission of data. Again, the exact characteristics depend on the user-rate chosen, with lower bit-rates offering more protection that higher bit rates. In this mode, the bit-rate specified by the user only corresponds to the bit rates employed between the two data adaptors. The bit-rate experienced by the end-user may be significantly below that rate.

Due to the strong protection, transmission errors usually increase delay but error rates are normally stable. Corrupted data is usually filtered out by the ARQ and FEC protocols. Peak delays may be experienced during handovers. Backpressure may built up due to the continuous re-transmission of corrupted data, which may lead the end-application to time-out and clear the call.

User Rate	Intermediate Rate (Data + Control)	FEC	Correction Power
0.3, 1.2, 2.4kbits/s	3.6kbits/s	19.2kbits/s	67% or 48 out of 240 bits
4.8kbits/s	6kbits/s	16.8kbits/s	40% or 48 out of 120bits
9.6kbits/s	12kbits/s	10.8kbits/s	20% or 48 out of 72bits

Figure 11: Forward Error Correction at various user bit-rates

Having available BCIEs and a separate signalling channel, mobile users as well as ISDN-customers may set up a connection using one telephone number and then configuring the channel as a voice channel or data channel afterwards. Callers from the PSTN network can not configure BCIE parameters to signal the system the intended use of the connection. PSTN access therefore requires a different number for each service, i.e. one number sets up the voice channel to a mobile phone, another number a fax-channel to the same phone, a third number a 9.6kbits/s transparent connection to the same phone.

3.1.3 GSM Short Messages Service

GSMs separate signalling channels hold resources beyond those needed for conventional signalling purposes. The SMS-service allows those surplus-resources to be used for the transmission of small packets of data; namely, SMS transmits junks of up to 160bytes. Unlike a conventional traffic-channel, no direct connection is established between the sender and the receiver. In fact, the network operates as a store-and-forward network with the message being held by the SMS-centre until the receiver is available, i.e. switched on.

Currently, no network in operation offers this service as yet. However, the method is being used by several networks for network administration and may be made available for public use at any one time. It is, however, yet uncertain in which way end-users can access the SMS-service. [Kai94]

3.2 Packet Radio

Among the various mobile communication techniques, packet radio networks are particularly interesting for the implementation of a wide range of applications like fleet management, pickup and delivery, steering of the mobile workforce, and many more. This is because packet radio networks are specifically optimised to transport data as opposed to speech. In addition, packet radio networks typically provide a high degree of reliability by applying extensive forward and backward error detection techniques and by employing recovery mechanisms such as acknowledgement and automatic retransmit mechanisms in the lower layer communication protocols. Furthermore, several networks employing various proprietary protocols are already installed in many countries inside and outside Europe and operational experience of several years has been built up in those countries as well.

From an economical point of view, packet radio systems also have the advantage that the connections are charged to the user per packet of traffic rather than per connection time.

In Europe, PTTs, and/or private consortia (where the laws permit), are currently bidding for the installation of mobile data networks in a number of countries. Competing networks will be set up at least in the bigger countries. The main competitors with respect to the provision of the mobile communications technology for the national networks are Motorola and Ericsson. A Motorola packet data network is currently being marketed in the US under the name *ARDIS* and in Germany as *MODACOM*. Ericsson technology with the *Mobitex* mobile data network is up and running in a number of European countries as well as in North America. [PA92]

3.2.1 ARDIS, MODACOM

In Germany, the MODACOM network was installed and made operational by Motorola for the German Telecom. The mobile technology was based on what is known as ARDIS in the US, but numerous modifications were made for the German environment. Now the network is operational in Germany and provides coverage of all metropolitan areas, industrial centres and the main motorways.

The Motorola-system transmits packetised data over the air interface. The geographical area is divided into cells, each of which is served by its

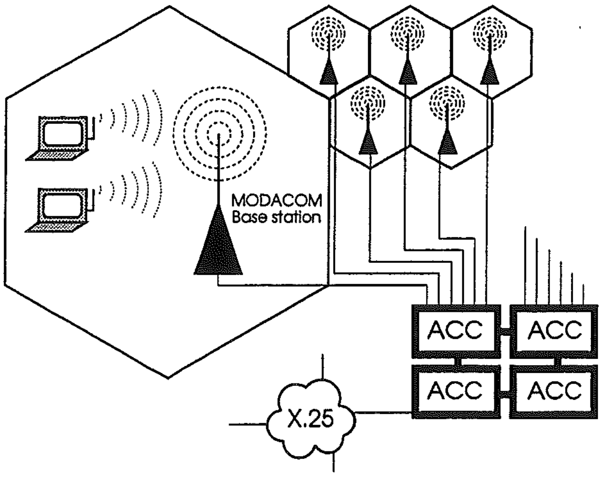

Figure 12: The network architecture of MODACOM

own base station. However, only two frequencies (one for the uplink, one for the downlink) are used in any one cell. Each base station sends and receives data over the air and also provides linkage to the MODACOM fixed network infrastructure. Figure 12 graphically illustrates the MODACOM/ARDIS architecture and its main infrastructure constituents.

The protocol (named RD-LAP) implemented over the air interface uses two frequencies per cell. On the downlink (i.e. from a base station to all mobile terminals within the respective cell) contention is prevented by exclusively reserving it for the base station. On the uplink (i.e. from all the mobile terminals in the respective cell to the base station) there is no administrative means to totally rule out contention among the mobile terminals. Collisions are dealt with by the RD-LAP protocol in a manner similar to the slotted ALOHA protocol. Details aside, there is, to a certain degree, an analogy to handling collisions in the better known Ethernet.

The gross transmission speed amounts to 9600 bits per second (per cell per link). Thus, being collision free, the net transmission rate on the downlink almost attains the gross value. Effective capacity on the uplink, of course, depends on packet length distribution and collision rate, but it is always smaller than on the downlink.

The mobile terminal is connected via an RS232 interface (serial line) to a MODACOM/ARDIS-modem either externally or internally. The modem is controlled through a mostly Hayes compatible command set, and data transfer between the modem and the terminal is asynchronous.

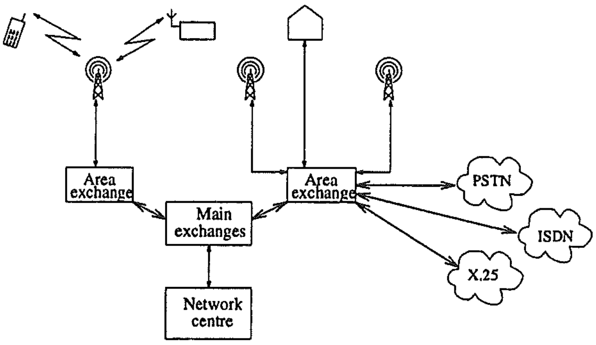

Figure 13: The network architecture of Mobitex

The fixed part of the network infrastructure, among others, consists of gateways into the X.25 network. Currently this is the only way to transfer data from the mobile terminal to a stationary (host) computer, apart from equipping the stationary side with a MODACOM/ARDIS modem as well. From the perspective of the network, a mobile-to-mobile store and forward link is then established [Tel92].

3.2.2 MOBITEX

The Mobitex-systems was developed by Ericsson in the early 80s and went into field trials in 1983. 1987 saw the beginning of its commercial operation in the largest cities and roads in Sweden. By the end of 1992, Mobitex-systems were up and running in Sweden, Norway, Finland, Canada, USA, France, the Netherlands and the UK [Ste91].

The basic network architecture is very similar to that used in cellular telephone networks, but many responsibilities have been moved downwards to the base stations. Mobitex is optimised for packets between 40 to 50 characters, even though shorter and longer packets (up to a *maximum transmission unit* or *MTU*) are acceptable. Due to the packet-oriented nature of the network many more users can be accommodated for in each cell using the same bandwidth as cellular telephones: in the latter, usually 20 to 25 simultaneous users can be supported. This figures is between 100 and 1000 for Mobitex. Speech channels are available but they dramatically drain resources from the network. Their usage should therefore be restricted to emergency situations, not at least by charging extremely high fees.

The network has three levels in its hierarchy. On the lowest level base stations receive packets from the mobile units and pass the traffic up to the medium level, the area exchanges. Most traffic handling is concentrated on this level, with each area exchange providing access into various wired data networks, in particular the wired telephone network PSTN, its digital counterpart ISDN and the data network X.25. The network centre is primarily concerned with network administration, billing and statistical tasks (see figure 13).

Error correction and packet re-transmissions are wholly handled by the network itself, providing the user with virtual-circuit quality packet connection [Ber89]. The gross transmission speed is around 8kbits/sec [Ste91]. In contrast to the Motorola-system, Mobitex is much more based on techniques originally developed for cellular telephony such as frequency reuse and multiple channels. All mobile stations periodically transmit their identities to the local base station, thus enabling the network to keep an up-to-date database of all mobile units and their locations.

4 Security issues

As with all communication systems, security in mobile telephony systems is primarily concerned with two issues - impersonation and secure transmission.

As for the latter, the early analogue systems up to TACS and AMPS make no provision: both voice and signalling channel are transmitted in analogue mode without any form of encryption, so it is very easy to listen to analogue cellular phones using conventional scanners and decode both the caller and the called number with simple hardware. Impersonation is slightly harder inasmuch as it requires hardware and expert knowledge: each mobile unit has its identification number, the so-called *Electronic Serial Number (ESN)*, programmed into it but it is possible to change it to another phone's ESN and thus to impersonate that user's identity, taking his/her calls and making calls on his/her bill. The network will eventually realise that two phones use the same ESN and will bar both phones to prevent fraud.

The German C-900 systems is slightly more security-aware. The signalling channel is digital and therefore harder to decode, and the voice channel, while still analogue, is encrypted using one of 256 keys which are assigned on a per-call bases. Whilst the system is far from being 'secure' inasmuch as it is possible to crack the key and decrypt calls, some privacy is being offered at least against the common scanner-attack.

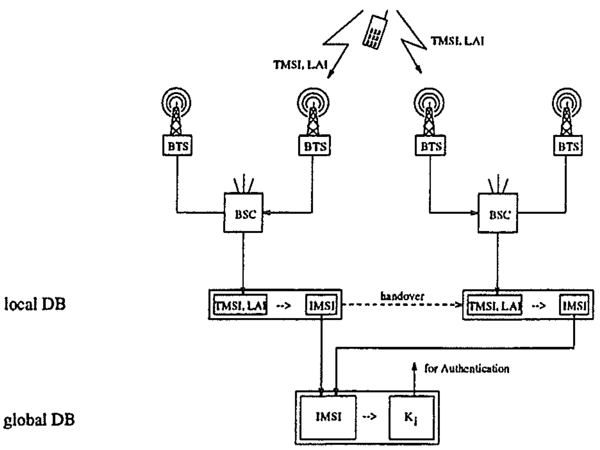

Figure 14: GSM security

GSM is the first mobile telephone system where security has been an issue right from the outset. Being a fully digital system, encryption can be fitted into the network more tightly. Since each mobile phone is identified with its unique IMSI, it is desirable not to transmit this information in clear text over the air interface. Instead, a *Temporary Mobile Station Identity (TMSI)* together with the *Location Area Identification (LAI)* is used. Databases are kept locally within the network and they will provide the actual IMSI for any two TMSI and LAI. The point is that this database is local and should, it is hoped, offer more security that the globally kept database (see figure 14). The TMSIs which are assigned by the base station servicing the call are changed regularly and passed on between base stations on trusted lines when mobiles move into another cell. Only if the mobile has moved outside the validity of its last TMSI or has not yet been assigned one, the base station will request the actual IMSI which is then sent in an encrypted form. It is important to protect the IMSI because for any given IMSI the secret station authentication key K_i can be queried from a network-wide database. The mobile phone has the reverse key k_i^{-1} hard-wired into its memory. K_i is used to authenticate the mobile station. This is done by the base station which encrypts a random number RAND using K_i, sends it to the mobile station, which in turn decrypts the number and sends it back to the base station. Authentication fails if the result does not match the initial RAND. If the mobile station has been authenticated, RAND together with K_i are used in a confidential algorithm (called *A8*) to produce the ciphering key K_c. K_c is used in another confidential algorithm (called *A5*) to encrypt all fu-

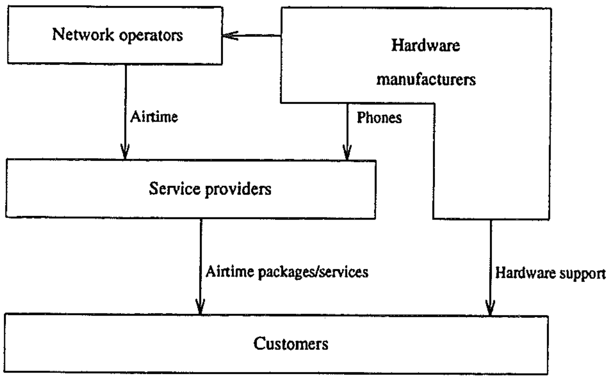

Figure 15: Market elements

ture transmissions, including the voice data transmissions. K_c can be created by the mobile station as well as the base station independently, so the key does not need to be transmitted.

The next step is to authenticate the subscriber which can be done easily using K_c and the SIM-ID. Security for the subscriber is provided by using SIMs which, it is claimed, can not be copied.

Authentication therefore happens on two levels. Firstly, the mobile phone itself has to be authenticated, and in the process encryption keys are assigned. This process ensures that mobile phones cannot be re-programmed to take on another IMSI, since it would be impossible the generate a matching key K_i^{-1}. Secondly, the subscriber is authenticated.

The level of security in the GSM-network is such that even federal intelligence agencies have to rely on the cooperation of network operators if they wish to monitor calls. The keys are held in network-wide databases, and calls placed in foreign countries are therefore always routed back to the home country to allow the home-system to check authentication, etc. The key therefore never leaves the home-country, which means that visited countries have normally no power to intercept calls placed on their territories. It is probably because of this fact that an astonishing number of German GSM-phones appear to be permanently in use by some extremely security-sensitive users in Sicily, Italy.

5 Mobile communication today

GSM distinguishes between *Network Operators* and *Service Providers.* The first is responsible to set up, maintain and operate the actual network and has therefore technical competence. In the UK, there are a number of networks being operated by *Cellnet, Vodafone, Mercury ('One2One')* and *Hutchinson Telecom ('Orange').*

In general those companies sell airtime to service providers who are responsible for the marketing of the network. In particular, service providers will compile packages including line rental, call charges, and other add-on services which they will then try to sell to customers. In highly competitive market such as the UK, network operators will assist service providers in creating suitable packages and advertise them jointly. In recent years networks have gone as far as significantly subsidising the price for the actual mobile phones. Analogue phones can now be bought for as little as £20, GSM phones for £50. This is clearly far below their actual value. In the UK, some network operators also act as service providers (Hutchinson and Mercury, in particular), but a fair number of service providers are established in the market, eg. *The Carphone Warehouse, People's Phones*, and others.

Hardware manufacturers such as Nokia, Motorola or Siemens are also competing fiercely for contracts with network operators. Figure 15 gives an outline of the interactions between the various market elements.

Unfortunately GSM has not yet achieved its initial goal: providing a worldwide standard for mobile telecommunication. To date, GSM has been adopted by 56 operators in 30 countries in Europe. GSM phones are, by and large, still significantly more expensive than conventional analogue phones, and though the situation is expected to change in favour of GSM phones, this still poses a serious obstacle to the successful introduction of GSM. Moreover, it can be observed that GSM is less successful in countries which already offer a fairly high-quality mobile telephone system, such as the UK, whereas in places like Germany, where mobile telephones were extremely expensive in the past, GSM has really taken off. There, the introduction of GSM coincided with the privatisation of the mobile communications market, with the introduction of a privately run network (*D2*) alongside the state-run *D1* network. This has caused a fierce competition which lead to a significant reduction in connection charges compared with the older C-system and significantly boosted the popularity of GSM.

Even less-developed countries such as Greece have started joint-ventures with leading European telecommunication operators in a quest to introduce high-quality GSM.

The American market suffers currently under the presence of numerous number of small network operators, few of which are financially able or willing to invest the enormous sums required to set up a GSM system [NW91]. In addition, two American standards have recently been defined: *IS54*, which employs digital transmission in a TDMA-mode - similar to GSM but significantly simplified and less expensive to implement, and the ambitious *IS95*, which employs digital transmission in a Code-Division Multiple Access schema [IRSD94]. The first has a user community of roughly 150,000 in North America, the latter is currently in field trials. This stands in competition with current 7,500,000 subscribers (world-wide) to the various GSM systems in operation.

Since the fall of the Soviet-Union mobile communications has also take off in Russia and several of the ex-member states; the situation there is best described as 'chaotic'. Systems where mainly bought up from scrap yards in Western countries and so some old NMT450 and AMPS systems achieved some belated fame. The networks are run by small private companies and operate locally, with phones working in one city being useless 20km down the road in another city.

5.1 Data communication

No clear picture can be drawn for the data communication market. Economies of scale are yet far off and unit prices for mobile modems are high. Another problem is created by the variety of network operators and the significant differences in legislation in different countries: some network operators are extremely reluctant to approve data modems for their networks, for there is not enough revenue created by data customers to justify the lengthy process of approval and risking network load created by those connections. Sometimes, the problem lies within the over-specification of the networks: the specification document for end-user systems in the NMT900 system, for example, specifies the presents of a numeric key-pad and also defines the exact functions for each key. When IBM introduced the *PCRadio*-product in 1991, which essentially was a simple but ruggedised PC with an integrated radio-modem, this specification could not be met. Consequently, Ericsson refused approval for the NMT900 network. The problem was eventually solved by running a TSR program that would provide a key-pad with the required functionality, if requested. The functioning of the modem was directly connected to the presents of the TSR, thus satisfying the strict specification.

Economies of scale can only be reached with a modem that can be used in a variety of countries, thus reaching a big user community. However, no manufacturer is currently willing to invest the sums required to develop such a device and to get approval for a network in many countries.

In addition, network operators consider the data-service an add-on and, while the overwhelming proportion of their revenue is created by the extremely profitable voice-services, have no incentive to support the development of such data adaptors.

Naturally, the situation is better for the data-only networks, but coverage is usually local and economies of scale can not be reached.

5.2 Charges

Each service provider is free in setting his own tariffs and charges, which consequently differ considerably from country to country. Typically, the subscriber will have to pay a fixed *connection charge* when the phone is registered with the network, and a regular *line rental* charge. Calls are normally charged to the next full 30 seconds, although some PCN services and all GSM services will charge to the second. Normally, different tariffs are available which offer different line-rental/call charges mixes for different users. Some typical figures for the UK market are listed in figure 16.

International calls placed from a mobile will automatically be charged at a higher rate. Calls that are placed in a foreign country are of a special nature. Because of the various security issues involved, the call is automatically routed back to the home-location of the mobile phone, thereby incurring the cost of an international call charged by the host-network at its own tariff. The home network will then process the call further and charge the call to the user accordingly.

As for mobile terminated calls, it must be assumed that the caller does not know the location of the mobile phone at the time of setting up the connection, and therefore a flat charge is made for these calls regardless of the location of the mobile. The issue of international roaming poses serious problems with respect to charging; the issue was first considered for the NMT450 system, when the first international roaming contract was agreed between Sweden, Norway and Finland. It was agreed that

Technology	Tariff	Connection charge	monthly line rental	call charge (per minute, peak time)
ETACS	Cellnet Lifetime, Vodafone LowCall	£25	£12.50	40p
ETACS	Cellnet Primetime, Vodafone Business	£50	£25	25p
GSM	Cellnet Primetime Plus, Vodafone EuroDigital	£50	£25	25p
PCN	Mercury One2One, Hutchinson orange	£30	£15	25p

Figure 16: Some typical UK tariffs (all excl. tax). *Source: The Carphone Warehouse, Jan. 1995*

the same flat rate would also apply to international calls. As for most GSM networks, calls to a mobile at a foreign location are charged as a national call to the caller; the forwarding to the actual international location is payed for by the callee (who is free to disable this feature).

6 Where does it all go from here?

Keeping in mind the pace at which development has taken place in the field of mobile communications over the last decade it is somewhat difficult to predict any future developments. GSM appears to be the sole winner in the race for a multi-national communication network. On the other hand, however, the development of code-division-multiple access in the USA, a derivative from technology already in use by the military, promises better spectrum efficiency than the TDMA systems used in GSM [HL91]. It is feared that, should the US market throw its weight behind IS95, GSM may find it difficult to survive in the long run, and it will certainly find it impossible to attain its position as a global standard [Rog91].

Personal communication networks also hold cues for the future. Cells will, over the next few months, decrease in size and will not only intrude offices and shopfloors but also private homes. Upcoming digital cordless standards such as *CT2*, *CT3* or *DECT* [Tut92] will then be integrated into PCN networks, such that phones can be used as cordless phones while at home and as cellular phones when on the road [Gro90].

Researchers all over the globe are currently trying to find a great unifying theory -so to speak- for the various communication technologies available today; it is hoped that e-mail, pagers, mobile phones, cordless phones and the like will one day all be unified under a *Global Intelligent Network* [DDF92] which will keep us connected with our loved ones, friends and colleagues day and night. What a wonderful vision of the future world!

References

[Ber89] Göran Berntson. Mobitex - a new network for mobile data communications. *Ericsson Review*, 1, 1989.

[BGMF94] G. Benelli, A. Ganzelli, P. Matteini, and A. Fioravanti. Some digital receivers for the gsm pan-european cellular communication system. *IEE Proceedings on Communications*, 141(3), June 1994.

[CMS93] James J. C. Chang, Richard A. Miska, and R. Anthony Shober. Wireless systems and technologies: An overview. *AT&T Technical Journal*, July/August 1993.

[DDF92] John H. Davis, Neil F. Dinn, and Warren E. Falconer. Technologies for global communications. *IEEE Communications Magazine*, October 1992.

[Goo94] Rupert Goodwins. Getting started - q&a. *What Mobile and cellphone magazine*, Winter 1994.

[Gro90] I. S. Groves. Personal mobile communications - a vision of the future. *Br Telecom Technological Journal*, 8(1), January 1990.

[HL91] Kurt Hellström and Åke Lundqvist. Trends in mobile communications. *Ericsson Review*, 3, 1991.

[Hod90] M. R. L. Hodges. The gsm radio interface. *Br Telecom Technological Journal*, 8(1), January 1990.

[IRSD94] Magnus Isaksson, Krister Raith, Anthony Sammarco, and John Diachina. A new standard for north american digital cellular. *Ericcson Review*, 2, 1994.

[JL90] Greger Jismalm and Jan-Olof Lejdal. Cell planning - products and services. *Ericsson Review*, 2, 1990.

[Kai94] Frank Kaiser. Jetzt funk's. *c't magazine für computer technik*, July 1994.

[Lam84] Richard Lambley. Developments in cellular radio. *Electronics & Wireless World*, June 1984.

[NW91] Roderick Nelson and Dan Westin. The north american cellular network. *Ericsson Review*, 4, 1991.

[PA92] Peter Peinl and Richard Adolf. An overview of mobile communication technologies. Technical report, IBM Science Centre Heidelberg, June 1992.

[PG89] J. Parsons and J. Gardiner. *Mobile Communications Systems*. Information Technology. Blackie, 1989.

[Pot92] A. Robin Potter. Implementation of pcns using dcs1800. *IEEE Communications Magazine*, December 1992.

[Rog91] Steve Rogerson. Talking confusion. *Electronics World & Wireless World*, January 1991.

[Ste91] Per Stein. Greater efficiency with mobile data communication. *Ericsson Review*, 4, 1991.

[Tel92] Telecom Mobilfunk. *Modacom. Mobile Datenkommunikation*, December 1992. KNr. 641 221 416-3.

[Tut92] W. Tuttlebee. Cordless personal communications. *IEEE Communications Magazine*, December 1992.

[You79] W. R. Young. Advanced mobile phone service: Introduction, background, and objectives. *The Bell System Technical Journal*, 58(1), January 1979.