3

CELLULAR NETWORK STANDARDS (GSM AND IS-95)

The cellular Mobile Systems of the early era were very much analog systems. These were mostly installed in vehicles and not at all prepared to be used as hand-held devices. The analog systems of that era had poor voice quality and suffered from interference of signals. Further, in analog system security of communication was found difficult to implement. These factors limited the massive use of mobile systems. The low-cost, hand-held mobile station with longer battery life, better quality of service, enhanced security was only possible after the introduction of digital communication systems. This chapter attempts to make the readers familiarized with the commercially available digital cellular communication standards such as GSM, IS-95.

3.1 DIGITAL CELLULAR COMMUNICATION

Digital radio was first introduced in the defence sector to ensure quality reception with a high level of security in an interference-prone zone. The demand of hand-held terminal with reduced size and power requirement, longer battery life, etc. has led to explore digital cellular system commercially. An improvement in VLSI technology coupled with the advancement in the Digital Signal Processing (DSP) has made it a reality. The digital cellular system outperforms an analog system in terms of:

- capacity
- quality of service
- security
- ability to support improved services such as wireless Internet
- battery life

Digital scrambled up the signals into bursts; so it is more secure than analog and, therefore, helps prevention of stealing phone account information. However, roaming is comparatively difficult with a digital phone than an analog. As there is no uniquely accepted industry standard in digital technology, roaming, i.e. using an operator's network other than the home network can be difficult. On the other hand, an analog system has better coverage and the initial cost for analog is usually less than digital. Therefore, to get the best voice quality as well as security, the terminal offering dual mode (digital/analog) feature can be a better choice. This allows automatic switching between the analog and digital modes depending on the controlling antennae of the region the mobile station (MS) currently belongs to.

3.2 MULTIPLE ACCESS TECHNIQUES

The frequency spectrum allocated to an application (e.g. cellular communication) is shared by the users in a digital system. The sharing mechanism is realized with any one of the multiple access techniques mentioned in Section 1.2.3.

In an effort to make the most efficient use of radio frequency spectrum, a finite natural resource, various technologies have been developed. The target is to simultaneously support as many users as possible by a finite range of spectrum. In a cellular network, transmission from the base station (BS) in forward (downlink) channel can be received by all the MSs under its control. This, in a word, is broadcasting. On the other hand, in the reverse (uplink) channel multiple users want to transmit information simultaneously to the serving base station (BS). Without proper coordination among the transmitting users, collisions occur. Therefore, when two or more users transmit simultaneously to the same serving base station, it needs to be conflict free. This necessitates the scheme for multiple accesses.

The key target in multiple accesses is to separate the transmitted signals from different users at the base station. Different techniques of user separation at receiver (BS) have given birth to different multiple access techniques, namely Frequency Division Multiple Access (FDMA), Time Division Multiple Access (TDMA) and Code Division Multiple Access (CDMA).

In FDMA, a user separation is implemented by providing separate frequency band to individual user whereas in TDMA, it is done by providing separate time slot to each user. On the other hand, CDMA uses separate code to differentiate transmitted signal from an individual user. The 1st generation wireless systems use FDMA while the 2nd generation systems use TDMA and CDMA.

3.2.1 FDMA

An FDMA system separates out the total available bandwidth into several non-overlapping smaller bands/channels. Each channel has the ability to support one user. Each user is assigned a unique frequency band or channel.

Figure 3.1 illustrates that each user is allowed a unique frequency band. These channels are assigned on demand. When an MS tries to communicate with a BS, it registers itself using control channel to the closest BS. During registration, the BS assigns the MS an available pair of channels, one to transmit (reverse channel) and the other to receive (forward channel).

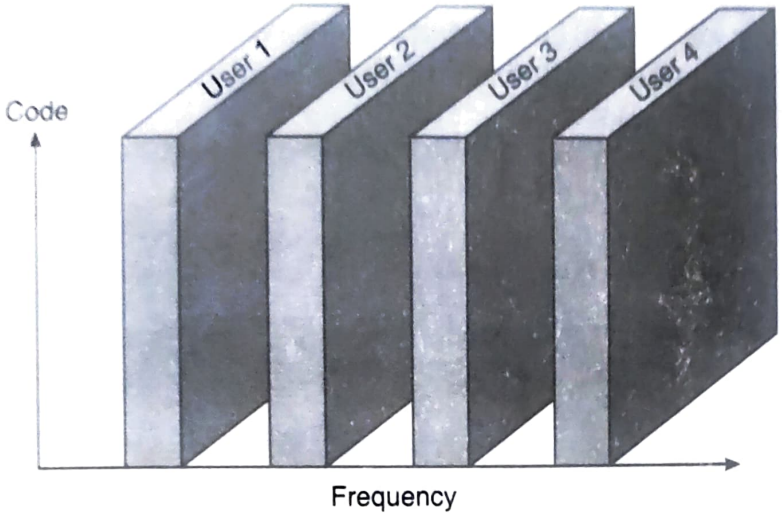

Figure 3.1 Frequency division multiple access.

To ensure interference-free transmission between two users or between the forward and reverse channel transmission of a user, it introduces the guard bands. The guard band separates out two such channels preventing interference. The guard bands are unused portions of the spectrum. For example, if 500 MHz bandwidth is provided for a system, entire bandwidth can be split into 12 channels each of 40 MHz. Each 40 MHz channel includes a 4 MHz guard channel. Effectively each channel is 36 MHz wide. In case of forward and reverse channels, either two antennas operating at different frequencies or one antenna with Frequency Division Duplexing (FDD) can provide the guard band.

As the multiple access is achieved by separating the user by the frequency allocated to them, frequency planning has to be done very carefully to avoid adjacent channel interference. This adjacent channel interference is an important factor in determining channel quality. However, frequency planning is complicated and difficult to achieve. Available frequency bands must be researched and analysed.

3.2.2 TDMA

In a system with a moderately large number of active users, the allocation of unique frequency band for each user cannot be a realistic option. As the demand increases to accommodate more users, spectrum efficiency[1] of

[1] Spectrum efficiency is measured as the amount of information that can be transmitted over a given bandwidth in a communication system. In the context of cellular mobile communication, spectrum efficiency means the maximum number of subscribers per cell that can be accommodated with an acceptable quality of service.

FDMA system becomes insufficient. A TDMA system splits users into an available pair of channels, but they also assign each user an available time slot within that channel. In each slot, only one user is allowed to either transmit or receive. Frequency division is still employed but these frequencies are now further subdivided into a defined number of time slots per frequency. Figure 3.2 shows that the entire spectrum is divided into four channels each of which is divided into three time slots. As the TDMA systems do not transmit all the time, the MSs gain an extended battery life as well as the talktime.

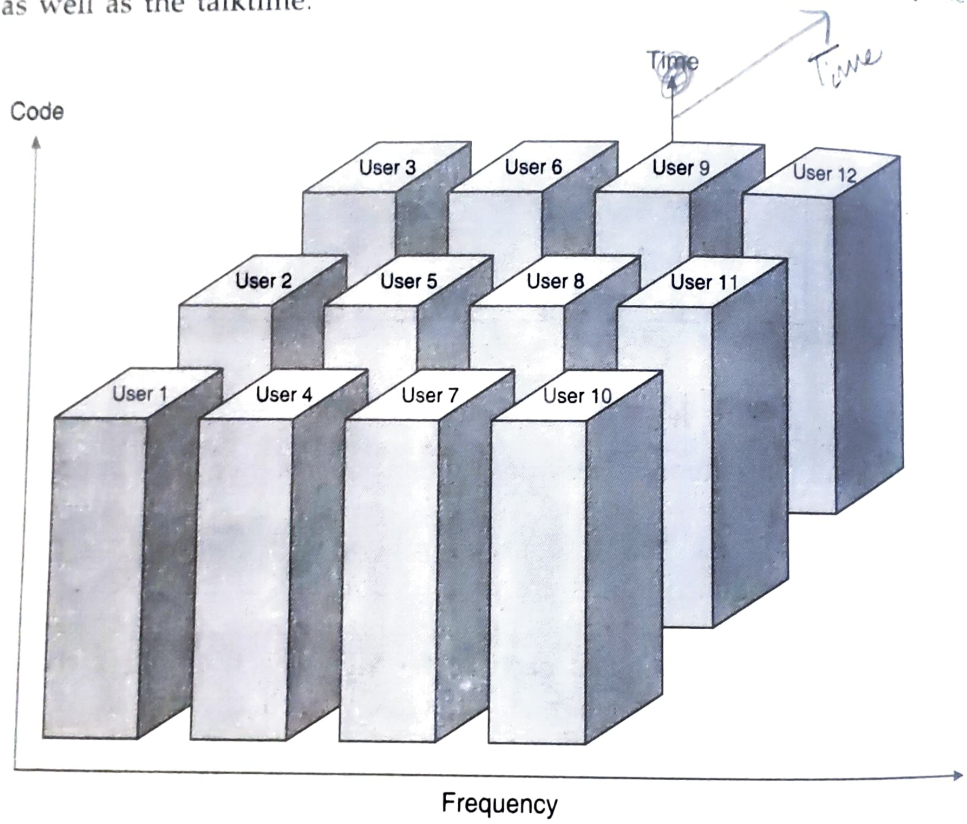

Figure 3.2 Time division multiple access.

In TDMA, like FDMA, the BS assigns the MS an available pair of channels. In addition to this, it assigns a time slot within the channel that must be available to the MS. The user can only be allowed to send or receive data for this time slot. The data bursts are fast reassembled at the receiving end, and therefore, appeared as continuous. Figures 3.1 and 3.2, point to the fact that while an FDMA system supports 4 users, the TDMA can support 12 users utilising the same bandwidth. As the TDMA systems first split an allotted portion of the frequency spectrum into smaller slots/channel, they require the same level of frequency planning as FDMA system.

3.2.3 CDMA

In code division multiple access (CDMA), signals from users (MSs) are differentiated at the receiver (BS) in code space whereas in FDMA it is

rentiated in frequency space and in TDMA, in time space. The main enge of CDMA is to find suitable code for a user's signal so that it is ue and can be differentiated from the signals of other users. The dle is mitigated in CDMA by finding mutually orthogonal codes having n autocorrelation[2] value for each user. Two codes are orthogonal to n other when their cross correlation[3] value is 0. However, the CDMA s not allocate unique channel for a user. Each user can utilize the ire block of allocated spectrum space to carry the messages, that is, all users can use the same carrier frequency and transmit simultaneously.

There are two popular coding techniques used in CDMA—walsh de[4] and PN (pseudo noise) code[5]. In the forward (BS to MSs) direction, ansmission for all the users originates from the same transmitter (BS) in perfectly coordinated manner. The orthogonal walsh code is used for rward channel transmission. In forward direction, transmission is riginated from a BS, a fixed station. It enables the generation of orthogonal odes for the user signals. On the other hand, the reverse transmission MS to BS) cannot be accurately coordinated due to the mobility of MSs. Generation of orthogonal codes for each user in such case is not at all possible for arbitrarily random locations of the MSs. Hence the PN code is used for the reverse channel transmission.

Spread spectrum

Spread spectrum is a technique that realizes the unique coding for signals from each user. This also improves the spectrum efficiency over that of FDMA and TDMA. The principle of spread spectrum communication is to spread the bandwidth of baseband information-carrying signals from different users to a much larger bandwidth. Ideally, the spreading signals used for different users are orthogonal to each other.

In spread spectrum, the narrow band information-carrying signal is multiplied by a very large bandwidth signal called the spreading signal. The bits of spreading signal are referred to as the chips. If T_c is the period of one chip, $1/T_c$ is the chip rate. The user data rate is $1/T_b$, where T_b is the period of one data bit (baseband information-carrying signal). The spreading factor is defined as the ratio of chip rate and user data rate (T_c/T_b).

One of the popular implementation of spread spectrum is the direct sequence spread spectrum (DSSS). In DSSS-based CDMA, the spreading signals with different chip rates are used to the different data signals. The chip rates used for the spreading signal are significantly greater than the

[2] Correlation is to determine how much similarity a code has with another code. It is computed by taking inner product of the codes. Autocorrelation is the correlation of a user code with itself.

[3] Cross correlation is correlation between two separately generated codes.

[4] Walsh codes are mathematically orthogonal codes.

[5] A PN code is a binary sequence that appears randomly but can be reproduced in a deterministic manner.

differentiated in frequency space and in TDMA, in time space. The main challenge of CDMA is to find suitable code for a user's signal so that it is unique and can be differentiated from the signals of other users. The hurdle is mitigated in CDMA by finding mutually orthogonal codes having high autocorrelation[2] value for each user. Two codes are orthogonal to each other when their cross correlation[3] value is 0. However, the CDMA does not allocate unique channel for a user. Each user can utilize the entire block of allocated spectrum space to carry the messages, that is, all the users can use the same carrier frequency and transmit simultaneously.

There are two popular coding techniques used in CDMA—walsh code[4] and PN (pseudo noise) code[5]. In the forward (BS to MSs) direction, transmission for all the users originates from the same transmitter (BS) in a perfectly coordinated manner. The orthogonal walsh code is used for forward channel transmission. In forward direction, transmission is originated from a BS, a fixed station. It enables the generation of orthogonal codes for the user signals. On the other hand, the reverse transmission (MS to BS) cannot be accurately coordinated due to the mobility of MSs. Generation of orthogonal codes for each user in such case is not at all possible for arbitrarily random locations of the MSs. Hence the PN code is used for the reverse channel transmission.

Spread spectrum

Spread spectrum is a technique that realizes the unique coding for signals from each user. This also improves the spectrum efficiency over that of FDMA and TDMA. The principle of spread spectrum communication is to spread the bandwidth of baseband information-carrying signals from different users to a much larger bandwidth. Ideally, the spreading signals used for different users are orthogonal to each other.

In spread spectrum, the narrow band information-carrying signal is multiplied by a very large bandwidth signal called the spreading signal. The bits of spreading signal are referred to as the chips. If T_c is the period of one chip, $1/T_c$ is the chip rate. The user data rate is $1/T_b$, where T_b is the period of one data bit (baseband information-carrying signal). The spreading factor is defined as the ratio of chip rate and user data rate (T_c/T_b).

One of the popular implementation of spread spectrum is the direct sequence spread spectrum (DSSS). In DSSS-based CDMA, the spreading signals with different chip rates are used to the different data signals. The chip rates used for the spreading signal are significantly greater than the

---

[2] Correlation is to determine how much similarity a code has with another code. It is computed by taking inner product of the codes. Autocorrelation is the correlation of a user code with itself.

[3] Cross correlation is correlation between two separately generated codes.

[4] Walsh codes are mathematically orthogonal codes.

[5] A PN code is a binary sequence that appears randomly but can be reproduced in a deterministic manner.

user data rates. Normally, the chip rate is in the order of "megachips per second" (Mcps), that is, millions of chips per second. For DSSS-based CDMA modulation (Refer to Figure 3.3(a)), the first task is to spread the information-carrying (user data) signal (digital modulation) and the second task is to modulate spread signal with carrier frequency. For example, spreading a user signal of 1 MHz bandwidth with 11 chip code generates a signal of 11-MHz bandwidth. Then the signal is converted to GHz bandwidth to make it ready for transmission. Demodulation (Refer to Figure 3.3(b)), the reverse process of modulation, is employed to retrieve user data at the receiver (BS) side.

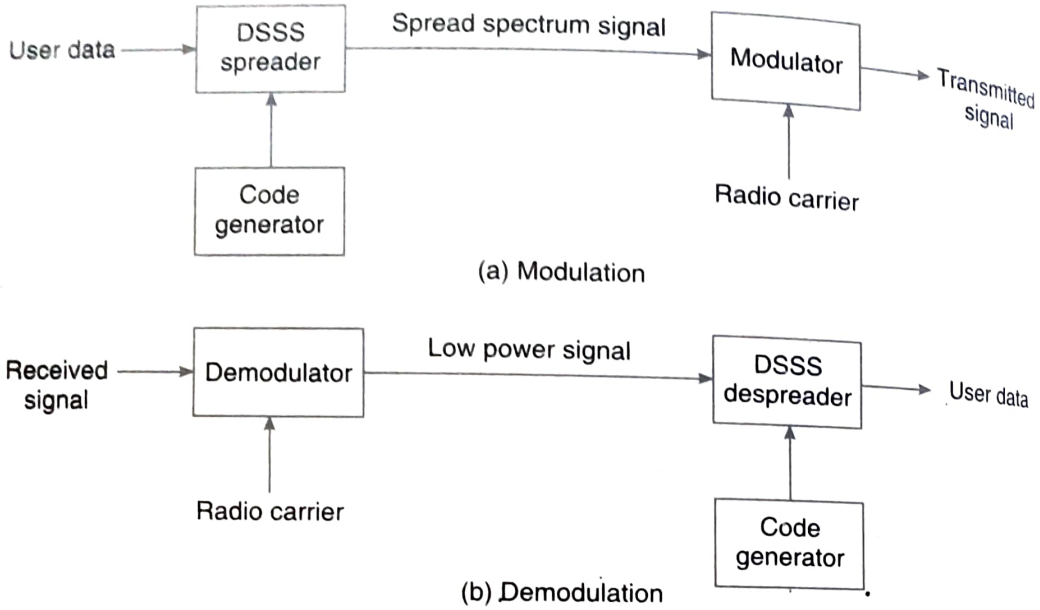

Figure 3.3 DSSS CDMA system.

If sufficient bandwidth is available, spreading is performed in two consecutive phases. In one phase data is spread by an orthogonal code to provide mutual orthogonality among all users in the same cell. In the next phase the resulting signal is further spread by a PN sequence to provide mutual orthogonality among the users in different cells. Such two phase spreading is called multiple spreading. IS-95 adopts this multiple spreading technique.

In CDMA for both the forward and reverse channel transmission, unique codes (walsh code and PN code respectively) are used for a user. A receiver would try to reconstruct only the specific desired user's data. And all other users' data would be sensed as noise. While detecting the user data signal, a receiver (BS) needs to know the code used by the transmitter (MS). An analogy can be drawn with a business party being held in a big hall. As it is shown in Figure 3.4, a number of persons are sharing views among themselves. Each pair of them is using a unique language for conversation. As a pair is using separate language (code), others conversations cannot be affected severely provided no one speaks much loudly compared to others. If one speaks so, it will add to noise.

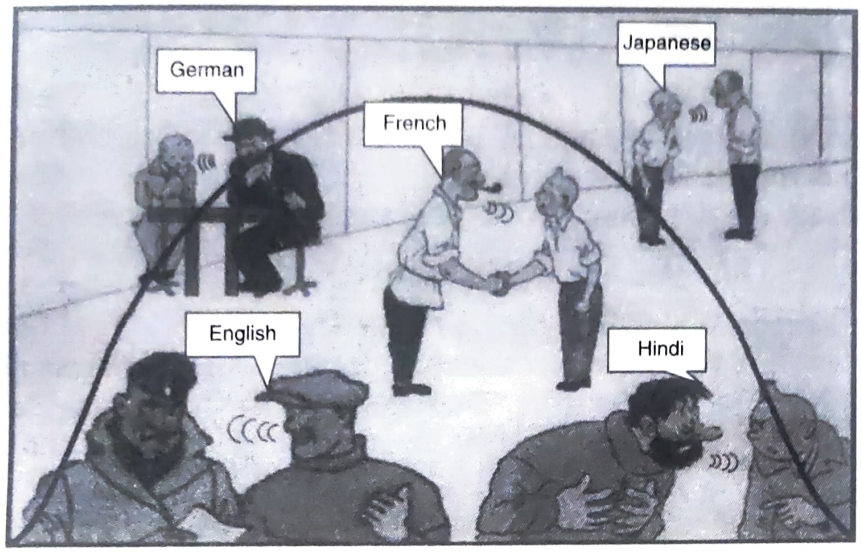

Figure 3.4 An analogy with CDMA.

It is mentioned earlier that the CDMA demands the codes should be orthogonal to each other and a code for a certain user should have a good autocorrelation. For example, let us consider the two simplest codes (1, 0) and (1, 1). Assuming 0 as –1 and 1 as +1, the codes become (+1, –1) and (+1, +1). The inner product of these two is $(+1) \cdot (+1) + (-1) \cdot (+1) = +1 - 1 = 0$ and hence the codes are orthogonal. A code with high autocorrelation means the absolute value of inner product of the code with itself is high. For example, for the code (+1, –1), the autocorrelation value is $(+1, -1) \cdot (+1, -1) = (+1) \cdot (+1) + (-1) \cdot (-1) = 1 + 1 = 2$. An example of a code, used for ISDN and IEEE 802.11 is 11 chip code (+1, –1, +1, +1, –1, +1, +1, +1, –1, –1, –1). The autocorrelation value of this code is 11.

How it works

Consider two MSs, MS_a and MS_b that want to transmit $data_a$ (1, 0) and $data_b$ (1, 1) respectively. CDMA assigns two unique and orthogonal codes—$code_a$ = (0, 1, 0, 1), $code_b$ = (0, 1, 1, 0). Both the MS_a and MS_b spread respective signal using their unique key as chip sequence (spreading signal). The spreading algorithm implies that if the data bit is 1, then send the corresponding code itself and for data bit 0, send the complement of the corresponding code. The resulting signals (encoded $data_a$ and $data_b$) are given below. For convenience binary 0 is assumed as –1, and binary 1 as +1.

Encode MS_a

$code_a = (-1, +1, -1, +1)$
$data_a = (+1, -1)$
$encoded\ data_a = code_a \cdot data_a$
$= (-1, +1, -1, +1) \cdot (+1, -1)$
$= (-1, +1, -1, +1), (+1, -1, +1, -1)$
$signal_a = (-1, +1, -1, +1, +1, -1, +1, -1)$

Encode MS_b

$code_b = (-1, +1, +1, -1)$
$data_b = (+1, +1)$
$encoded\ data_b = code_b \cdot data_b$
$= (-1, +1, +1, -1) \cdot (+1, +1)$
$= (-1, +1, +1, -1), (-1, +1, +1, -1)$
$signal_b = (-1, +1, +1, -1, -1, +1, +1, -1)$

The received signal at the BS is the sum of these two signals; signal$_a$ and signal$_b$, if there is no other MSs under the serving BS. That is, $\text{signal}_{\text{received-at-BS}} = \text{signal}_a + \text{signal}_b = (-2, +2, 0, 0, 0, 0, +2, -2)$. As the code$_a$ and code$_b$ are orthogonal to each other, the data$_a$ and data$_b$ can easily be extracted from the $\text{signal}_{\text{received-at-BS}}$. The decoding performed at BS to extract the data sent by the MS$_a$ and MS$_b$ are described below.

The extraction of data sent by MS$_i$ means:

(i) $\text{decode}_i = \text{signal}_{\text{received-at-BS}} \cdot \text{code}_i$

(ii) replace values greater than 0 of decode$_a$ as 1 and values less than 0 as 0 to get the data$_i$. Therefore

Decode MS$_a$

$\text{code}_a = (-1, +1, -1, +1)$

$\text{signal}_{\text{received-at-BS}} = (-2, +2, 0, 0, 0, 0, +2, -2)$

(i) $\text{decode}_a = \text{signal}_{\text{received-at-BS}} \cdot \text{code}_a$

$= ((-2, +2, 0, 0), (0, 0, +2, -2)) \cdot (-1, +1, -1, +1)$

$= (2 + 2 + 0 + 0), (0 + 0 - 2 - 2)$

$= (+4, -4)$

$\text{decode}_a = (+4, -4)$

(ii) $\text{data}_a = (1, 0) = (+1, -1)$

Decode MS$_b$

$\text{code}_b = (-1, +1, +1, -1)$

$\text{signal}_{\text{received-at-BS}} = (-2, +2, 0, 0, 0, 0, +2, -2)$

(i) $\text{decode}_b = \text{signal}_{\text{received-at-BS}} \cdot \text{code}_b$

$= ((-2, +2, 0, 0), (0, 0, +2, -2)) \cdot (-1, +1, +1, -1)$

$= (2 + 2 + 0 + 0), (0 + 0 + 2 + 2)$

$= (+4, +4)$

$\text{decode}_b = (+4, +4)$

(ii) $\text{data}_b = (1, 1) = (+1, +1)$

A CDMA system provides more privacy than FDMA or TDMA systems. Due to the wide bandwidth of a spread-spectrum signal in CDMA, it is very difficult to cause jamming or interference. It appears as nothing more than a slight rise in the noise floor[6] or interference level. In other technologies, the power of signal is concentrated in a narrower band and, therefore, easier to detect or decode.

Increase in number of users in a CDMA system linearly raises the noise floor. Thus there is no absolute limit on the number of users in CDMA. Rather, the system performance gradually degrades for all users as the number of users is increased. Hence the CDMA has a soft capacity limit[7]. Moreover, there is a problem of self-jamming[8]. Self-jamming arises while spreading sequences of different users are not exactly orthogonal resulting in one user disrupting the transmission to the other, and vice versa. Further the users of CDMA share the same channel. It may lead to the near-far problem (Section 2.5). To combat this problem, power control

[6] Noise floor is the measure of sum of signal generated from all the noise sources.

[7] Soft capacity limit is a limit of accommodating number of users beyond which system performance degrades to an unacceptable extent.

[8] Transmission of radio signals that disrupts communications by decreasing the signal-to-noise ratio.

technique is employed in most CDMA implementations so that each MS under a BS coverage can provide the same signal level to the BS receiver.

The well-designed filters employed in FDMA ensure zero or a minimal spectral overlap whereas employment of slot synchronization in TDMA reduces timing jitter. This makes both FDMA and TDMA conflict-free multiple access schemes. On the other hand, CDMA is a spread spectrum technique. Although orthogonality between transmitted signals from different users is the key element of CDMA, in reality the generated wideband spreading functions are not truly orthogonal and cause some interference. Hence CDMA is interference limited multiple access strategy.

The following discussion summarizes the merits and demerits of the three multiple access techniques—FDMA, TDMA and CDMA.

FDMA

Merits

- simple to implement
- fairly efficient with a small population and when traffic is almost constant
- fewer bits are needed as overhead (synchronization, etc.) as compared to TDMA

Demerits

- frequency planning is difficult
- if a channel is not used by the designated user, it cannot be used by others to increase or share capacity—a wastage of resources

TDMA

Merits

- data transmission for users is not continuous leading to low battery consumption
- can accommodate much more users in the same spectrum space as compared to FDMA; capacity increases in high traffic areas
- can allocate different number of time slots per frame to different users. Thus bandwidth can be supplied on demand to different users based on priority.

Demerits

- as high synchronization overhead is required, receivers are to be synchronized for each data burst and, therefore, requires additional overheads in comparison to FDMA
- frequency guard bands lead to spectrum inefficiency
- security is much lower than that in CDMA
- frequency planning is critical
- the number of users accommodated is less than CDMA in the same spectrum space

CDMA

Merits

- best spectrum efficiency: capacity increases 8 to 10 times than that of an analog system and 4 to 5 times than the other digital systems, making it most useful in high traffic areas with a large number of users and limited spectrum.
- simplified frequency planning as all users utilize the same radio frequency spectrum
- random walsh codes enhance user privacy/security

Demerits

- base station equipment is expensive
- difficult to implement in comparison to FDMA and TDMA
- continuous power adjustment of MS is essential to avoid near-far problem
- self-jamming arises from the fact that the spreading sequences of different users are not exactly orthogonal

Legacy commercial telecommunication networks such as analog networks based on Advanced Mobile Phone System (AMPS) are built around the FDMA. Currently, there are several digital cellular standards available worldwide. The example of TDMA-based standards are Global System for Mobile (GSM), Digital Cellular Systems (DCS), etc. On the other hand, IS-95 is the CDMA-based standard. The details of GSM (TDMA) and IS-95 (CDMA) standards are provided in the following sections.

3.3 GSM

Global system for Mobile (GSM) is a 2nd generation cellular system. It was developed in the 1990s to solve the fragmentation problems of first cellular systems and introduced in Europe. By 1993, GSM and its technically equivalent offshoot, DCS (Digital Cellular System) 1800, were adopted outside Europe, too.

GSM is the world's first cellular system to specify digital modulation and network level architectures and services. It has two objectives—pan-European roaming and interaction with the ISDN. Roaming yields compatibility around Europe and interaction with the ISDN offers capability to extend the single-subscriber-line system to a multi-service system. System capacity was not a primary issue at the initial development phase, however, to keep pace with the rapid growth in cellular services, many revisions have been made to GSM.

3.3.1 System Architecture

The GSM consists of three major interconnected components—Base Station Subsystem (BSS), Network and Switching Subsystem (NSS) and Operation

Support Subsystem (OSS). A mobile station (MS) interacts with BSS through air interface. The important interfaces are Um, A_{bis} and A. Figure 3.5 illustrates the block diagram of GSM system.

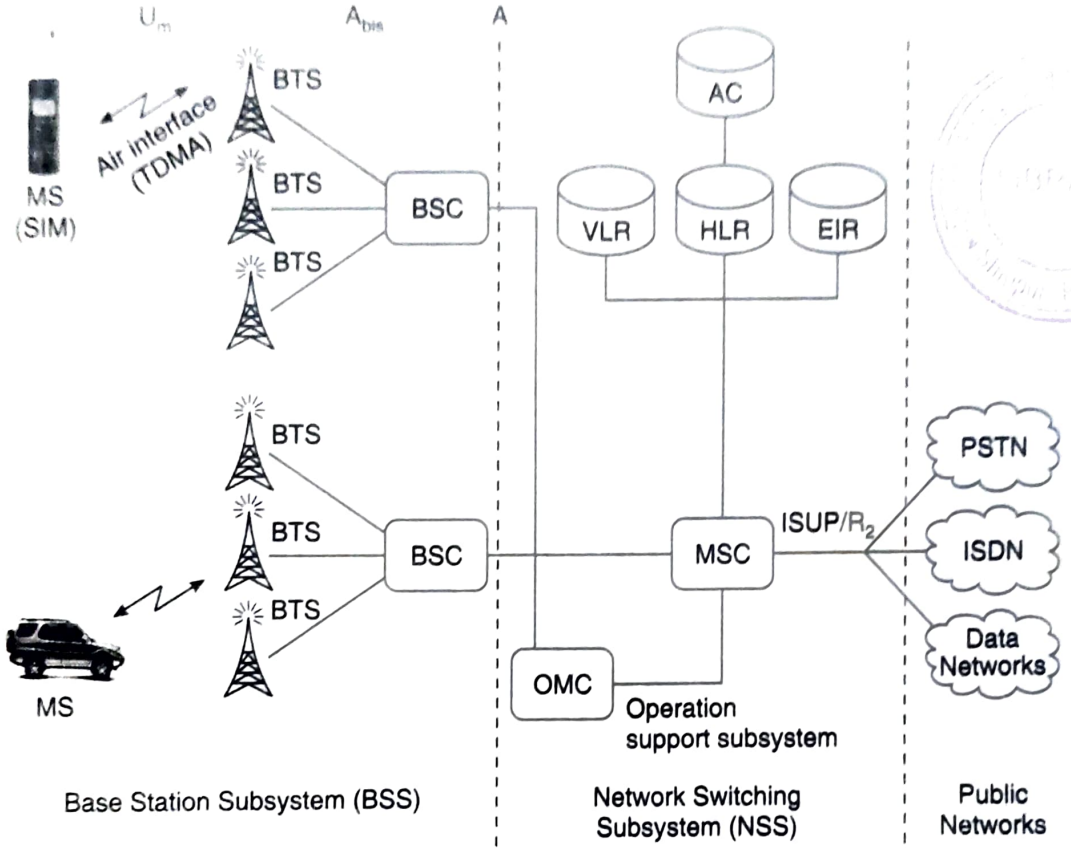

Figure 3.5 GSM architecture.

Mobile station: A mobile station comprises of device-dependent hardware/software and the subscriber-dependent SIM (subscriber identity module). The device is identified by IMEI (international mobile equipment identity) whereas a subscriber is identified by SIM. The device dependent hardware are transceiver, antenna, etc. and the softwares are for theft protection, etc. The SIM is a small detachable card and securely stores user specific information such as IMSI (international mobile subscriber identity), PIN (personal identity number), PUK (PIN unlocking key) and an authentication key to run user validation algorithm. The IMSI is similar to MIN whereas the IMEI is equivalent to ESN in AMPS. At any instant of time, an MS is either in idle state or in dedicated state. In the idle state, the MS only listens to the network but neither transmits nor receives any data. In this state it does not use any dedicated resources. On the other hand, in dedicated mode the MS transmits/receives data using dedicated resources. From an idle state, if an MS desires to access the network, it enters into the dedicated state and starts using resources.

BSS modules: A BSS consists of Base Station Controllers (BSCs), each of which controls a number of Base Transceiver Stations (BTSs). The BSCs

are connected to an MSC. An MS communicates BSS via the BTS through TDMA air interface (Figure 3.5).

A BSS provides radio transmission paths between an MS and the MSCs. It also manages the radio interface between the MSs and all other subsystems of GSM. Handoffs (Handovers in GSM terminology) between two BTSs, under the control of a BSC, are handled by the BSC. As a result, switching burden on MSC reduces drastically.

NSS modules: The MSC is the central component of NSS. The other components of NSS are Home Location Register (HLR), Visitor Location Register (VLR), Equipment Identity Register (EIR) and Authentication Centre (AC). There is logically one HLR per GSM network. It is also implemented as a distributed database.

The NSS manages switching function of the system. It allows MSCs to communicate with other networks such as PSTN, ISDN, etc. In addition, an NSS is responsible for all the functionalities needed to handle a mobile subscriber, such as registration, authentication, location updating, handovers and call routing to a roaming user. An MSC can connect data networks, e.g. X.25 with the help of IWF (inter-working functions), an additional module attached to it. However, the IWF is not shown in Figure 3.5. On the other hand, an MSC handles all signalling needed for connection set-up, release and handover to other MSCs using SS7[9] (standard signalling system No. 7) used in ISDN and current public networks.

The HLR is a huge database capable of managing to the tune of few million subscribers' data. It stores IMSI and corresponding list of subscribed services such as call forwarding, call waiting, roaming limitation, etc. In addition to this, the HLR maintains location information for each subscriber registered in the GSM network. If the system has to establish a call to an MS, the MSC seeks routing information from the HLR.

The VLR temporarily stores the IMSI and information for each subscriber, visiting the coverage area of an MSC. The VLR temporarily assigns TMSI (temporary mobile subscriber identity) to each roaming subscriber for concealing IMSI, a user identity. The VLR may change TMSI dynamically. The TMSI remains valid within the coverage of the VLR. A VLR is capable of managing up to one million subscribers. Although each functional entity can be implemented as an independent unit, most manufacturers of the switching equipment implement one VLR together with one MSC, so that the geographical area controlled by the MSC corresponds to that controlled by the VLR, simplifying the signalling required. When an MS becomes roaming that is, comes to a new MSC it registers itself to the VLR of the MSC. Once it is registered is in the VLR, the MSC sends the necessary information to the visiting subscriber's HLR so that calls to the roaming mobile can be appropriately routed over the PSTN by the roaming user's HLR. The database in VLR is

[9] A set of telephony signalling protocols which are used to set up, maintain and release PSTN calls.

temporary, that is, the subscriber data is stored as long as the subscriber is within the service area. The HLR and VLR, together with the MSC, tackle the call routing and roaming facilities of GSM. Both the database (HLR and VLR) contains a large volume of data and, therefore, maintains suitable database organizations to retrieve the desired subscriber's information in real time.

Both the AC and EIR database are used for providing security. When EIR checks the validity of the equipment by checking IMEI, the AC provides information for verifying the subscribers SIM cards. The EIR database contains a list of all invalid MSs of the network. That is, the list contains the IMEI of MSs that are either to be banned or to be monitored. EIR marks an IMEI as invalid if the MS is stolen/blacklisted or its type is not approved. When such an MS seeks registration to a GSM network, the MS is requested to provide the IMEI for equipment verification. If the MS is found in the list, access to the network is denied to that MS. In some implementations, EIR is integrated with the HLR.

The AC is a database that keeps a copy of the secret-key stored in each subscriber's SIM card. When an MS keeps its power on, it attempts to connect to the GSM network. The SIM of the MS is authenticated at this stage. The AC combines the secret-key with IMSI to produce a challenge/response type of identification and supplies it to the MSC. If the SIM can generate the same identification with the secret-key and IMSI stored in it, the SIM can gain access to the network thereby authentication of SIM is done. This method of SIM authentication will be elaborated further during discussion on user validation in Section 3.3.8. In addition to SIM authentication, the AC participates in securing radio communication between MS and the network. Both the AC and the SIM produce a cipher key separately. With the available cipher key at both the MS and the network side, radio transmission from the MS is encrypted whereas at the network side it is decrypted and vice versa.

OSS modules: It supports one or more Operation Maintenance Centres (OMCs) and is solely accessed by the staff of GSM operating company. An OSS maintains all the telecommunication hardware and network operations within a particular coverage area. It is the manager of all sorts of charging and billing procedures and allows system engineers to monitor, diagnose and troubleshoot all the aspects of the GSM.

3.3.2 OSI Layers in GSM

The GSM is an OSI (Open System Interconnection) network as shown in Figure 3.6. It has three layers, namely physical layer (Layer 1), data link layer (Layer 2) and network layer (Layer 3). As the air interface between MS and BTS is provided by U_m, the protocol layers connected by U_m are given special attention. The other interfaces (A_{bis}, A, ISUP/R2) lie between the modules in a fixed network.

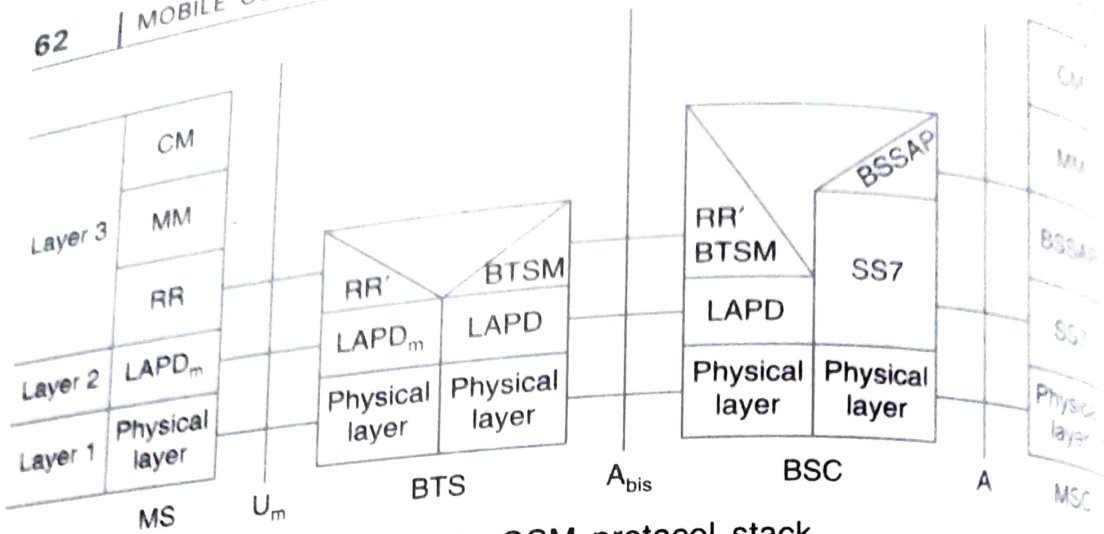

Figure 3.6 GSM protocol stack.

Physical layer: This layer is responsible for creation of bursts, multiplexing of bursts into a TDMA frame and finally for actual transmission of the data. The layer is responsible for channel coding and error detection/correction. Further, a physical layer includes some synchronization features that do not have any significance to the higher layers, since those features are purely hardware related. A physical layer is also responsible for detection of idle channels and the channel quality of downlink. However, it cannot identify the data types or data formats and cannot differentiate the control and user data. The data packets received from data link layer are transmitted without additional verification.

The physical layer at U_m interface performs encryption/decryption. The encryption in GSM is performed only between MS and BSS over the air interface. For the air interface of GSM systems, the GMSK (Gaussian Minimum Shift Keying) modulation is used. Over the terrestrial interfaces, i.e. in A_{bis} and A interfaces, data transmission at the physical layer uses pulse code modulation (PCM). The implementation of physical layer depends greatly on the type of interface.

Data link layer: This layer is responsible for the packaging of data. The data are combined into packets or frames and then handed to the physical layer for synchronous and asynchronous transmission. The main tasks of layer 2 are to detect and correct the errors. Data frames are formed by introducing start/stop marks and the check sums. When a receiver detects an error, it tries to correct the error or requests retransmission.

Layer 2 protocols might change from interface to interface. For example, layer 2 protocols for data exchange between an MS and the BTS is $LAPD_m$ (Link Access Procedure on the D channel). $LAPD_m$ is a modified version of the LAPD over U_m interface. Data exchange from the BTS to the BSC

[6] ITU (International Telecommunication Union) specifies the data link layer protocol LAPD in ISDN protocol stack on the D (data) channel referring to the ISDN channel that carries control and signalling information.

follows LAPD whereas data exchange between the BSC and the MSC follows MTP2/SS7 (message transfer part, layer 2 of SS7) as the layer 2 protocols.

Network layer: It prescribes the path a message has to take and the recipient of that message. All the information necessary to route a data packet is provided at this layer (layer 3). A network layer is divided into three sublayers—Radio Resource (RR) management, Mobility Management (MM) and Call Management (CM) units. A part of RR called RR′ is implemented in BTS, and the remainder is installed in BSC. The functions of RR′ are controlled by the BSC via the BTS management (BTSM) unit. Most of the RR functions are performed at the BSS.

The time duration when a mobile is in dedicated mode and busy with the configuration of radio channels is called RR-session. The responsibility of the RR sublayer is to manage this session. The main tasks of this layer, therefore, are to set up, maintain, and release of radio channels. In addition to this, it manages power control, discontinuous transmission/reception and handovers that are elaborated below:

- *Power control:* It is to minimize co-channel interference and to conserve power level that maintains an acceptable signal quality at MS. Mobile Station decides the acceptable power level measuring the bit error ratio (BER)[11].
- *Discontinuous transmission:* This is a power-saving mechanism. It is possible to implement in GSM exploiting the fact that a person speaks less than 40% of time during a normal conversation. By turning off the transmitter for rest of the 60% time can save power. However, in order to distinguish voice and background noise, very accurate voice activity detector should be used. While transmitter is off, the receiving end will hear a total silence. To avoid this, comfort noise is generated matching the characteristic of background noise.
- *Discontinuous reception:* This is also a power-saving mechanism. While in idle mode, mobile station has to listen only to paging channel that does not consume significant power.
- *Handover:* When an MS is engaged to a call and moves it may go away from one BTS (BTS_1) and come closer to another BTS (BTS_2). As a result the signal strength received at MS from BTS_1 decreases whereas the signal strength from BTS_2 increases. When the signal strength from BTS_1 falls below a threshold, the control of the ongoing call goes from BTS_1 to BTS_2. The event of such switching of control of a call in progress from one BTS to the other is called handover. It will be elaborated in Section 3.3.4.

[11] BER is the ratio of the number of bits incorrectly received to the total number of bits transmitted during a specified time interval. For example, if 3 bits are erroneous out of total 10^5 bits, the BER is 3×10^{-5}.

The MM sublayer manages the problem arises out of mobility of a user. To keep track of a subscriber, the MM layer notes the location data of the user. It also takes care of the task of authentication and secured communication. Whenever a call is made for a mobile user, it is desirable to locate the user correctly. In one extreme, to solve this problem, the system has to page the whole network to locate a user. It requires a huge number of paging messages leading to wastage of scarce bandwidth. In the other extreme, the MS has to update the system on every move. This causes wastage of bandwidth due to a lot of obsolete update messages. The trade-off between these two extremes sometimes gives an optimal solution in some cases. GSM has implemented an optimal solution introducing a concept of location area. The location area is defined as a group of neighbouring cells. The critical issue is to select number of cells forming a location area. The solutions are mostly based on statistical data.

Whenever a user moves from one location area to the other, irrespective of the MS's state (idle/dedicated) it sends this update to the current VLR responsible for the MS. The VLR in turn informs the HLR of the MS about the update. While a call is in progress (MS is in dedicated state) and if the location area is changed the update is performed after the call is terminated. When an incoming call arrives for an MS, the current location area is paged to track the MS.

The CM sublayer manages circuit-oriented services such as call set-up, call maintenance and call termination. In addition to these services, it manages short messaging service.

The additional protocols used at A interface (Figure 3.6) are Signalling System No. 7 (SS7) and BSS Application Part (BSSAP). The SS7 is used between BSC & MSC, MSC & HLR/VLR and MSC & another MSC. This protocol transfers all management information among MSCs, HLR, VLRs, AC, EIR and OMC. This information exchange enables location update, handover, authentication, incoming call routing, etc. The BSSAP protocol is used for communication between an MSC and a BSC. RR messages are sent between BSC and MSC using the BSSAP. It manages the allocation of suitable radio resources to the MSs and tackles mobility management.

The additional protocol used at A_{bis} interface is BTS management (BTSM). The BTSM works between BTS and BSC and allows control of the radio equipment and radio frequency allocation to the BTS. However, the A_{bis} interface has not yet been standardized.

ISUP/R2: Referring Figure 3.5, ISUP (Integrated Services Digital Network User Part) is a part of the SS7 signalling protocol stack. ISUP/R2 is a variant of ISUP. It is used to establish call set-up involving PSTN. During a call set-up, using the ISUP/R2 signalling service, a switch sends the necessary call set-up information, e.g. the caller and called subscriber numbers, to the next switching point en-route.

3.3.3 Services and Features

GSM avails three classes of services—bearer services, teleservices and supplementary services. Other than these three, GSM provides a number of high data rate services. Two such services, namely HSCSD (high speed circuit switch data) and GPRS (General Packet Radio Service), are discussed elaborately in Section 6.2.

Bearer services: The services provide a reliable data transport connection and are called lower level services. Some of the bearer services are data, packet, etc. Such services can be (i) Transparent (T) or (ii) Non-Transparent (NT). The transparent services ensure constant bit rate with changing bit error probabilities whereas the non-transparent services activate an additional protocol between an MS and the MSC for resending blocks with observed errors.

Teleservices: The services like voice, SMS and Facsimile are the teleservices and use bearer services for transport. For example, the facsimile teleservice requires bearer service for error correction during facsimile transmission.

Supplementary services: Primarily supplementary services are of two types—call offering and call restriction. The common call offering service is Call Forwarding Unconditional (CFU). On the other hand, the example of call restriction service is Barring of All Outgoing Calls (BAOC).

3.3.4 Handover

The Handoff mechanism elaborated with respect to cellular mobile communication (Section 2.7) is referred to as Handover in GSM. Handover occurs due to switching of an ongoing call to a different channel. Primarily there are two categories of handover—(i) internal handover and (ii) external handover. The internal handover is initiated due to switching channels in the same cell or switching to a channel in different cell under the same BSC. Such handovers are internal to a BSC and involve only one BSC. The MSC is notified only on completion of the handover.

The external handover takes place in two cases—while switching channels from the cells under the control of two different BSCs but belonging to the same MSC, or while switching channels from the cells that are under the control of different MSCs.

Handover may be initiated by MSC or by an MS. The MS always scans broadcast control channels of up to 16 neighbouring cells and forms a list of the six best candidates based on the received signal strength for possible handover. This information is then transmitted to the current BTS at least once per second. The BSC and MSC use this information for making handover decision.

The received signal strength at MS is measured as low due to either physical interference or movement of MS to another cell. This makes difficult in taking handover decision. In GSM, although there is no recommendation

about when to perform handover, it is implemented by the following two techniques.

- If signal received by an MS from the servicing BTS degrades beyond some threshold, transmission power is to be increased. If power increase does not lead to improved signal quality, a handover is initiated. However, the increase in transmission power may cause interference with neighbouring cells.
- Each MS constantly measures signal strength from the servicing BTS as well as from neighbouring cell BTSs. Whenever the signal level of a neighbouring BTS is higher than the serving BTS, a handover is initiated. This technique avoids neighbouring cell interference but it is quite complicated.

3.3.5 GSM Channels

GSM uses two bands, each of 25 MHz (890–915 MHz and 935–960 MHz), for system use. The 890–915 MHz band is for uplink/reverse link and 935–960 MHz is for forward link. The available frequency bands are divided into 250 kHz wide channels, that is, the 25 MHz bandwidth can create 100 channels. If guard band of 125 kHz is considered on both sides of the designated spectrum, there can be $(25000 - 125 \times 2)/250 = 99$ available channels. In GSM, the bandwidth is divided among as many users as possible following a scheme that crossbreeds TDMA and FDMA. Each BTS is assigned one or more channels or carrier frequencies. Every channel is time shared among the eight users. The introduction of TDMA scheme enables partitioning of a channel in time, forming logical channels.

GSM views its channels as physical and logical channels. A physical channel corresponds to a segment of one or more radio frequency channels used to carry information. Such a segment is defined in terms of frequency, time, code, etc. depending on the multiple access technique used by the system. In GSM TDMA, the physical channels are determined by the time slot of a carrier frequency. On the other hand, logical channels are identified by the category of information carried within the physical channel. Each logical channel is mapped or multiplexed onto one or more physical channels.

Logical channels are used to handle either traffic data, signalling data or control data. The different categories of logical channels are shown in Figure 3.7. Accordingly logical channels are either traffic channel or control channel. The traffic channels and some of the control channels are dedicated. Rest of the control channels are referred to as common control channel. A dedicated channel provides a bi-directional point-to-point transmission link.

The common control channels are accessed by MSs both in idle and dedicated modes. An MS in idle mode accesses these channels to switch to dedicated mode whereas the MS in dedicated mode accesses common control channel to monitor surrounding BSs for handover information.

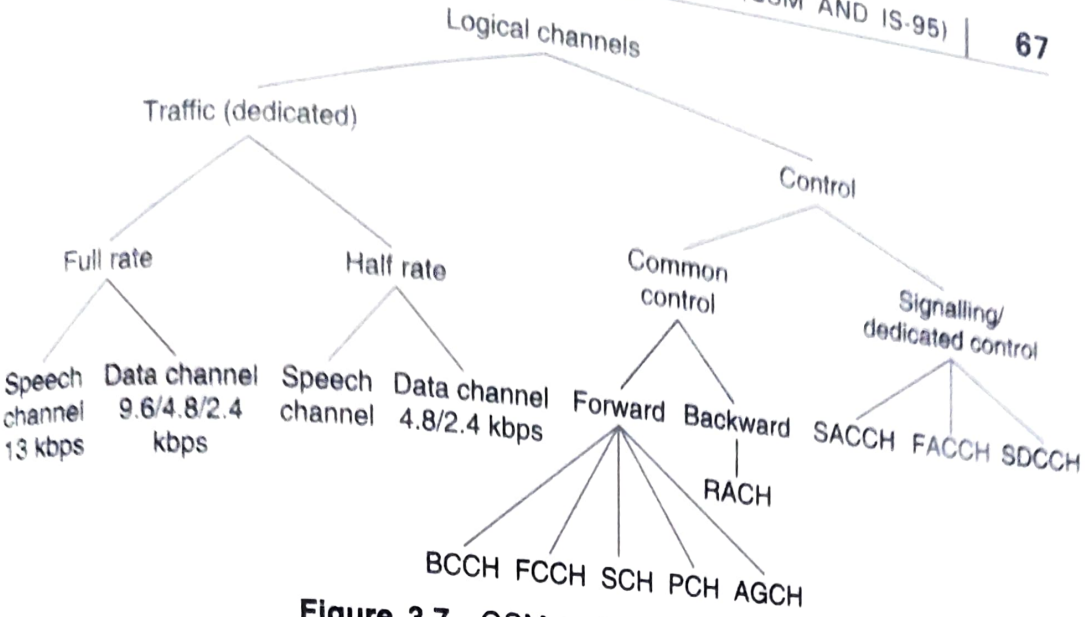

Figure 3.7 GSM logical channels.

The dedicated control channels are for maintenance of the call as well as for enabling a call set-up. These are utilized for managing handover, when the call is in progress, and finally for termination of the call. On the other hand, the traffic channels handle actual payload.

All the GSM logical channels are elaborated further below.

Traffic channels

Traffic channels carry digitally encoded user speech or data. It supports identical functions and formats on both the forward and reverse link. There are three types of traffic channels—TCH/F (full rate), TCH/H (half rate) and TCH/8 (one-eighth rate).

TCH/F: It transmits speech code of 13 kbps or three data-mode rates 12, 6 and 3.6 kbps.

TCH/H: It transmits speech code of 7 kbps or two data modes, 6 and 3.6 kbps.

TCH/8: It is used for low rate signalling channels, common channels and data channels.

Common control channels

A common control channel is unidirectional, that is, either from BTS to MS(s) or from MS to BTS. The tasks of all the common control channels are described below.

BCCH (Broadcast Control Channel): This is a forward control channel. BCCHs are for continuous information broadcasting about BTS identification, channel allocation, and list of adjacent BTSs. The channel allocation information helps an MS to know channel availability. On the other hand, with the help of the list of adjacent BTS information an MS can select one as its serving BTS. The BCCH is a point-to-multipoint channel from BSS to MSs.

FCCH (Frequency Correction Channel) and SCH (Synchronization Channel): These two classes of forward channels are also for broadcasting information. Each cell/BTS broadcasts one FCCH and one SCH flags in the first time slot of a TDMA frame. This broadcast helps an MS to synchronize its internal frequency standard to the exact frequency of the serving BTS. The broadcast also helps an MS to identify exactly when each time slot sequence begins.

PCH (Paging Channel): When a call comes for the MS, the BTS tracks the MS by the paging channel and alerts the MS of incoming call. It is a forward control channel.

RACH (Random Access Channel): It is a reverse control channel. When the MS desires to originate or set up a call, the MS uses RACH to send a request for accessing the network. In case of setting up a call in response to an alert received by PCH, the MS acknowledges the page received from the PCH by using RACH.

AGCH (Access Grant Channel): The AGCH, a forward control channel, is the final common control channel message sent by a BTS before the MS leaves all common control channels and uses dedicated channels. In response to RACH sent by an MS, the BTS uses AGCH to grant the request of the MS for accessing the network.

Dedicated control channels

All the channels in this group are bi-directional.

SDCCH (Stand-alone Dedicated Control Channel): In between granting access of network to the MS by BTS and assigning TCH to the MS, the BTS and MSC check the authentication of the MS. During this intermediate period when the MS waits for TCH, the BTS assigns SDCCH to the MS. If any information (e.g. authentication) to set up a call is sought from BTS during this period, the MS sends it over SDCCH.

SACCH (Slow Associated Control Channel): The SACCH is always associated with a traffic channel or a SDCCH. It provides a comparatively slow signalling connection. If the SACCH is associated with TCH, it is used for sending short message service (SMS).

FACCH (Fast Associated Control Channel): When a call is in progress, system information such as channel quality, power level, etc. that is necessary for call maintenance and handover are exchanged between BTS and MS over the FACCH. This channel works by stealing slots from a traffic channel.

Other than the three groups of channels mentioned above, Cell Broadcast Channel (CBCH) is an additional feature of a GSM system. It is a forward, point-to-multipoint logical channel. The CBCH is to support SMS broadcast service by means of which an MS can receive data/message broadcast, e.g. traffic, weather reports, etc. from the network service centre. It allows a limited short text as broadcast message.