Telecommunication systems

4

Digital cellular networks are the segment of the market for mobile and wireless devices which are growing most rapidly. They are the wireless extensions of traditional PSTN or ISDN networks and allow for seamless roaming with the same mobile phone nation or even worldwide. Today, these systems are mainly used for voice traffic. However, data traffic is continuously growing and, therefore, this chapter presents several technologies for wireless data transmission using cellular systems.[1]

The systems presented fit into the traditional telephony architecture and do not originate from computer networks. The basic versions typically implement a circuit-switched service, focused on voice, and only offer data rates of up to, e.g., 9.6 kbit/s. However, service is provided up to a speed of 250 km/h (e.g., using GSM in a car) where most other wireless systems fail.

The **worldwide market** figures for cellular networks are as follows (GSM Association, 2002). The most popular digital system is GSM, with approximately 70 per cent market share. (This system will be presented in section 4.1.) The analog AMPS system still holds three per cent, whereas the Japanese PDC holds five per cent (60 million users). The remainder is split between CDMA (12 per cent) and TDMA (10 per cent) systems, and other technologies. In **Europe** almost everyone uses the digital GSM system (over 370 million) with almost no analog systems left. The situation is different in the **US** and some other countries that have adopted US technology (e.g., South Korea, Canada). Here, the digital market is split into TDMA, CDMA, and GSM systems with 107 million TDMA, 135 million CDMA, and only 16 million GSM users (North America only). While only one digital system exists in Europe, the US market is divided into several systems. This leads to severe problems regarding coverage and service availability, and is one of the examples where market forces did not ensure improved services (compared to the common standard in Europe).

Figure 4.1 shows the worldwide number of subscribers to different mobile phone technologies (GSM Association, 2002). The figure combines different versions of the same technology (e.g., GSM working on 900, 1,800, and 1,900 MHz).

1 All systems presented here are digital, for older analog systems such as the US AMPS (advanced mobile phone system) the reader is referred to, e.g., Goodman (1997).

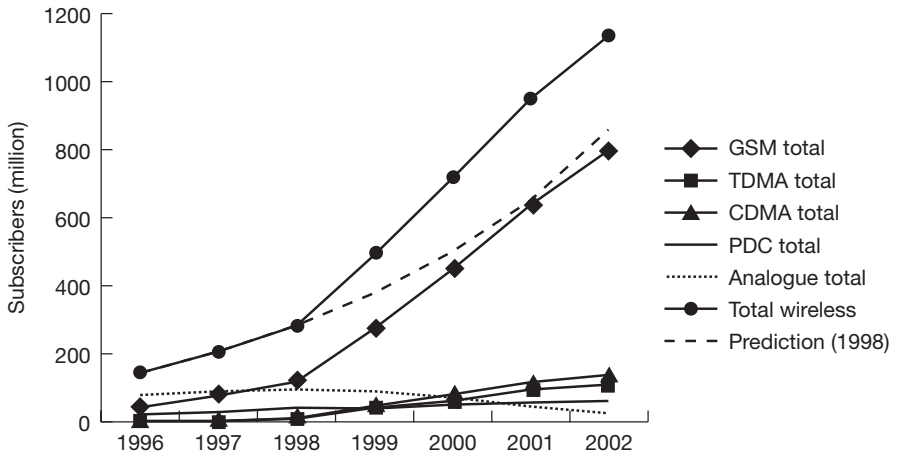

Figure 4.1
Worldwide subscribers of different mobile phone technologies

The two upper lines in the graph show the total number of users and the predictions from 1998. It is interesting that no one foresaw the tremendous success of the mobile communication technology. The graph shows, too, that the time for analog systems is over and GSM is heavily dominating the current market. GSM, TDMA, CDMA, and PDC are all second generation systems. It is important to note that today more people use mobile phone systems than fixed telephones! The graphs of mobile and fixed users crossed in March 2002.

The following sections present prominent examples for second generation (2G) mobile phone networks, cordless telephones, trunked radio systems, and third generation (3G) mobile phone networks. This chapter uses GSM as the main example for a 2G fully digital mobile phone system, not only because of market success, but also due to the system architecture that served many other systems as an early example. Other systems adopted mobility management, mobile assisted handover and other basic ideas (Goodman, 1997), (Stallings, 2002), (Pahlavan, 2002). While US systems typically focus on the air interface for their specification, a system like GSM has many open interfaces and network entities defined in the specification. While the first approach enables companies to have their own, proprietary and possibly better solutions, the latter enables network providers to choose between many different products from different vendors. DECT and TERTA are used, respectively as examples for cordless telephony and trunked radio systems. One reason for this is their system architecture which is similar to GSM. This is not very surprising as all three systems have been standardized by ETSI. The main focus is always on data service, so the evolution of GSM offering higher data rates and packet-oriented transfer is also presented. The chapter concludes with UMTS as a prominent example for 3G mobile telecommunication networks. UMTS is Europe's and Japan's proposal

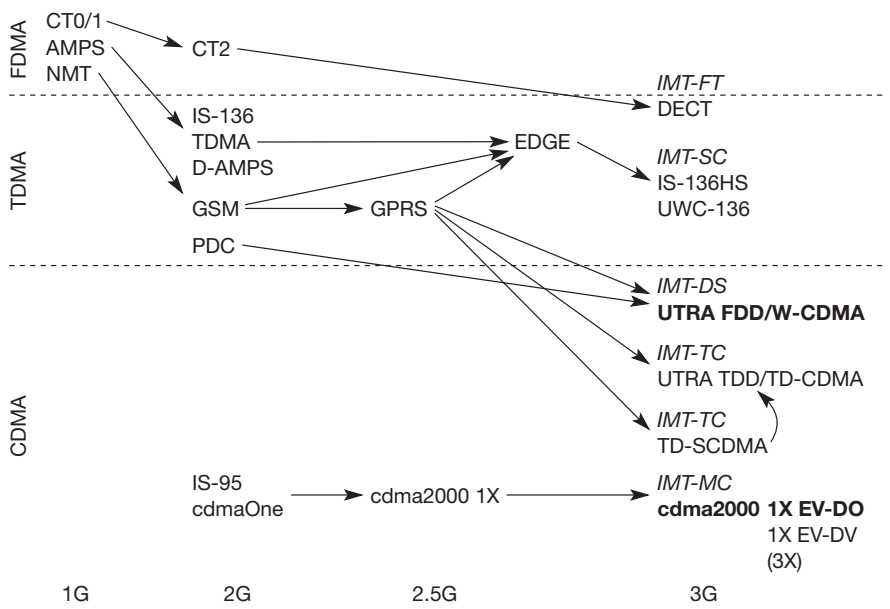

Figure 4.2 Development of different generations of mobile telecommunication systems

for the next generation mobile and wireless system within the ITU IMT-2000 framework. The early phases of UMTS show the evolutionary path from GSM via GSM with higher data rates to UMTS, which allows for saving a lot of investment into the infrastructure. Later phases of UMTS development show more and more the integration of Internet technology that simplifies service creation and offers a migration path to more advanced networks.

Figure 4.2 shows several development and migration paths for different mobile telecommunication systems presented in this chapter. The diagram is divided into the three main multiplexing schemes, FDMA, TDMA, and CDMA. The figure classifies the technologies into three generations. The first generation comprises analog systems, which typically rely on FDMA. The first 2G systems hit the market in the early nineties. In the US **D-AMPS** was a digital successor of **AMPS**, in Europe **GSM** was developed as a replacement for several versions of **NMT**, and **PDC** was introduced in Japan. All these 2G systems introduced a TDMA mechanism in addition to FDMA, which is still used for channel separation. With cdmaOne the first CDMA technology was available in the US as a competitor to the TDMA technologies. Between the second and third generation there is no real revolutionary step. The systems evolved over time: **GPRS** introduced a packet-oriented service and higher data rates to GSM (but can also be used for TDMA systems in general), **EDGE** proposes a new modulation scheme, and **cdmaOne** was enhanced to **cdma2000 1x** offering higher data rates. These systems are often called 2.5G systems.[2] Most, but not all, systems

2 Note that cdma2000 1x, the first version of cdma2000, was not accepted as 3G system by the ITU.

added CDMA technology to become 3G systems. Cordless telephone systems started with CT0 and CT1, became digital with CT2, and ended in Europe in the fully digital standard **DECT**. This standard has even been chosen as one of the candidates for a 3G system (IMT-FT).

While the number of different systems might be confusing, there are some "natural" development paths. Most network providers offering GSM service today will deploy UMTS, while cdmaOne users will choose cdma2000 for simpler migration. The reasons for this are quite simple. With the introduction of GPRS in GSM networks, the core of the network was already enhanced in a way that it can be directly used for UMTS with the radio technologies **UTRA FDD** and **UTRA TDD**. A similar path for evolution exists for **TD-SCDMA**, the Chinese proposal for a 3G system (which has been integrated into UTRA TDD). With some simplification it can be said that UMTS mainly adds a new radio interface but relies in its initial phase on the same core network as GSM/GPRS. Also for cdmaOne the evolution to cdma2000 technologies is quite natural, as the new standard is backward compatible and can reuse frequencies. Cdma2000 1x still uses the same 1.25 MHz channels as cdmaOne does, but offers data rates of up to 153 kbit/s. The **cdma2000 3x** standard uses three 1.25 MHz channels to fit into ITU's frequency scheme for 3G. However, this standard is not pushed as much as the following enhancements of cdma2000 1x. These enhancements are:

- **cdma2000 1x EV-DO** (evolution-data optimized, also known as high data rate (HDR), some call it data only) promising peak data rates of 2.4 Mbit/s using a second 1.25 MHz channel; and
- **cdma2000 1x EV-DV** (evolution-data and voice) aiming at 1.2 Mbit/s for mobile and 5.2 Mbit/s for stationary users.

Cdma2000 1x EV-DO was the first version of cdma2000 accepted by the ITU as 3G system. More information about the technologies and acronyms used in the diagram is provided in the following sections.

4.1 GSM

GSM is the most successful digital mobile telecommunication system in the world today. It is used by over 800 million people in more than 190 countries. In the early 1980s, Europe had numerous coexisting analog mobile phone systems, which were often based on similar standards (e.g., NMT 450), but ran on slightly different carrier frequencies. To avoid this situation for a second generation fully digital system, the **groupe spéciale mobile (GSM)** was founded in 1982. This system was soon named the **global system for mobile communications (GSM)**, with the specification process lying in the hands of ETSI (ETSI, 2002), (GSM Association, 2002). In the context of UMTS and the creation of 3GPP (Third generation partnership project, 3GPP, 2002a) the whole development process of GSM was transferred to 3GPP and further development is combined with 3G development. 3GPP assigned new numbers to all GSM stan-

dards. However, to remain consistent with most of the GSM literature, this GSM section stays with the original numbering (see 3GPP, 2002a, for conversion). Section 4.4 will present the ongoing joint specification process in more detail.

The primary goal of GSM was to provide a mobile phone system that allows users to roam throughout Europe and provides voice services compatible to ISDN and other PSTN systems. The specification for the initial system already covers more than 5,000 pages; new services, in particular data services, now add even more specification details. Readers familiar with the ISDN reference model will recognize many similar acronyms, reference points, and interfaces. GSM standardization aims at adopting as much as possible.

GSM is a typical second generation system, replacing the first generation analog systems, but not offering the high worldwide data rates that the third generation systems, such as UMTS, are promising. GSM has initially been deployed in Europe using 890–915 MHz for uplinks and 935–960 MHz for downlinks – this system is now also called **GSM 900** to distinguish it from the later versions. These versions comprise GSM at 1800 MHz (1710–1785 MHz uplink, 1805–1880 MHz downlink), also called **DCS (digital cellular system) 1800**, and the GSM system mainly used in the US at 1900 MHz (1850–1910 MHz uplink, 1930–1990 MHz downlink), also called **PCS (personal communications service) 1900**. Two more versions of GSM exist. **GSM 400** is a proposal to deploy GSM at 450.4–457.6/478.8–486 MHz for uplinks and 460.4–467.6/488.8–496 MHz for downlinks. This system could replace analog systems in sparsely populated areas.

A GSM system that has been introduced in several European countries for railroad systems is **GSM-Rail** (GSM-R, 2002), (ETSI, 2002). This system does not only use separate frequencies but offers many additional services which are unavailable using the public GSM system. GSM-R offers 19 exclusive channels for railroad operators for voice and data traffic (see section 4.1.3 for more information about channels). Special features of this system are, e.g., emergency calls with acknowledgements, voice group call service (VGCS), voice broadcast service (VBS). These so-called advanced speech call items (ASCI) resemble features typically available in trunked radio systems only (see section 4.3). Calls are prioritized: high priority calls pre-empt low priority calls. Calls have very short set-up times: emergency calls less than 2 s, group calls less than 5 s. Calls can be directed for example, to all users at a certain location, all users with a certain function, or all users within a certain number space. However, the most sophisticated use of GSM-R is the control of trains, switches, gates, and signals. Trains going not faster than 160 km/h can control all gates, switches, and signals themselves. If the train goes faster than 160 km/h (many trains are already capable of going faster than 300 km/h) GSM-R can still be used to maintain control.

The following section describes the architecture, services, and protocols of GSM that are common to all three major solutions, **GSM 900**, **GSM 1800**, and **GSM 1900**. GSM has mainly been designed for this and voice services and this still constitutes the main use of GSM systems. However, one can foresee that many future applications for mobile communications will be data driven. The relationship of data to voice traffic will shift more and more towards data.

4.1.1 Mobile services

GSM permits the integration of different voice and data services and the interworking with existing networks. Services make a network interesting for customers. GSM has defined three different categories of services: bearer, tele, and supplementary services. These are described in the following subsections. Figure 4.3 shows a reference model for GSM services. A **mobile station MS** is connected to the **GSM public land mobile network (PLMN)** via the U_m interface. (GSM-PLMN is the infrastructure needed for the GSM network.) This network is connected to transit networks, e.g., **integrated services digital network (ISDN)** or traditional **public switched telephone network (PSTN)**. There might be an additional network, the source/destination network, before another **terminal TE** is connected. **Bearer services** now comprise all services that enable the transparent transmission of data between the interfaces to the network, i.e., S in case of the mobile station, and a similar interface for the other terminal (e.g., S_0 for ISDN terminals). Interfaces like U, S, and R in case of ISDN have not been defined for all networks, so it depends on the specific network which interface is used as a reference for the transparent transmission of data. In the classical GSM model, bearer services are connection-oriented and circuit- or packet-switched. These services only need the lower three layers of the ISO/OSI reference model.

Within the mobile station MS, the **mobile termination (MT)** performs all network specific tasks (TDMA, FDMA, coding etc.) and offers an interface for data transmission (S) to the terminal TE which can then be network independent. Depending on the capabilities of TE, further interfaces may be needed, such as R, according to the ISDN reference model (Halsall, 1996). **Tele services** are application specific and may thus need all seven layers of the ISO/OSI reference model. These services are specified end-to-end, i.e., from one terminal TE to another.

4.1.1.1 Bearer services

GSM specifies different mechanisms for data transmission, the original GSM allowing for data rates of up to 9600 bit/s for non-voice services. Bearer services permit transparent and non-transparent, synchronous or asynchronous data transmission. **Transparent bearer services** only use the functions of the physical layer (layer 1) to transmit data. Data transmission has a constant delay and throughput if no transmission errors occur. The only mechanism to increase

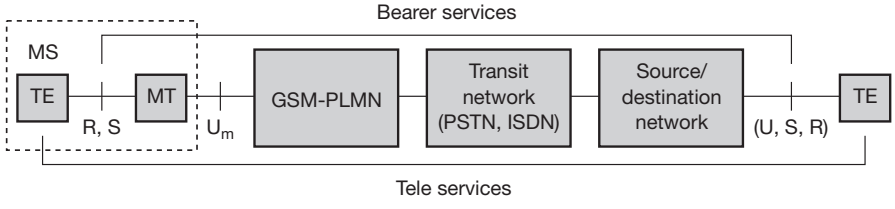

Figure 4.3 Bearer and tele services reference model

transmission quality is the use of **forward error correction (FEC)**, which codes redundancy into the data stream and helps to reconstruct the original data in case of transmission errors. Depending on the FEC, data rates of 2.4, 4.8, or 9.6 kbit/s are possible. Transparent bearer services do not try to recover lost data in case of, for example, shadowing or interruptions due to handover.

Non-transparent bearer services use protocols of layers two and three to implement error correction and flow control. These services use the transparent bearer services, adding a **radio link protocol (RLP)**. This protocol comprises mechanisms of **high-level data link control (HDLC)**, (Halsall, 1996) and special selective-reject mechanisms to trigger retransmission of erroneous data. The achieved bit error rate is less than 10^{-7}, but now throughput and delay may vary depending on transmission quality.

Using transparent and non-transparent services, GSM specifies several bearer services for interworking with PSTN, ISDN, and packet switched public data networks (PSPDN) like X.25, which is available worldwide. Data transmission can be full-duplex, synchronous with data rates of 1.2, 2.4, 4.8, and 9.6 kbit/s or full-duplex, asynchronous from 300 to 9,600 bit/s (ETSI, 1991a). Clearly, these relatively low data rates reflect the assumption that data services will only constitute some small percentage of the overall traffic. While this is still true of GSM networks today, the relation of data and voice services is changing, with data becoming more and more important. This development is also reflected in the new data services (see section 4.1.8).

4.1.1.2 Tele services

GSM mainly focuses on voice-oriented tele services. These comprise encrypted voice transmission, message services, and basic data communication with terminals as known from the PSTN or ISDN (e.g., fax). However, as the main service is **telephony**, the primary goal of GSM was the provision of high-quality digital voice transmission, offering at least the typical bandwidth of 3.1 kHz of analog phone systems. Special codecs (coder/decoder) are used for voice transmission, while other codecs are used for the transmission of analog data for communication with traditional computer modems used in, e.g., fax machines.

Another service offered by GSM is the **emergency number**. The same number can be used throughout Europe. This service is mandatory for all providers and free of charge. This connection also has the highest priority, possibly pre-empting other connections, and will automatically be set up with the closest emergency center.

A useful service for very simple message transfer is the **short message service (SMS)**, which offers transmission of messages of up to 160 characters. SMS messages do not use the standard data channels of GSM but exploit unused capacity in the signalling channels (see section 4.1.3.1). Sending and receiving of SMS is possible during data or voice transmission. SMS was in the GSM standard from the beginning; however, almost no one used it until millions of young people discovered this service in the mid-nineties as a fun service. SMS

can be used for "serious" applications such as displaying road conditions, e-mail headers or stock quotes, but it can also transferr logos, ring tones, horoscopes and love letters. Today more than 30 billion short messages are transferred worldwide per month! SMS is big business today, not only for the network operators, but also for many content providers. It should be noted that SMS is typically the only way to reach a mobile phone from within the network. Thus, SMS is used for updating mobile phone software or for implementing so-called push services (see chapter 10).

The successor of SMS, the **enhanced message service (EMS)**, offers a larger message size (e.g., 760 characters, concatenating several SMs), formatted text, and the transmission of animated pictures, small images and ring tones in a standardized way (some vendors offered similar proprietary features before). EMS never really took off as the **multimedia message service (MMS)** was available. (Nokia never liked EMS but pushed Smart Messaging, a proprietary system.) MMS offers the transmission of larger pictures (GIF, JPG, WBMP), short video clips etc. and comes with mobile phones that integrate small cameras. MMS is further discussed in the context of WAP in chapter 10.

Another non-voice tele service is **group 3 fax**, which is available worldwide. In this service, fax data is transmitted as digital data over the analog telephone network according to the ITU-T standards T.4 and T.30 using modems. Typically, a transparent fax service is used, i.e., fax data and fax signaling is transmitted using a transparent bearer service. Lower transmission quality causes an automatic adaptation of the bearer service to lower data rates and higher redundancy for better FEC.

4.1.1.3 Supplementary services

In addition to tele and bearer services, GSM providers can offer **supplementary services**. Similar to ISDN networks, these services offer various enhancements for the standard telephony service, and may vary from provider to provider. Typical services are user **identification**, call **redirection**, or **forwarding** of ongoing calls. Standard ISDN features such as **closed user groups** and **multiparty** communication may be available. Closed user groups are of special interest to companies because they allow, for example, a company-specific GSM sub-network, to which only members of the group have access.

4.1.2 System architecture

As with all systems in the telecommunication area, GSM comes with a hierarchical, complex system architecture comprising many entities, interfaces, and acronyms. Figure 4.4 gives a simplified overview of the GSM system as specified in ETSI (1991b). A GSM system consists of three subsystems, the **radio sub system (RSS)**, the **network and switching subsystem (NSS)**, and the **operation subsystem (OSS)**. Each subsystem will be discussed in more detail in the following sections. Generally, a GSM customer only notices a very small fraction of the whole network – the mobile stations (MS) and some antenna masts of the base transceiver stations (BTS).

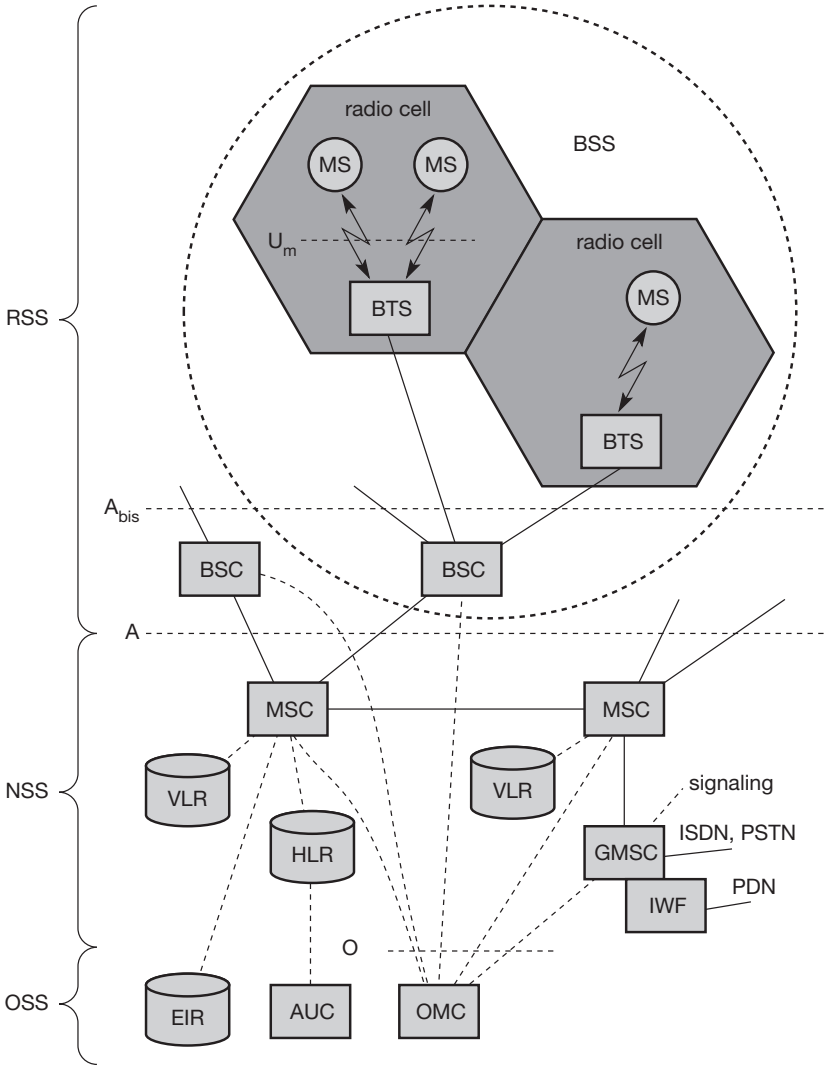

Figure 4.4
Functional architecture of a GSM system

4.1.2.1 Radio subsystem

As the name implies, the **radio subsystem (RSS)** comprises all radio specific entities, i.e., the **mobile stations (MS)** and the **base station subsystem (BSS)**. Figure 4.4 shows the connection between the RSS and the NSS via the **A interface** (solid lines) and the connection to the OSS via the **O interface** (dashed lines). The A interface is typically based on circuit-switched PCM-30 systems (2.048 Mbit/s), carrying up to 30 64 kbit/s connections, whereas the O interface uses the Signalling System No. 7 (SS7) based on X.25 carrying management data to/from the RSS.

- **Base station subsystem (BSS)**: A GSM network comprises many BSSs, each controlled by a base station controller (BSC). The BSS performs all functions necessary to maintain radio connections to an MS, coding/decoding of voice, and rate adaptation to/from the wireless network part. Besides a BSC, the BSS contains several BTSs.
- **Base transceiver station (BTS)**: A BTS comprises all radio equipment, i.e., antennas, signal processing, amplifiers necessary for radio transmission. A BTS can form a radio cell or, using sectorized antennas, several cells (see section 2.8), and is connected to MS via the **U_m interface** (ISDN U interface for mobile use), and to the BSC via the **A_{bis} interface**. The U_m interface contains all the mechanisms necessary for wireless transmission (TDMA, FDMA etc.) and will be discussed in more detail below. The A_{bis} interface consists of 16 or 64 kbit/s connections. A GSM cell can measure between some 100 m and 35 km depending on the environment (buildings, open space, mountains etc.) but also expected traffic.
- **Base station controller (BSC)**: The BSC basically manages the BTSs. It reserves radio frequencies, handles the handover from one BTS to another within the BSS, and performs paging of the MS. The BSC also multiplexes the radio channels onto the fixed network connections at the A interface.

Table 4.1 gives an overview of the tasks assigned to the BSC and BTS or of tasks in which these entities support other entities in the network.

- **Mobile station (MS)**: The MS comprises all user equipment and software needed for communication with a GSM network. An MS consists of user independent hard- and software and of the **subscriber identity module (SIM)**, which stores all user-specific data that is relevant to GSM.[3] While an MS can be identified via the **international mobile equipment identity (IMEI)**, a user can personalize any MS using his or her SIM, i.e., user-specific mechanisms like charging and authentication are based on the SIM, not on the device itself. Device-specific mechanisms, e.g., theft protection, use the device specific IMEI. Without the SIM, only emergency calls are possible. The SIM card contains many identifiers and tables, such as card-type, serial number, a list of subscribed services, a **personal identity number (PIN)**, a **PIN unblocking key (PUK)**, an **authentication key K_i**, and the **international mobile subscriber identity (IMSI)** (ETSI, 1991c). The PIN is used to unlock the MS. Using the wrong PIN three times will lock the SIM. In such cases, the PUK is needed to unlock the SIM. The MS stores dynamic information while logged onto the GSM system, such as, e.g., the **cipher key K_c** and the location information consisting of a **temporary mobile subscriber identity (TMSI)** and the **location area identification (LAI)**. Typical MSs for GSM 900 have a transmit power of up to 2 W, whereas for GSM 1800 1 W is enough due to the smaller cell size. Apart from the telephone interface, an

3 Many additional items can be stored on the mobile device. However, this is irrelevant to GSM.

Function	BTS	BSC
Management of radio channels		X
Frequency hopping	X	X
Management of terrestrial channels		X
Mapping of terrestrial onto radio channels		X
Channel coding and decoding	X	
Rate adaptation	X	
Encryption and decryption	X	X
Paging	X	X
Uplink signal measurement	X	
Traffic measurement		X
Authentication		X
Location registry, location update		X
Handover management		X

Table 4.1 Tasks of the BTS and BSC within a BSS

MS can also offer other types of interfaces to users with display, loudspeaker, microphone, and programmable soft keys. Further interfaces comprise computer modems, IrDA, or Bluetooth. Typical MSs, e.g., mobile phones, comprise many more vendor-specific functions and components, such as cameras, fingerprint sensors, calendars, address books, games, and Internet browsers. Personal digital assistants (PDA) with mobile phone functions are also available. The reader should be aware that an MS could also be integrated into a car or be used for location tracking of a container.

4.1.2.2 Network and switching subsystem

The "heart" of the GSM system is formed by the **network and switching subsystem (NSS)**. The NSS connects the wireless network with standard public networks, performs handovers between different BSSs, comprises functions for worldwide localization of users and supports charging, accounting, and roaming of users between different providers in different countries. The NSS consists of the following switches and databases:

- **Mobile services switching center (MSC)**: MSCs are high-performance digital ISDN switches. They set up connections to other MSCs and to the BSCs via the A interface, and form the fixed backbone network of a GSM system. Typically, an MSC manages several BSCs in a geographical region. A **gateway MSC (GMSC)** has additional connections to other fixed networks, such as **PSTN** and **ISDN**. Using additional **interworking functions (IWF)**, an MSC

can also connect to **public data networks (PDN)** such as X.25. An MSC handles all signaling needed for connection setup, connection release and handover of connections to other MSCs. The **standard signaling system No. 7 (SS7)** is used for this purpose. SS7 covers all aspects of control signaling for digital networks (reliable routing and delivery of control messages, establishing and monitoring of calls). Features of SS7 are number portability, free phone/toll/collect/credit calls, call forwarding, three-way calling etc. An MSC also performs all functions needed for supplementary services such as call forwarding, multi-party calls, reverse charging etc.

- **Home location register (HLR)**: The HLR is the most important database in a GSM system as it stores all user-relevant information. This comprises static information, such as the **mobile subscriber ISDN number (MSISDN)**, subscribed services (e.g., call forwarding, roaming restrictions, GPRS), and the **international mobile subscriber identity (IMSI)**. Dynamic information is also needed, e.g., the current **location area (LA)** of the MS, the **mobile subscriber roaming number (MSRN)**, the current VLR and MSC. As soon as an MS leaves its current LA, the information in the HLR is updated. This information is necessary to localize a user in the worldwide GSM network. All these user-specific information elements only exist once for each user in a single HLR, which also supports charging and accounting. The parameters will be explained in more detail in section 4.1.5. HLRs can manage data for several million customers and contain highly specialized data bases which must fulfill certain real-time requirements to answer requests within certain time-bounds.
- **Visitor location register (VLR)**: The VLR associated to each MSC is a dynamic database which stores all important information needed for the MS users currently in the LA that is associated to the MSC (e.g., IMSI, MSISDN, HLR address). If a new MS comes into an LA the VLR is responsible for, it copies all relevant information for this user from the HLR. This hierarchy of VLR and HLR avoids frequent HLR updates and long-distance signaling of user information. The typical use of HLR and VLR for user localization will be described in section 4.1.5. Some VLRs in existence, are capable of managing up to one million customers.

4.1.2.3 Operation subsystem

The third part of a GSM system, the **operation subsystem (OSS)**, contains the necessary functions for network operation and maintenance. The OSS possesses network entities of its own and accesses other entities via SS7 signaling (see Figure 4.4). The following entities have been defined:

- **Operation and maintenance center (OMC)**: The OMC monitors and controls all other network entities via the O interface (SS7 with X.25). Typical OMC management functions are traffic monitoring, status reports of network entities, subscriber and security management, or accounting and billing. OMCs use the concept of **telecommunication management network (TMN)** as standardized by the ITU-T.

- **Authentication centre (AuC)**: As the radio interface and mobile stations are particularly vulnerable, a separate AuC has been defined to protect user identity and data transmission. The AuC contains the algorithms for authentication as well as the keys for encryption and generates the values needed for user authentication in the HLR. The AuC may, in fact, be situated in a special protected part of the HLR.
- **Equipment identity register (EIR)**: The EIR is a database for all IMEIs, i.e., it stores all device identifications registered for this network. As MSs are mobile, they can be easily stolen. With a valid SIM, anyone could use the stolen MS. The EIR has a blacklist of stolen (or locked) devices. In theory an MS is useless as soon as the owner has reported a theft. Unfortunately, the blacklists of different providers are not usually synchronized and the illegal use of a device in another operator's network is possible (the reader may speculate as to why this is the case). The EIR also contains a list of valid IMEIs (white list), and a list of malfunctioning devices (gray list).

4.1.3 Radio interface

The most interesting interface in a GSM system is U_m, the radio interface, as it comprises many mechanisms presented in chapters 2 and 3 for multiplexing and media access. GSM implements SDMA using cells with BTS and assigns an MS to a BTS. Furthermore, FDD is used to separate downlink and uplink as shown in Figures 3.3 and 4.5. Media access combines TDMA and FDMA. In GSM 900, 124 channels, each 200 kHz wide, are used for FDMA, whereas GSM 1800 uses, 374 channels. Due to technical reasons, channels 1 and 124 are not used for transmission in GSM 900. Typically, 32 channels are reserved for organizational data; the remaining 90 are used for customers. Each BTS then manages a single channel for organizational data and, e.g., up to 10 channels for user data. The following example is based on the GSM 900 system, but GSM works in a similar way at 1800 and 1900 MHz.

While Figure 3.3 in chapter 3 has already shown the FDM in GSM, Figure 4.5 also shows the TDM used. Each of the 248 channels is additionally separated in time via a **GSM TDMA frame**, i.e., each 200 kHz carrier is subdivided into frames that are repeated continuously. The duration of a frame is 4.615 ms. A frame is again subdivided into 8 **GSM time slots**, where each slot represents a physical TDM channel and lasts for 577 μs. Each TDM channel occupies the 200 kHz carrier for 577 μs every 4.615 ms.

Data is transmitted in small portions, called **bursts**. Figure 4.5 shows a so-called **normal burst** as used for data transmission inside a time slot (user and signaling data). In the diagram, the burst is only 546.5 μs long and contains 148 bits. The remaining 30.5 μs are used as **guard space** to avoid overlapping with other bursts due to different path delays and to give the transmitter time to turn on and off. Filling the whole slot with data allows for the transmission of

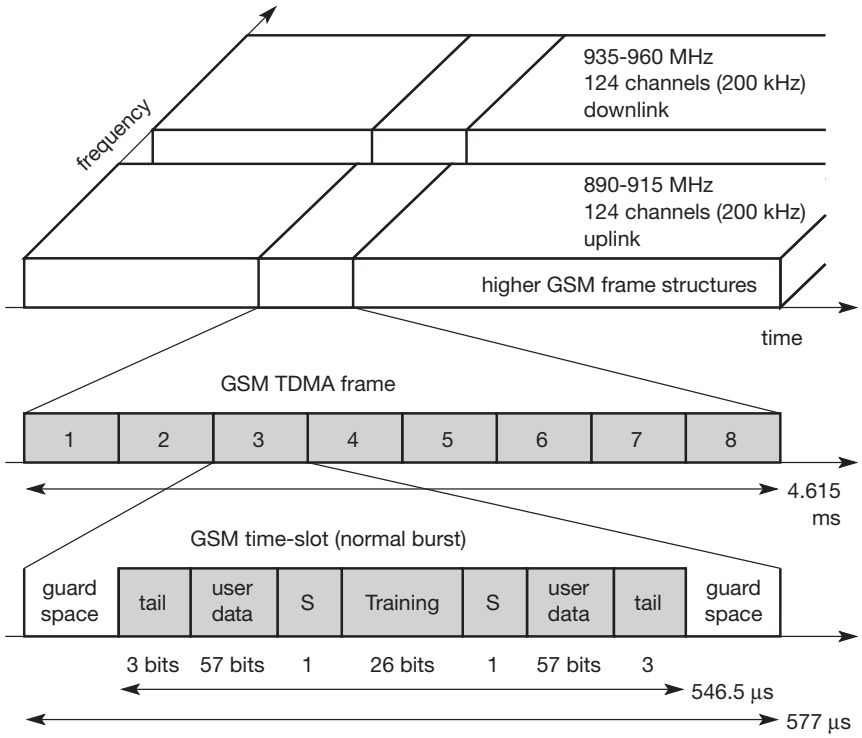

Figure 4.5 GSM TDMA frame, slots, and bursts

156.25 bit within 577 μs. Each physical TDM channel has a raw data rate of about 33.8 kbit/s, each radio carrier transmits approximately 270 kbit/s over the U_m interface.

The first and last three bits of a normal burst (**tail**) are all set to 0 and can be used to enhance the receiver performance. The **training** sequence in the middle of a slot is used to adapt the parameters of the receiver to the current path propagation characteristics and to select the strongest signal in case of multi-path propagation. A flag **S** indicates whether the **data** field contains user or network control data. Apart from the normal burst, ETSI (1993a) defines four more bursts for data transmission: a **frequency correction** burst allows the MS to correct the local oscillator to avoid interference with neighboring channels, a **synchronization burst** with an extended training sequence synchronizes the MS with the BTS in time, an **access burst** is used for the initial connection setup between MS and BTS, and finally a **dummy burst** is used if no data is available for a slot.

Two factors allow for the use of simple transmitter hardware: on the one hand, the slots for uplink and downlink of a physical TDM channel are separated in frequency (45 MHz for GSM 900, 95 MHz for GSM 1800 using FDD). On the other hand, the TDMA frames are shifted in time for three slots, i.e., if the BTS sends data at time t_0 in slot one on the downlink, the MS accesses slot

one on the uplink at time t_0+3·577 μs. An MS does not need a full-duplex transmitter, a simpler half-duplex transmitter switching between receiving and sending is enough.

To avoid frequency selective fading, GSM specifies an optional **slow frequency hopping** mechanism. MS and BTS may change the carrier frequency after each frame based on a common hopping sequence. An MS changes its frequency between up and downlink slots respectively.

4.1.3.1 Logical channels and frame hierarchy

While the previous section showed the physical separation of the medium into 8*124 duplex channels, this section presents logical channels and a hierarchy of frames based on the combination of these physical channels. A physical channel consists of a slot, repeated every 4.615 ms. Think of a logical channel C_1 that only takes up every fourth slot and another logical channel C_2 that uses every other slot. Both logical channels could use the same physical channel with the pattern $C_1C_2xC_2C_1C_2xC_2C_1$ etc. (The x indicates that the physical channel still has some capacity left.)

GSM specifies two basic groups of logical channels, i.e., traffic channels and control channels:[4]

- **Traffic channels (TCH)**: GSM uses a TCH to transmit user data (e.g., voice, fax). Two basic categories of TCHs have been defined, i.e., **full-rate TCH (TCH/F)** and **half-rate TCH (TCH/H)**. A TCH/F has a data rate of 22.8 kbit/s, whereas TCH/H only has 11.4 kbit/s. With the voice codecs available at the beginning of the GSM standardization, 13 kbit/s were required, whereas the remaining capacity of the TCH/F (22.8 kbit/s) was used for error correction (**TCH/FS**). Improved codes allow for better voice coding and can use a TCH/H. Using these TCH/HSs doubles the capacity of the GSM system for voice transmission. However, speech quality decreases with the use of TCH/HS and many providers try to avoid using them. The standard codecs for voice are called **full rate** (FR, 13 kbit/s) and **half rate** (HR, 5.6 kbit/s). A newer codec, **enhanced full rate** (EFR), provides better voice quality than FR as long as the transmission error rate is low. The generated data rate is only 12.2 kbit/s. New codecs, which automatically choose the best mode of operation depending on the error rate (AMR, adaptive multi-rate), will be used together with 3G systems. An additional increase in voice quality is provided by the so-called **tandem free operation (TFO)**. This mode can be used if two MSs exchange voice data. In this case, coding to and from PCM encoded voice (standard in ISDN) can be skipped and the GSM encoded voice data is directly exchanged. Data transmission in GSM is possible at many different data rates, e.g., **TCH/F4.8** for 4.8 kbit/s, **TCH/F9.6** for 9.6 kbit/s, and, as a newer specification, **TCH/F14.4** for 14.4 kbit/s. These logical channels differ in terms of their coding schemes and error correction capabilities.

4 More information about channels can be found in Goodman (1997) and ETSI (1993a).

- **Control channels (CCH)**: Many different CCHs are used in a GSM system to control medium access, allocation of traffic channels or mobility management. Three groups of control channels have been defined, each again with subchannels (maybe you can imagine why the initial specification already needed over 5,000 pages):
  - **Broadcast control channel (BCCH)**: A BTS uses this channel to signal information to all MSs within a cell. Information transmitted in this channel is, e.g., the cell identifier, options available within this cell (frequency hopping), and frequencies available inside the cell and in neighboring cells. The BTS sends information for frequency correction via the **frequency correction channel (FCCH)** and information about time synchronization via the **synchronization channel (SCH)**, where both channels are subchannels of the BCCH.
  - **Common control channel (CCCH)**: All information regarding connection setup between MS and BS is exchanged via the CCCH. For calls toward an MS, the BTS uses the **paging channel (PCH)** for paging the appropriate MS. If an MS wants to set up a call, it uses the **random access channel (RACH)** to send data to the BTS. The RACH implements multiple access (all MSs within a cell may access this channel) using slotted Aloha. This is where a collision may occur with other MSs in a GSM system. The BTS uses the **access grant channel (AGCH)** to signal an MS that it can use a TCH or SDCCH for further connection setup.
  - **Dedicated control channel (DCCH)**: While the previous channels have all been unidirectional, the following channels are bidirectional. As long as an MS has not established a TCH with the BTS, it uses the **stand-alone dedicated control channel (SDCCH)** with a low data rate (782 bit/s) for signaling. This can comprise authentication, registration or other data needed for setting up a TCH. Each TCH and SDCCH has a **slow associated dedicated control channel (SACCH)** associated with it, which is used to exchange system information, such as the channel quality and signal power level. Finally, if more signaling information needs to be transmitted and a TCH already exists, GSM uses a **fast associated dedicated control channel (FACCH)**. The FACCH uses the time slots which are otherwise used by the TCH. This is necessary in the case of handovers where BTS and MS have to exchange larger amounts of data in less time.

However, these channels cannot use time slots arbitrarily – GSM specifies a very elaborate multiplexing scheme that integrates several hierarchies of frames. If we take a simple TCH/F for user data transmission, each TCH/F will have an associated SACCH for slow signaling. If fast signaling is required, the FACCH uses the time slots for the TCH/F. A typical usage pattern of a physical channel for data transmission now looks like this (with T indicating the user traffic in the TCH/F and S indicating the signalling traffic in the SACCH):

TTTTTTTTTTTTSTTTTTTTTTTTTx

TTTTTTTTTTTTSTTTTTTTTTTTTx

Twelve slots with user data are followed by a signalling slot. Again 12 slots with user data follow, then an unused slot. This pattern of 26 slots is repeated over and over again. In this case, only 24 out of 26 physical slots are used for the TCH/F. Now recall that each normal burst used for data transmission carries 114 bit user data and is repeated every 4.615 ms. This results in a data rate of 24.7 kbit/s. As the TCH/F only uses 24/26 of the slots, the final data rate is 22.8 kbit/s as specified for the TCH/F. The SACCH thus has a capacity of 950 bit/s.

This periodic pattern of 26 slots occurs in all TDMA frames with a TCH. The combination of these frames is called **traffic multiframe**. Figure 4.6 shows the logical combination of 26 frames (TDMA frames with a duration of 4.615 ms) to a multiframe with a duration of 120 ms. This type of multiframe is used for TCHs, SACCHs for TCHs, and FACCHs. As these logical channels are all associated with user traffic, the multiframe is called traffic multiframe. TDMA frames containing (signaling) data for the other logical channels are combined to a **control multiframe**. Control multiframes consist of 51 TDMA frames and have a duration of 235.4 ms.

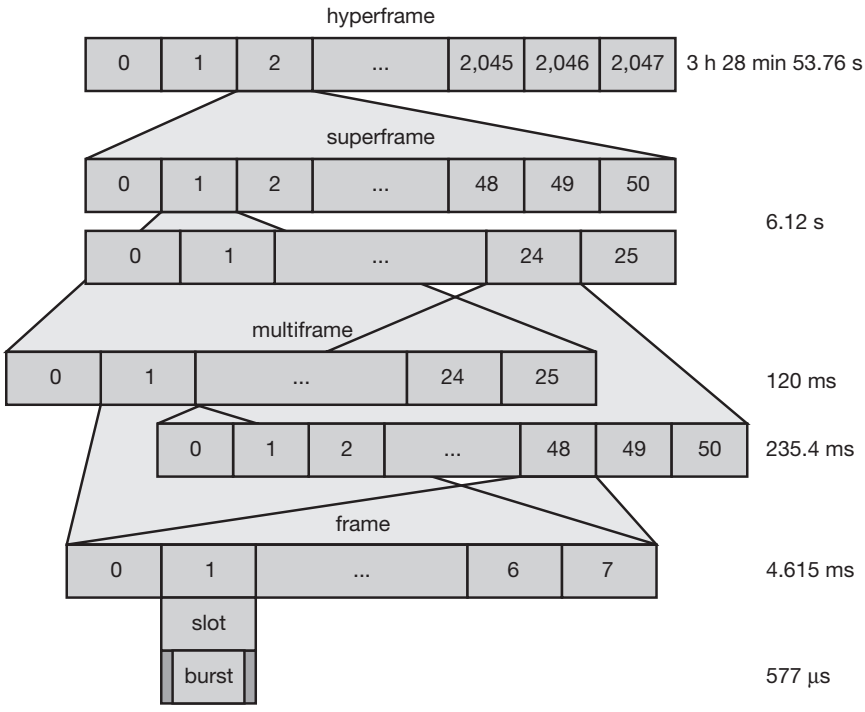

Figure 4.6
GSM structuring of time using a frame hierarchy

This logical frame hierarchy continues, combining 26 multiframes with 51 frames or 51 multiframes with 26 frames to form a **superframe**. 2,048 superframes build a **hyperframe** with a duration of almost 3.5 hours. Altogether, 2,715,648 TDMA frames form a hyperframe. This large logical structure is needed for encryption – GSM counts each TDMA frame, with the frame number forming input for the encryption algorithm. The frame number plus the slot number uniquely identify each time slot in GSM.

4.1.4 Protocols

Figure 4.7 shows the protocol architecture of GSM with signaling protocols, interfaces, as well as the entities already shown in Figure 4.4. The main interest lies in the U_m interface, as the other interfaces occur between entities in a fixed network. **Layer 1**, the physical layer, handles all **radio**-specific functions. This includes the creation of bursts according to the five different formats, **multiplexing** of bursts into a TDMA frame, **synchronization** with the BTS, detection of idle channels, and measurement of the **channel quality** on the downlink. The physical layer at U_m uses GMSK for digital **modulation** and performs **encryption/decryption** of data, i.e., encryption is not performed end-to-end, but only between MS and BSS over the air interface.

Synchronization also includes the correction of the individual path delay between an MS and the BTS. All MSs within a cell use the same BTS and thus must be synchronized to this BTS. The BTS generates the time-structure of frames, slots etc. A problematic aspect in this context are the different round trip times (RTT). An MS close to the BTS has a very short RTT, whereas an MS 35 km away already exhibits an RTT of around 0.23 ms. If the MS far away used the slot structure with-

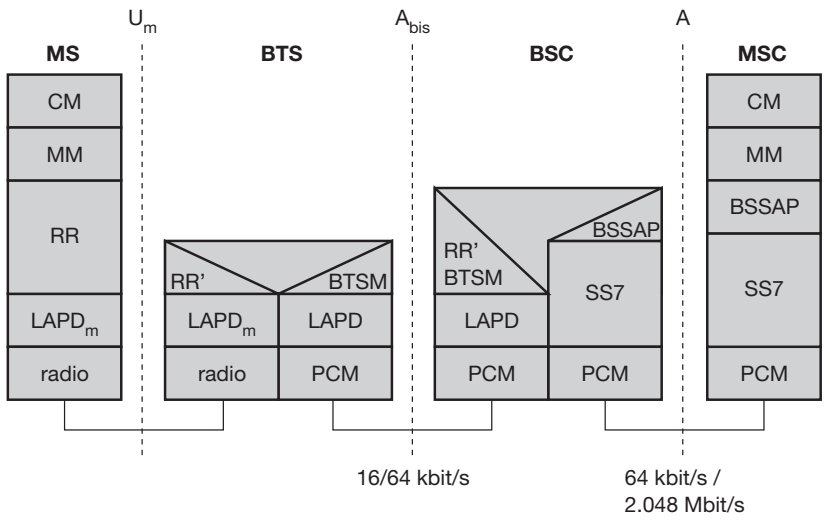

Figure 4.7 Protocol architecture for signaling

out correction, large guard spaces would be required, as 0.23 ms are already 40 per cent of the 0.577 ms available for each slot. Therefore, the BTS sends the current RTT to the MS, which then adjusts its access time so that all bursts reach the BTS within their limits. This mechanism reduces the guard space to only 30.5 μs or five per cent (see Figure 4.5). Adjusting the access is controlled via the variable **timing advance**, where a burst can be shifted up to 63 bit times earlier, with each bit having a duration of 3.69 μs (which results in the 0.23 ms needed). As the variable timing advance cannot be extended a burst cannot be shifted earlier than 63 bit times. This results in the 35 km maximum distance between an MS and a BTS. It might be possible to receive the signals over longer distances; to avoid collisions at the BTS, access cannot be allowed.[5]

The main tasks of the physical layer comprise **channel coding** and **error detection/correction**, which is directly combined with the coding mechanisms. Channel coding makes extensive use of different **forward error correction (FEC)** schemes. FEC adds redundancy to user data, allowing for the detection and correction of selected errors. The power of an FEC scheme depends on the amount of redundancy, coding algorithm and further interleaving of data to minimize the effects of burst errors. The FEC is also the reason why error detection and correction occurs in layer one and not in layer two as in the ISO/OSI reference model. The GSM physical layer tries to correct errors, but it does not deliver erroneous data to the higher layer.

Different logical channels of GSM use different coding schemes with different correction capabilities. Speech channels need additional coding of voice data after analog to digital conversion, to achieve a data rate of 22.8 kbit/s (using the 13 kbit/s from the voice codec plus redundancy, CRC bits, and interleaving (Goodman, 1997). As voice was assumed to be the main service in GSM, the physical layer also contains special functions, such as **voice activity detection (VAD)**, which transmits voice data only when there is a voice signal. This mechanism helps to decrease interference as a channel might be silent approximately 60 per cent of the time (under the assumption that only one person speaks at the same time and some extra time is needed to switch between the speakers). During periods of silence (e.g., if a user needs time to think before talking), the physical layer generates a **comfort noise** to fake a connection (complete silence would probably confuse a user), but no actual transmission takes place. The noise is even adapted to the current background noise at the communication partner's location.

All this interleaving of data for a channel to minimize interference due to burst errors and the recurrence pattern of a logical channel generates a **delay** for transmission. The delay is about 60 ms for a TCH/FS and 100 ms for a TCH/F9.6

5 A special trick allows for larger cells. If the timing advance for MSs that are further away than 35 km is set to zero, the bursts arriving from these MSs will fall into the following time slot. Reception of data is simply shifted one time slot and again the timing advance may be used up to a distance of 70 km (under simplified assumptions). Using this special trick, the capacity of a cell is decreased (near and far MSs cannot be mixed arbitrarily), but coverage of GSM is extended. Network operators may choose this approach, e.g., in coastal regions.

(within 100 ms signals in fixed networks easily travel around the globe). These times have to be added to the transmission delay if communicating with an MS instead of a standard fixed station (telephone, computer etc.) and may influence the performance of any higher layer protocols, e.g., for computer data transmission (see chapter 9).

Signaling between entities in a GSM network requires higher layers (see Figure 4.7). For this purpose, the **$LAPD_m$** protocol has been defined at the U_m interface for **layer two**. $LAPD_m$, as the name already implies, has been derived from link access procedure for the D-channel (**LAPD**) in ISDN systems, which is a version of HDLC (Goodman, 1997), (Halsall, 1996). $LAPD_m$ is a lightweight LAPD because it does not need synchronization flags or checksumming for error detection. (The GSM physical layer already performs these tasks.) $LAPD_m$ offers reliable data transfer over connections, re-sequencing of data frames, and flow control (ETSI, 1993b), (ETSI, 1993c). As there is no buffering between layer one and two, $LAPD_m$ has to obey the frame structures, recurrence patterns etc. defined for the U_m interface. Further services provided by $LAPD_m$ include segmentation and reassembly of data and acknowledged/unacknowledged data transfer.

The network layer in GSM, **layer three**, comprises several sublayers as Figure 4.7 shows. The lowest sublayer is the **radio resource management (RR)**. Only a part of this layer, **RR'**, is implemented in the BTS, the remainder is situated in the BSC. The functions of RR' are supported by the BSC via the **BTS management (BTSM)**. The main tasks of RR are setup, maintenance, and release of radio channels. RR also directly accesses the physical layer for radio information and offers a reliable connection to the next higher layer.

Mobility management (MM) contains functions for registration, authentication, identification, location updating, and the provision of a **temporary mobile subscriber identity (TMSI)** that replaces the **international mobile subscriber identity (IMSI)** and which hides the real identity of an MS user over the air interface. While the IMSI identifies a user, the TMSI is valid only in the current location area of a VLR. MM offers a reliable connection to the next higher layer.

Finally, the **call management (CM)** layer contains three entities: **call control (CC)**, **short message service (SMS)**, and **supplementary service (SS)**. SMS allows for message transfer using the control channels SDCCH and SACCH (if no signaling data is sent), while SS offers the services described in section 4.1.1.3. CC provides a point-to-point connection between two terminals and is used by higher layers for call establishment, call clearing and change of call parameters. This layer also provides functions to send in-band tones, called **dual tone multiple frequency (DTMF)**, over the GSM network. These tones are used, e.g., for the remote control of answering machines or the entry of PINs in electronic banking and are, also used for dialing in traditional analog telephone systems. These tones cannot be sent directly over the voice codec of a GSM MS, as the codec would distort the tones. They are transferred as signals and then converted into tones in the fixed network part of the GSM system.

Additional protocols are used at the A_{bis} and A interfaces (the internal interfaces of a GSM system not presented here). Data transmission at the physical layer typically uses **pulse code modulation (PCM)** systems. While PCM systems offer transparent 64 kbit/s channels, GSM also allows for the submultiplexing of four 16 kbit/s channels into a single 64 kbit/s channel (16 kbit/s are enough for user data from an MS). The physical layer at the A interface typically includes leased lines with 2.048 Mbit/s capacity. LAPD is used for layer two at A_{bis}, BTSM for BTS management.

Signaling system No. 7 (SS7) is used for signaling between an MSC and a BSC. This protocol also transfers all management information between MSCs, HLR, VLRs, AuC, EIR, and OMC. An MSC can also control a BSS via a **BSS application part (BSSAP)**.

4.1.5 Localization and calling

One fundamental feature of the GSM system is the automatic, worldwide localization of users. The system always knows where a user currently is, and the same phone number is valid worldwide. To provide this service, GSM performs periodic location updates even if a user does not use the mobile station (provided that the MS is still logged into the GSM network and is not completely switched off). The HLR always contains information about the current location (only the location area, not the precise geographical location), and the VLR currently responsible for the MS informs the HLR about location changes. As soon as an MS moves into the range of a new VLR (a new location area), the HLR sends all user data needed to the new VLR. Changing VLRs with uninterrupted availability of all services is also called **roaming**. Roaming can take place within the network of one provider, between two providers in one country (national roaming is, often not supported due to competition between operators), but also between different providers in different countries (international roaming). Typically, people associate international roaming with the term roaming as it is this type of roaming that makes GSM very attractive: one device, over 190 countries!

To locate an MS and to address the MS, several numbers are needed:

- **Mobile station international ISDN number (MSISDN):**[6] The only important number for a user of GSM is the phone number. Remember that the phone number is not associated with a certain device but with the SIM, which is personalized for a user. The MSISDN follows the ITU-T standard E.164 for addresses as it is also used in fixed ISDN networks. This number consists of the **country code (CC)** (e.g., +49 179 1234567 with 49 for Germany), the **national destination code (NDC)** (i.e., the address of the network provider, e.g., 179), and the **subscriber number (SN)**.

6 In other types of documentation, this number is also called 'Mobile Subscriber ISDN Number' or 'Mobile Station ISDN Number'. Even the original ETSI standards use different wordings for the same acronym. However, the term 'subscriber' is much better suited as it expresses the independence of the user related number from the device (station).

- **International mobile subscriber identity (IMSI):** GSM uses the IMSI for internal unique identification of a subscriber. IMSI consists of a **mobile country code (MCC)** (e.g., 240 for Sweden, 208 for France), the **mobile network code (MNC)** (i.e., the code of the network provider), and finally the **mobile subscriber identification number (MSIN)**.
- **Temporary mobile subscriber identity (TMSI):** To hide the IMSI, which would give away the exact identity of the user signaling over the air interface, GSM uses the 4 byte TMSI for local subscriber identification. TMSI is selected by the current VLR and is only valid temporarily and within the location area of the VLR (for an ongoing communication TMSI and LAI are sufficient to identify a user; the IMSI is not needed). Additionally, a VLR may change the TMSI periodically.
- **Mobile station[7] roaming number (MSRN):** Another temporary address that hides the identity and location of a subscriber is MSRN. The VLR generates this address on request from the MSC, and the address is also stored in the HLR. MSRN contains the current **visitor country code (VCC)**, the **visitor national destination code (VNDC)**, the identification of the current MSC together with the subscriber number. The MSRN helps the HLR to find a subscriber for an incoming call.

All these numbers are needed to find a subscriber and to maintain the connection with a mobile station. The interesting case is the **mobile terminated call (MTC)**, i.e., a situation in which a station calls a mobile station (the calling station could be outside the GSM network or another mobile station). Figure 4.8 shows the basic steps needed to connect the calling station with the mobile user. In step 1, a user dials the phone number of a GSM subscriber. The fixed network (PSTN) notices (looking at the destination code) that the number belongs to a user in the GSM network and forwards the call setup to the Gateway MSC (2). The GMSC identifies the HLR for the subscriber (which is coded in the phone number) and signals the call setup to the HLR (3). The HLR now checks whether the number exists and whether the user has subscribed to the requested services, and requests an MSRN from the current VLR (4). After receiving the MSRN (5), the HLR can determine the MSC responsible for the MS and forwards this information to the GMSC (6). The GMSC can now forward the call setup request to the MSC indicated (7).

From this point on, the MSC is responsible for all further steps. First, it requests the current status of the MS from the VLR (8). If the MS is available, the MSC initiates paging in all cells it is responsible for (i.e. the location area, LA, 10), as searching for the right cell would be too time consuming (but this approach puts some load on the signaling channels so optimizations exist). The

7 Here, a discrepancy exists between ITU-T standards and ETSI's GSM. MS can denote mobile station or mobile subscriber. Typically, almost all MS in GSM refer to subscribers, as identifiers are not dependent on the station, but on the subscriber identity (stored in the SIM).

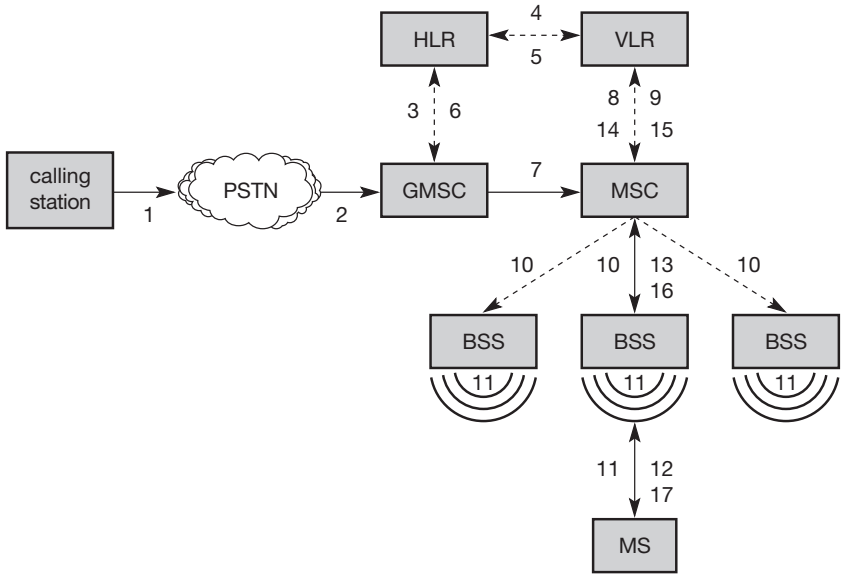

Figure 4.8
Mobile terminated call (MTC)

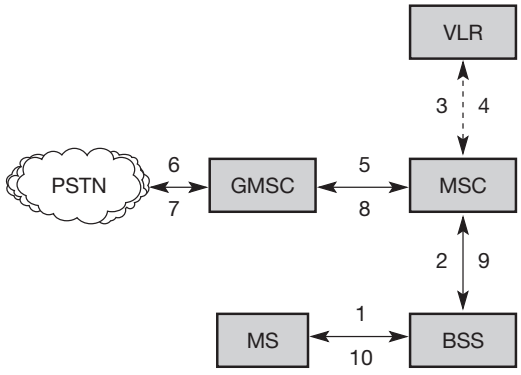

Figure 4.9
Mobile originated call (MOC)

BTSs of all BSSs transmit this paging signal to the MS (11). If the MS answers (12 and 13), the VLR has to perform security checks (set up encryption etc.). The VLR then signals to the MSC to set up a connection to the MS (steps 15 to 17).

It is much simpler to perform a **mobile originated call (MOC)** compared to a MTC (see Figure 4.9). The MS transmits a request for a new connection (1), the BSS forwards this request to the MSC (2). The MSC then checks if this user is allowed to set up a call with the requested service (3 and 4) and checks the availability of resources through the GSM network and into the PSTN. If all resources are available, the MSC sets up a connection between the MS and the fixed network.

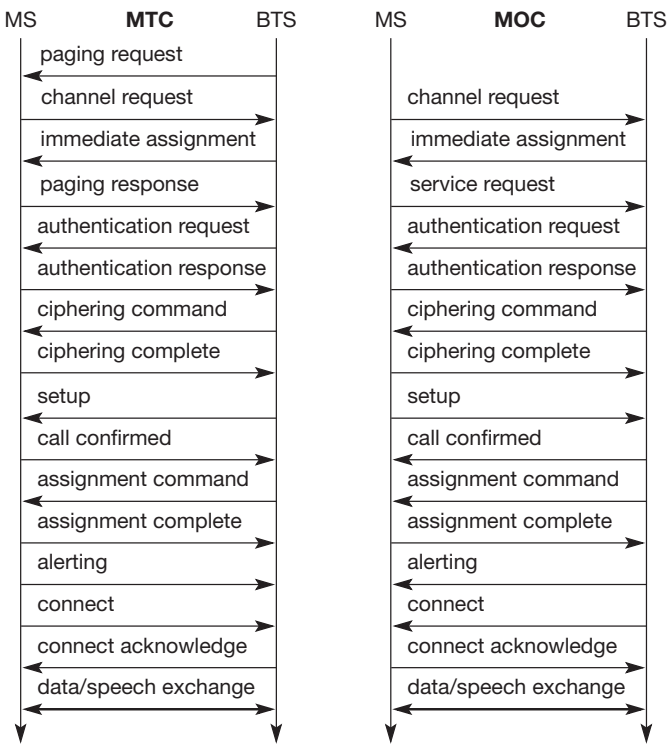

Figure 4.10 Message flow for MTC and MOC

In addition to the steps mentioned above, other messages are exchanged between an MS and BTS during connection setup (in either direction). These messages can be quite often heard in radios or badly shielded loudspeakers as crackling noise before the phone rings. Figure 4.10 shows the messages for an MTC and MOC. Paging is only necessary for an MTC, then similar message exchanges follow. The first step in this context is the channel access via the random access channel (RACH) with consecutive channel assignment; the channel assigned could be a traffic channel (TCH) or a slower signalling channel SDCCH.

The next steps, which are needed for communication security, comprise the authentication of the MS and the switching to encrypted communication. The system now assigns a TCH (if this has not been done). This has the advantage of only having to use an SDCCH during the first setup steps. If the setup fails, no TCH has been blocked. However, using a TCH from the beginning has a speed advantage.

The following steps depend on the use of MTC or MOC. If someone is calling the MS, it answers now with 'alerting' that the MS is ringing and with 'connect' that the user has pressed the connect button. The same actions

happen the other way round if the MS has initiated the call. After connection acknowledgement, both parties can exchange data.

Closing the connection comprises a user-initiated disconnect message (both sides can do this), followed by releasing the connection and the radio channel.

4.1.6 Handover

Cellular systems require **handover** procedures, as single cells do not cover the whole service area, but, e.g., only up to 35 km around each antenna on the countryside and some hundred meters in cities (Tripathi, 1998). The smaller the cell size and the faster the movement of a mobile station through the cells (up to 250 km/h for GSM), the more handovers of ongoing calls are required. However, a handover should not cause a cut-off, also called **call drop**. GSM aims at maximum handover duration of 60 ms.

There are two basic reasons for a handover (about 40 have been identified in the standard):

- The mobile station **moves out of the range** of a BTS or a certain antenna of a BTS respectively. The received **signal level** decreases continuously until it falls below the minimal requirements for communication. The **error rate** may grow due to interference, the distance to the BTS may be too high (max. 35 km) etc. – all these effects may diminish the **quality of the radio link** and make radio transmission impossible in the near future.
- The wired infrastructure (MSC, BSC) may decide that the **traffic in one cell is too high** and shift some MS to other cells with a lower load (if possible). Handover may be due to **load balancing**.

Figure 4.11 shows four possible handover scenarios in GSM:

- **Intra-cell handover:** Within a cell, narrow-band interference could make transmission at a certain frequency impossible. The BSC could then decide to change the carrier frequency (scenario 1).
- **Inter-cell, intra-BSC handover:** This is a typical handover scenario. The mobile station moves from one cell to another, but stays within the control of the same BSC. The BSC then performs a handover, assigns a new radio channel in the new cell and releases the old one (scenario 2).
- **Inter-BSC, intra-MSC handover:** As a BSC only controls a limited number of cells; GSM also has to perform handovers between cells controlled by different BSCs. This handover then has to be controlled by the MSC (scenario 3). This situation is also shown in Figure 4.13.
- **Inter MSC handover:** A handover could be required between two cells belonging to different MSCs. Now both MSCs perform the handover together (scenario 4).

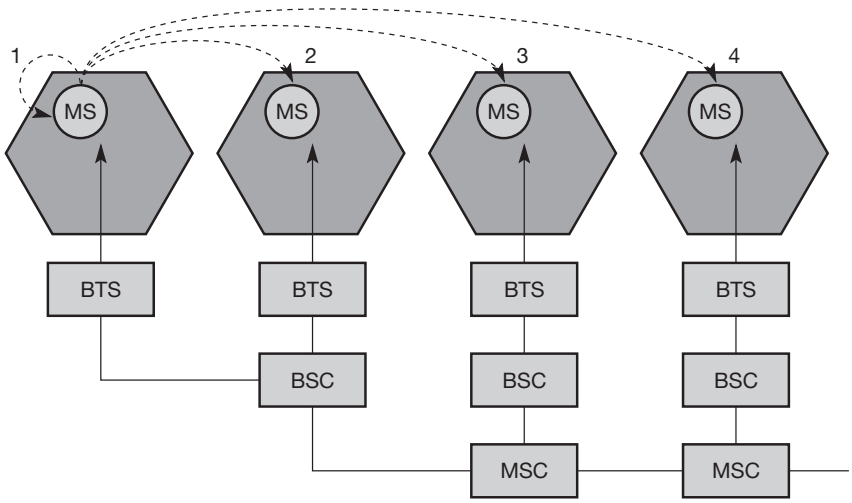

Figure 4.11 Types of handover in GSM

To provide all the necessary information for a handover due to a weak link, MS and BTS both perform periodic measurements of the downlink and uplink quality respectively. (Link quality comprises signal level and bit error rate.) Measurement reports are sent by the MS about every half-second and contain the quality of the current link used for transmission as well as the quality of certain channels in neighboring cells (the BCCHs).

Figure 4.12 shows the typical behavior of the received signal level while an MS moves away from one BTS (BTS_{old}) closer to another one (BTS_{new}). In this case, the handover decision does not depend on the actual value of the received signal level, but on the average value. Therefore, the BSC collects all values (bit error rate and signal levels from uplink and downlink) from BTS and MS and calculates average values. These values are then compared to thresholds, i.e., the handover margin (HO_MARGIN), which includes some hysteresis to avoid a ping-pong effect (Wong, 1997). (Without hysteresis, even short-term interference, e.g., shadowing due to a building, could cause a handover.) Still, even with the HO_MARGIN, the ping-pong effect may occur in GSM – a value which is too high could cause a cut-off, and a value which is too low could cause too many handovers.

Figure 4.13 shows the typical signal flow during an inter-BSC, intra-MSC handover. The MS sends its periodic measurements reports, the BTS_{old} forwards these reports to the BSC_{old} together with its own measurements. Based on these values and, e.g., on current traffic conditions, the BSC_{old} may decide to perform a handover and sends the message HO_required to the MSC. The task of the MSC then comprises the request of the resources needed for the handover from the new BSC, BSC_{new}. This BSC checks if enough resources (typically frequencies or time slots) are available and activates a physical channel at the BTS_{new} to prepare for the arrival of the MS.

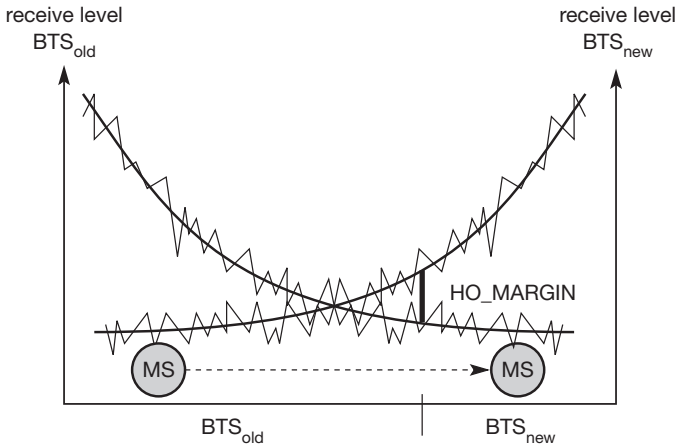

Figure 4.12 Handover decision depending on receive level

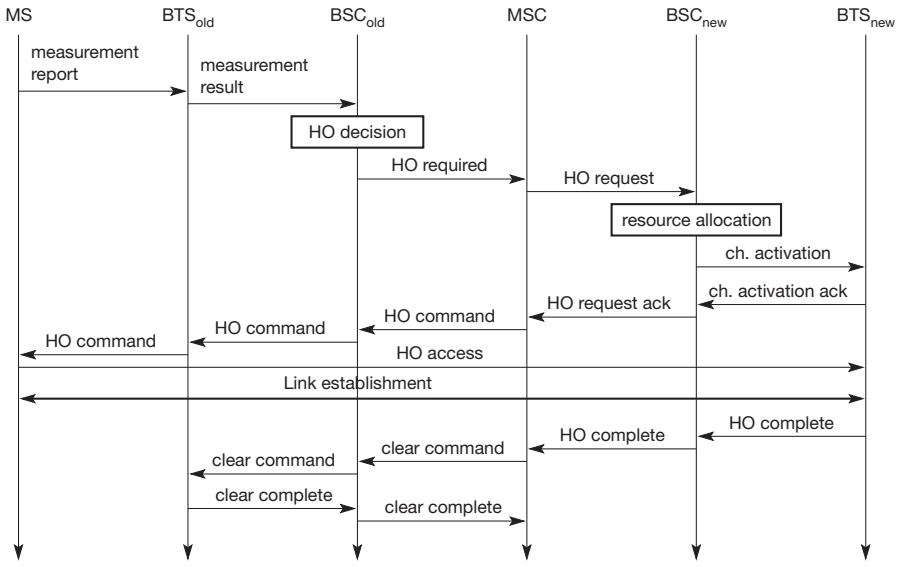

Figure 4.13 Intra-MSC handover

The BTS_{new} acknowledges the successful channel activation, BSC_{new} acknowledges the handover request. The MSC then issues a handover command that is forwarded to the MS. The MS now breaks its old radio link and accesses the new BTS. The next steps include the establishment of the link (this includes layer two link establishment and handover complete messages from the MS). Basically, the MS has then finished the handover, but it is important to release the resources at the old BSC and BTS and to signal the successful handover using the handover and clear complete messages as shown.

More sophisticated handover mechanisms are needed for seamless handovers between different systems. For example, future 3G networks will not cover whole countries but focus on cities and highways. Handover from,

e.g., UMTS to GSM without service interruption must be possible. Even more challenging is the seamless handover between wireless LANs (see chapter 7) and 2G/3G networks. This can be done using multimode mobile stations and a more sophisticated roaming infrastructure. However, it is still not obvious how these systems may scale for a large number of users and many handovers, and what handover quality guarantees they can give.

4.1.7 Security

GSM offers several security services using confidential information stored in the AuC and in the individual SIM (which is plugged into an arbitrary MS). The SIM stores personal, secret data and is protected with a PIN against unauthorized use. (For example, the secret key K_i used for authentication and encryption procedures is stored in the SIM.) The security services offered by GSM are explained below:

- **Access control and authentication:** The first step includes the authentication of a valid user for the SIM. The user needs a secret PIN to access the SIM. The next step is the subscriber authentication (see Figure 4.10). This step is based on a challenge-response scheme as presented in section 4.1.7.1.
- **Confidentiality:** All user-related data is encrypted. After authentication, BTS and MS apply encryption to voice, data, and signaling as shown in section 4.1.7.2. This confidentiality exists only between MS and BTS, but it does not exist end-to-end or within the whole fixed GSM/telephone network.
- **Anonymity:** To provide user anonymity, all data is encrypted before transmission, and user identifiers (which would reveal an identity) are not used over the air. Instead, GSM transmits a temporary identifier (TMSI), which is newly assigned by the VLR after each location update. Additionally, the VLR can change the TMSI at any time.

Three algorithms have been specified to provide security services in GSM. **Algorithm A3** is used for **authentication**, **A5** for **encryption**, and **A8** for the **generation of a cipher key**. In the GSM standard only algorithm A5 was publicly available, whereas A3 and A8 were secret, but standardized with open interfaces. Both A3 and A8 are no longer secret, but were published on the internet in 1998. This demonstrates that security by obscurity does not really work. As it turned out, the algorithms are not very strong. However, network providers can use stronger algorithms for authentication – or users can apply stronger end-to-end encryption. Algorithms A3 and A8 (or their replacements) are located on the SIM and in the AuC and can be proprietary. Only A5 which is implemented in the devices has to be identical for all providers.

4.1.7.1 Authentication

Before a subscriber can use any service from the GSM network, he or she must be authenticated. Authentication is based on the SIM, which stores the **individual authentication key K_i**, the **user identification IMSI**, and the algorithm used for authentication **A3**. Authentication uses a challenge-response method: the access

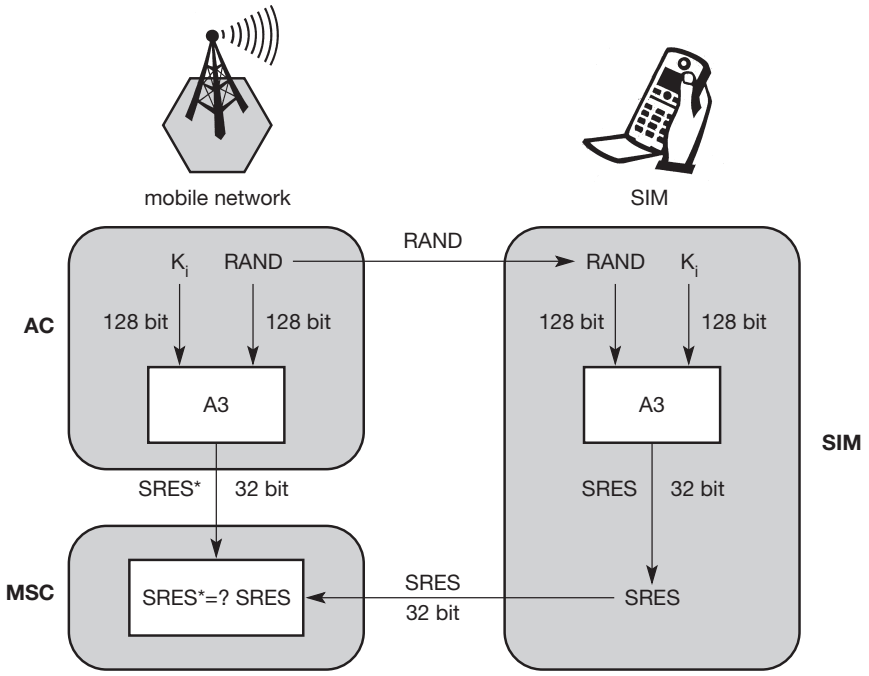

Figure 4.14 Subscriber authentication

control AC generates a random number **RAND** as challenge, and the SIM within the MS answers with **SRES** (signed response) as response (see Figure 4.14). The AuC performs the basic generation of random values RAND, signed responses SRES, and cipher keys K_c for each IMSI, and then forwards this information to the HLR. The current VLR requests the appropriate values for RAND, SRES, and K_c from the HLR.

For authentication, the VLR sends the random value RAND to the SIM. Both sides, network and subscriber module, perform the same operation with RAND and the key K_i, called A3. The MS sends back the SRES generated by the SIM; the VLR can now compare both values. If they are the same, the VLR accepts the subscriber, otherwise the subscriber is rejected.

4.1.7.2 Encryption

To ensure privacy, all messages containing user-related information are encrypted in GSM over the air interface. After authentication, MS and BSS can start using encryption by applying the cipher key K_c (the precise location of security functions for encryption, BTS and/or BSC are vendor dependent). K_c is generated using the individual key K_i and a random value by applying the algorithm A8. Note that the SIM in the MS and the network both calculate the same K_c based on the random value RAND. The key K_c itself is not transmitted over the air interface.

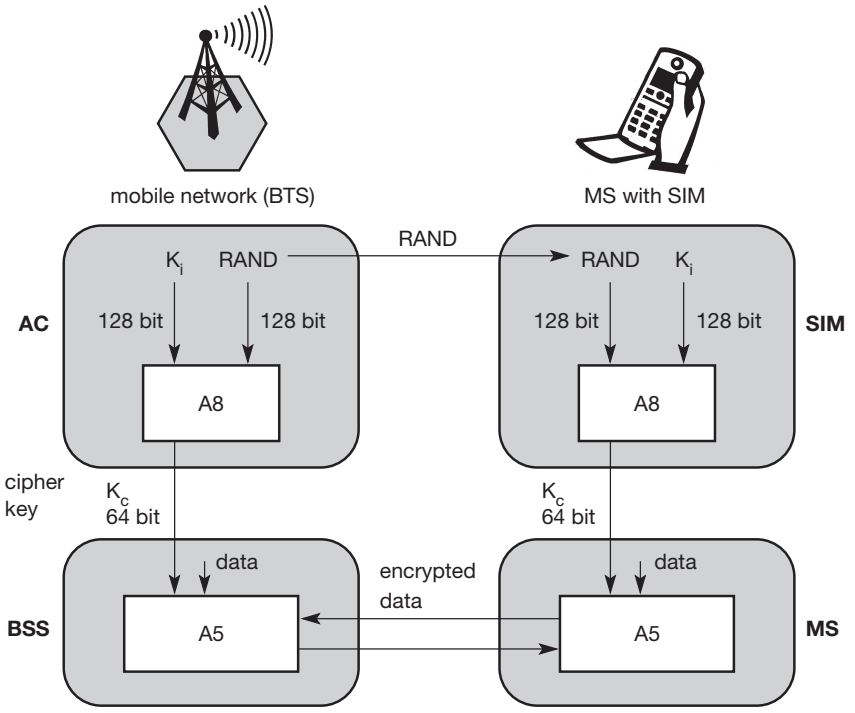

Figure 4.15
Data encryption

MS and BTS can now encrypt and decrypt data using the algorithm A5 and the cipher key K_c. As Figure 4.15 shows, K_c should be a 64 bit key – which is not very strong, but is at least a good protection against simple eavesdropping. However, the publication of A3 and A8 on the internet showed that in certain implementations 10 of the 64 bits are always set to 0, so that the real length of the key is thus only 54 consequently, the encryption is much weaker.

4.1.8 New data services

As mentioned above, the standard bandwidth of 9.6 kbit/s (14.4 kbit/s with some providers) available for data transmission is not sufficient for the requirements of today's computers. When GSM was developed, not many people anticipated the tremendous growth of data communication compared to voice communication. At that time, 9.6 kbit/s was a lot, or at least enough for standard group 3 fax machines. But with the requirements of, e.g., web browsing, file download, or even intensive e-mail exchange with attachments, this is not enough.

To enhance the data transmission capabilities of GSM, two basic approaches are possible. As the basic GSM is based on connection-oriented traffic channels, e.g., with 9.6 kbit/s each, several channels could be combined to increase bandwidth. This system is called HSCSD and is presented in the following section. A

more progressive step is the introduction of packet-oriented traffic in GSM, i.e., shifting the paradigm from connections/telephone thinking to packets/internet thinking. The system, called GPRS, is presented in section 4.1.8.2.

4.1.8.1 HSCSD

A straightforward improvement of GSM's data transmission capabilities is **high speed circuit switched data (HSCSD)**, which is available with some providers. In this system, higher data rates are achieved by bundling several TCHs. An MS requests one or more TCHs from the GSM network, i.e., it allocates several TDMA slots within a TDMA frame. This allocation can be asymmetrical, i.e., more slots can be allocated on the downlink than on the uplink, which fits the typical user behavior of downloading more data compared to uploading. Basically, HSCSD only requires software upgrades in an MS and MSC (both have to be able to split a traffic stream into several streams, using a separate TCH each, and to combine these streams again).

In theory, an MS could use all eight slots within a TDMA frame to achieve an **air interface user rate (AIUR)** of, e.g., 8 TCH/F14.4 channels or 115.2 kbit/s (ETSI, 1998e). One problem of this configuration is that the MS is required to send and receive at the same time. Standard GSM does not require this capability – uplink and downlink slots are always shifted for three slots. ETSI (1997a) specifies the AIUR available at 57.6 kbit/s (duplex) using four slots in the uplink and downlink (Table 4.2 shows the permitted combinations of traffic channels and allocated slots for non-transparent services).

Although it appears attractive at first glance, HSCSD exhibits some major disadvantages. It still uses the connection-oriented mechanisms of GSM. These are not at all efficient for computer data traffic, which is typically bursty and asymmetrical. While downloading a larger file may require all channels reserved, typical web browsing would leave the channels idle most of the time. Allocating channels is reflected directly in the service costs, as, once the channels have been reserved, other users cannot use them.

AIUR	TCH / F4.8	TCH / F9.6	TCH / F14.4
4.8 kbit/s	1	–	–
9.6 kbit/s	2	1	–
14.4 kbit/s	3	–	1
19.2 kbit/s	4	2	–
28.8 kbit/s	–	3	2
38.4 kbit/s	–	4	–
43.2 kbit/s	–	–	3
57.6 kbit/s	–	–	4

Table 4.2 Available data rates for HSCSD in GSM

For *n* channels, HSCSD requires *n* times signaling during handover, connection setup and release. Each channel is treated separately. The probability of blocking or service degradation increases during handover, as in this case a BSC has to check resources for *n* channels, not just one. All in all, HSCSD may be an attractive interim solution for higher bandwidth and rather constant traffic (e.g., file download). However, it does not make much sense for bursty internet traffic as long as a user is charged for each channel allocated for communication.

4.1.8.2 GPRS

The next step toward more flexible and powerful data transmission avoids the problems of HSCSD by being fully packet-oriented. The **general packet radio service (GPRS)** provides packet mode transfer for applications that exhibit traffic patterns such as frequent transmission of small volumes (e.g., typical web requests) or infrequent transmissions of small or medium volumes (e.g., typical web responses) according to the requirement specification (ETSI, 1998a). Compared to existing data transfer services, GPRS should use the existing network resources more efficiently for packet mode applications, and should provide a selection of QoS parameters for the service requesters. GPRS should also allow for broadcast, multicast, and unicast service. The overall goal in this context is the provision of a more efficient and, thus, cheaper packet transfer service for typical internet applications that usually rely solely on packet transfer. Network providers typically support this model by charging on volume and not on connection time as is usual for traditional GSM data services and for HSCSD. The main benefit for users of GPRS is the 'always on' characteristic – no connection has to be set up prior to data transfer. Clearly, GPRS was driven by the tremendous success of the packet-oriented internet, and by the new traffic models and applications. However, GPRS, as shown in the following sections, needs additional network elements, i.e., software and hardware. Unlike HSCSD, GPRS does not only represent a software update to allow for the bundling of channels, it also represents a big step towards UMTS as the main internal infrastructure needed for UMTS (in its initial release) is exactly what GPRS uses (see section 4.4).

The main concepts of GPRS are as follows (ETSI, 1998b). For the new GPRS radio channels, the GSM system can allocate between one and eight time slots within a TDMA frame. Time slots are not allocated in a fixed, pre-determined manner but on demand. All time slots can be shared by the active users; up- and downlink are allocated separately. Allocation of the slots is based on current load and operator preferences. Depending on the coding, a transfer rate of up to 170 kbit/s is possible. For GPRS, operators often reserve at least a time slot per cell to guarantee a minimum data rate. The GPRS concept is independent of channel characteristics and of the type of channel (traditional GSM traffic or control channel), and does not limit the maximum data rate (only the GSM transport system limits the rate). All GPRS services can be used in parallel to conventional services. Table 4.3 shows the typical data rates available with GPRS if it is used together with GSM (GPRS can also be used for other TDMA systems).

Coding scheme	1 slot	2 slots	3 slots	4 slots	5 slots	6 slots	7 slots	8 slots
CS-1	9.05	18.2	27.15	36.2	45.25	54.3	63.35	72.4
CS-2	13.4	26.8	40.2	53.6	67	80.4	93.8	107.2
CS-3	15.6	31.2	46.8	62.4	78	93.6	109.2	124.8
CS-4	21.4	42.8	64.2	85.6	107	128.4	149.8	171.2

Table 4.3 GPRS data rates in kbit/s

In the beginning, only coding schemes CS-1 and CS-2 are available. The system chooses a coding scheme depending on the current error rate (CS-4 provides no error correction capabilities).

It should be noted that the real available data rate heavily depends on the current load of the cell as GPRS typically only uses idle time slots. The transfer rate depends on the capabilities of the MS as not all devices are able to send and receive at the same time. Table 4.4 gives examples for device classes together with their ability to use time slots for sending and receiving data. The maximum possible number of slots limits the transfer rate even more. For example, a class 12 device may receive data using 4 slots within a GSM time frame or it may send data using 4 slots. However, a maximum number of 5 slots may be used altogether. Using all 8 slots for data encoded using CS-4 yields the maximum rate of 171.2 kbit/s. Today, a typical MS is a class 10 device using CS-2, which results in a receiving rate of 53.6 kbit/s and a sending rate of 26.8 kbit/s.

In phase 1, GPRS offers a **point-to-point (PTP)** packet transfer service (ETSI, 1998c). One of the PTP versions offered is the **PTP connection oriented network service (PTP-CONS)**, which includes the ability of GPRS to maintain a virtual circuit upon change of the cell within the GSM network. This type of

Class	Receiving slots	Sending slots	Maximum number of slots
1	1	1	2
2	2	1	3
3	2	2	3
5	2	2	4
8	4	1	5
10	4	2	5
12	4	4	5

Table 4.4 Examples for GPRS device classes

Table 4.5 Reliability classes in GPRS according to ETSI (1998c)

Reliability class	Lost SDU probability	Duplicate SDU probability	Out of sequence SDU probability	Corrupt SDU probability
1	10^{-9}	10^{-9}	10^{-9}	10^{-9}
2	10^{-4}	10^{-5}	10^{-5}	10^{-6}
3	10^{-2}	10^{-5}	10^{-5}	10^{-2}

service corresponds to **X.25**, the typical circuit-switched packet-oriented transfer protocol available worldwide. The other PTP version offered is the **PTP connectionless network service (PTP-CLNS)**, which supports applications that are based on the Internet Protocol **IP**. Multicasting, called **point-to-multipoint (PTM)** service, is left for GPRS phase 2.

Users of GPRS can specify a **QoS-profile**. This determines the **service precedence** (high, normal, low), **reliability class** and **delay class** of the transmission, and **user data throughput**. GPRS should adaptively allocate radio resources to fulfill these user specifications. Table 4.5 shows the three reliability classes together with the maximum probabilities for a lost service data unit (SDU), a duplicated SDU, an SDU out of the original sequence, and the probability of delivering a corrupt SDU to the higher layer. Reliability class 1 could be used for very error-sensitive applications that cannot perform error corrections themselves. If applications exhibit greater error tolerance, class 2 could be appropriate. Finally, class 3 is the choice for error-insensitive applications or applications that can handle error corrections themselves.

Delay within a GPRS network is incurred by channel access delay, coding for error correction, and transfer delays in the fixed and wireless part of the GPRS network. The delay introduced by external fixed networks is out of scope. However, GPRS does not produce additional delay by buffering packets as store-and-forward networks do. If possible, GPRS tries to forward packets as fast as possible. Table 4.6 shows the specified maximum mean and 95 percentile delay values for packet sizes of 128 and 1,024 byte. As we can clearly see, no matter which class, all delays are orders of magnitude higher than fixed network delays. This is a very important characteristic that has to be taken into account when implementing higher layer protocols such as TCP on top of GPRS networks (see chapter 9). Typical round trip times (RTT) in fixed networks are in the order of 10 to 100 ms. Using real unloaded GPRS networks round trip times of well above 1 s for even small packets (128–512 byte) are common. Additionally, GPRS exhibits a large jitter compared to fixed networks (several 100 ms are not uncommon). This characteristic has a strong impact on user experience when, e.g., interactive Internet applications are used on top of GPRS.

	SDU size 128 byte		SDU size 1,024 byte	
Delay Class	**Mean**	**95 percentile**	**Mean**	**95 percentile**
1	<0.5 s	<1.5 s	<2 s	<7 s
2	<5 s	<25 s	<15 s	<75 s
3	<50 s	<250 s	<75 s	<375 s
4		Unspecified		

Table 4.6 Delay classes in GRPS according to ETSI (1998c)

Finally, GPRS includes several **security services** such as authentication, access control, user identity confidentiality, and user information confidentiality. Even a completely **anonymous service** is possible, as, e.g., applied for road toll systems that only charge a user via the MS independent of the user's identity.

The **GPRS architecture** introduces two new network elements, which are called **GPRS support nodes (GSN)** and are in fact routers. All GSNs are integrated into the standard GSM architecture, and many new interfaces have been defined (see Figure 4.16). The **gateway GPRS support node (GGSN)** is the interworking unit between the GPRS network and external **packet data networks (PDN)**. This node contains routing information for GPRS users, performs address conversion, and tunnels data to a user via encapsulation. The GGSN is connected to external networks (e.g., IP or X.25) via the G_i interface and transfers packets to the SGSN via an IP-based GPRS backbone network (G_n interface).

The other new element is the **serving GPRS support node (SGSN)** which supports the MS via the G_b interface. The SGSN, for example, requests user addresses from the **GPRS register (GR)**, keeps track of the individual MSs' location, is responsible for collecting billing information (e.g., counting bytes), and performs several security functions such as access control. The SGSN is connected to a BSC via frame relay and is basically on the same hierarchy level as an MSC. The GR, which is typically a part of the HLR, stores all GPRS-relevant data. GGSNs and SGSNs can be compared with home and foreign agents, respectively, in a mobile IP network (see chapter 8).

As shown in Figure 4.16, packet data is transmitted from a PDN, via the GGSN and SGSN directly to the BSS and finally to the MS. The MSC, which is responsible for data transport in the traditional circuit-switched GSM, is only used for signaling in the GPRS scenario. Additional interfaces to further network elements and other PLMNs can be found in ETSI (1998b).

Before sending any data over the GPRS network, an MS must attach to it, following the procedures of the **mobility management**. The attachment procedure includes assigning a temporal identifier, called a **temporary logical link identity (TLLI)**, and a **ciphering key sequence number (CKSN)** for data encryption. For each MS, a **GPRS context** is set up and stored in the MS and in

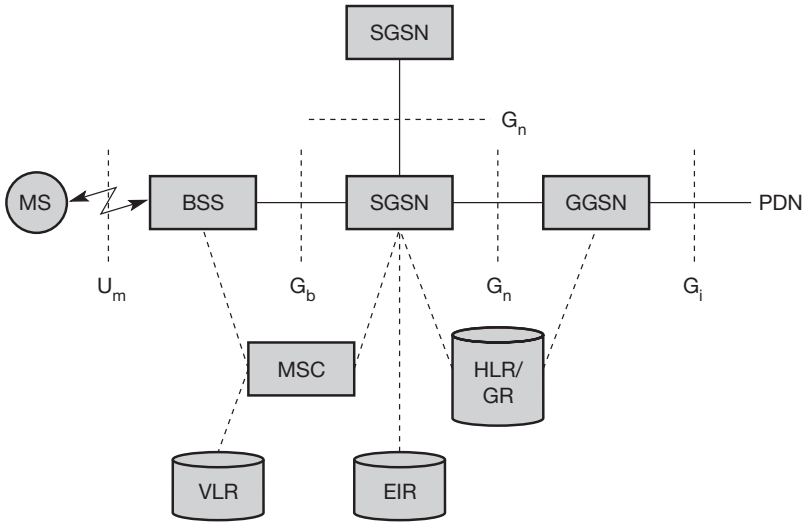

Figure 4.16
GPRS architecture reference model

the corresponding SGSN. This context comprises the status of the MS (which can be ready, idle, or standby; ETSI, 1998b), the CKSN, a flag indicating if compression is used, and routing data (TLLI, the routing area RA, a cell identifier, and a packet data channel, PDCH, identifier). Besides attaching and detaching, mobility management also comprises functions for authentication, location management, and ciphering (here, the scope of ciphering lies between MS and SGSN, which is more than in standard GSM). In **idle** mode an MS is not reachable and all context is deleted. In the **standby** state only movement across routing areas is updated to the SGSN but not changes of the cell. Permanent updating would waste battery power, no updating would require system-wide paging. The update procedure in standby mode is a compromise. Only in the **ready** state every movement of the MS is indicated to the SGSN.

Figure 4.17 shows the protocol architecture of the transmission plane for GPRS. Architectures for the signaling planes can be found in ETSI (1998b). All data within the GPRS backbone, i.e., between the GSNs, is transferred using the **GPRS tunnelling protocol (GTP)**. GTP can use two different transport protocols, either the reliable **TCP** (needed for reliable transfer of X.25 packets) or the non-reliable **UDP** (used for IP packets). The network protocol for the GPRS backbone is **IP** (using any lower layers). To adapt to the different characteristics of the underlying networks, the **subnetwork dependent convergence protocol (SNDCP)** is used between an SGSN and the MS. On top of SNDCP and GTP, user packet data is tunneled from the MS to the GGSN and vice versa. To achieve a high reliability of packet transfer between SGSN and MS, a special LLC is used, which comprises ARQ and FEC mechanisms for PTP (and later PTM) services.

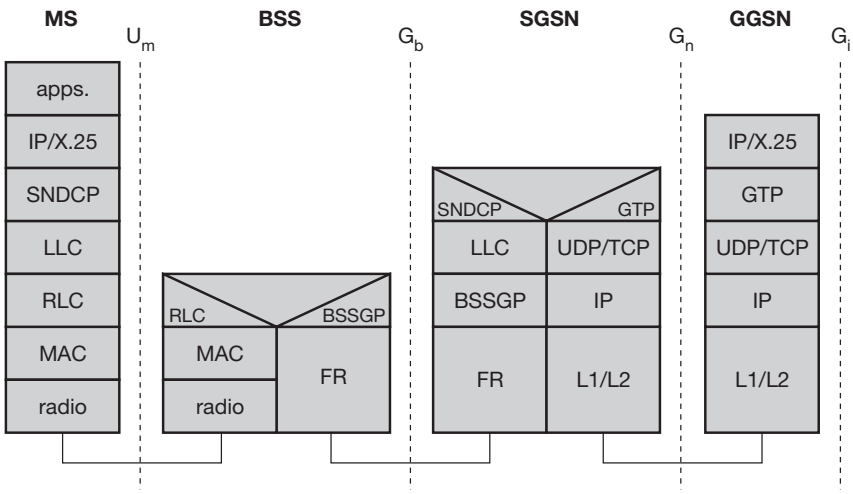

Figure 4.17 GPRS transmission plane protocol reference model

A **base station subsystem GPRS protocol (BSSGP)** is used to convey routing and QoS-related information between the BSS and SGSN. BSSGP does not perform error correction and works on top of a **frame relay (FR)** network. Finally, radio link dependent protocols are needed to transfer data over the U_m interface. The **radio link protocol (RLC)** provides a reliable link, while the **MAC** controls access with signaling procedures for the radio channel and the mapping of LLC frames onto the GSM physical channels. The **radio interface** at U_m needed for GPRS does not require fundamental changes compared to standard GSM (Brasche, 1997), (ETSI, 1998d). However, several new logical channels and their mapping onto physical resources have been defined. For example, one MS can allocate up to eight **packet data traffic channels (PDTCHs)**. Capacity can be allocated on demand and shared between circuit-switched channels and GPRS. This allocation can be done dynamically with load supervision or alternatively, capacity can be pre-allocated.

A very important factor for any application working end-to-end is that it does not 'notice' any details from the GSM/GPRS-related infrastructure. The application uses, e.g., TCP on top of IP, IP packets are tunneled to the GGSN, which forwards them into the PDN. All PDNs forward their packets for a GPRS user to the GGSN, the GGSN asks the current SGSN for tunnel parameters, and forwards the packets via SGSN to the MS. Although MSs using GPRS may be considered as part of the internet, one should know that operators typically perform an address translation in the GGSN using NAT. All MSs are assigned private IP addresses which are then translated into global addresses at the GGSN. The advantage of this approach is the inherent protection of MSs from attacks (the subscriber typically has to pay for traffic even if it originates from an attack!) – private addresses are not routed through the internet so it is not possible to

reach an MS from the internet. This is also a disadvantage if an MS wants to offer a service using a fixed, globally visible IP address. This is difficult with IPv4 and NAT and it will be interesting to see how IPv6 is used for this purpose (while still protecting the MSs from outside attacks as air traffic is expensive).

4.2 DECT

Another fully digital cellular network is the **digital enhanced cordless telecommunications (DECT)** system specified by ETSI (2002, 1998j, k), (DECT Forum, 2002). Formerly also called **digital European cordless telephone and digital European cordless telecommunications**, DECT replaces older analog cordless phone systems such as CT1 and CT1+. These analog systems only ensured security to a limited extent as they did not use encryption for data transmission and only offered a relatively low capacity. DECT is also a more powerful alternative to the digital system CT2, which is mainly used in the UK (the DECT standard works throughout Europe), and has even been selected as one of the 3G candidates in the IMT-2000 family (see section 4.4). DECT is mainly used in offices, on campus, at trade shows, or in the home. Furthermore, access points to the PSTN can be established within, e.g., railway stations, large government buildings and hospitals, offering a much cheaper telephone service compared to a GSM system. DECT could also be used to bridge the last few hundred meters between a new network operator and customers. Using this 'small range' local loop, new companies can offer their service without having their own lines installed in the streets. DECT systems offer many different interworking units, e.g., with GSM, ISDN, or data networks. Currently, over 100 million DECT units are in use (DECT, 2002).

A big difference between DECT and GSM exists in terms of cell diameter and cell capacity. While GSM is designed for outdoor use with a cell diameter of up to 70 km, the range of DECT is limited to about 300 m from the base station (only around 50 m are feasible inside buildings depending on the walls). Due to this limited range and additional multiplexing techniques, DECT can offer its service to some 10,000 people within one km^2. This is a typical scenario within a big city, where thousands of offices are located in skyscrapers close together. DECT also uses base stations, but these base stations together with a mobile station are in a price range of €100 compared to several €10,000 for a GSM base station. GSM base stations can typically not be used by individuals for private networks. One reason is licensing as all GSM frequencies have been licensed to network operators. DECT can also handle handover, but it was not designed to work at a higher speed (e.g., up to 250 km/h like GSM systems). Devices handling GSM and DECT exist but have never been a commercial success.

DECT works at a frequency range of 1880–1990 MHz offering 120 full duplex channels. Time division duplex (TDD) is applied using 10 ms frames. The frequency range is subdivided into 10 carrier frequencies using FDMA, each frame being divided into 24 slots using TDMA. For the TDD mechanism,

12 slots are used as uplink, 12 slots as downlink (see Figure 3.4). The digital modulation scheme is GMSK – each station has an average transmission power of only 10 mW with a maximum of 250 mW.

4.2.1 System architecture

A DECT system, may have various different physical implementation depending on its actual use. Different DECT entities can be integrated into one physical unit; entities can be distributed, replicated etc. However, all implementations are based on the same logical reference model of the system architecture as shown in Figure 4.18. A **global network** connects the local communication structure to the outside world and offers its services via the interface D_1. Global networks could be integrated services digital networks (ISDN), public switched telephone networks (PSTN), public land mobile networks (PLMN), e.g., GSM, or packet switched public data network (PSPDN). The services offered by these networks include transportation of data and the translation of addresses and routing of data between the local networks.

Local networks in the DECT context offer local telecommunication services that can include everything from simple switching to intelligent call forwarding, address translation etc. Examples for such networks are analog or digital private branch exchanges (PBXs) or LANs, e.g., those following the IEEE 802.x family of LANs. As the core of the DECT system itself is quite simple, all typical network functions have to be integrated in the local or global network, where the databases **home data base (HDB)** and **visitor data base (VDB)** are also located. Both databases support mobility with functions that are similar to those in the HLR and VLR in GSM systems. Incoming calls are automatically forwarded to the current subsystem responsible for the DECT user, and the current VDB informs the HDB about changes in location.

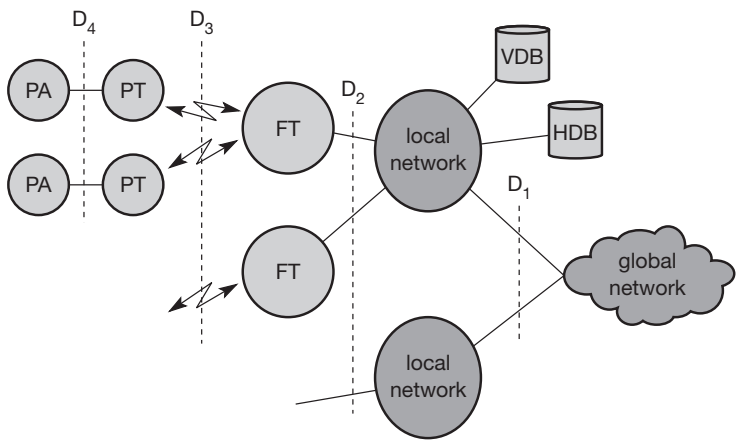

Figure 4.18 DECT system architecture reference model

The DECT core network consists of the **fixed radio termination (FT)** and the **portable radio termination (PT)**, and basically only provides a multiplexing service. FT and PT cover layers one to three at the fixed network side and mobile network side respectively. Additionally, several portable applications (PA) can be implemented on a device.

4.2.2 Protocol architecture

The DECT protocol reference architecture follows the OSI reference model. Figure 4.19 shows the layers covered by the standard: the physical layer, medium access control, and data link control[8] for both the **control plane (C-Plane)** and the **user plane (U-Plane)**. An additional network layer has been specified for the C-Plane, so that user data from layer two is directly forwarded to the U-Plane. A management plane vertically covers all lower layers of a DECT system.

4.2.2.1 Physical layer

As in all wireless networks, the **physical layer** comprises all functions for modulation/demodulation, incoming signal detection, sender/receiver synchronization, and collection of status information for the management plane. This layer generates the physical channel structure with a certain, guaranteed throughput. On request from the MAC layer, the physical layer assigns a channel for data transmission.

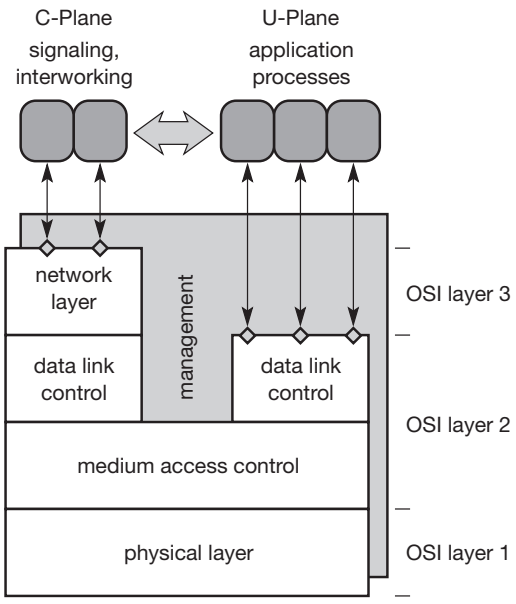

Figure 4.19 DECT protocol layers

8 Strictly speaking, the name "data link control" for the upper part of layer two is wrong in this architecture. According to the OSI reference model, the data link control (layer two) comprises the logical link control (layer 2b) and the medium access control (layer 2a).

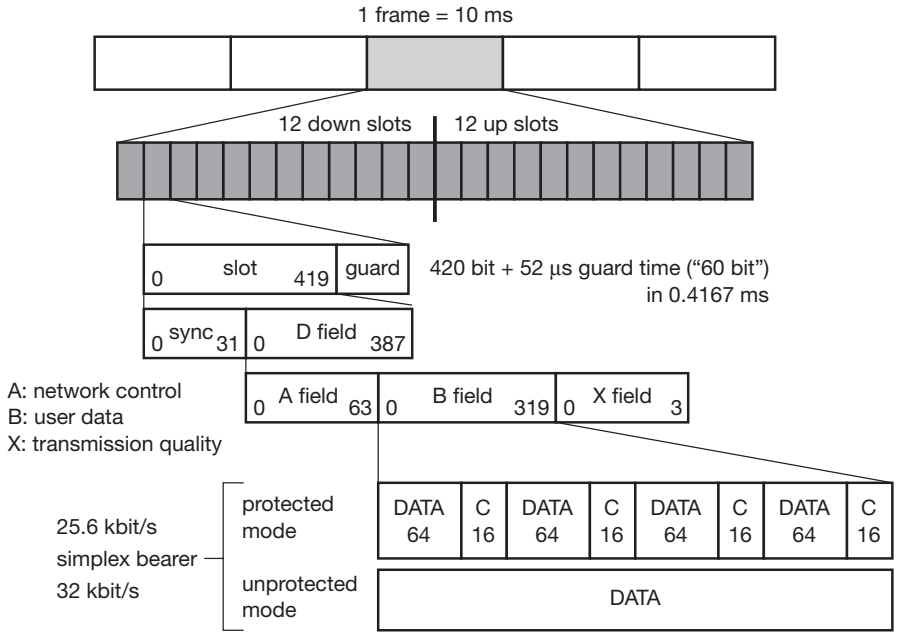

Figure 4.20
DECT multiplex and frame structure

Figure 4.20 shows the standard TDMA frame structure used in DECT and some typical data packets. Each frame has a duration of 10 ms and contains 12 slots for the downlink and 12 slots for the uplink in the **basic connection** mode. If a mobile node receives data in slot s, it returns data in slot s+12. An **advanced connection** mode allows different allocation schemes. Each slot has a duration of 0.4167 ms and can contain several different physical packets. Typically, 420 bits are used for data; the remaining 52 μs are left as **guard space**. The 420 data bits are again divided into a 32 bit **synchronization pattern** followed by the **data** field D.

The fields for data transmission now use these remaining 388 bits for **network control** (A field), **user data** (B field), and the transfer of the **transmission quality** (X field). While network control is transmitted with a data rate of 6.4 kbit/s (64 bit each 10 ms), the user data rate depends on additional error correction mechanisms. The **simplex bearer** provides a data rate of 32 kbit/s in an **unprotected mode**, while using a 16 bit CRC **checksum** C for a data block of 64 bit in the **protected mode** reduces the data rate to 25.6 kbit/s. A **duplex bearer** service is produced by combining two simplex bearers. DECT also defines bearer types with higher throughputs by combining slots, e.g., the **double duplex bearer** offers 80 kbit/s full-duplex.

4.2.2.2 Medium access control layer

The **medium access control (MAC)** layer establishes, maintains, and releases channels for higher layers by activating and deactivating physical channels. MAC multiplexes several logical channels onto physical channels. Logical channels exist for signaling network control, user data transmission, paging, or sending broadcast messages. Additional services offered include segmentation/reassembly of packets and error control/error correction.

4.2.2.3 Data link control layer

The **data link control (DLC)** layer creates and maintains reliable connections between the mobile terminal and the base station. Two services have been defined for the **C-Plane**: a **connectionless broadcast** service for paging (called **Lb**) and a **point-to-point** protocol similar to LAPD in ISDN, but adapted to the underlying MAC (called **LAPC+Lc**).

Several services exist for the **U-Plane**, e.g., a transparent unprotected service (basically a null service), a forward error correction service, rate adaptation services, and services for future enhancements. If services are used, e.g., to transfer ISDN data at 64 kbit/s, then DECT also tries to transfer 64 kbit/s. However, in case of errors, DECT raises the transfer rate to 72 kbit/s, and includes FEC and a buffer for up to eight blocks to perform ARQ. This buffer then introduces an additional delay of up to 80 ms.

4.2.2.4 Network layer

The **network layer** of DECT is similar to those in ISDN and GSM and only exists for the **C-Plane**. This layer provides services to request, check, reserve, control, and release resources at the fixed station (connection to the fixed network, wireless connection) and the mobile terminal (wireless connection). The **mobility management (MM)** within the network layer is responsible for identity management, authentication, and the management of the location data bases. **Call control (CC)** handles connection setup, release, and negotiation. Two message services, the **connection oriented message service (COMS)** and the **connectionless message service (CLMS)** transfer data to and from the interworking unit that connects the DECT system with the outside world.

4.3 TETRA

Trunked radio systems constitute another method of wireless data transmission. These systems use many different radio carriers but only assign a specific carrier to a certain user for a short period of time according to demand. While, for example, taxi services, transport companies with fleet management systems and rescue teams all have their own unique carrier frequency in traditional systems, they can share a whole group of frequencies in trunked radio systems for better frequency reuse via FDM and TDM techniques. These types of radio systems typically offer

interfaces to the fixed telephone network, i.e., voice and data services, but are not publicly accessible. These systems are not only simpler than most other networks, they are also reliable and relatively cheap to set up and operate, as they only have to cover the region where the local users operate, e.g., a city taxi service.

To allow a common system throughout Europe, ETSI standardized the **TETRA** system (**terrestrial trunked radio**)[9] in 1991 (ETSI, 2002), (TETRA MoU, 2002). This system should replace national systems, such as MODACOM, MOBITEX and COGNITO in Europe that typically connect to an X.25 packet network. (An example system from the US is ARDIS.) TETRA offers two standards: the **Voice+Data (V+D)** service (ETSI, 1998l) and the **packet data optimized (PDO)** service (ETSI, 1998m). While V+D offers circuit-switched voice and data transmission, PDO only offers packet data transmission, either connection-oriented to connect to X.25 or connectionless for the ISO CLNS (connectionless network service). The latter service can be point-to-point or point-to-multipoint, the typical delay for a short message (128 byte) being less than 100 ms. V+D connection modes comprise unicast and broadcast connections, group communication within a certain protected group, and a direct ad hoc mode without a base station. However, delays for short messages can be up to 500 ms or higher depending on the priority.

TETRA also offers bearer services of up to 28.8 kbit/s for unprotected data transmission and 9.6 kbit/s for protected transmission. Examples for end-to-end services are call forwarding, call barring, identification, call hold, call priorities, emergency calls and group joins. The system architecture of TETRA is very similar to GSM. Via the radio interface U_m, the **mobile station (MS)** connects to the **switching and management infrastructure (SwMI)**, which contains the user data bases (HDB, VDB), the base station, and interfaces to PSTN, ISDN, or PDN. The system itself, however, is much simpler in real implementation compared to GSM, as typically no handover is needed. Taxis usually remain within a certain area which can be covered by one TETRA cell.

Several frequencies have been specified for TETRA which uses FDD (e.g., 380–390 MHz uplink/390–400 MHz downlink, 410–420 MHz uplink/420–430 MHz downlink). Each channel has a bandwidth of 25 kHz and can carry 36 kbit/s. Modulation is DQPSK. While V+D uses up to four TDMA voice or data channels per carrier, PDO performs statistical multiplexing. For accessing a channel, slotted Aloha is used.

Figure 4.21 shows the typical **TDMA frame structure** of TETRA. Each **frame** consists of four slots (four channels in the V+D service per carrier), with a frame duration of 56.67 ms. Each **slot** carries 510 bits within 14.17 ms, i.e., 36 kbit/s. 16 frames together with one **control frame** (CF) form a **multiframe**, and finally, a **hyperframe** contains 60 multiframes. To avoid sending and receiving at the same time, TETRA shifts the uplink for a period of two slots compared to the downlink.

9 Formerly known as trans-European trunked radio, but worldwide marketing is better without "Europe" in the name (see DECT).

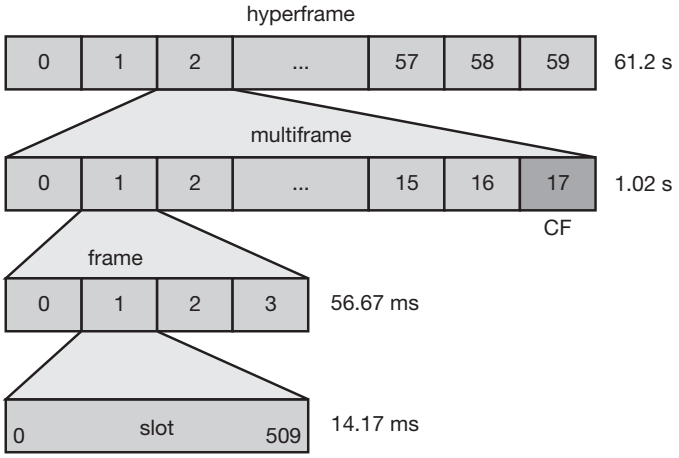

Figure 4.21 TETRA frame structure

TETRA offers **traffic channels (TCH)** and **control channels (CCH)** similar to GSM. Typical TCHs are TCH/S for voice transmission, and TCH/7.2, TCH/4.8, TCH/2.4 for data transmission (depending on the FEC mechanisms required).

However, in contrast to GSM, TETRA offers additional services like group call, acknowledged group call, broadcast call, and discreet listening. Emergency services need a sub-second group-call setup in harsh environments which possibly lack all infrastructure. These features are currently not available in GSM or other typical mobile telephone networks, so TETRA is complementary to other systems. TETRA has been chosen by many government organizations in Europe and China.

4.4 UMTS and IMT-2000

A lot has been written about third generation (or 3G) networks in the last few years. After a lot of hype and frustration these networks are currently deployed in many countries around the world. But how did it all start? First of all, the International Telecommunication Union (ITU) made a request for proposals for radio transmission technologies (RTT) for the **international mobile telecommunications (IMT) 2000** program (ITU, 2002), (Callendar, 1997), (Shafi, 1998). IMT-2000, formerly called future public land mobile telecommunication system (FPLMTS), tried to establish a common worldwide communication system that allowed for terminal and user mobility, supporting the idea of universal personal telecommunication (UPT). Within this context, ITU has created several recommendations for FPLMTS systems, e.g., network architectures for FPLMTS (M.817), Requirements for the Radio Interface(s) for FPLMTS (M.1034), or Framework for Services Supported by FPLMTS (M.816). The number 2000 in IMT-2000 should indicate the start of the system (year 2000+x) and the spec-

trum used (around 2000 MHz). IMT-2000 includes different environments such as indoor use, vehicles, satellites and pedestrians. The world radio conference (WRC) 1992 identified 1885–2025 and 2110–2200 MHz as the frequency bands that should be available worldwide for the new IMT-2000 systems (Recommendation ITU-R M.1036). Within these bands, two times 30 MHz have been reserved for mobile satellite services (MSS).

Figure 4.22 shows the ITU frequency allocation (from the world administrative radio conference, 1992) together with examples from several regions that already indicate the problem of worldwide common frequency bands. In Europe, some parts of the ITU's frequency bands for IMT-2000 are already allocated for DECT (see section 4.2). The remaining frequencies have been split into bands for UTRA-FDD (uplink: 1920–1980 MHz, downlink: 2110–2170 MHz) and UTRA-TDD (1900–1920 MHz and 2010–2025 MHz). The technology behind UTRA-FDD and –TDD will subsequently be explained in more detail as they form the basis of UMTS. Currently, no other system is planned for IMT-2000 in Europe. More bandwidth is available in China for the Chinese 3G system TD-SCDMA or possibly other 3G technologies (such as W-CDMA or cdma2000 – it is still open which system will dominate the Chinese market; Chen, 2002). Again slightly different frequencies are used by the 3G services in Japan, which are based on W-CDMA (like UTRA-FDD) or cdma2000. An open question is the future of 3G in the US as the ITU's frequency bands have already been allocated for 2G networks or are reserved for other use. In addition to the original frequency allocations, the world radio conference (WRC) allocated new terrestrial IMT-2000 bands in the range of 800–1000 MHz, 1700–1900 MHz and 2500–2700 MHz in 2000. This approach includes the reuse of 2G spectrum (Evci, 2001).

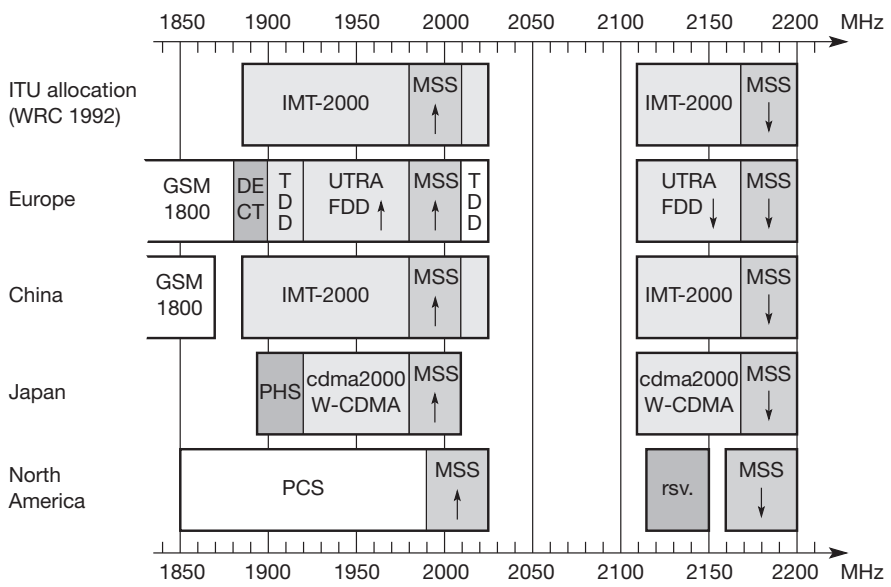

Figure 4.22 IMT-2000 frequencies

Now the reader might be confused by all the different technologies mentioned in the context of IMT-2000. Wasn't the plan to have a common global system? This was the original plan, but after many political discussions and fights about patents this idea was dropped and a so-called family of 3G standards was adopted.

For the RTT, several proposals were received in 1998 for indoor, pedestrian, vehicular, and satellite environments. These came from many different organizations, e.g., **UWC-136** from the Universal Wireless Communications Consortium (US) that extends the IS-136 standard into the third generation systems, **cdma2000** that is based on the IS-95 system (US), and wideband packet CDMA (WP-CDMA) which tries to align to the European UTRA proposal. Basically, three big regions were submitting proposals to the ITU: ETSI for Europe, ARIB (Association of Radio Industries and Broadcasting) and TTC (Telecommunications Technology Council) for Japan, and ANSI (American National Standards Institute) for the US.

The European proposal for IMT-2000 prepared by ETSI is called **universal mobile telecommunications system (UMTS)** (Dasilva, 1997), (Ojanperä, 1998), the specific proposal for the radio interface RTT is **UMTS (now: universal) terrestrial radio access (UTRA)** (ETSI, 1998n), (UMTS Forum, 2002). UMTS as initially proposed by ETSI rather represents an evolution from the second generation GSM system to the third generation than a completely new system. In this way, many solutions have been proposed for a smooth transition from GSM to UMTS, saving money by extending the current system rather than introducing a new one (GSMMoU, 1998).

One initial enhancement of GSM toward UMTS was **enhanced data rates for global (or: GSM) evolution (EDGE)**, which uses enhanced modulation schemes (8 PSK instead of GSM's GMSK, see chapter 2) and other techniques for data rates of up to 384 kbit/s using the same 200 kHz wide carrier and the same frequencies as GSM (i.e., a data rate of 48 kbit/s per time slot is available). EDGE can be introduced incrementally offering some channels with EDGE enhancement that can switch between EDGE and GSM/GPRS. In Europe, EDGE was never used as a step toward UMTS but operators directly jumped onto UMTS. However, EDGE can also be applied to the US IS-136 system and may be a choice for operators that want to enhance their 2G systems (3G Americas, 2002).

Besides enhancing data rates, new additions to GSM, like **customized application for mobile enhanced logic (CAMEL)** introduce intelligent network support. This system supports, for example, the creation of a **virtual home environment (VHE)** for visiting subscribers. GSMMoU (1998) provides many proposals covering QoS aspects, roaming, services, billing, accounting, radio aspects, core networks, access networks, terminal requirements, security, application domains, operation and maintenance, and several migration aspects.

UMTS fits into a bigger framework developed in the mid-nineties by ETSI, called **global multimedia mobility (GMM)**. GMM provides an architecture to integrate mobile and fixed **terminals**, many different **access network**s (GSM BSS, DECT, ISDN, UMTS, LAN, WAN, CATV, MBS), and several **core transport**

networks (GSM NSS+IN, ISDN+IN, B-ISDN+TINA, TCP/IP) (ETSI, 2002). Within this framework, ETSI developed **basic requirements** for UMTS and for UTRA, the radio interface (ETSI, 1998h). Key requirements are minimum data rates of 144 kbit/s for rural outdoor access (with the goal of 384 kbit/s) at a maximum speed of 500 km/h.[10] For suburban outdoor use a minimum of 384 kbit/s should be achieved with the goal of 512 kbit/s at 120 km/h. For indoor or city use with relatively short ranges, up to 2 Mbit/s are required at 10 km/h (walking).

UMTS should also provide several bearer services, real-time and non real-time services, circuit and packet switched transmission, and many different data rates. Handover should be possible between UMTS cells, but also between UMTS and GSM or satellite networks. The system should be compatible with GSM, ATM, IP, and ISDN-based networks. To reflect the asymmetric bandwidth needs of typical users, UMTS should provide a variable division of uplink and downlink data rates. Finally, UMTS has to fit into the IMT-2000 framework (this is probably the decisive factor for its success). As the global UMTS approach is rather ambitious, a more realistic alternative for the initial stages would be UMTS cells in cities providing a subset of services.

Several companies and interest groups have handed in proposals for UTRA (ETSI, 1998i), of which ETSI selected two for UMTS in January 1998. For the **paired band** (using FDD as a duplex mechanism), ETSI adopted the **W-CDMA** (Wideband CDMA) proposal, for the **unpaired band** (using TDD as duplex mechanism) the **TD-CDMA** (Time Division CDMA) proposal is used (Adachi, 1998), (Dahlman, 1998), (ETSI, 1998n). The paired band is typically used for public mobile network providers (wide area, see GSM), while the unpaired band is often used for local and indoor communication (see DECT). The following sections will present key properties of the initial UMTS system.

What happened to the IMT-2000 family? Figure 4.23 gives an overview. As a single standard could not be found, the ITU standardized five groups of 3G radio access technologies.

- **IMT-DS:** The **direct spread** technology comprises wideband CDMA (**W-CDMA**) systems. This is the technology specified for UTRA-FDD and used by all European providers and the Japanese NTT DoCoMo for 3G wide area services. To avoid complete confusion ITU's name for the technology is IMT-DS, ETSI called it UTRA-FDD in the UMTS context, and technology used is called W-CDMA (in Japan this is promoted as FOMA, freedom of mobile multimedia access). Today, standardization of this technology takes place in 3GPP (Third generation partnership project, 3GPP, 2002a). Section 4.4.1 provides more detail about the standardization process.
- **IMT-TC:** Initially, this family member, called **time code**, contained only the UTRA-TDD system which uses time-division CDMA (**TD-CDMA**). Later on, the Chinese proposal, TD-synchronous CDMA (**TD-SCDMA**) was added.

10 This speed is a problem as currently, only DAB can provide higher bit rates at high speeds.

Both standards have been combined and 3GPP fosters the development of this technology. It is unclear when and to what extent this technology will be introduced. The initial UMTS installations are based on W-CDMA.

- **IMT-MC:** cdma2000 is a **multi-carrier** technology standardized by 3GPP2 (Third generation partnership project 2, 3GPP2, 2002), which was formed shortly after 3GPP to represent the second main stream in 3G technology. Version cdma2000 EV-DO has been accepted as the 3G standard.
- **IMT-SC:** The enhancement of the US TDMA systems, UWC-136, is a **single carrier** technology originally promoted by the Universal Wireless Communications Consortium (UWCC). It is now integrated into the 3GPP efforts. This technology applies EDGE, among others, to enhance the 2G IS-136 standard.
- **IMT-FT**: As **frequency time** technology, an enhanced version of the cordless telephone standard DECT has also been selected for applications that do not require high mobility. ETSI is responsible for the standardization of DECT.

The main driving forces in the standardization process are 3GPP and 3GPP2. ETSI has moved its GSM standardization process to 3GPP and plays a major role there. 3GPP tends to be dominated by European and Japanese manufacturers and standardization bodies, while 3GPP2 is dominated by the company Qualcomm and CDMA network operators. The quarrels between Qualcomm and European manufacturers (e.g., Nokia, Ericsson) regarding CDMA patents (UMTS and cdma2000 use CDMA) even escalated into the political arena back in 1998 (US vs EU). Everything cooled down when hundreds of patents had been exchanged and the systems had been harmonized (e.g., CDMA chipping rates).

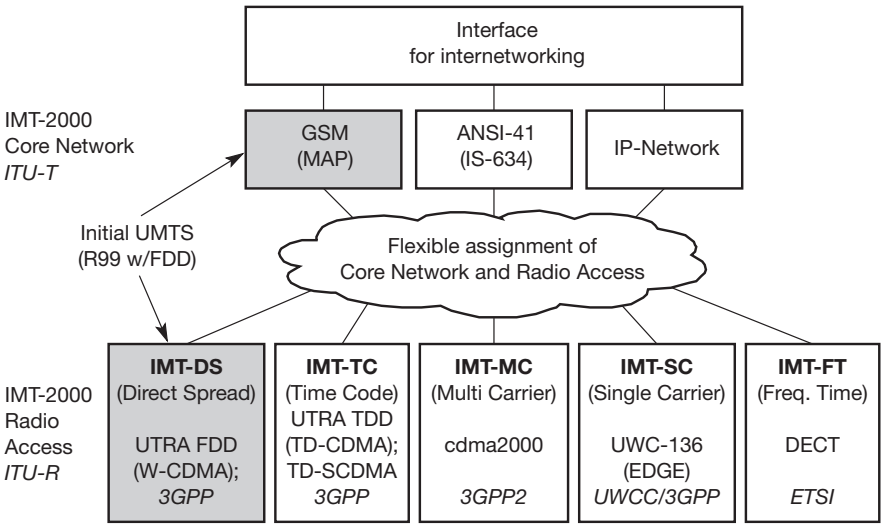

Figure 4.23 The IMT-2000 family

Figure 4.23 shows more than just the radio access technologies. One idea of IMT-2000 is the flexible assignment of a core network to a radio access system. The classical core network uses SS7 for signaling which is enhanced by ANSI-41 (cdmaOne, cdma2000, TDMA) or MAP (GSM) to enable roaming between different operators. The evolution toward 4G systems is indicated by the use of all-IP core networks (see Chapter 11). Obviously, internet-working functions have to be provided to enable cross-system data transfer, roaming, billing etc.

4.4.1 UMTS releases and standardization

UMTS as discussed today and introduced in many countries relies on the initial release of the UMTS standard called **release 99** or **R99** for short. This release of the specification describes the new radio access technologies UTRA FDD and UTRA TDD, and standardizes the use of a GSM/GPRS network as core within 440 separate specifications. This enables a cost effective migration from GSM to UMTS. The initial installations will even offer the FDD mode only as indicated in Figure 4.23. This release was (almost) finalized in 1999 – hence the name R99. The following sections will focus on this release as it is unclear when, and to what extent, the following releases will be realized.

After R99 the release 2000 or R00 followed. However, in September 2000 3GPP realized that it would be impossible to finalize the standard within the year 2000. 3GPP decided to split R2000 into two standards and call them release 4 (Rel-4) and release 5 (Rel-5). The version of all standards finalized for R99 start with 3.x.y (a reason for renaming R99 into Rel-3), Rel-4 and Rel-5 versions start with 4.x.y and 5.x.y, respectively. The standards are grouped into series. For example, radio aspects are specified in series 25, technical realization in series 23, and codecs in series 26. The complete standard number (e.g., TS 25.401 V3.10.0) then identifies the series (25), the standard itself (401), the release (3), and the version within the release (10.0). All standards can be downloaded from `www.3gpp.org` (the example given is the UTRAN overall description, release 99, from June 2002).

Release 4 introduces quality of service in the fixed network plus several execution environments (e.g., MExE, mobile execution environment, see chapter 10) and new service architectures. Furthermore, the Chinese proposal, TD-SCDMA was added as low chiprate option to UTRA-TDD (only 1.28 Mchip/s occupying only 1.6 MHz bandwidth). This release already consists of over 500 specifications and was frozen in March 2001.

Release 5 specifies a radically different core network. The GSM/GPRS based network will be replaced by an almost all-IP-core. While the radio interfaces remain the same, the changes in the core are tremendous for telecommunication network operators who have used traditional telephone technologies for many years. The content of this specification was frozen March 2002. This standard integrates IP-based multimedia services (IMS) controlled by the IETF's session initiation protocol (SIP, RFC 3261; Rosenberg, 2002; SIP Forum, 2002). A high speed downlink packet access (HSDPA) with speeds in the order of

8–10 Mbit/s was added as well as a wideband 16 kHz AMR codec for better audio quality. Additional features are end-to-end QoS messaging and several data compression mechanisms.

3GPP is currently working on **release 6** (and thinking of release 7) which is expected to be frozen in March 2003. This release comprises the use of multiple input multiple output (MIMO) antennas, enhanced MMS, security enhancements, WLAN/UMTS interworking, broadcast/multicast services, enhanced IMS, IP emergency calls, and many more management features (3GPP, 2002a).

The reader should not forget that many companies still have to make any money from, release 99, so it is not clear at what time and to what extent the new releases will be implemented. The following describes the initial UMTS standard, release 99, which is currently deployed.

4.4.2 UMTS system architecture

Figure 4.24 shows the very simplified UMTS reference architecture which applies to both UTRA solutions (3GPP, 2000). The **UTRA network (UTRAN)** handles cell level mobility and comprises several **radio network subsystems (RNS)**. The functions of the RNS include radio channel ciphering and deciphering, handover control, radio resource management etc. The UTRAN is connected to the **user equipment (UE)** via the radio interface $\mathbf{U_u}$ (which is comparable to the U_m interface in GSM). Via the $\mathbf{I_u}$ interface (which is similar to the A interface in GSM), UTRAN communicates with the **core network (CN)**. The CN contains functions for inter-system handover, gateways to other networks (fixed or wireless), and performs location management if there is no dedicated connection between UE and UTRAN.

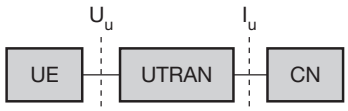

Figure 4.24 Main components of the UMTS reference architecture

UMTS further subdivides the above simplified architecture into so-called **domains** (see Figure 4.25). The **user equipment** domain is assigned to a single user and comprises all the functions that are needed to access UMTS services. Within this domain are the USIM domain and the mobile equipment domain. The **USIM** domain contains the SIM for UMTS which performs functions for encryption and authentication of users, and stores all the necessary user-related data for UMTS. Typically, this USIM belongs to a service provider and contains a micro processor for an enhanced program execution environment (USAT, UMTS SIM application toolkit). The end device itself is in the **mobile equipment** domain. All functions for radio transmission as well as user interfaces are located here.

The **infrastructure** domain is shared among all users and offers UMTS services to all accepted users. This domain consists of the **access network** domain, which contains the radio access networks (RAN), and the core network domain, which contains access network independent functions. The **core network** domain can be separated into three domains with specific tasks. The **serving network** domain comprises all functions currently used by a user for accessing

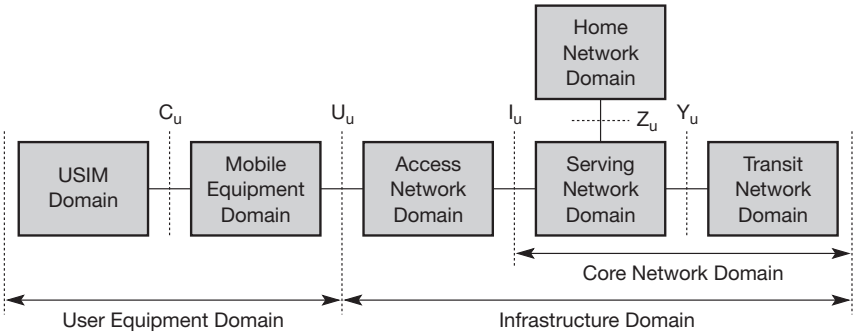

Figure 4.25
UMTS domains and interfaces

UMTS services. All functions related to the home network of a user, e.g., user data look-up, fall into the **home network** domain. Finally, the **transit network** domain may be necessary if, for example, the serving network cannot directly contact the home network. All three domains within the core network may be in fact the same physical network. These domains only describe functionalities.

4.4.3 UMTS radio interface

The biggest difference between UMTS and GSM comes with the new radio interface (U_u). The duplex mechanisms are already well known from GSM (FDD) and DECT (TDD). However, the direct sequence (DS) CDMA used in UMTS is new (for European standards, not in the US where CDMA technology has been available since the early nineties). DS-CDMA was introduced in chapters 2 and 3. This technology multiplies a stream of bits with a chipping sequence. This spreads the signal and, if the chipping sequence is unique, can separate different users. All signals use the same frequency band (in UMTS/IMT-2000 5 MHz-wide bands have been specified and licensed to network operators). To separate different users, the codes used for spreading should be (quasi) orthogonal, i.e., their cross-correlation should be (almost) zero.

UMTS uses a constant **chipping rate** of 3.84 Mchip/s. Different user data rates can be supported using different spreading factors (i.e., the number of chips per bit). Figure 4.26 shows the basic ideas of spreading and separation of different senders in UMTS. The first step in a sender is spreading of user data ($data_i$) using orthogonal **spreading** codes. Using orthogonal codes separates the different data streams of a sender. UMTS uses so-called **orthogonal variable spreading factor (OVSF)** codes. Figure 4.27 shows the basic idea of OVSF. Orthogonal codes are generated by doubling a chipping sequence X with and without flipping the sign of the chips. This results in X and –X, respectively. Doubling the chipping sequence also results in spreading a bit twice as much as before. The spreading factor SF=n becomes 2n. Starting with a spreading factor of 1, Figure 4.27 shows the generation of orthogonal codes with different spreading factors. Two codes are orthogonal as long as one code is never a part of the other code. Looking at the coding tree in Figure 4.27 and considering the construction of the codes, orthogonality is guaranteed if one code has not been generated based on another. For example, if a sender uses the code (1,–1)

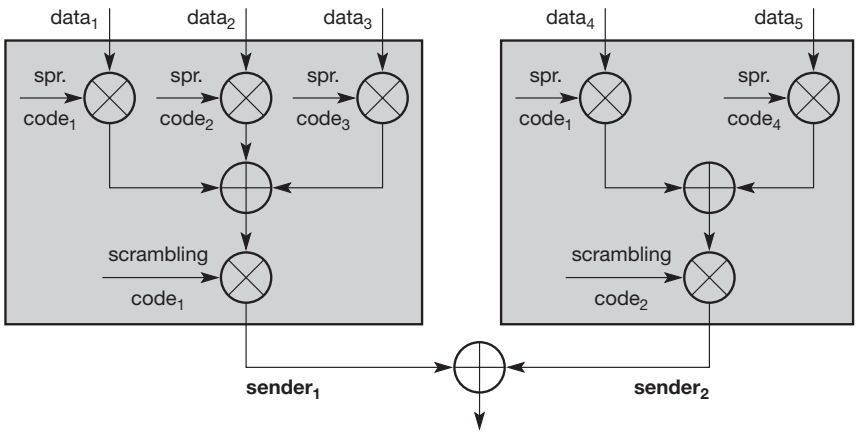

Figure 4.26 Spreading and scrambling of user date

with spreading factor 2, it is not allowed to use any of the codes located in the subtrees generated out of (1,–1). This means that, e.g., (1,–1,1,–1), (1,–1,–1,1,–1,1,1,–1), or (1,–1,–1,1,–1,1,1,–1,–1,1,1,–1,1,–1,–1,1) cannot be used anymore. However, it is no problem to use codes with different spreading factors if one code has not been generated using the other. Thus, (1,–1) block only the lower subtree in Figure 4.27, many other codes from the upper part can still be used. An example for a valid combination in OVSF is (1,–1), (1,1,–1,–1), (1,1,1,1,1,1,1,1), (1,1,1,1,–1,–1,–1,–1, 1,1,1,1,–1,–1,–1,–1), (1,1,1,1,–1,–1,–1,–1,–1,–1,–1,–1,1,1,1,1). This combination occupies the whole code spaces and allows for the transmission of data with different spreading factors (2, 4, 8, and 2*16). This example shows the tight coupling of available spreading factors and orthogonal codes.

Now remember that UMTS uses a constant chipping rate (3.84 Mchip/s). Using different spreading factors this directly translates into the support of different data rates. If the chipping rate is constant, doubling the spreading factor means dividing the data rate by two. But this also means that UMTS can only support a single data stream with SF=1 as then no other code may be used. Using the example combination above, a stream with half the maximum data rate, one with a fourth, one with an eighth, and two with a sixteenth are supported at the same time.

Each sender uses OVSF to spread its data streams as Figure 4.26 shows. The spreading codes chosen in the senders can be the same. Using different spreading codes in all senders within a cell would require a lot of management and would increase the complexity. After spreading all chip streams are added and scrambled. **Scrambling** does not spread the chip sequence any further but XORs chips based on a code. In the **FDD** mode, this scrambling code is unique for each sender and separates all senders (UE and base station) in a cell. After scrambling, the signals of different senders are quasi orthogonal. Quasi-orthogonal signals have the nice feature that they stay quasi-orthogonal even if they are not synchronized. Using orthogonal codes would require chip-synchronous reception and tight synchronization (this is done in other CDMA networks). For **TDD** the scrambling code is cell specific, i.e., all stations in a cell use the same scrambling code and cells are separated using different codes. The scrambled chips are **QPSK** modulated and transmitted.

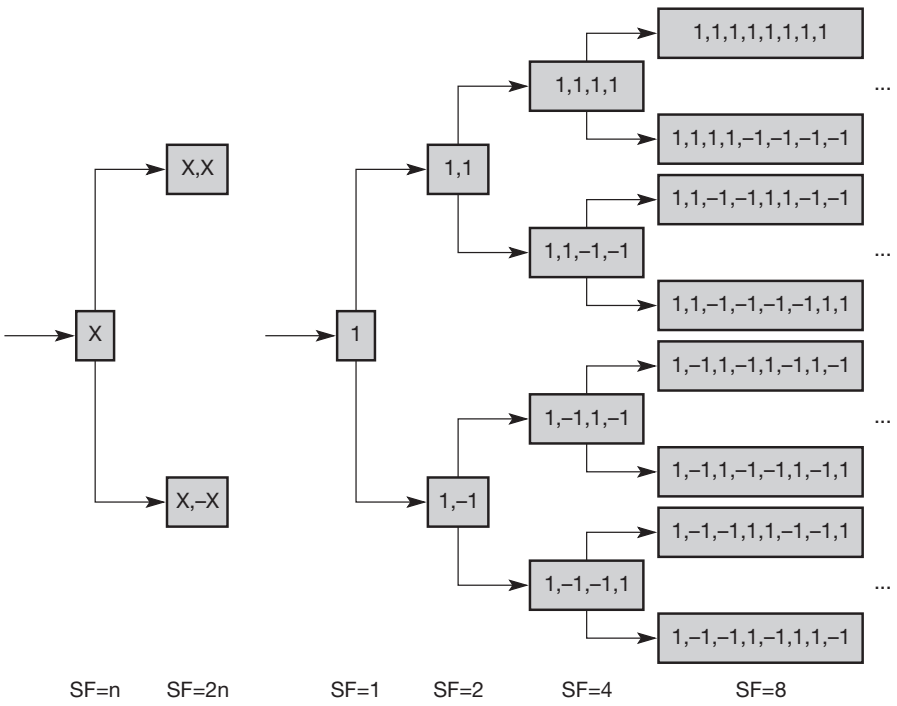

Figure 4.27
OVSF code tree used for orthogonal spreading

4.4.3.1 UTRA-FDD (W-CDMA)

The FDD mode for UTRA uses **wideband CDMA (W-CDMA)** with direct sequence spreading. As implied by FDD, uplink and downlink use different frequencies. A mobile station in Europe sends via the uplink using a carrier between 1920 and 1980 MHz, the base station uses 2110 to 2170 MHz for the downlink (see Figure 4.22). Figure 4.28 shows a radio frame comprising 15 time slots. Time slots in W-CDMA are not used for user separation but to support periodic functions (note that this is in contrast to, e.g., GSM, where time slots are used to separate users!). A radio frame consists of 38,400 chips and has a duration of 10 ms. Each time slot consists of 2,560 chips, which roughly equals 666.6 μs.[11] The occupied bandwidth per W-CDMA channel is 4.4 to 5 MHz (channel spacing can be varied to avoid interference between channels of different operators). These 5 MHz bands of the spectrum have been sold in many countries using an auction or a beauty contest. In Germany, the FDD spectrum was sold for over 50 billion Euros during an auction! But that was at a time when marketing people tried to convince everyone that UMTS would bring

11 Early version of W-CDMA specified a chipping rate of 4.096 M chip/s and 16 time slots per frame. This was changed during the harmonization process which was necessary to avoid patent conflicts and to enable devices that can handle different CDMA standards. The harmonization process is fostered by the operators harmonization group (OHG), which is an informal steering group of wireless operator companies promoting 3G harmonization. The OHG was founded in 1999.

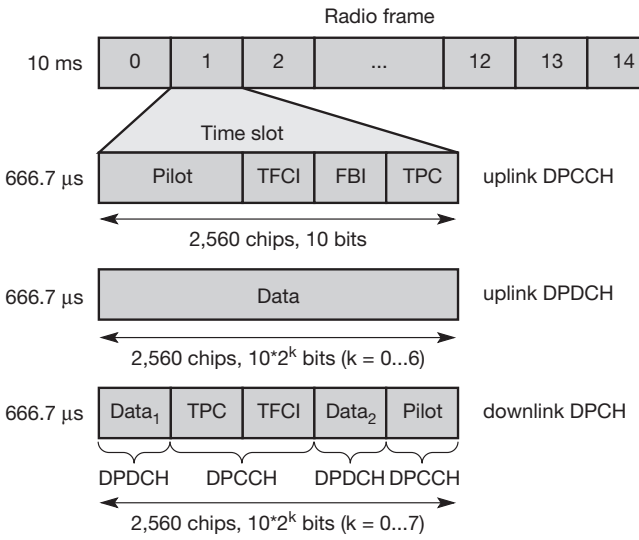

Figure 4.28
UTRA FDD (W-CDMA) frame structure

high-bandwidth applications to any mobile device with high profits for all. Today, most people are much more realistic and know that data rates will be quite low in the beginning (150 kbit/s per user are realistic, 2 Mbit/s are not). The capacity of a cell under realistic assumptions (interference etc.), i.e., the sum of all data rates, will rather be 2 Mbit/s. To provide high data rates a lot of money has to be invested in the infrastructure: UTRA FDD requires at least twice as many base stations as GSM; cell diameters of 500 m will be commonplace. This shows clearly that this technology will not cover whole countries in the near future but cities and highways only. People in the countryside will have to rely on GSM/GPRS for many more years to come.

Back to the frame structure shown in Figure 4.28. Similar to GSM, UMTS defines many logical and physical channels, and their mapping. The figure shows three examples of physical channels as they are used for data transmission. Two physical channels are shown for the uplink.

- **Dedicated physical data channel (DPDCH):** This channel conveys user or signaling data. The spreading factor of this channel can vary between 4 and 256. This directly translates into the data rates this channel can offer: 960 kbit/s (spreading factor 4, 640 bits per slot, 15 slots per frame, 100 frames per second), 480, 240, 120, 60, 30, and 15 kbit/s (spreading factor 256). This also shows one of the problems of using OVSF for spreading: only certain multiples of the basic data rate of 15 kbit/s can be used. If, for example, 250 kbit/s are needed the device has to choose 480 kbit/s, which wastes bandwidth. In each connection in layer 1 it can have between zero and six DPDCHs. This results in a theoretical maximum data rate of 5,740 kbit/s (UMTS describes UEs with a maximum of 1,920 kbit/s only). Table 4.7 shows typical user data rates together with the required data rates on the physical channels.

User data rate [kbit/s]	12.2 (voice)	64	144	384
DPDCH [kbit/s]	60	240	480	960
DPCCH [kbit/s]	15	15	15	15
Spreading	64	16	8	4

Table 4.7 Typical UTRA–FDD uplink data rates

- **Dedicated physical control channel (DPCCH):** In each connection layer 1 needs exactly one DPCCH. This channel conveys control data for the physical layer only and uses the constant spreading factor 256. The **pilot** is used for channel estimation. The **transport format combination identifier (TFCI)** specifies the channels transported within the DPDCHs. Signaling for a soft handover is supported by the **feedback information field (FBI)**. The last field, **transmit power control (TPC)** is used for controlling the transmission power of a sender. Power control is performed in each slot, thus 1,500 power control cycles are available per second. Tight power control is necessary to mitigate near-far-effects as explained in chapter 2. Six different DPCCH bursts have been defined which differ in the size of the fields.
- **Dedicated physical channel (DPCH):** The downlink time multiplexes control and user data. Spreading factors between 4 and 512 are available. Again, many different burst formats (17 altogether) have been defined which differ in the size of the field shown in Figure 4.28. The available data rates for data channels (DPDCH) within a DPCH are 6 (SF=512), 24, 51, 90, 210, 432, 912, and 1,872 kbit/s (SF=4).

While no collisions can occur on the downlink (only the base station sends on the downlink), medium access on the uplink has to be coordinated. A **physical random access channel (PRACH)** is used for this purpose. UTRA-FDD defines 15 random access slots within 20 ms; within each access slot 16 different access preambles can be used for random access. Using slotted Aloha, a UE can access an access slot by sending a preamble. The UE starts with the lowest available transmission power to avoid interfering with other stations. If no positive acknowledgement is received, the UE tries another slot and another preamble with the next higher power level (power ramping). The number of available access slots can be defined per cell and is transmitted via a broadcast channel to all UEs.

A UE has to perform the following steps during the **search for a cell** after power on:

- **Primary synchronization:** A UE has to synchronize with the help of a 256 chip primary synchronization code. This code is the same for all cells and helps to synchronize with the time slot structure.

- **Secondary synchronization:** During this second phase the UE receives a secondary synchronization code which defines the group of scrambling codes used in this cell. The UE is now synchronized with the frame structure.
- **Identification of the scrambling code:** The UE tries all scrambling codes within the group of codes to find the right code with the help of a correlator. After these three steps the UE can receive all further data over a broadcast channel.

4.4.3.2 UTRA-TDD (TD-CDMA)

The second UTRA mode, UTRA-TDD, separates up and downlink in time using a radio frame structure similar to FDD. 15 slots with 2,560 chips per slot form a radio frame with a duration of 10 ms. The chipping rate is also 3.84 Mchip/s. To reflect different user needs in terms of data rates, the TDD frame can be **symmetrical** or **asymmetrical**, i.e., the frame can contain the same number of uplink and downlink slots or any arbitrary combination. The frame can have only one **switching point** from uplink to downlink or several switching points. However, at least one slot must be allocated for the uplink and downlink respectively.

The system can change the spreading factor (1, 2, 4, 8, 16) as a function of the desired data rate. Using the burst type shown in Figure 4.29 results in data rates of 6,624, 3,312, 1,656, 828, and 414 kbit/s respectively (if all slots are used for data transmission). The figure shows a burst of type 2 which comprises two **data** fields of 1,104 chips each. Spreading is applied to these data fields only. Additionally, a **midample** is used for training and channel estimation. As TDD uses the same scrambling codes for all stations, the stations must be tightly synchronized and the spreading codes are available only once per slot. This results in a maximum number of 16 simultaneous sending stations. To loosen the tight synchronization a little bit, a **guard period (GP)** has been introduced at the end of each slot. Due to the tight synchronization and the use of orthogonal codes, a simpler power control scheme with less power control cycles (e.g., 100 per second) is sufficient.

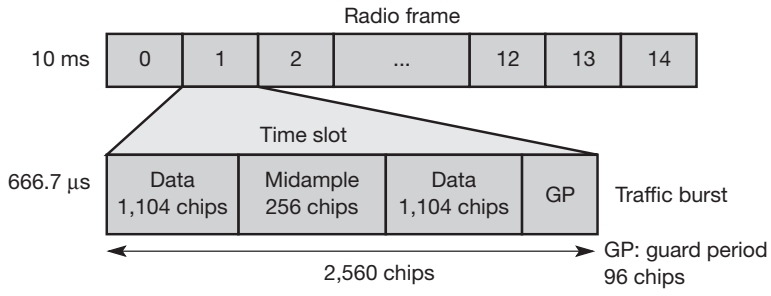

Figure 4.29 UTRA TDD (TD–CDMA) frame structure

UTRA TDD occupies 5 MHz bandwidth per channel as UTRA FDD does per direction (FDD needs 2x 5 MHz). Compared to the license for FDD, TDD was quite cheap. Germany paid less than €300 million. Figure 4.22 shows the location of the spectrum for this UMTS mode, but it is unclear to what extend this system will be deployed. The coverage per cell is even less than using FDD, UEs must not move too fast – this sounds like the characteristics of WLANs which are currently deployed in many places.

4.4.4 UTRAN

Figure 4.30 shows the basic architecture of the UTRA network (UTRAN; 3GPP, 2002b). This consists of several **radio network subsystems (RNS)**. Each RNS is controlled by a **radio network controller (RNC)** and comprises several components that are called node B. An RNC in UMTS can be compared with the BSC; a node B is similar to a BTS. Each **node B** can control several antennas which make a radio cell. The mobile device, UE, can be connected to one or more antennas as will subsequently be explained in the context of handover. Each RNC is connected with the core network (CN) over the interface $\mathbf{I_u}$ (similar to the role of the A interface in GSM) and with a node B over the interface $\mathbf{I_{ub}}$. A new interface, which has no counterpart in GSM, is the interface $\mathbf{I_{ur}}$ connecting two RNCs with each other. The use of this interface is explained together with the UMTS handover mechanisms.

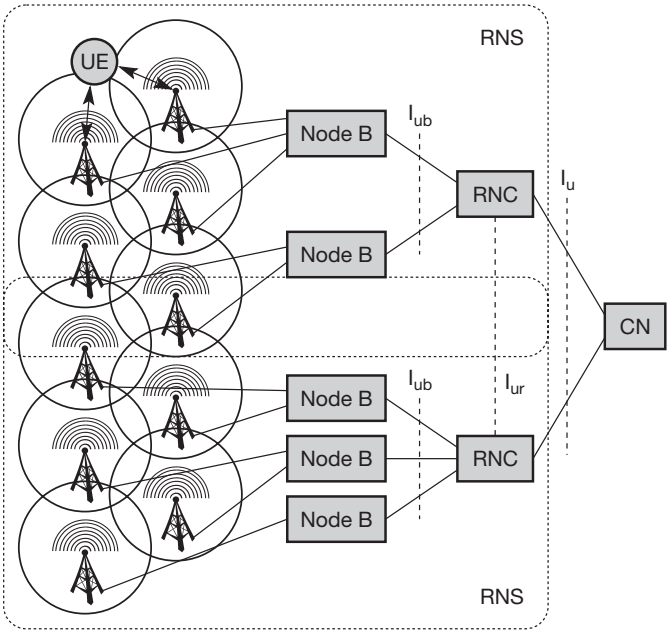

Figure 4.30
Basic architecture of the UTRA network

4.4.4.1 Radio network controller

An **RNC** in UMTS has a broad spectrum of tasks as listed in the following:

- **Call admission control:** It is very important for CDMA systems to keep the interference below a certain level. The RNC calculates the traffic within each cell and decides, if additional transmissions are acceptable or not.
- **Congestion control:** During packet-oriented data transmission, several stations share the available radio resources. The RNC allocates bandwidth to each station in a cyclic fashion and must consider the QoS requirements.
- **Encryption/decryption:** The RNC encrypts all data arriving from the fixed network before transmission over the wireless link (and vice versa).
- **ATM switching and multiplexing, protocol conversion:** Typically, the connections between RNCs, node Bs, and the CN are based on ATM. An RNC has to switch the connections to multiplex different data streams. Several protocols have to be converted – this is explained later.
- **Radio resource control:** The RNC controls all radio resources of the cells connected to it via a node B. This task includes interference and load measurements. The priorities of different connections have to be obeyed.
- **Radio bearer setup and release:** An RNC has to set-up, maintain, and release a logical data connection to a UE (the so-called UMTS radio bearer).
- **Code allocation:** The CDMA codes used by a UE are selected by the RNC. These codes may vary during a transmission.
- **Power control:** The RNC only performs a relatively loose power control (the outer loop). This means that the RNC influences transmission power based on interference values from other cells or even other RNCs. But this is not the tight and fast power control performed 1,500 times per second. This is carried out by a node B. This outer loop of power control helps to minimize interference between neighbouring cells or controls the size of a cell.
- **Handover control and RNS relocation:** Depending on the signal strengths received by UEs and node Bs, an RNC can decide if another cell would be better suited for a certain connection. If the RNC decides for handover it informs the new cell and the UE as explained in subsection 4.4.6. If a UE moves further out of the range of one RNC, a new RNC responsible for the UE has to be chosen. This is called RNS relocation.
- **Management:** Finally, the network operator needs a lot of information regarding the current load, current traffic, error states etc. to manage its network. The RNC provides interfaces for this task, too.

4.4.4.2 Node B

The name node B was chosen during standardization until a new and better name was found. However, no one came up with anything better so it remained. A node B connects to one or more antennas creating one or more cells (or sectors in GSM speak), respectively. The cells can either use FDD or TDD

or both. An important task of a node B is the inner loop power control to mitigate near-far effects. This node also measures connection qualities and signal strengths. A node B can even support a special case of handover, a so-called softer handover which takes place between different antennas of the same node B (see section 4.4.6).

4.4.4.3 User equipment

The UE shown in Figure 4.30 is the counterpart of several nodes of the architecture.

- As the counterpart of a node B, the UE performs signal quality measurements, inner loop power control, spreading and modulation, and rate matching.
- As a counterpart of the RNC, the UE has to cooperate during handover and cell selection, performs encryption and decryption, and participates in the radio resource allocation process.
- As a counterpart of the CN, the UE has to implement mobility management functions, performs bearer negotiation, or requests certain services from the network.

This list of tasks of a UE, which is not at all exhaustive, already shows the complexity such a device has to handle. Additionally, users also want to have organizers, games, cameras, operating systems etc. and the stand-by time should be high.

4.4.5 Core network

Figure 4.31 shows a high-level view of the UMTS release 99 core network architecture together with a UTRAN RNS and a GSM BSS (see section 4.1). This shows the evolution from GSM/GPRS to UMTS. The core network (CN) shown here is basically the same as already explained in the context of GSM (see Figure 4.4) and GPRS (see Figure 4.16). The **circuit switched domain (CSD)** comprises the classical circuit switched services including signaling. Resources are reserved at connection setup and the GSM components MSC, GMSC, and VLR are used. The CSD connects to the RNS via a part of the I_u interface called $\mathbf{I_uCS}$. The CSD components can still be part of a classical GSM network connected to a BSS but need additional functionalities (new protocols etc.).

The **packet switched domain (PSD)** uses the GPRS components SGSN and GGSN and connects to the RNS via the $\mathbf{I_uPS}$ part of the I_u interface. Both domains need the data-bases EIR for equipment identification and HLR for location management (including the AuC for authentication and GR for user specific GPRS data).

Reusing the existing infrastructure helps to save a lot of money and may convince many operators to use UMTS if they already use GSM. The UMTS industry pushes their technology with the help of the market dominance of GSM. This is basically the same as cdma2000, which is a evolution of cdmaOne. The real flexible core network comes with releases 5 and 6, where the GSM

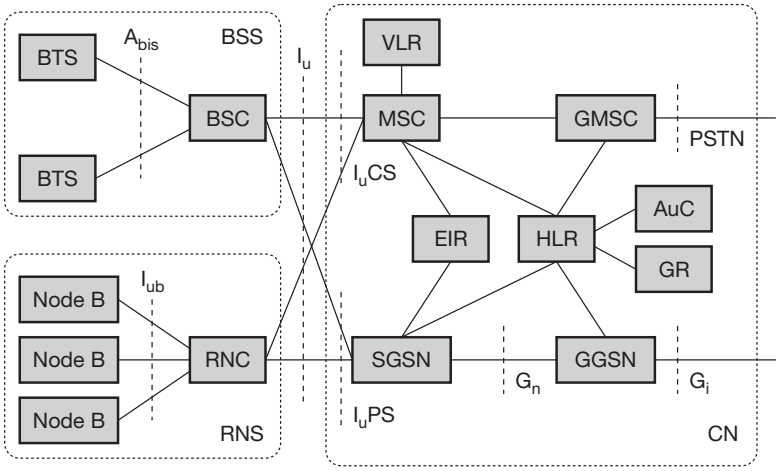

Figure 4.31
UMTS core network together with a 3G RNS and a 2G BSS

circuit switched part is being replaced by an all-IP core. Chapter 11 presents this idea in the context of 4G networks. It is not yet clear when this replacement of GSM will take place as many questions are still open (quality of service and security being the most important).

Figure 4.32 shows the protocol stacks of the users planes of the circuit switched and packet switched domains, respectively. The **CSD** uses the **ATM adaptation layer 2** (AAL2) for user data transmission on top of ATM as transport technology. The RNC in the UTRAN implements the radio link control (RLC) and the MAC layer, while the physical layer is located in the node B. The AAL2 **segmentation and reassembly** layer (SAR) is, for example, used to segment data packets received from the RLC into small chunks which can be transported in ATM. AAL2 and ATM has been chosen, too, because these protocols can transport and multiplex low bit rate voice data streams with low jitter and latency (compared to the protocols used in the PSD).

In the **PSD** several more protocols are needed. Basic data transport is performed by different lower layers (e.g., ATM with AAL5, frame relay). On top of these lower layers UDP/IP is used to create a UMTS internal IP network. All packets (e.g., IP, PPP) destined for the UE are encapsulated using the **GPRS tunneling protocol (GTP)**. The RNC performs protocol conversion from the combination GTP/UDP/IP into the **packet data convergence protocol (PDCP)**. This protocol performs header compression to avoid redundant data transmission using scarce radio resources. Comparing Figure 4.32 with Figure 4.17 (GPRS protocol reference model) shows a difference with respect to the tunnel. In UMTS the RNC handles the tunneling protocol GTP, while in GSM/GPRS GTP is used between an SGSN and GGSN only. The BSC in GSM is not involved in IP protocol processing.

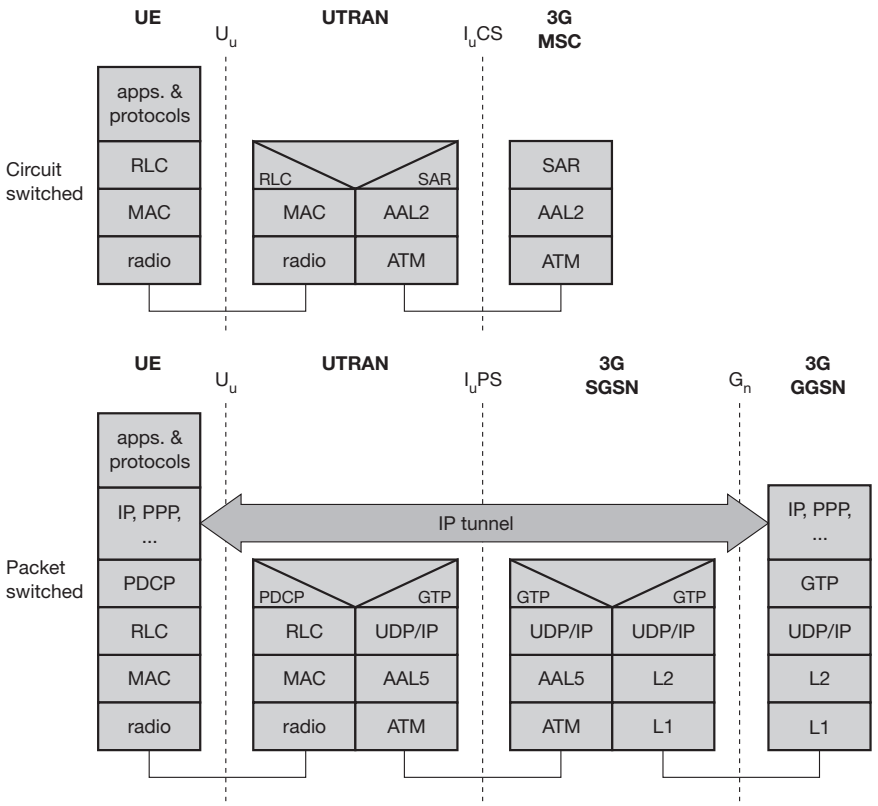

Figure 4.32
User plane protocol stacks (circuit and packet switched)

The **radio layer** (physical layer) depends on the UTRA mode (see sections 4.4.3.1 and 4.4.3.2). The **medium access control (MAC)** layer coordinates medium access and multiplexes logical channels onto transport channels. The MAC layers also help to identify mobile devices and may encrypt data. The **radio link control (RLC)** layer offers three different transport modes. The **acknowledged mode** transfer uses ARQ for error correction and guarantees one-time in-order delivery of data packets. The **unacknowledged mode** transfer does not perform ARQ but guarantees at least one-time delivery of packets with the help of sequence numbers. The **transparent mode** transfer simply forwards MAC data without any further processing. The system then has to rely on the FEC which is always used in the radio layer. The RLC also performs segmentation and reassembly and flow control. For certain services the RLC also encrypts.

4.4.6 Handover

UMTS knows two basic classes of handovers:

- **Hard handover:** This handover type is already known from GSM and other TDMA/FDMA systems. Switching between different antennas or different systems is performed at a certain point in time. **UTRA TDD** can only use this type. Switching between TDD cells is done between the slots of different frames. **Inter frequency handover**, i.e., changing the carrier frequency, is a hard handover. Receiving data at different frequencies at the same time requires a more complex receiver compared to receiving data from different sources at the same carrier frequency. Typically, all **inter system handovers** are hard handovers in UMTS. This includes handovers to and from GSM or other IMT-2000 systems. A special type of handover is the handover to a satellite system (inter-segment handover), which is also a hard handover, as different frequencies are used. However, it is unclear what technology will be used for satellite links if it will ever come. To enable a UE to listen into GSM or other frequency bands, UMTS specifies a **compressed mode** transmission for UTRA FDD. During this mode a UE stops all transmission. To avoid data loss, either the spreading factor can be lowered before and after the break in transmission (i.e., more data can be sent in shorter time) or less data is sent using different coding schemes.
- **Soft handover:** This is the real new mechanism in UMTS compared to GSM and is only available in the FDD mode. Soft handovers are well known from traditional CDMA networks as they use **macro diversity**, a basic property of CDMA. As shown in Figure 4.33, a UE can receive signals from up to three different antennas, which may belong to different node Bs. Towards the UE the RNC splits the data stream and forwards it to the node Bs. The UE combines the received data again. In the other direction, the UE simply sends its data which is then received by all node Bs involved. The RNC combines the data streams received from the node Bs. The fact that a UE receives data from different antennas at the same time makes a handover soft. Moving from one cell to another is a smooth, not an abrupt process.

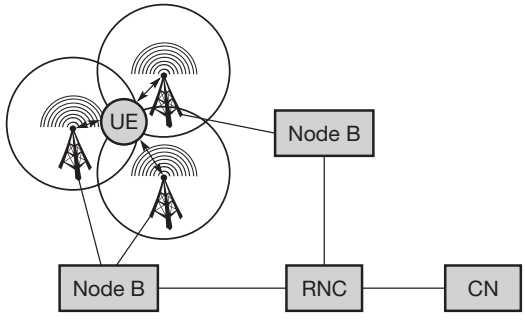

Figure 4.33 Marco-diversity supporting soft handovers

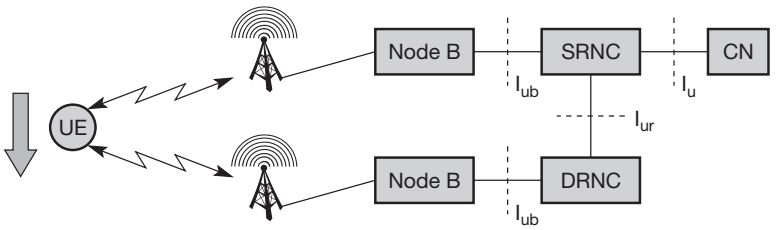

Figure 4.34 Serving RNC and drift RNC

Macro-diversity makes the transmission more robust with respect to fast fading, multi-path propagation, and shading. If one path is blocked by an obstacle the chances are good that data can still be received using another antenna. During a soft handover a UE receives power control commands from all involved node Bs. The UE then lowers transmission power as long as it still receives a command to lower the power. This avoids interference if, for example, the UE is in the transmission area of two antennas, one close, one further away. Without the above mechanism the UE's signal may be too strong when listening to the antenna further away. The lower the interference a UE introduces into a cell, the higher the capacity. Without this control, cell breathing would be even more problematic than it already is in CDMA networks.

As soft handover is not supported by the CN, all mechanisms related to soft handover must be located within UTRAN. Figure 4.34 shows a situation where a soft handover is performed between two node Bs that do not belong to the same RNC. In this case one RNC controls the connection and forwards all data to and from the CN. If the UE moves in the example from the upper cell to the lower cell, the upper RNC acts as a **serving RNC (SRNC)** while the other is the **drift RNC (DRNC)**. (If the whole RNS is considered, the terms are serving RNS and drift RNS, respectively.) The SRNC forwards data received from the CN to its node B and to the DRNC via the I_{ur} interface (splitting). This mechanism does not exist in, e.g., GSM. Data received by the lower node B is forwarded by the DRNC to the SRNC. The SRNC combines both data streams and forwards a single stream of data to the CN. The CN does not notice anything from the simultaneous reception. If the UE moves further down and drops out of the transmission area of the upper node B, two RNCs reserve resources for data transmission, SRNC and DRNC, although none of SRNC's node Bs transmit data for this UE. To avoid wasting resources, SRNC relocation can be performed. This involves the CN so is a hard handover.

Figure 4.35 gives an overview of several common handover types in a combined UMTS/GSM network (UMTS specifies ten different types which include soft and hard handover). The combination of a UTRA-FDD/GSM device will be the most common case in the beginning as coverage of 3G networks will be poor.

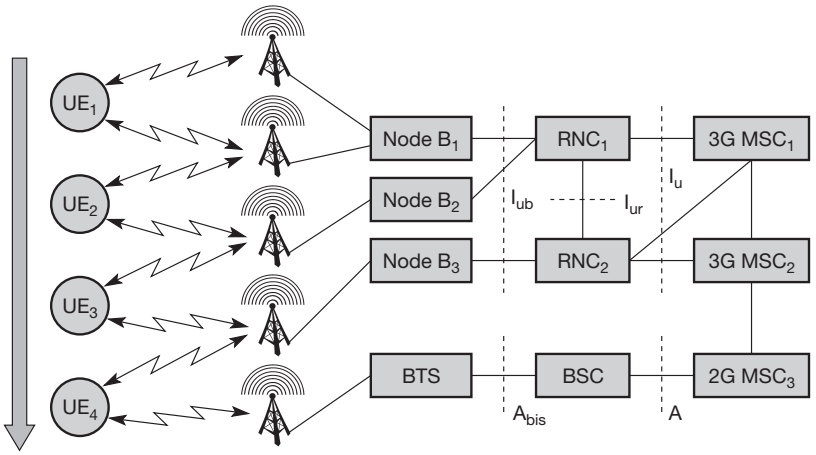

Figure 4.35 Overview of different handover types

- **Intra-node B, intra-RNC:** UE_1 moves from one antenna of node B_1 to another antenna. This type of handover is called **softer handover**. In this case node B_1 performs combining and splitting of the data streams.
- **Inter-node B, intra-RNC:** UE_2 moves from node B_1 to node B_2. In this case RNC_1 supports the soft handover by combining and splitting data.
- **Inter-RNC:** When UE_3 moves from node B_2 to node B_3 two different types of handover can take place. The **internal inter-RNC** handover is not visible for the CN, as described in Figure 4.34. RNC_1 can act as SRNC, RNC_2 will be the DRNC. The CN will communicate via the same interface I_u all the time. As soon as a relocation of the interface I_u takes place (relocation of the controlling RNC), the handover is called an **external inter-RNC** handover. Communication is still handled by the same MSC_1, but the external handover is now a hard handover.
- **Inter-MSC:** It could be also the case that MSC_2 takes over and performs a hard handover of the connection.
- **Inter-system:** UE_4 moves from a 3G UMTS network into a 2G GSM network. This hard handover is important for real life usability of the system due to the limited 3G coverage in the beginning.

4.5 Summary

This chapter has, for the most part, presented GSM as the most successful second generation digital cellular network. Although GSM was primarily designed for voice transmission, the chapter showed the evolution toward a more data-oriented transfer via HSCSD and GPRS. This evolution also includes the transition from a circuit-switched network to a packet-switched system that comes closer to the internet model. Other systems presented include DECT, the

digital standard for cordless phones, and TETRA, a trunked radio system. DECT can be used for wireless data transmission on a campus or indoors, but also for wireless local loops (WLL). For special scenarios, e.g., emergencies, trunked radio systems such as TETRA can be the best choice. They offer a fast connection setup (even within communication groups) and can work in an ad hoc network, i.e., without a base station.

The situation in the US is different from Europe. Based on the analog AMPS system, the US industry developed the TDMA system IS-54 that adds digital traffic channels. IS-54 uses dual mode mobile phones and incorporates several GSM ideas, such as, associated control channels, authentication procedures using encryption, and mobile assisted handover (called handoff). The Japanese PDC system was designed using many ideas in IS-54.

The next step, IS-136, includes digital control channels (IS-54 uses analog AMPS control channels) and is more efficient. Now fully digital phones can be used, several additional services are offered, e.g., voice mail, call waiting, identification, group calling, or SMS. IS-136 is also called North American TDMA (NA-TDMA) or Digital AMPS (D-AMPS) and operates at 800 and 1,900 MHz. Enhancements of D-AMPS/IS-136 toward IMT-2000 include advanced modulation techniques for the 30 kHz radio carrier, shifting data rates up to 64 kbit/s (first phase, called 136+). The second phase, called 136HS (High Speed) comprises a new air interface specification based on the EDGE technology.

IS-95 (promoted as cdmaOne) is based on CDMA, which is a completely different medium access method. Before deployment, the system was proclaimed as having many advantages over TDMA systems, such as its much higher capacity of users per cell, e.g., 20 times the capacity of AMPS. Today, CDMA providers are making more realistic estimates of around five times as many users. IS-95 offers soft handover, avoiding the GSM ping-pong effect (Wong, 1997). However, IS-95 needs precise synchronization of all base stations (using GPS satellites which are military satellites, so are not under control of the network provider), frequent power control, and typically, dual mode mobile phones due to the limited coverage. The basic ideas of CDMA have been integrated into most 3G systems.

This chapter also presented an overview of current and future third generation systems. UMTS, a proposal of operators and companies involved in the GSM business, was discussed in more detail. This standard is more an evolutionary approach than a revolution. To avoid even higher implementation costs, UMTS tries to reuse as much infrastructure as possible while introducing new services and higher data rates based on CDMA technology. The initial installations will basically use the GSM/GPRS infrastructure and offer only moderate data rates. The initial capacity of a UMTS cell is approximately 2 Mbit/s; cell diameters are in the order of 500 m. UMTS will be used to offload GSM networks and to offer enhanced data rates in cities as a first step. Future releases aim to replace the infrastructure by an (almost) all-IP network. These ideas will be presented together with a look at fourth generation systems in chapter 11. It

is quite clear that it will take a long time before 3G services are available in many places. It took GSM 10 years to become the most successful 2G mobile communication system. A similar period of time will be needed for 3G systems to succeed. Meanwhile, customers will need multiple mode phones offering, e.g., GSM 900/1800/1900 and UMTS UTRA-FDD services. It is not clear if and when UTRA-TDD will succeed. Providers already using cdmaOne will take the evolutionary path via cdma2000 1x toward the 3G system cdma2000 1x EV-DO. Several tests have already been conducted for 3G satellite services in the MSS spectrum (e.g., satellite based multicast, Nussli, 2002). However, right now many companies will wait before investing money in satellite services (see chapter 5). The main problem of multi-mode systems is the inter-system handover. While this chapter introduces handover scenarios within UMTS and GSM, and between GSM and UMTS, even more complex scenarios could comprise wireless LANs (see chapter 7) or other packet-oriented networks (Pahlavan, 2000).

4.6 Review exercises

1. Name some key features of the GSM, DECT, TETRA, and UMTS systems. Which features do the systems have in common? Why have the three older different systems been specified? In what scenarios could one system replace another? What are the specific advantages of each system?
2. What are the main problems when transmitting data using wireless systems that were made for voice transmission? What are the possible steps to mitigate the problems and to raise efficiency? How can this be supported by billing?
3. Which types of different services does GSM offer? Give some examples and reasons why these services have been separated.
4. Compared to the TCHs offered, standard GSM could provide a much higher data rate (33.8 kbit/s) when looking at the air interface. What lowers the data rates available to a user?
5. Name the main elements of the GSM system architecture and describe their functions. What are the advantages of specifying not only the radio interface but also all internal interfaces of the GSM system?
6. Describe the functions of the MS and SIM. Why does GSM separate the MS and SIM? How and where is user-related data represented/stored in the GSM system? How is user data protected from unauthorized access, especially over the air interface? How could the position of an MS (not only the current BTS) be localized? Think of the MS reports regarding signal quality.
7. Looking at the HLR/VLR database approach used in GSM – how does this architecture limit the scalability in terms of users, especially moving users?
8. Why is a new infrastructure needed for GPRS, but not for HSCSD? Which components are new and what is their purpose?

9 What are the limitations of a GSM cell in terms of diameter and capacity (voice, data) for the traditional GSM, HSCSD, GPRS? How can the capacity be increased?

10 What multiplexing schemes are used in GSM and for what purpose? Think of other layers apart from the physical layer.

11 How is synchronization achieved in GSM? Who is responsible for synchronization and why is it so important?

12 What are the reasons for the delays in a GSM system for packet data traffic? Distinguish between circuit-switched and packet-oriented transmission.

13 Where and when can collisions occur while accessing the GSM system? Compare possible collisions caused by data transmission in standard GSM, HSCSD, and GPRS.

14 Why and when are different signaling channels needed? What are the differences?

15 How is localization, location update, roaming, etc. done in GSM and reflected in the data bases? What are typical roaming scenarios?

16 Why are so many different identifiers/addresses (e.g., MSISDN, TMSI, IMSI) needed in GSM? Give reasons and distinguish between user-related and system-related identifiers.

17 Give reasons for a handover in GSM and the problems associated with it. What are the typical steps for handover, what types of handover can occur? Which resources need to be allocated during handover for data transmission using HSCSD or GPRS respectively? What about QoS guarantees?

18 What are the functions of authentication and encryption in GSM? How is system security maintained?

19 How can higher data rates be achieved in standard GSM, how is this possible with the additional schemes HSCSD, GPRS, EDGE? What are the main differences of the approaches, also in terms of complexity? What problems remain even if the data rate is increased?

20 What limits the data rates that can be achieved with GPRS and HSCSD using real devices (compared to the theoretical limit in a GSM system)?

21 Using the best delay class in GPRS and a data rate of 115.2 kbit/s – how many bytes are in transit before a first acknowledgement from the receiver could reach the sender (neglect further delays in the fixed network and receiver system)? Now think of typical web transfer with 10 kbyte average transmission size – how would a standard TCP behave on top of GPRS (see chapters 9 and 10)? Think of congestion avoidance and its relation to the round-trip time. What changes are needed?

22 How much of the original GSM network does GPRS need? Which elements of the network perform the data transfer?

23 What are typical data rates in DECT? How are they achieved considering the TDMA frames? What multiplexing schemes are applied in DECT and for what purposes? Compare the complexity of DECT with that of GSM.

24 Who would be the typical users of a trunked radio system? What makes trunked radio systems particularly attractive for these user groups? What are the main differences to existing systems for that purpose? Why are trunked radio systems cheaper compared to, e.g., GSM systems for their main purposes?

25 Summarize the main features of third generation mobile phone systems. How do they achieve higher capacities and higher data rates? How does UMTS implement asymmetrical communication and different data rates?

26 Compare the current situation of mobile phone networks in Europe, Japan, China, and North America. What are the main differences, what are efforts to find a common system or at least interoperable systems?

27 What disadvantage does OVSF have with respect to flexible data rates? How does UMTS offer different data rates (distinguish between FDD and TDD mode)?

28 How are different DPDCHs from different UEs within one cell distinguished in UTRA FDD?

29 Which components can perform combining/splitting at what handover situation? What is the role of the interface I_{ur}? Why can CDMA systems offer soft handover?

30 How does UTRA-FDD counteract the near-far effect? Why is this not a problem in GSM?

4.7 References

3G Americas (2002) http://www.3gamericas.org/.

3GPP (2000) *General UMTS architecture*, 3rd Generation Partnership Project, 3G TS 23.101 3.1.0 (2000-12).

3GPP (2002a) 3rd Generation Partnership Project, http://www.3gpp.org/.

3GPP (2002b) *UTRAN overall description*, 3rd Generation Partnership Project, 3GPP TS 25.401 V3.10.0 (2002-06).

3GPP2 (2002) 3rd Generation Partnership Project 2, http://www.3gpp2.org/.

Adachi, F., Sawahashi, M., Suda, H. (1998) 'Wideband DS-CDMA for next-generation mobile communications systems,' *IEEE Communications Magazine*, 36(9).

Brasche, G., Walke, B. (1997) 'Concepts, services, and protocols of the new GSM phase 2+ General Packet Radio Service,' *IEEE Communications Magazine*, 35(8).

Callendar, M. (1997) 'International Mobile Telecommunications-2000 standards efforts of the ITU,' collection of articles in *IEEE Personal Communications*, 4(4).

Chen, H.-H., Fan, C.-X., Lu, W. (2002) 'China's Perspectives on 3G Mobile Communications and Beyond: TD-SCDMA Technology,' *IEEE Wireless Communications*, 9(2).

Dahlman, E., Gudmundson, B., Nilsson, M., Sköld, J. (1998) 'UMTS/IMT-2000 based on wideband CDMA,' *IEEE Communications Magazine*, 36(9).

Dasilva, J., Ikonomou, D., Erben, H. (1997) 'European R&D programs on third-generation mobile communication systems,' *IEEE Personal Communications*, 4(1).

DECT (2002), http://www.dect.ch/, http://www.dectweb.com/.
ETSI (1991a) *Bearer services supported by a GSM PLMN*, European Telecommunications Standards Institute, GSM recommendations 02.02.
ETSI (1991b) *General description of a GSM PLMN*, European Telecommunications Standards Institute, GSM recommendations 01.02.
ETSI (1991c) *Subscriber Identity Modules, Functional Characteristics*, European Telecommunications Standards Institute, GSM recommendations 02.17.
ETSI (1993a) *Multiplexing and multiple access on the radio path*, European Telecommunications Standards Institute, GSM recommendations 05.02.
ETSI (1993b) *MS-BSS data link layer – general aspects*, European Telecommunications Standards Institute, GSM recommendations 04.05.
ETSI (1993c) *MS-BSS data link layer specification*, European Telecommunications Standards Institute, GSM recommendations 04.06.
ETSI (1997a) *High Speed Circuit Switched Data (HSCSD)*, Stage 1, European Telecommunications Standards Institute, GSM 02.34, V5.2.1.
ETSI (1998a) *General Packet Radio Service (GPRS)*; *Requirements specification of GPRS*, European Telecommunications Standards Institute, TR 101 186, V6.0.0 (1998–04).
ETSI (1998b) *General Packet Radio Service (GPRS); Service description; Stage 2*, European Telecommunications Standards Institute, EN 301 344, V6.1.1 (1998–08).
ETSI (1998c) *General Packet Radio Service (GPRS); Service description; Stage 1*, European Telecommunications Standards Institute, EN 301 113, V6.1.1 (1998–11).
ETSI (1998d) *General Packet Radio Service (GPRS); Overall description of the GPRS radio interface; Stage 2*, European Telecommunications Standards Institute, TS 101 350, V6.0.1 (1998–08).
ETSI (1998e) *High Speed Circuit Switched Data (HSCSD); Stage 2*, European Telecommunications Standards Institute, TS 101 038, V5.1.0 (1998–07).
ETSI (1998f) *Universal Mobile Telecommunications System (UMTS); Concept groups for the definition of the UMTS Terrestrial Radio Access (UTRA)*, European Telecommunications Standards Institute, TR 101 397, V3.0.1 (1998–10).
ETSI (1998h) *High level requirements relevant for the definition of the UMTS Terrestrial Radio Access (UTRA) concept*, European Telecommunications Standards Institute, TR 101 398, V3.0.1 (1998–10).
ETSI (1998i) *Concept groups for the definition of the UMTS Terrestrial Radio Access (UTRA)*, European Telecommunications Standards Institute, TR 101 397, V3.0.1 (1998–10).
ETSI (1998j) *Digital Enhanced Cordless Telecommunications (DECT), Generic Access Profile (GAP)*, European Telecommunications Standards Institute, EN 300 444, V1.3.2 (1998–03).

ETSI (1998k) *Digital Enhanced Cordless Telecommunications (DECT), Common Interface (CI)*, European Telecommunications Standards Institute, EN 300 175, V1.4.1 (1998–02).

ETSI (1998l) *Terrestrial Trunked Radio (TETRA), Voice plus Data (V+D)*, European Telecommunications Standards Institute, ETS 300 392 series of standards.

ETSI (1998m) *Terrestrial Trunked Radio (TETRA), Packet Data Optimized (PDO)*, European Telecommunications Standards Institute, ETS 300 393 series of standards.

ETSI (1998n) *The ETSI UMTS Terrestrial Radio Access (UTRA) ITU-R Radio Transmission Technologies (RTT) Candidate Submission*, European Telecommunications Standards Institute.

ETSI (2002) European Telecommunications Standards Institute, http://www.etsi.org/.

Evci, C. (2001) 'Optimizing and licensing the radio frequency spectrum for terrestrial 3G users,' *Alcatel Telecommunications Review*, 1/2001.

Goodman, D. (1997) *Wireless Personal Communications Systems*. Addison-Wesley Longman.

GSM Association (2002), http://www.gsmworld.com/.

GSMMoU (1998) *Vision for the evolution from GSM to UMTS*, GSM MoU Association, Permanent Reference Document, V 3.0.0.

GSM-R (2002), The GSM-R Industry Group, http://www.gsm-rail.com/.

Halsall, F. (1996) *Data communications, computer networks and open systems*. Addison-Wesley Longman.

ITU (2002) *International Mobile Telecommunications*, International Telecommunication Union, http://www.itu.int/imt/.

Nussli, C; Bertout, A. (2002) 'Satellite-based multicast architecture for multimedia services in 3G mobile networks,' *Alcatel Telecommunications Review*, 2/2002.

Ojanperä, T.; Prasad, R. (1998) 'An overview of third-generation wireless personal communications: A European perspective,' *IEEE Personal Communications*, 5(6).

Pahlavan, K., Krishnamurthy, P., Hatami, A., Ylianttila, M., Makela, J.-P., Pichna, R., Vallström, J. (2000) 'Handoff in Hybrid Mobile Data Networks,' *IEEE Personal Communications*, 7(2).

Pahlavan, K., Krishnamurthy, P. (2002) *Principles of Wireless Networks*. Prentice Hall.

Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., Schooler, E. (2002) SIP: *Session Initiation Protocol*, RFC 3261, updated by RFC 3265.

SIP Forum (2002) http://www.sipforum.com/.

Shafi, M., Sasaki, A., Jeong, D. (1998) 'IMT-2000 developments in the Asia Pacific region,' collection of articles, *IEEE Communications Magazine*, 36(9).

Stallings, W. (2002) *Wireless Communications and Networks*. Prentice Hall.

TETRA MoU (2002) TETRA Memorandum of Understanding, http://www.tetramou.com/.

Tripathi, N.D., Reed, J.H., VanLandingham, H.F. (1998) 'Handoffs in cellular systems,' *IEEE Personal Communications*, 5(6).
UMTS Forum (2002) http://www.umts-forum.org/.
Wong, D., Lim, T. (1997) 'Soft handoffs in CDMA mobile systems,' *IEEE Personal Communications*, 4(6).