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

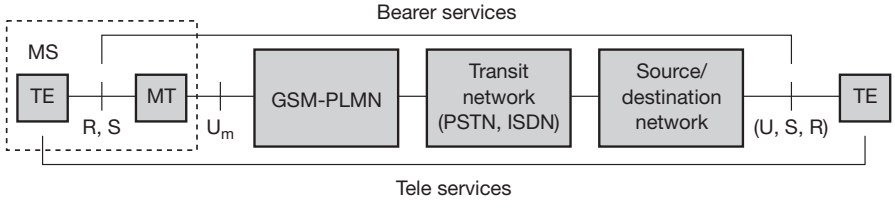

Figure 4.3 Bearer and tele services reference model

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).

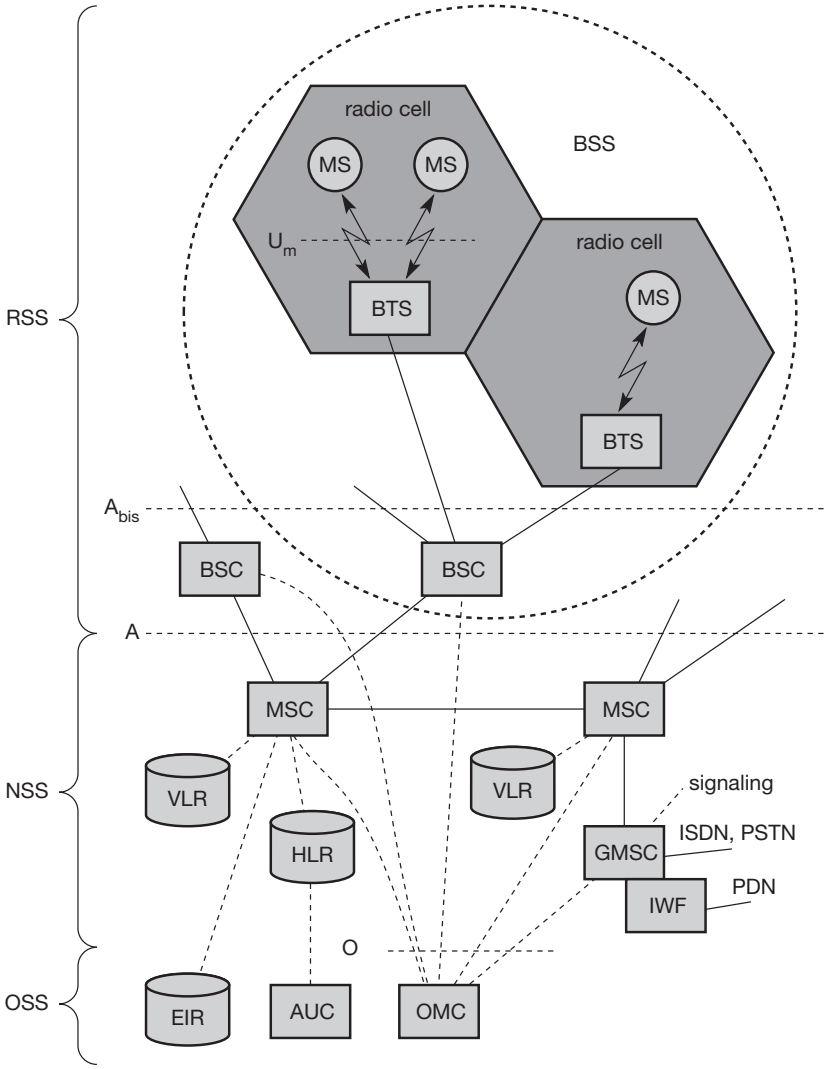

Figure 4.4
Functional architecture of a GSM system

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

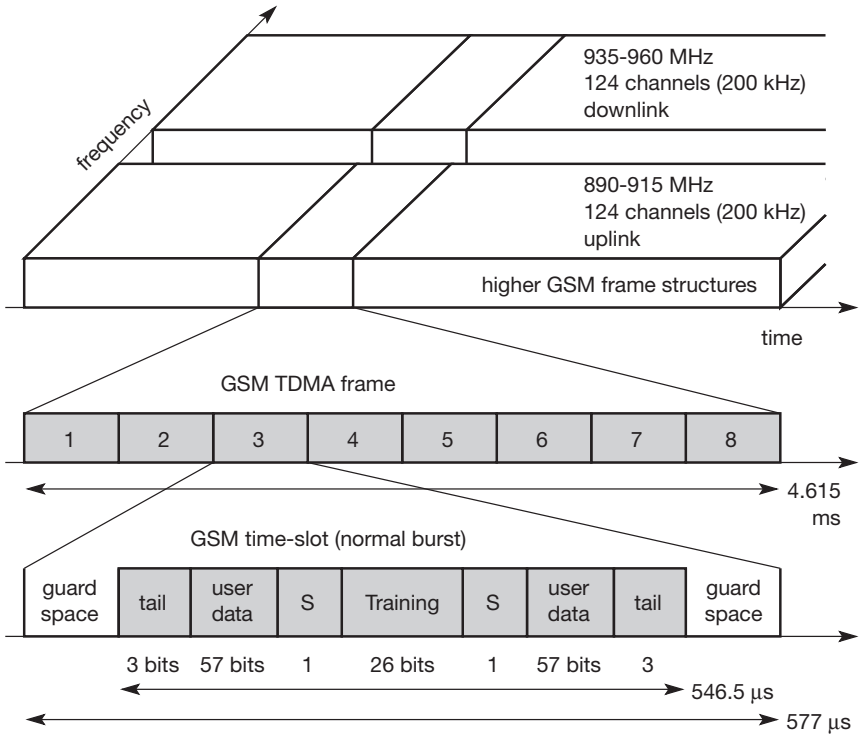

Figure 4.5 GSM TDMA frame, slots, and bursts

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.

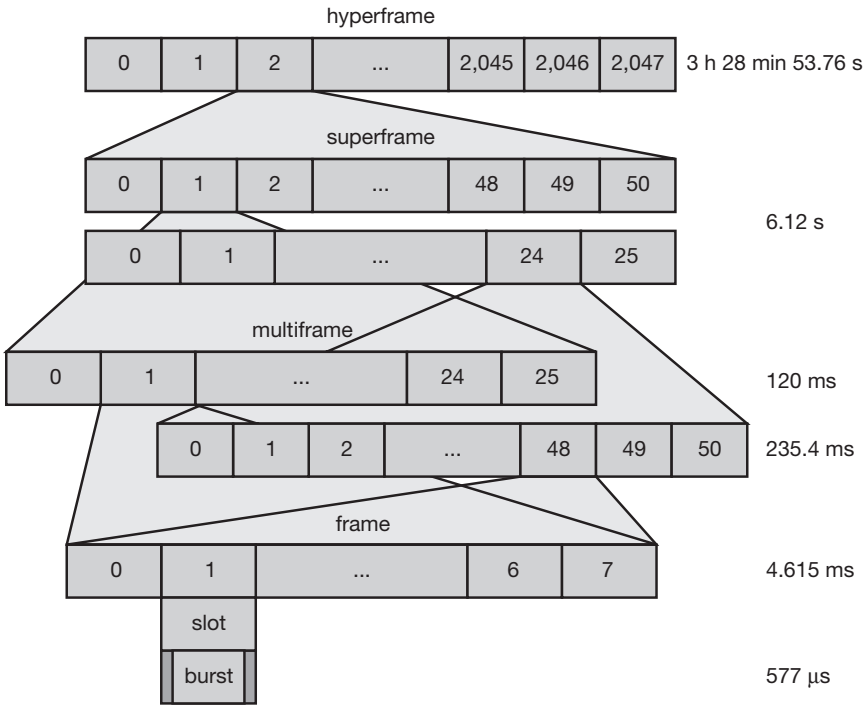

Figure 4.6
GSM structuring of time using a frame hierarchy

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-

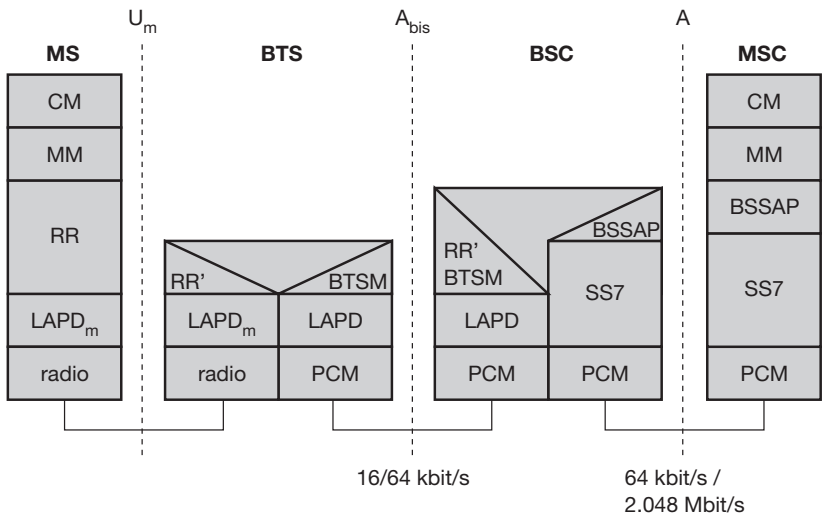

Figure 4.7 Protocol architecture for signaling

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).

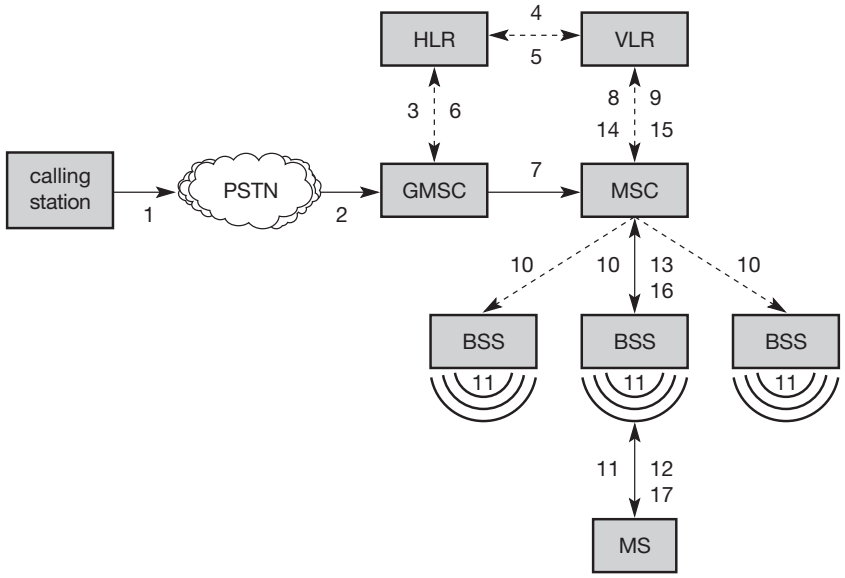

Figure 4.8
Mobile terminated call (MTC)

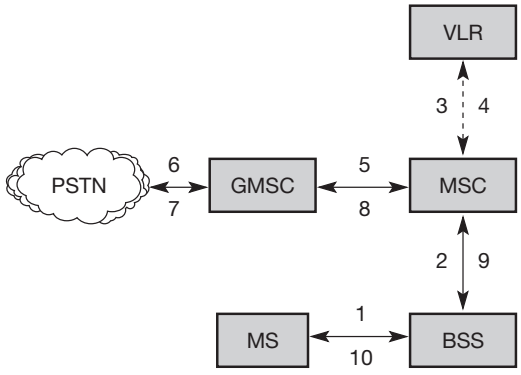

Figure 4.9
Mobile originated call (MOC)

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.

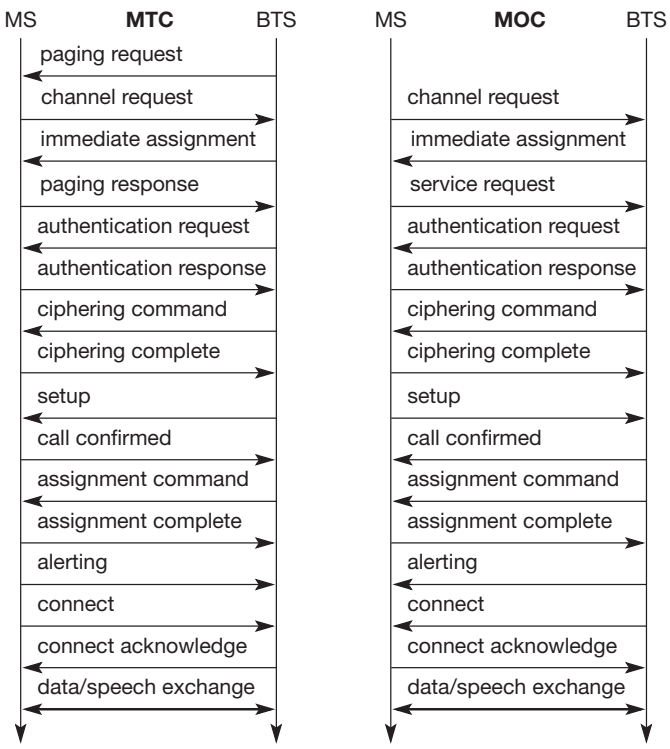

Figure 4.10 Message flow for MTC and MOC

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).

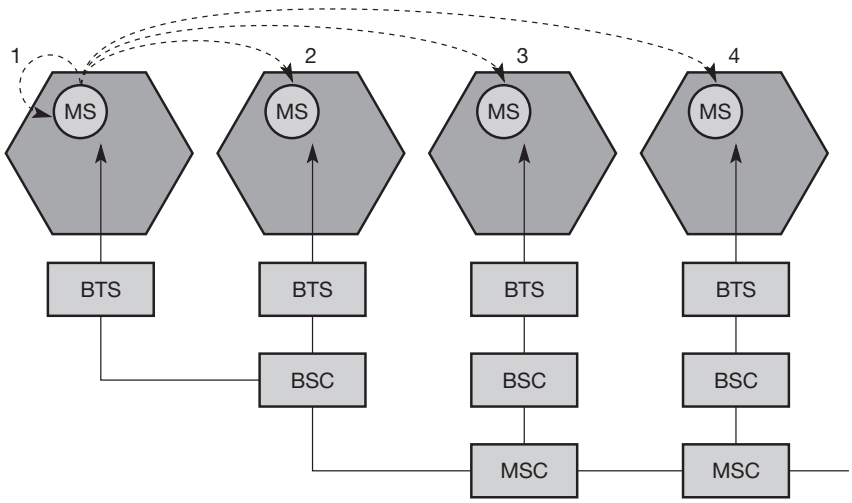

Figure 4.11 Types of handover in GSM

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.

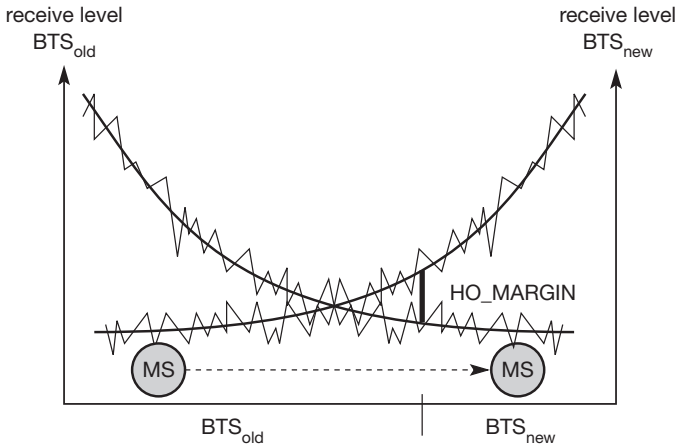

Figure 4.12 Handover decision depending on receive level

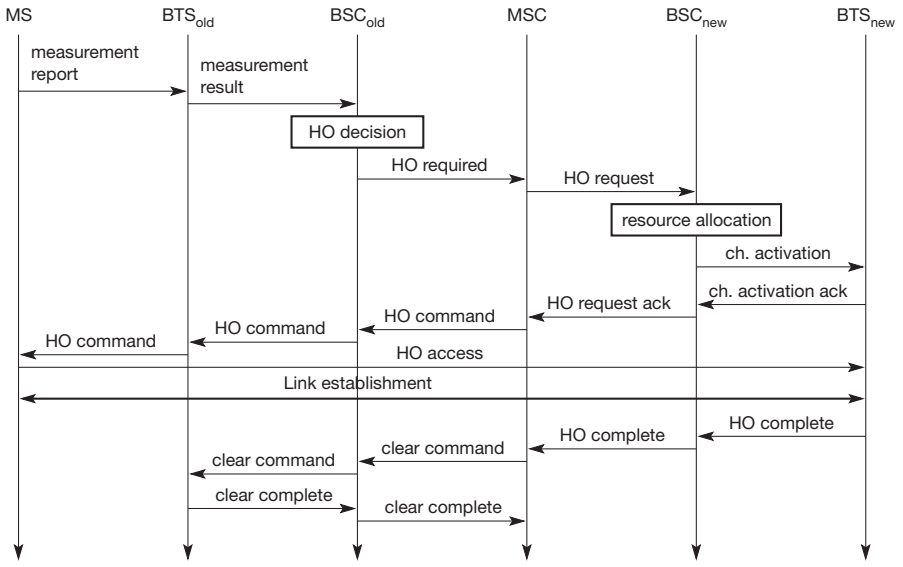

Figure 4.13 Intra-MSC handover

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

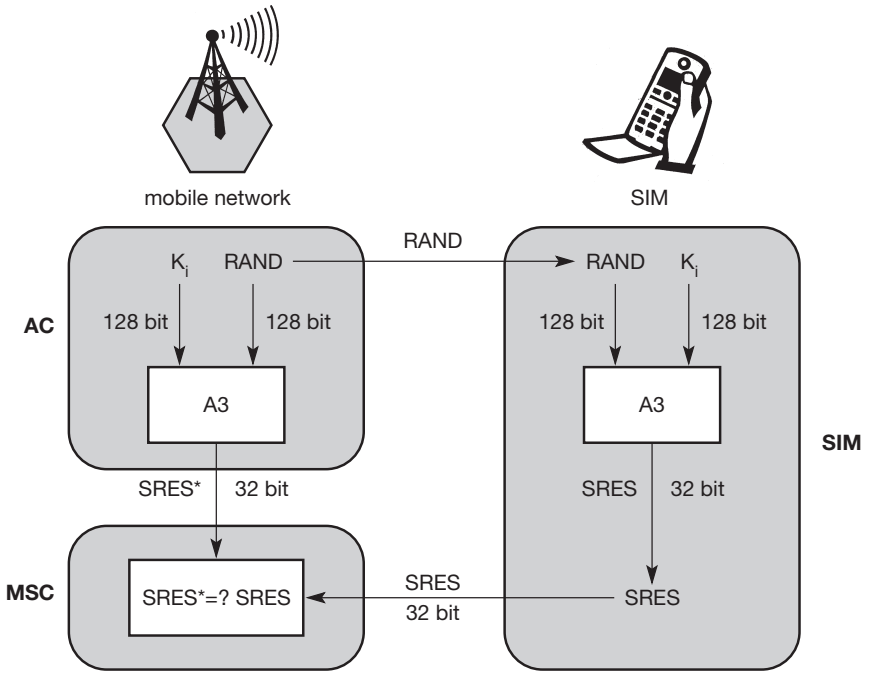

Figure 4.14 Subscriber authentication

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.

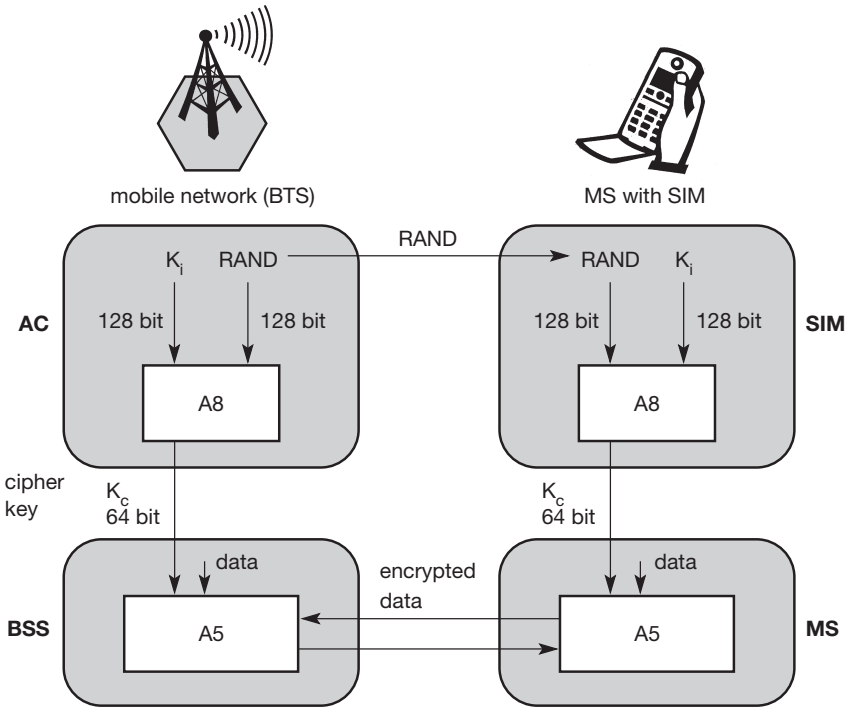

Figure 4.15
Data encryption

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

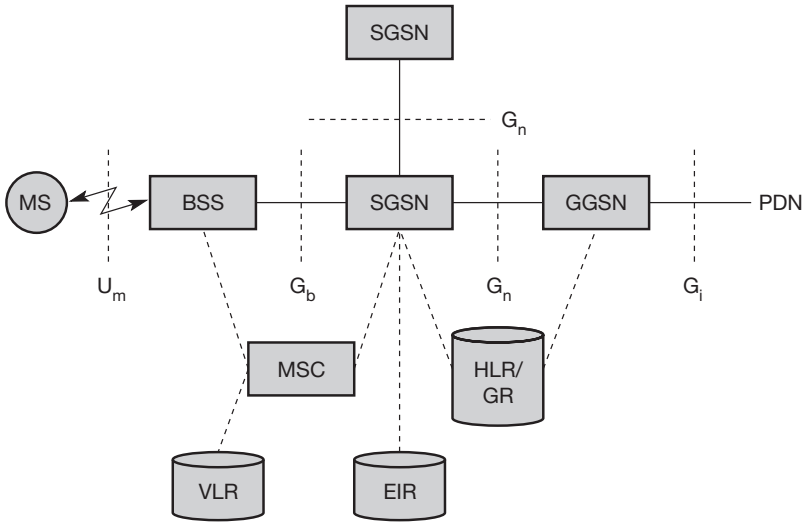

Figure 4.16
GPRS architecture reference model

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.

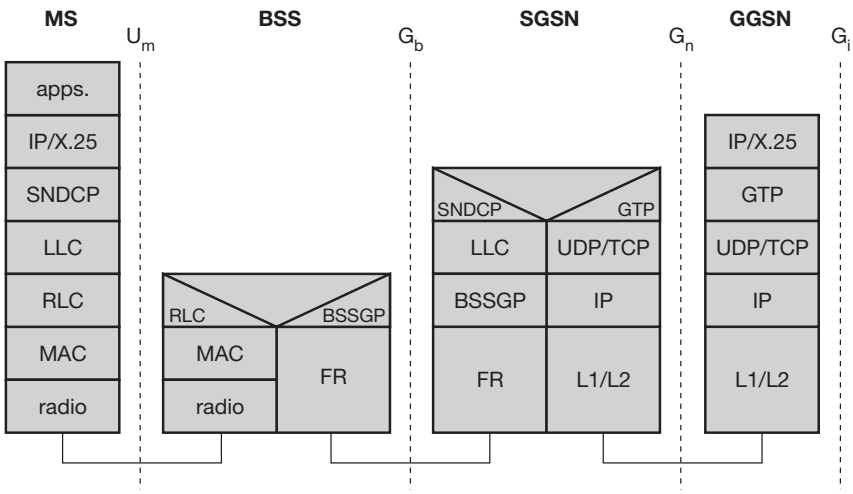

Figure 4.17 GPRS transmission plane protocol reference model

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,