Note that two basic network functions had to be performed during the call release operation. The first operation was a connection management function and the second operation was a radio resource management function.

These three examples are but a few of the possible traffic cases that can exist. Location updating and handoff management cases will be looked at in Chapter 4. These last few examples have been presented with the goal of increasing the reader's understanding of the various individual operations needed by a wireless mobile network and the overall system-level functions that occur.

QUESTIONS AND PROBLEMS

1. Which two elements of a wireless cellular system perform the "air interface" function?
2. What is the function of the transcoder controller?
3. What is the function of the visitor location register?
4. What is the function of the home location register?
5. What is the function of the mobile switching center?
6. What wireless cellular network element or elements provide security functions for the system?
7. What does a cell global identity number correspond to?
8. The LAI is used for what purpose?
9. What is the function of a radio network controller?
10. Name the two core networks associated with 3G cellular networks.
11. What is the difference between an MSISDN number and an IMSI number?
12. What is the purpose of a global title?
13. What is a mobile global title?
14. What is global title translation?
15. Using the Internet, determine the mobile country code for Mexico.
16. Explain the function of a mobile station roaming number.
17. During a mobile-originated call, when is authentication and encryption performed?
18. What is the first step performed by the mobile during a call release operation?
19. What is the last step performed during a call release operation?
20. What wireless cellular network elements are involved in a mobile-originated call?

CHAPTER 4

Wireless Network Architecture and Operation

Upon completion of this chapter, the student should be able to:

- Discuss the cellular concept and explain the advantages of frequency reuse.
- Draw a diagram of a typical cellular cluster and explain the meaning of frequency reuse number.
- Discuss how the capacity of a cellular system may be expanded.
- Explain the difference between cell splitting and sectoring.
- Discuss the use of backhaul networks for cellular systems.
- Explain the concept of mobility management and discuss the operations it supports.
- Discuss the concepts of power management and network security.

The cellular concept and its potential for increasing the number of wireless users in a certain geographic area had been proposed many years before it was ever put into practical use. The analog technology used by the first cellular systems dictated a certain type of cellular architecture. As time has past, newer digital technologies and the public's very rapid acceptance of cellular telephones has caused the architectures of today's cellular systems to change in an effort to adjust to the new technologies and the added demand for capacity.

Capacity expansion techniques include the splitting or sectoring of cells and the overlay of smaller cell clusters over larger clusters as demand and technology changes warrant. As demand for newer data services has increased, cellular operators have turned toward the development of their own private data networks to backhaul traffic from their cell sites to a common point of presence where a connection can be made to the PSTN or the PDN.

As cellular systems have matured and become nationwide wireless networks, mobility management has taken on an even more important role in the operation of wireless cellular networks. Mobility management is used to keep track of the current location of a cellular subscriber and to assist in the implementation of cellular handoff. Although not as glamorous as mobility management, power management and wireless network security have become more important issues as the cellular industry heads into its third decade of operation and wireless system engineers fine-tune their designs to build more secure systems and achieve even greater efficiencies of operation.

This chapter will examine all of the abovementioned issues and present several examples of typical cellular architectures and network operations.

4.1 THE CELLULAR CONCEPT

As briefly outlined in Chapter 2, the concept of cellular telephone service was first proposed in the 1940s. The cellular concept would provide a method by which frequency reuse could be maximized thus in essence multiplying the number of available channels in a particular geographic location. The concept of frequency reuse itself was not new at the time for it had been the guiding principle of the licensing of AM commercial broadcasting stations for years and is still used today to determine the granting of licenses for new stations in the broadcasting bands (AM, FM, and TV) and other radio services. However, in broadcasting (a **simplex** or single-direction transmission operation) the goal is to reach as many receivers as possible with a single broadcasting transmitter. This usually entails the use of a high-power transmitter to provide coverage of some particular geographic or trading area. However, there is nothing to prevent the same frequency assignment or cochannel from being used in another area of the country where the signals from distant cochannel stations do not extend to it. Since most users of the radio frequency spectrum recognize it as a limited resource, attempts are usually made to use it as efficiently as possible.

For **duplex** or two-way radio operation, where a system design goal is to allow as many simultaneous users of the available radio spectrum as possible, the reuse of that spectrum is crucial to maximizing the number of potential users. The cellular concept provides a means of maximizing radio spectrum usage. Another benefit of cellular radio systems is that the amount of mobile output power required is not as large due to the smaller cells used and therefore the power requirements for the mobile are reduced, which allows for longer battery life and smaller mobile station form factors.

Introduction

The first mobile telephone service, offered by AT&T and the Bell Southwestern Telephone Company in St. Louis, Missouri, consisted of several colocated transmitters on the top of Southwestern Bell's headquarters. A 250-watt FM transmitter paged mobiles when there was an incoming call for the mobile. This system's high-powered base station transmitters and elevated antennas provided a large coverage area and enough signal power to penetrate the urban canyons of the city. At the same time, however, the frequencies used by the system could not be used by any other services or similar systems for approximately a seventy-five-mile radius around the base station.

The first proposed cellular system would use many low-power transmitters with antennas mounted on shorter towers, to provide a much shorter frequency reuse distance. The area served by each transmitter would be considered a cell. The first cellular systems used omnidirectional antennas and therefore produced cells that tended to be circular in shape. As the technology used to create more efficient cellular mobile systems has evolved, so has the design and implementation of the cellular concept. These changes will be outlined in this chapter.

The Cellular Advantage

The deployment of a large number of low-power base stations to create an effective cellular mobile system is a large and expensive task. The acquisition of land for cell sites; the associated hardware; radio base station transceivers and controllers; antennas and towers; the communications links between the base stations, base station controllers, and mobile switching centers; and finally, the cost of the radio frequency spectrum needed to implement the system can be enormous. Mobile service providers can only recover their costs and make a profit if they can support a sufficient number of mobile subscribers. The cellular concept allows a large enough increase in capacity to make these operations economically feasible.

The implementation of the basic cellular architecture consists of dividing up the coverage area into a number of smaller areas or cells that will be served by their own base stations. The radio channels must be allocated to these smaller cells in such a way as to minimize interference but at the same time provide the necessary system performance to handle the traffic load within the cells. Cells are grouped into **clusters**

that make use of all the available radio spectrum. Since adjacent cells cannot use the same frequency channels, the total frequency allocation is divided up over the cluster and then repeated for other clusters in the system. The number of cells in a cluster is known as the cluster size or the **frequency reuse** factor.

For cellular architecture planning one must be concerned with interference from radio transmitters in other cells using the same radio channel and from interference from other transmitters on nearby channels. The first type of interference is known as cochannel and the latter is known as first-adjacent channel, second-adjacent channel, and so on. Using the cellular concept and careful design techniques can increase the maximum number of system users substantially. The following example will illustrate this point.

Example 4–1

Consider the following case: a service provider wants to provide cellular communications to a particular geographic area. The total bandwidth the service provider is licensed for is 5 MHz. Each system subscriber requires 10 kHz of bandwidth when using the system. If the service provider was to provide coverage from only one transmitter site, the total theoretical number of possible simultaneous users is 500 (5 MHz/10 kHz/user = 500 users). If, however, the service provider implements a cellular system with thirty-five transmitter sites, located to minimize interference and provide total coverage of the area, determine the new system capacity.

Solution: Using a cluster size of 7, the total system bandwidth is divided by 7 yielding approximately 714 kHz of bandwidth per cell (5000 kHz/ 7 = 714 kHz), and this is repeated over the 5 clusters (35/7 = 5). Now each cell has a capacity of 71 simultaneous users (714 kHz/10 kHz/user = 71 users) or a total system capacity of 2485 users (35 cells × 71 users/cell = 2485 users). This is a system capacity increase of approximately 5 times.

Cellular Hierarchy

Before examining the technical characteristics of frequency reuse and reuse number, it is helpful to define the hierarchical structure of today's cell sizes. The wireless industry has more or less settled on some particular names to indicate the size of a cell. Going from the smallest to the largest, cells that are less than 100 meters in diameter are known as **picocells,** cells with a diameter between 100 meters and 1000 meters (1 km) are known as **microcells,** and cells greater than 1000 meters in diameter are known as **macrocells.** These definitions are also related to the various possible operating environments that one might find oneself in. Picocells are usually found in the indoor environment (e.g., inside of buildings), microcells are found in the outdoor-to-indoor and pedestrian environment (urban), and macrocells are found in the vehicular and high- antenna environment (suburban). Each of these particular environments presents a different type of radio link propagation scenario that affects the required equipment and other technical aspects of the hardware used to implement the particular type of cell.

Newer technologies have expanded our concept of cells to include the global environment served by a variety of satellite systems and smaller cells for personal area networks (PANs) usually considered being less than ten meters in diameter. Although the terms have not become universal yet, cells with global coverage have been referred to as **megacells** and very small cells have been referred to as **femtocells.** Figure 4–1 illustrates the relative coverage areas of the various cell sizes. It is entirely possible to have mixed environments that are served by several different types of cell structures simultaneously.

4.2 CELL FUNDAMENTALS

Since the first cellular systems usually employed omnidirectional antennas and thus theoretically produced circular-shaped cells, the reader might be puzzled by the cellular industry's de facto choice of a hexagon as

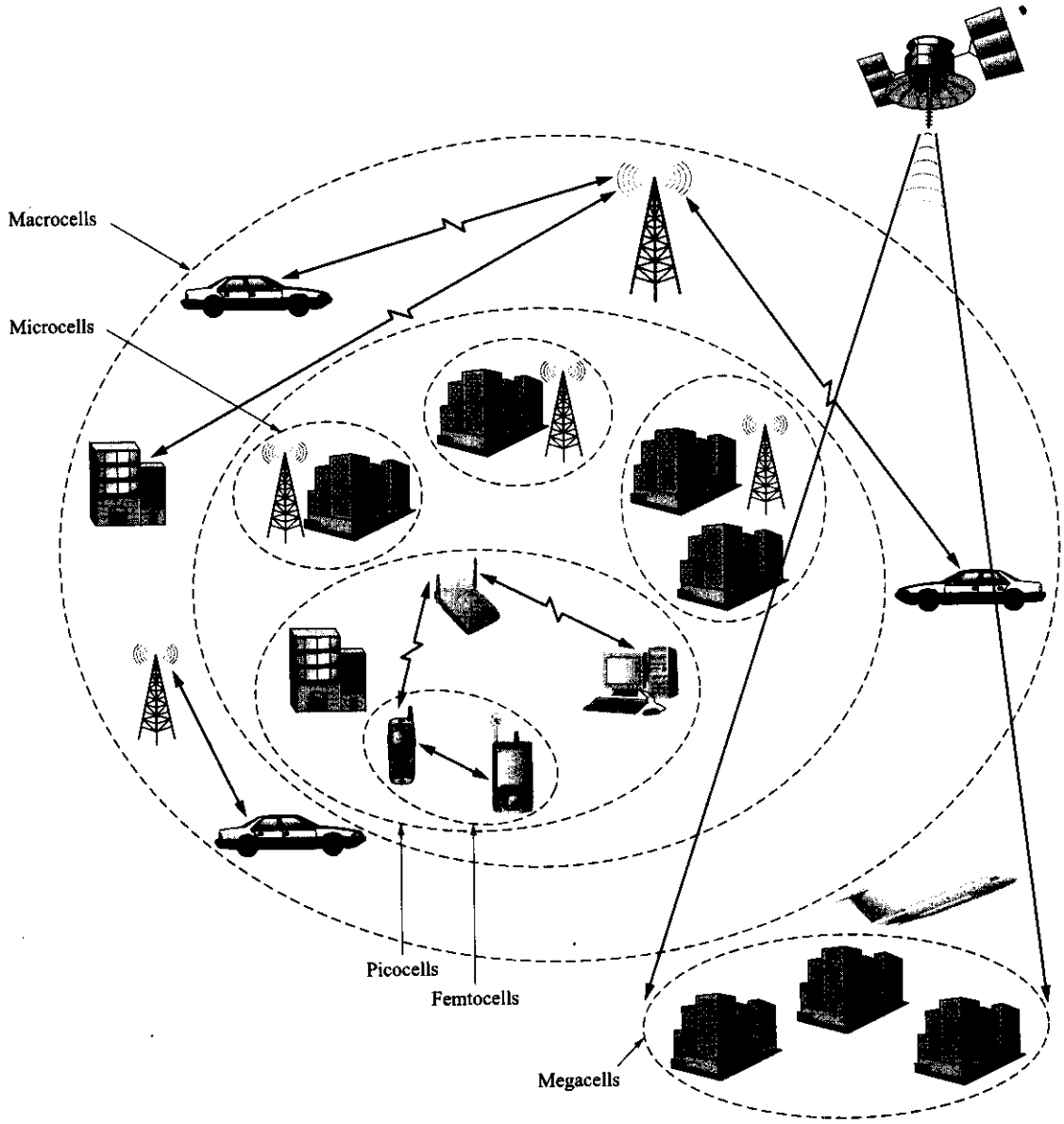


Figure 4–1 Relative coverage areas of different size cells.

shown in Figure 4–2 to represent a typical cell's coverage area in a service provider's network. Any initial consideration of the shape to use for a typical cell must be concerned with the fact that a true circular coverage area is rarely obtained in practice. Propagation conditions, terrain, and the environment (urban, suburban, etc.) all contribute to the distortion of an antenna's radiation pattern and hence coverage area. Furthermore, using circles to lay out a network's coverage area leaves gaps between adjacent tangent circles or ambiguous areas if the circles are overlapped. Referring to Figure 4–2, one can see that the use of a hexagon, however, allows for the complete theoretical coverage of an area without any overlapping cells or gaps in the coverage. Squares or equilateral triangles could also be used but the hexagon is the closest approximation to a circle. The use of hexagons also makes the theoretical calculation of several system parameters much easier.

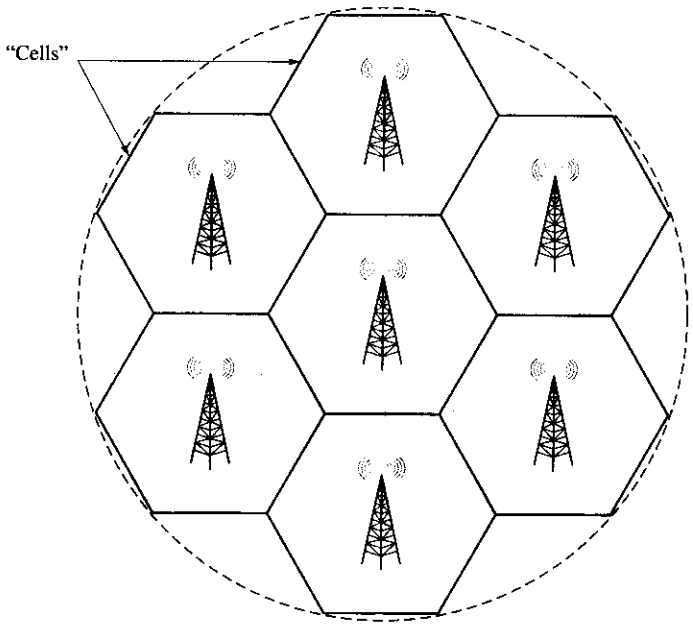


Figure 4–2 Use of hexagons to represent cellular coverage.

Reuse Number

In an attempt to gain the maximum reuse of frequencies for a cellular system, cells are arranged in clusters. To determine the minimum-size cluster that can be used it is necessary to calculate the interference levels generated by cochannel cells. Since there are several options to the size of cell clusters (see Figure 4–3 for several examples), a relationship to determine the reuse distance has been determined that relates cluster size, cell radius, and the reuse distance. The frequency reuse distance can be calculated by:

$$D = R(3N)^{1/2} \qquad \textbf{4-1}$$

where R = cell radius and N = reuse pattern.

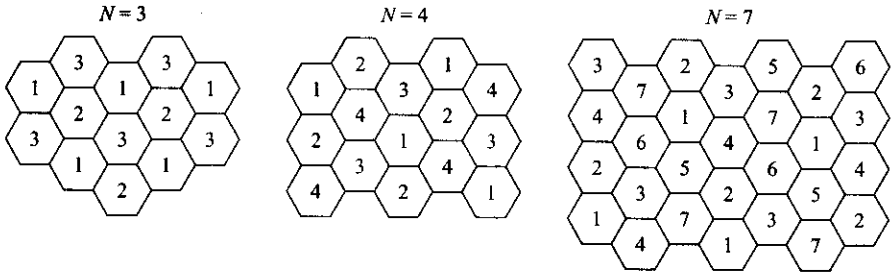


Figure 4–3 Various cellular reuse patterns.

Values of N can only take on numbers calculated from the following expression: $i^2 + ij + j^2$ where i and j are integers.

As can be seen from Equation 4-1, the smaller the value of N the closer the reuse distance and therefore the larger the system capacity or total number of possible users. It should be pointed out that reducing the size of the reuse distance D may provide the ability to handle more subscribers but it also increases network

costs in terms of the required hardware and acquisition of cell sites, increases the complexity of the network, and increases the number of operations required to provide mobility. The following example will illustrate the relationship between cluster size and reuse distance.

Example 4–2

For a mobile system cluster size of 7, determine the frequency reuse distance if the cell radius is five kilometers. Repeat the calculation for a cluster size of 4.

Solution: Figure 4–4 shows the typical arrangement for a cluster size of $N = 7$ and the reuse distance for cell 3. This is the cluster size typically used for the first-generation AMPS system used in the United States. .

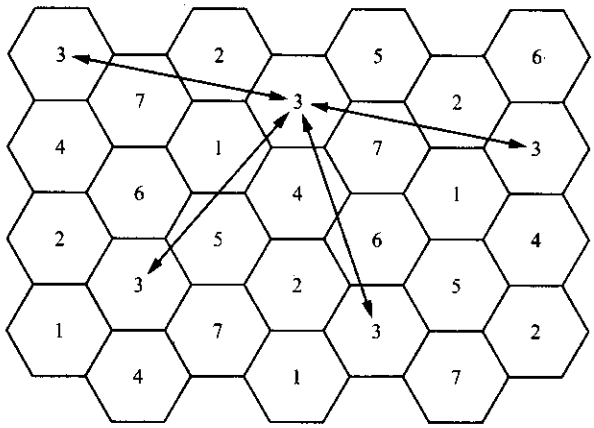


Figure 4–4 A frequency reuse diagram with the reuse distance, D, indicated (cluster size $N = 7$).

As mentioned earlier, using the expression $i^2 + ij + j^2$, one can show that a possible value for N is 7. As shown in Figure 4–4, the hexagons (cells) are arranged with one hexagon in the center of a cluster and six other hexagons surrounding the middle hexagon. Adjacent clusters repeat the previous pattern. The reuse distance is found from the following equation:

$$D = R(3N)^{1/2}$$

Therefore, for a cluster size of 7,

$$D = 5(3 \times 7)^{1/2} = 5(21)^{1/2} = 5(4.5823) = 22.913 \text{ km}$$

For a cluster size of 4, the reuse distance is given by:

$$D = 5(3 \times 4)^{1/2} = 5(12)^{1/2} = 5(3.464) = 17.32 \text{ km}$$

As can be seen, a smaller cluster size results in a smaller reuse distance.

Cellular Interference Issues

As already covered in the previous section, the frequency reuse distance can be calculated from Equation 4-1. Additionally, more complex calculations can yield the signal-to-interference ratio for a particular cluster size, N. The **signal-to-interference ratio** (S/I or SIR) gives an indication of the quality of the received signal much like the time-honored signal-to-noise ratio (SNR) measurement. Using a fairly simple mathematical model for S/I ratio calculations involving omnidirectional cells yields the results tabulated in Table 4–1 for several common values of N:

Table 4–1 Signal-to-interference ratio for various cluster sizes.

Cluster Size, N	*S/I Ratio*
3	11.3 dB
4	13.8 dB
7	18.7 dB
12	23.3 dB

The reader should be reminded that smaller cluster sizes will yield a larger possible subscriber base but as shown in Table 4–1 the trade-off is a lowered S/I ratio and the corresponding decrease in radio link quality. As a practical example of this fact consider the AMPS mobile system. The AMPS system did not yield usable voice-quality radio links unless an S/I ratio exceeding 18 dB was available. This value of S/I was only possible for a cluster of size 7 and up. Therefore, the typical AMPS system was deployed with a cluster size of $N = 7$ as shown previously in Figure 4–4.

Example 4–3

Show a possible distribution of channels for an AMPS system with a cluster size of $N = 7$.

Solution: For this situation, the 416 radio channels are divided by the 7 cells per cluster to yield 59+ channels per cell site. Each cell can have three control channels and some 56+ traffic channels. Table 4–2 shows one possible channel assignment scheme.

Table 4–2 A possible assignment of AMPS channels for a cluster size of 7.

Cell 1	*Cell 2*	*Cell 3*	*Cell 4*	*Cell 5*	*Cell 6*	*Cell 7*
Control Channels						
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
Traffic Channels						
22	23	24	25	26	27	28
29	30	31	32	33	34	35
36	37	...	...	...	...	...
...	...	...	...	...	...	401
402	403	404	405	406	407	408
409	410	411	412	413	414	415
416						

Note how each cell has a channel spacing of 7 × 30 kHz = 210 kHz and that this channel allocation is repeated in each cluster of 7 cells. Another way of assigning channels when the cluster size is 7 will be introduced later.

4.3 CAPACITY EXPANSION TECHNIQUES

As cellular mobile telephone service grew in popularity during the 1990s, the need to expand system capacity also grew. Most cellular providers will initially implement their systems by providing service in a coverage area with the least amount of initial investment (i.e., the least number of cell sites). As demand grows the system is usually expanded with additional cell sites to handle the increased traffic. There are several ways in which a service provider may increase capacity. The first and simplest method is to obtain additional frequency spectrum. Although this sounds like a fairly straightforward approach, it has proven to be one of the most expensive. Government auctions have sold frequency spectrum to service providers in countries all around the world. The fairly recent auctions of the PCS bands in the United States by the FCC in the mid-1990s yielded approximately $20 billion. The results of those high prices caused several of the top bidders for that spectrum to eventually declare bankruptcy. Another problem with this approach is that in many instances there is no frequency spectrum available to be auctioned off. In the United States as in many countries worldwide, previous spectrum allocations and incumbent radio services or applications are inhibiting and in some cases preventing the expansion of new advanced wireless mobile technologies. This topic will be treated more fully in other chapters.

The other approaches to capacity expansion are either architecturally or technologically enabled. Changes in cellular architecture like cell sectoring, cell splitting, and using various overlaid cell schemes can all provide increased system capacity. Another technique is to employ different channel allocation schemes that effectively increase cell capacity to meet changes in traffic patterns. Lastly, the adoption of next-generation technology implementations tends to provide an inherent capacity expansion within the new technology itself. The next few sections will provide more detail about these different methods.

Cell Splitting

If a cellular service provider initially deploys a network with fairly large cells, the coverage area will be large but the maximum number of subscribers will be limited. If a portion or portions of the system experience an increasing traffic load that is pushing the system to its limit (subscribers experience a high rate of unavailable service or blocking) then the service provider can use a technique known as **cell splitting** to increase capacity in the overburdened areas of the system. Consider the following example of cell splitting shown in Figure 4–5. Assume that Cell A has become saturated and is unable to support its traffic load. Using cell splitting, six new smaller cells with approximately one-quarter the area of the larger cells are inserted into the system around A in such a way as to be halfway between two cochannel cells. These

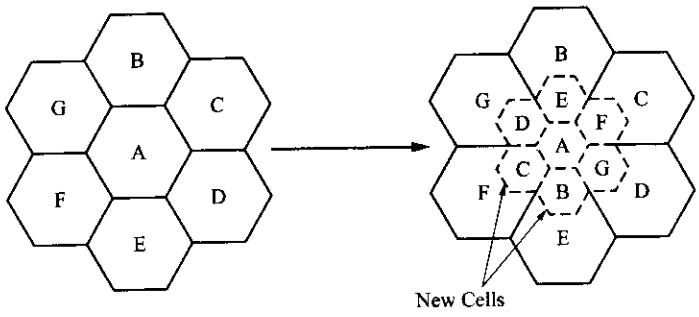


Figure 4–5 Increasing capacity by cell splitting.

smaller cells will use the same channels as the corresponding pair of larger cochannel cells. In order that the overall system frequency reuse plan be preserved, the transmit power of these cells must be reduced by a factor of approximately 16 or 12 dB.

Cell splitting will work quite well on paper; however, in practice many times the process is not as smooth as one would desire. Very often, due to the difficulty of acquiring appropriately located cell sites, the conversion process will be prolonged and different size cells will exist in the same area. In these cases, it is necessary to form two groups of channels in the old cell; one group that corresponds to the small-cell frequency reuse requirements and another group that corresponds to the old-cell reuse requirements. Usually the larger cell channels are reserved for highly mobile traffic and therefore will have fewer handoffs than the smaller cells. As the splitting process moves toward completion the number of channels in the small cells will increase until eventually all the channels in the area are used by the lower-power group of cells and the original Cell A has had its power reduced and also joins the new smaller cluster. As traffic increases in other areas of the system this process may be repeated over again. Eventually the entire system will be rescaled with smaller cells in the high-traffic areas and larger cells on the outskirts of the system or in areas of low traffic or low population density.

Cell splitting effectively increases system capacity by reducing the cell size and therefore reducing the frequency reuse distance thus permitting the use of more channels.

Cell Sectoring

Another popular method to increase cellular system capacity is to use **cell sectoring.** Cell sectoring uses directional antennas to effectively split a cell into three or sometimes six new cells. The vast majority of cellular providers use this technique for any of the cellular systems presently in operation. As shown in Figure 4–6, the new cell structure now uses three-directional antennas with 120-degree beamwidths to "illuminate" the entire area previously serviced by a single omnidirectional antenna. Now the channels allocated to a cell are further divided and only used in one sector of the cell.

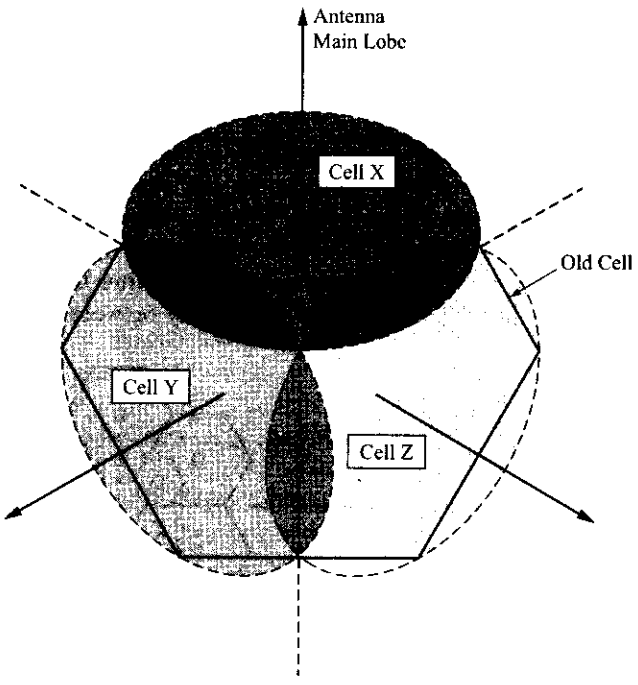


Figure 4–6 Increasing capacity by cell sectoring.

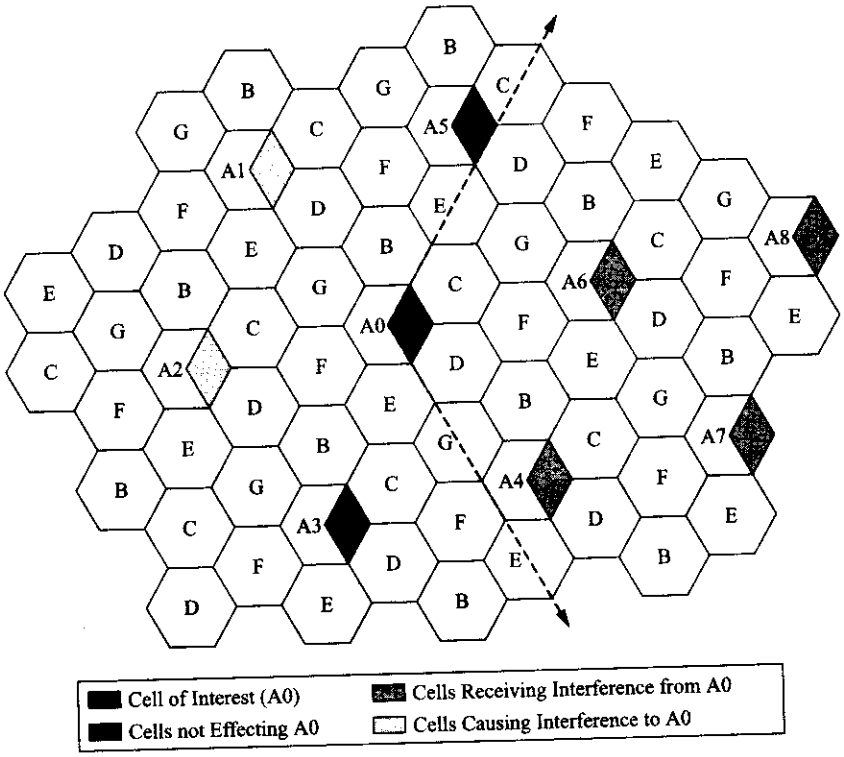


Figure 4–7 Interference reduction due to cell sectoring.

As shown in Figure 4–7, the sectoring of a cell results in a reduction in the amount of interference that the sector experiences from its cochannel neighbors in adjacent clusters and conversely the amount of interference that the sector supplies to its cochannel neighbors. Before sectoring, for a cluster size of 7, a cell receives and gives interference to six other nearest cochannel cells in other clusters. Now, as shown by Figure 4–7, for Cell A0, the number of interfering cells has been reduced to two (A1 and A2). This results in a higher S/I ratio for that sector and its companion sectors in other clusters. Table 4–3 tabulates these new values for a three-sector scheme for some common values of cluster size.

Table 4–3 Signal-to-interference ratio for three sector schemes.

Cluster Size, N	*S/I Ratio*
3	16.08 dB
4	18.58 dB
7	23.44 dB
12	28.12 dB

Note that these results indicate that for AMPS service, if a three-sector-per-cell site scheme is used, one can reduce the reuse cluster size down to 4 and gain more system capacity! If the sector scheme uses 60-degree beamwidth antennas yielding six sectors per cell site, then it is possible to employ a reuse cluster size of 3 for AMPS service. Note that the sectoring process does not require new cell sites, only additional

directional antennas and triangular mounting platforms or other mounting schemes to create the desired sectors with the appropriate directional antennas. The term *cell site* has now taken on new meaning as the typical cellular system architecture has increased in complexity. Example 4–4 illustrates the channel distribution employed in sectored systems.

Example 4–4

For a system with frequency reuse number $N = 7$, show a possible distribution of the channels over the sectored system. Recall that without sectoring, as shown in Example 4–3, each cell had some 59+ channels available.

Solution: When a system like this is sectored, now each of the three sectors uses the frequencies previously allocated to the entire cell. Now each sector can have 416/21 = 19+ frequencies. This type of configuration is typically known as a 7/21 reuse plan. Figure 4–8 illustrates a possible implementation of this 7/21 reuse plan. Also shown are the cellular architectures of several other reuse plans (4/12 and 3/9). Table 4–4 shows a possible way to assign channels for a 7/21 frequency reuse plan.

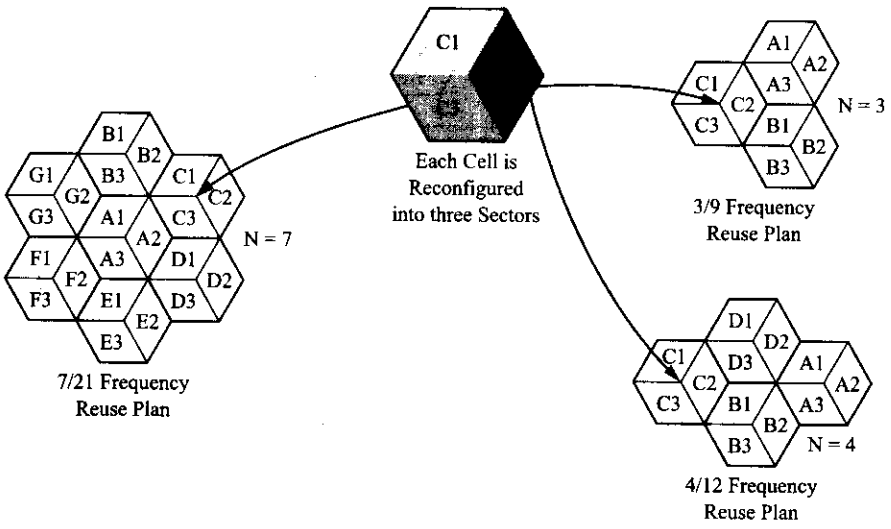


Figure 4–8 7/21, 3/9, and 4/12 frequency reuse plans for sectored cells (Example 4–4).

Note that for this assignment scheme, each sector has a single control channel (channels 333 – 313) and then the other 395 traffic channels (channels 1 – 312, 667 – 716, and 991 – 1023) are distributed over the twenty-one sectors before they are repeated again in other clusters.

Overlaid Cells

The use of **overlaid cells** was first introduced in the section on cell splitting. This method can be used to expand the capacity of cellular systems in two ways. The first method explained here may be applied to what are known as split-band analog systems. However, since analog FM modulation cellular systems are at the end of their life cycle in the United States, only a brief coverage will be given to this topic.

The reader may recall from Chapter 2 the description of several follow-on first-generation analog systems that used a form of narrowband FM, such as the NAMPS or NTACS systems. Using overlaid cells an operational wideband analog system could be upgraded to increase its capacity by overlaying another

Table 4–4 Channel assignment scheme for a 7/21 frequency reuse plan.

Frequency Group	*A1*	*B1*	*C1*	*D1*	*E1*	*F1*	*G1*	*A2*	*B2*	*C2*	*D2*	*E2*	*F2*	*G2*	*A3*	*B3*	*C3*	*D3*	*E3*	*F3*	*G3*
Analog Control Channels	333	332	331	330	329	328	327	326	325	324	323	322	321	320	319	318	317	316	315	314	313
Traffic Channels	312	311	310	309	308	307	306	305	304	303	302	301	300	299	298	...					
	291	...																			
	270	...																			
	249	...																			
	228	...																			
	...																				
	39	...																			
	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1			
										716	715	714	713	712	711	710	709	708	707	706	705
	704	703	702	701	700	699	698	697	696	695	694	...	...	...	...	...	...	...	...	...	...
	...	...	...	...	...	...	...	676	675	674	673	672	671	670	669	668	667				
	999	998	997	996	995	994	993	992	991												
	1020	1019																	1023	1022	1021

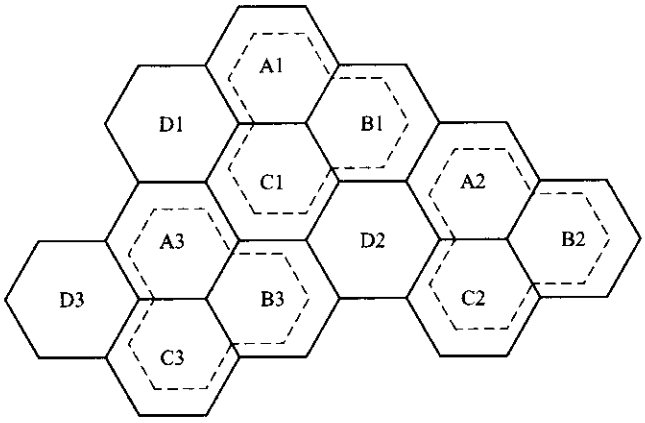


Figure 4–9 Overlaid cells in a split-band system.

analog system with narrower bandwidth requirements over it. In such a split-band overlay system, channels are divided between a larger macrocell (using AMPS or TACS) and the overlaid microcell (using NAMPS or NTACS) that is contained in its entirety within the macrocell. This type of situation is shown in Figure 4–9. The channels assigned to the macrocell are used to service users in the area between the microcells, and the channels assigned to the microcells service the microcells. With correct system design the two areas just mentioned will be equal in size. The net effect of this design is an increase in the total number of system channels since now the entire system bandwidth is allocated to both the original wideband system and the newer, more efficient narrowband system. This type of system migration requires the use of dual-mode mobile stations.

The second method of using overlaid cells may be applied to GSM or NA-TDMA systems. As an example of this method, consider a system with a cluster size of $N = 4$. On top of this system, a cluster of overlaid cells is applied with a cluster size of 3. If the channels for the overlaid cell cluster are taken from the underlaid cluster, the system capacity increases since the area needed for the overlaid cells is only 75% of that needed for the underlaid cells. The more channels borrowed from the underlaid system for the overlay system, the greater the increase in system capacity. This type of expansion allows operators to migrate their systems using the same base station and mobile station equipment. See Figure 4–10 for an illustration of this technique.

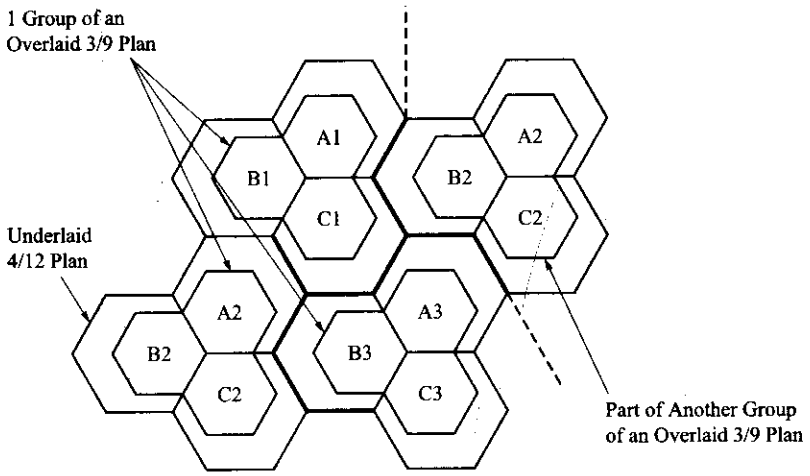


Figure 4–10 Overlaid cells in a reduced cluster size system.

The use of overlay schemes that have subcells within larger cells has also been known as "**tiering**" by some segments of the cellular industry.

Channel Allocation

The methods of capacity expansion presented so far have relied on changes to the cellular system's architecture to gain additional system resources. The designs of these systems have used an equal distribution of channels for each equal-size cell or sector within a system. This would be all right if the number of users within a cell or sector was constant, or better still if the amount of traffic offered to each cell in the system was constant. However, neither of these cases is true. In practice, the amount of traffic offered to a wireless cellular system and to individual cells of that system is dynamic, with certain times of the day and week commanding widely different levels of usage.

One can come up with many different scenarios of activities that might cause the amount of traffic to change. For example, during events like rock concerts and sporting events, the amount of traffic offered to cellular systems can change drastically for short periods on the scale of hours. Other events like professional golf tournaments or state fairs could change traffic intensity for a week or longer. The business district of a metropolitan area may experience changing levels of traffic over the course of a workday, with a certain average level of activity during the workweek, but see a decline in that activity on the weekend. The first scenario is so extraordinary that it is very difficult to design anything into the system to handle the extremely large increase in traffic offered to the system. In many such cases, cellular providers will bring in portable cellular sites (sometimes known as "cells on wheels" or COWs) to handle the increased demand. A national cellular service provider may have dozens of COWs that are deployed all over the country at any given time. COWs are also deployed during natural disasters to restore disrupted communications. On the other hand, the traffic scenario within the business district can be dealt with to some degree through channel allocation techniques.

Cellular service providers are very sensitive to the issue of nonavailability of service or what is known as **call blocking.** Since there is a great deal of competition for subscribers within the industry, it behooves the various cellular providers to configure the capacity of their systems so that there is a minimal amount of blocking. Most cellular operators attempt to keep the probability of call blockage below 2%, believing that subscribers will remain satisfied with their service at this level.

The goal of channel allocation is to attempt to stabilize the temporal fluctuations of call blockage over both the short and long term throughout the entire mobile network. Presently, there are several methods that can be used to achieve a more optimal channel allocation across a system. Also, it should be noted that if service providers can reduce the average system call blocking probability, it allows them to accept more subscribers, which effectively expands the system capacity.

There are three main methods used for achieving a more efficient system channel allocation scheme. Fixed channel schemes examine systemwide traffic patterns over time and then "fine-tune" the system by allocating additional channels where needed. This means that instead of equally dividing up the channels over the cells, some cells will receive larger channel allocations than others. Usually, very complex algorithms are needed to determine the final allocation of the channels, and these allocations are periodically updated as a traffic usage database grows and new patterns of use emerge. The second allocation method is known as channel borrowing. In this scheme, high-traffic cells can borrow channels from low-traffic cells and keep them as needed or until the offered traffic returns to normal. For this scheme, a borrowed channel from another cell will effectively cause additional cells in other clusters to lose the use of that particular channel since the reuse distance of the borrowed cell has decreased relative to these cells. After the traffic over the borrowed channel is complete, the channel is returned to use in its original cell. The third channel allocation technique is known as dynamic channel allocation (DCA). In any number of DCA schemes, all the available channels are placed in a channel pool. A channel is assigned to a new call by virtue of the systemwide signal-to-interference statistics. Each channel can be used in each cell as long as the necessary signal-to-interference ratio is met. This is an extremely complex system that uses many network resources

to accomplish its operation. Another downside to the scheme is that every cell site must be capable of transmitting every one of the system's assigned channels; this is an expensive proposition.

The use of sophisticated channel allocation techniques will continue to grow as better algorithms are developed and the wireless cellular networks become more intelligent.

Other Capacity Expansion Schemes

There are several other methods that can be used for capacity expansion of a cellular system. This section will offer an overview of several of these techniques.

Lee's Microcell Technology

A downside to the sectoring concept is the need for an increased number of handoffs and the resulting increased load on the network's switching elements. A technique known as Lee's microcell method has been proposed that uses zones instead of sectors to reduce the number of handoffs required as a mobile station moves from one zone to another within the microcell. As shown in Figure 4–11, this technique employs three antennas that provide coverage by "looking" into the microcell. All three antennas are connected to the same base station by high-speed microwave or fiber links. The antenna with the best reception of the mobile is used for both the uplink and downlink. As the mobile travels within the microcell the same channel can be used and there is no need for handoff operations. As the mobile moves into another zone the base station simply switches the channel to a different zone.

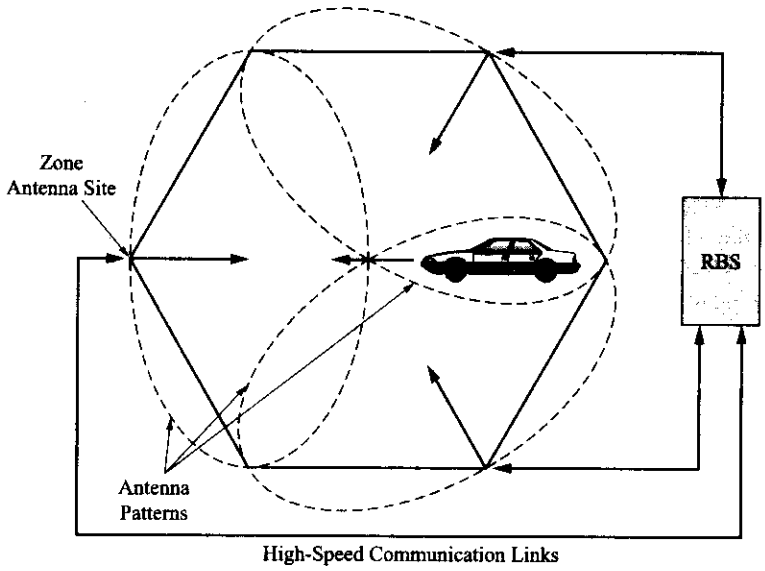


Figure 4–11 Lee's microcell concept for capacity expansion.

Smart Antenna Technology

Although not implemented as of yet, the specifications for 3G cellular systems call for the support of **smart antenna** technology. Using smart antennas, a base station could direct a narrow beam of radio waves at a particular mobile station and then reuse the same channel over another narrow beam aimed at yet another mobile in another location. Smart antennas use phased array technology. This technology allows for the creation of directional antenna patterns that may be sequentially switched to other patterns at high speed. This technique is sometimes referred to as space division multiple access (SDMA).

Many presently installed systems use a form of **space diversity** to enhance system operation by using two or more receiving antennas. In these systems, the best signal from the receiving antennas is chosen for use by the system. This is sometimes mistakenly referred to as smart antenna technology.

Migration to Digital Technology

The last expansion method consists of the migration to a newer-generation (digital modulation based) system. Most of the technologically advanced countries in the world have already gone through this process and are poised to continue evolving their systems to 3G technologies that offer more digital services. Some of the less advanced countries still use first-generation systems. In fact, used first-generation equipment continues to be resold to these countries as they expand their current systems. Second-generation systems use time division multiple access (TDMA) and code division multiple access (CDMA) technologies to achieve greater capacity than analog systems. For TDMA systems, multiple timeslots are used per channel allowing for multiple users per channel. For CDMA systems, multiple users may use the same channel simultaneously. TDMA systems are much more immune to noise and interference and can therefore provide service with much lower values of S/I than an analog system. The NA-TDMA IS-136 system only needs an S/I ratio of 12 dB whereas a GSM system can operate with an S/I ratio of only 9 dB. This translates into frequency reuse factors of 4 and 3, respectively (refer back to Table 4–1). CDMA systems with their inherent interference handling capabilities may use the same frequencies in adjacent cells and thereby lower the frequency reuse factor to 1. In all cases these newer systems provide increased system capacity.

4.4 CELLULAR BACKHAUL NETWORKS

As cellular systems have evolved from voice-only to both voice and data services systems, the requirements for connectivity to the PSTN and PDN have changed. First-generation systems provided a voice connection to the PSTN. A subscriber could make a voice call over the PSTN or, if desired, send data through the PSTN by the use of a voiceband modem (a circuit-switched data transmission). The infrastructure of first-generation cellular systems was typically connected together using T-carrier, E-carrier, or J-carrier facilities. T1/E1/J1 lines were used between the mobile switching center (MSC) and the base station (BS) and later the base station controller (BSC). Recall that T1s are used in the United States, E1s in Europe and other parts of the world, and J1s in Japan. The connection between the MSC and the BS carried PCM-encoded voiceband signals at 64 kbps. A T1/J1 can handle twenty-four voiceband calls and an E1 can handle thirty.

For second-generation cellular systems the voiceband signals are usually transcoded (compressed and reformatted) at the BSC and sent over T1/E1/J1 facilities at either 8 kbps or 16 kbps allowing as many as 192 voice channels over a single T1 or J1 line. If the capacity of a single T1/E1/J1 was insufficient, additional facilities would have to be used. Between the MSC and the PSTN, traffic was typically aggregated and, if warranted, usually sent over a larger T3 facility that could provide room for growth. The use of fiber facilities to perform this function is commonplace now. The cellular service provider has to rent these lines from the local telephone company on a monthly or yearly basis. Therefore, anything that can be done to reduce or minimize these costs is welcomed by the cellular operator.

With newer (2.5G+ and 3G) systems, cellular operators are starting to install their own private wideband networks to backhaul both voice and data from the BSs to the BSCs and finally to the MSCs in an effort to reduce costs.

When mobile data services like CDPD were introduced to first-generation cellular systems, the connection to the public data network was completed through separate facilities and kept independent from the voice network. At the time, the amount of data traffic was light enough to be carried over leased lines. When second-generation systems using TDMA and CDMA technologies were introduced, several changes occurred within the existing wireless networks. CDMA systems maintained the connection between the MSC and the BSC for voice traffic but introduced the interworking function/packet data service node

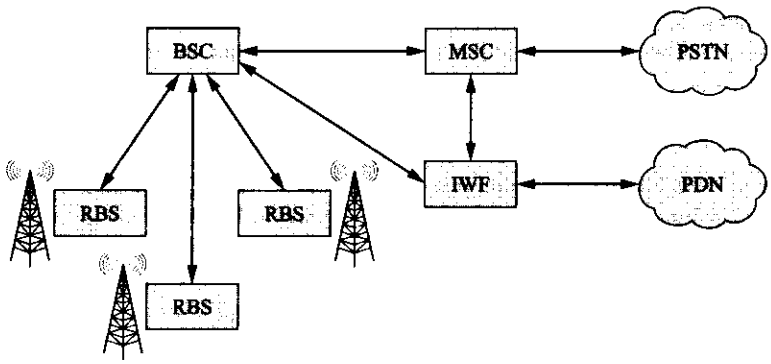


Figure 4–12 CDMA cellular system data network connection.

(IWF/PDSN) network element that connects directly to the external packet network and the BSC. As shown in Figure 4–12, this node is responsible for proper protocol conversion and mapping between the wireless network and the external packet network.

GSM cellular systems introduced packet-switched data services through general packet radio service (GPRS). In this system, in addition to the traditional GSM network components, a GPRS public land mobile network (PLMN) has been added that interfaces to packet data networks (X.25, IP, etc.) as shown in Figure 4–13. Through this GPRS PLMN, the GSM subscriber is able to access Web sites through public servers or corporate intranets through private enterprise servers. Voice services are supplied through the traditional GSM PLMN as indicated by Figure 4–13. For both CDMA and GSM cellular systems there are interconnections between the two networks to provide location information about the mobile station.

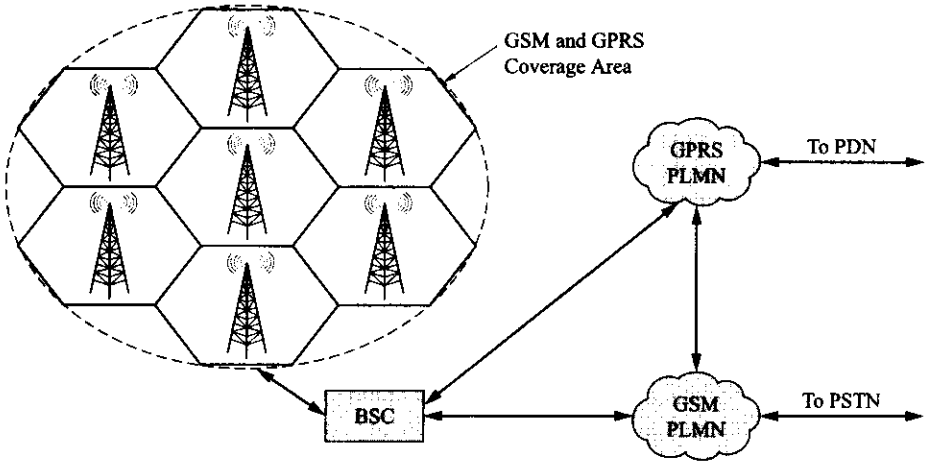


Figure 4–13 GSM cellular system data network connection.

As 3G networks with high-speed data services are deployed, with the ultimate goal of an all-IP wireless network (4G cellular), service providers are looking increasingly toward the building of their own packet networks to provide connectivity between the wireless network and the core voice and data networks. To accomplish this goal, wireless cellular equipment manufacturers are providing network solutions to service providers. Often these solutions involve advanced digital transport technologies like asynchronous transfer mode (ATM) and SONET/SDH (synchronous optical network/synchronous digital hierarchy) used to transmit data over fiber or digital microwave radio link facilities. Many cell sites do not have access to fiber

facilities; therefore, digital microwave radio is an attractive option to be used in the building of these networks. More detail about the deployment of these types of cellular backhaul networks will be presented in Chapters 7, 8, and 12.

4.5 MOBILITY MANAGEMENT

The most important characteristic of wireless telecommunications systems is the ability to provide mobility to the user. The general public has demonstrated its desire for "untethered" electronic communications many times since the first radio signals were transmitted over 100 years ago. Whether it has been the acceptance of car radios, cordless telephones, or cellular phones, the early rate of adoption of each of these innovations has been exponential in nature. The number of worldwide mobile telephone subscribers is predicted to pass two billion by the year 2007. As stated before, wireless mobile telephones are evolving into more than just voice-oriented devices. Modern mobiles or subscriber devices have the ability to provide data services and access to the Internet with ever increasing data transmission rates. The functionality of these subscriber devices is being enhanced by multimedia capabilities that support voice, high-quality audio, and video messaging. Cell phones with built-in video cameras are in fact here and no longer just a futuristic invention made popular by the Dick Tracy comic strips of so many years ago.

With the likelihood of the cellular subscriber base exceeding more than two billion within this decade, one might pause for a moment to consider the complexity of the systems needed to manage all these users. Certainly there is a need for a physical infrastructure to support the operations mentioned earlier but there also is a need for a radio network that manages the countless operations needed to make the entire system function correctly.

Contrast a wireless system with a traditional wireline system where the physical infrastructure is connected to the fixed subscriber device and therefore the signals are guided by the transmission media to the correct destination. Indeed, although the PSTN needs a switching "fabric" at the core of the network to direct one's call to the correct telephone, once the connection is made, there is an end-to-end physical path for the signal to propagate over. A wireless system does not have the luxury of knowing where the mobile subscriber is at all times and therefore must incorporate a means to determine this information and subsequently infuse this data into the system. At the same time, a mobile station should have the ability to be able to continuously access or use the services of the system that it is connected to. Wireless network functionalities necessary for efficient system operation are achieved through the use of programmable information processing systems and information data-bases built into the major system components (e.g., MSC and BSC) and the radio signal measurement capabilities built into the air interface components (i.e., base and mobile stations). The next several sections will discuss the concept of mobility management for cellular systems. The goal of these sections is to explain how the network knows where the subscriber is (location management) and how it keeps track of and in contact with the mobile station as the subscriber moves around from cell to cell (handoff management).

Mobility management for wireless LANs and other wireless data networks covered by the IEEE 802.XX standards will be covered in the chapters devoted to those topics.

Location Management

Location management is the process of keeping track of the present or last known location of a mobile station and the delivery of both voice and data to it as it moves around. Since there are literally hundreds of thousands of worldwide cell sites, there needs to be functionality built into every cellular system that will provide the system with the ability to locate one particular mobile station out of the billion plus in existence.

This process is best explained, in the case of a voice call, as follows: When a call is made that passes through the PSTN, a dedicated traffic channel must be set up from the BS to the MS for a call to be completed. The PSTN sets up the circuit over the fixed part of the network and the wireless network will

allocate a pair of radio channels for the air interface connection. Naturally, for this process to be successful, the location of the MS must be known. Additionally, if the mobile moves during the time span of the conversation, a process must be in place to provide for a continuous radio link even though the mobile might move into another cell. For the case of a data transfer, packets are typically addressed to an end terminal or destination device. The packets are directed through the data network by routers to a particular device. For a fixed device this corresponds to a fixed location. For the mobile device it is necessary to know the location of the device before the data packet can be delivered to it. Furthermore, the system must know the availability of the called party. In a fixed system, busy signals are used to denote a telephone already in use. For a mobile system, the mobile may be in use or may not even be turned on. In both of these cases, the network must be able to determine the status of the mobile and take the necessary action to deal with the incoming call or data transfer. This action might be the playing of a recorded message indicating that the mobile is busy and then implementing an answering machine function or the storage of the data transfer information on some type of network storage device for later delivery.

In general, there are three basic functions performed by location management: location updating, sending paging messages, and the transmission of location information to other network elements. The next several sections will examine these generic network operations in more detail. Later chapters will provide system-specific details.

Location Updating

The location updating function is performed by the mobile station. Recall that when the MS is first turned on, it performs an initial system registration or "attach" with the base station of the cell that it is located in and thereafter this information is periodically checked to verify its accuracy and prevent an accidental detach of the mobile from the system. If the mobile does not change location, the access point to the fixed network remains unchanged and the fixed portion of the wireless network delivers information to the mobile using this particular access point.

The system is designed so that the mobile station will send an update message every time it changes its point of access to the fixed network. As stated earlier, after the initial power-up registration the mobile station and base station will periodically exchange their respective identification information. If the MS receives the ID of a BS or a location area (LA) that is different from the value stored in its memory (this could happen through a handoff during a call or simply be due to the mobile's change of location), the MS will send a location updating request message to the fixed network through this new access point and also provide information about the mobile's previous access point. This information will be entered into a VLR database maintained by the fixed portion of the wireless network and be used by the network to locate the MS. The motion of the MS can therefore be tracked by this process to a specific LA or base station. This process has its drawbacks because updates are periodic and therefore introduce some uncertainty into the exact location of the mobile. In an extreme case, a mobile may be turned off and transported across the country by the subscriber. In this instance, an incoming call to the mobile while it is out of service will result in a page being sent to the last known access point, which would produce a no response or failed page. After that failed attempt, the system might possibly page a group or groups of surrounding cells, which will also fail. The system would then enter its voice message mode indicating the unavailability of the mobile. When the mobile is turned on again, this problem will be resolved by the system when new registration information is received and the mobile's location is updated within the fixed portion of the system. Now any calls will be directed to the mobile's new location.

A balance needs to be achieved by the wireless network involving the number of update messages and the number of cells that must be paged by the system to locate the mobile. If updating is performed very frequently, the location of the mobile will be known with a greater degree of certainty; however, the system resources (both radio and network) used to accomplish this task will be excessive. On the other hand, if updating is performed infrequently, the number of access points that need to be paged to find the mobile increases and may have the adverse effect of causing too many calls to be dropped or data packets lost due

to long delays in the determination of the mobile's location. A forthcoming example will illustrate a typical system that provides a compromise between these two conflicting goals.

There are usually two types of updating schemes used by wireless networks—static and dynamic. For static schemes, the cellular network's geographic layout determines when the location updating needs to be initiated. For dynamic schemes, the user's mobility and the cellular system layout both contribute to the initiation of the location updating algorithm.

Today, most cellular systems use the static method of location updating (see Figure 4–14). In this approach, a group of cells is assigned a location area identification value (LAI) (refer back to Figure 3–9). As shown in Figure 4–14, each BS in the LA broadcasts its ID number in a periodic fashion over a control channel. The MSs that are attached to the base stations within the LA are required to listen to the control channel for the LA ID. If the LA ID changes, the MS will have to send a location update message to the new base station. The BS will forward the updated information to the VLR database located in the fixed portion of the wireless network. Now, if there is an incoming message for an MS, a paging message will be sent to all the cells in the LA where the MS is listed as being present. The MS, unless it has moved into another LA, will respond to the paging message. One problem faced by a static location area ID scheme is known as the "ping-pong" effect. This effect can occur if the mobile is moving in a path that takes it back and forth between the borders of two LAs. This problem can also affect the handoff process. Practical solutions used to prevent this effect will be presented when handoff is discussed.

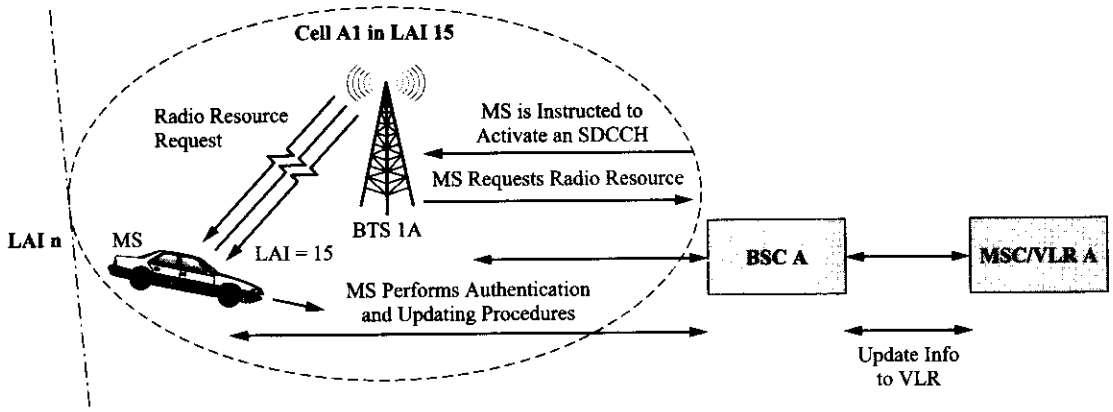


Figure 4–14 Cellular location updating.

Dynamic location updating plans are not as popular within the wireless industry. These schemes are typically based on the status or state of the mobile. Some of the typical measures used to determine the mobile's status and hence determine the need to perform the updating algorithm are elapsed time, total distance traveled, call patterns, number of different LAs entered, and so on.

Paging Messages

An incoming call or message to a mobile station will initiate the paging of the mobile. Paging consists of the broadcasting of a message either to a cell or to a group of cells that is meant to bring a response from a single particular mobile. This response will start the process by which communications between the PSTN or the PDN will be established with the mobile. The paging of a mobile is more efficient if the exact cell the mobile is registered in is known. However, as pointed out, this information is not always available. Therefore, several different strategies for paging exist. Sometimes a scheme known as blanket paging is employed. This type of a page will be broadcast to all cells in a particular location area. If successful, the mobile will respond after the first paging cycle and delays will be kept to a minimum. Otherwise, a scheme of sequential paging is used. In this paging strategy, the cell where the mobile was last registered is paged

first. If not successful, the next group of surrounding cells is paged. If this attempt to reach the mobile is still not successful, another larger ring of surrounding cells is paged and so on until the page is successful or a paging cycle timer expires and the MS is declared unreachable by the system. Depending upon several system variables, both paging schemes offer various advantages and disadvantages.

Transmission of the Location Information between Network Elements

For location updating to work correctly in a wireless network, there must exist several databases where mobile station information can be stored and accessed by the network as needed. When a subscriber enters into a service contract with a service provider, the subscriber's mobile device is registered (i.e., mobile ID numbers are stored) in a home location register (HLR) maintained by the subscriber's home network. This HLR database is usually colocated with the mobile switching center (MSC) and also stores the user's profile, which includes permanent data about subscribers, including call plan supplementary services, location information, and authentication parameters. Another database known as the visitor location register (VLR) is also maintained by the home network and also usually colocated with the MSC (MSC/VLR). The home VLR will temporarily store information about any MS that has registered itself with the home network. Therefore, if an MS is turned on by a subscriber in the user's home network area, the home VLR will temporarily store that user's information.

Within a particular network there are usually several to many MSCs used to support the network's operation. Depending upon the particular mobile network topology, each MSC may contain HLR and VLR database functions or, alternately, single HLRs (configured as an MSC/HLR/VLR) might service a group of MSC/VLRs (see Figure 4–15). For a small system another possibility is that a Gateway MSC (GMSC) might house the HLR function for a group of integrated MSC/VLRs. A gateway MSC is an MSC that interfaces the mobile network with other networks such as the PSTN.

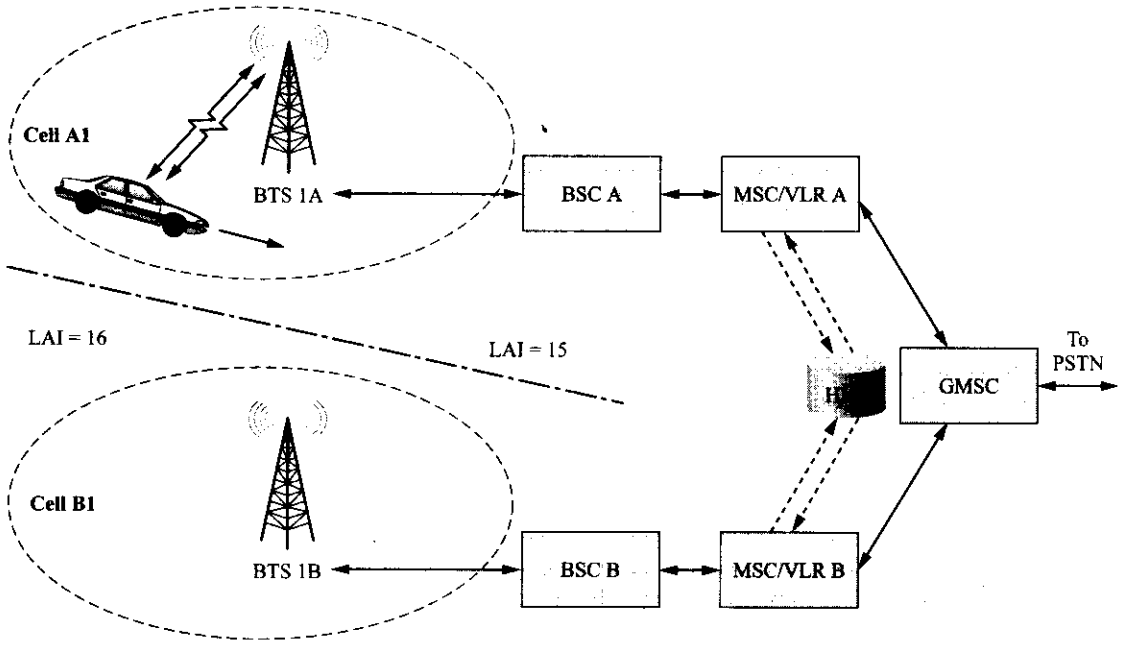


Figure 4–15 A typical cellular system.

At this time, let us examine several possible scenarios that could occur during the operation of a wireless network. The first possibility has already been mentioned; the user turns on a mobile within his or her home area. The mobile registers with the VLR for the home area. The colocated or system HLR confirms that the

subscriber has network privileges. Communication between a remote HLR and MSC/VLR occurs using a particular signaling protocol over an SS7 network. The second case would occur when the user is away from his or her home location. Now the mobile registers with the VLR of another MSC or a "foreign" network. The first possibility refers to the fact that the subscriber is still connecting to his or her own service provider's network but a different MSC/VLR is covering the area where the subscriber is now located. Whereas, a foreign network belongs to a different service provider (this type of connection is called roaming). In these situations, the MSC/VLR must send a message to the subscriber's HLR to verify authentication information about the mobile. The HLR will respond to the request by transmitting the information back to the requesting MSC/VLR over the SS7 signaling network.

A few comments about the communications between MSC/VLRs and HLRs are appropriate here. For a GSM cellular system and most other modern systems, the SS7 system is used to communicate these messages. The signaling done over this network is accomplished using message transfer part (MTP) as the common platform and signaling connection control part (SCCP) to provide the additional functionality to connect network databases (HLRs and MSC/VLRs) without any speech connection occurring during this operation. More detail about these operations will be given elsewhere in this text.

Handoff Management

It addition to the location management functions already described, a cellular system needs to be able to track the location of a subscriber as that subscriber moves within a coverage area and to be able to maintain the subscriber's connection to the system. If the subscriber moves from one cell to another, the cellular system must have the ability to reconfigure the connection to the mobile from the current base station to the new BS in the new cell. This connection handover process is known as **handoff.**

For first-generation cellular systems, the handoff process for voice calls could cause a noticeable interruption of the conversation (a hard handoff) and in some severe cases dropped calls. Second-generation cellular systems using digital technology have mitigated some of these problems with seamless handoffs, and CDMA systems have incorporated soft handoffs into their systems thus all but eliminating interrupted calls. For data transmissions, handoff can result in dropped packets, but this is not as severe a problem for bursty or packet data traffic since this type of traffic only needs intermittent connectivity and retransmission can be employed to counteract lost packets.

As shown in Figure 4–16, handoff basically consists of a two-step process. First, a handoff management algorithm determines that handoff is required and initiates the process. The second step consists of actually physically restructuring the connection and then updating the network databases about the new connection and location of the MS. For the handoff process to be successful the network elements involved in the delivery of either voice or data services to the mobile must be aware of all changes to the mobile's point of access. On the air interface side of the system, the former serving point has to be informed about the change or dissociation of the mobile while the mobile is reassociated with the system through the new serving point. On the network side, the various databases must be updated to reflect the correct location of the MS. This is all necessary for the correct routing of data packets or voice calls. The next sections will provide more detail about these operations.

Handoff Control

The algorithm used to determine when to make a handoff can be located in a network element or in a mobile terminal. For cellular systems the network controls the handoff for voice calls and this is known as network-controlled handoff or NCHO. If the mobile terminal controls the handoff, this is known as mobile-controlled handoff or MCHO, and if information supplied by the mobile helps determine when handoff should occur, this is known as mobile-assisted handoff or MAHO. In all cases, the handoff-controlling entity uses some particular algorithm that employs various measures of system performance to make a decision about the need for handoff.

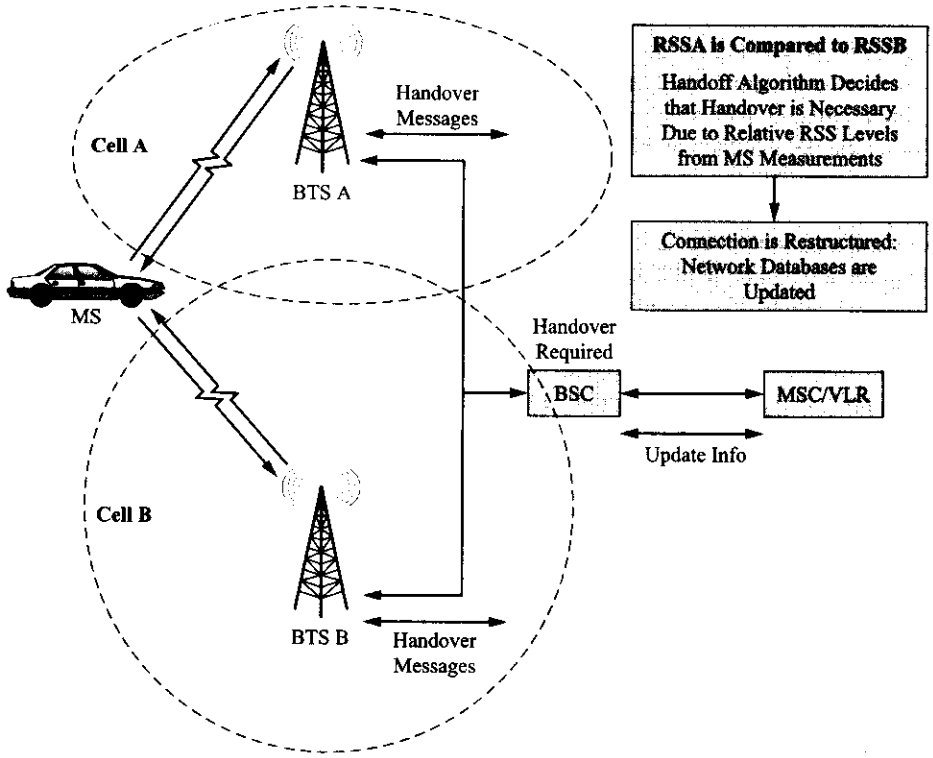


Figure 4–16 Typical cellular handoff operations.

The most common measurement used in this process is the received signal strength (RSS) from the mobile's point of attachment and the RSS of the nearest other possible points of attachment (i.e., radio base stations in adjacent cells). Other associated measurements that might be included in the process are system path loss, carrier- and signal-to-interference ratios and measures of bit error rate (BER), symbol or block error rate, and so on. A problem with using signal-strength measurements is that received signal strength can undergo extreme fluctuations due to signal fading effects that are completely random in nature. Error rates are also similarly affected by the randomness of propagation conditions.

Traditional handoff algorithms would initiate handoffs when the power received from the current RBS dropped below that received by another nearby RBS. Additional fine-tuning of the algorithm has incorporated threshold levels and hysteresis to prevent erroneous handoff requests and to mitigate the ping-pong effect mentioned earlier. As an example, with both threshold and hysteresis, handoff will only occur if the received power from a nearby RBS is above that received from the current RBS by a certain hysteresis value and the power from the current RBS is also below a certain threshold power level. Figure 4–17 shows some examples of the possible different algorithms used for handoff decision making in conjunction with the signal power being received by the current RBS and the signal power from a RBS that the MS is approaching.

Cellular service provider engineers are continually fine-tuning system handoff algorithms to improve system performance. Measures of system performance might include such things as call blocking and call dropping probability, required time to complete a handoff, and system handoff rate. Although these performance measures are typically used to improve the delivery of voice calls and the efficiency of the network, they might not necessarily result in higher data throughput rates or provide for required QoS continuity during a handoff, all important issues in the delivery of data services.

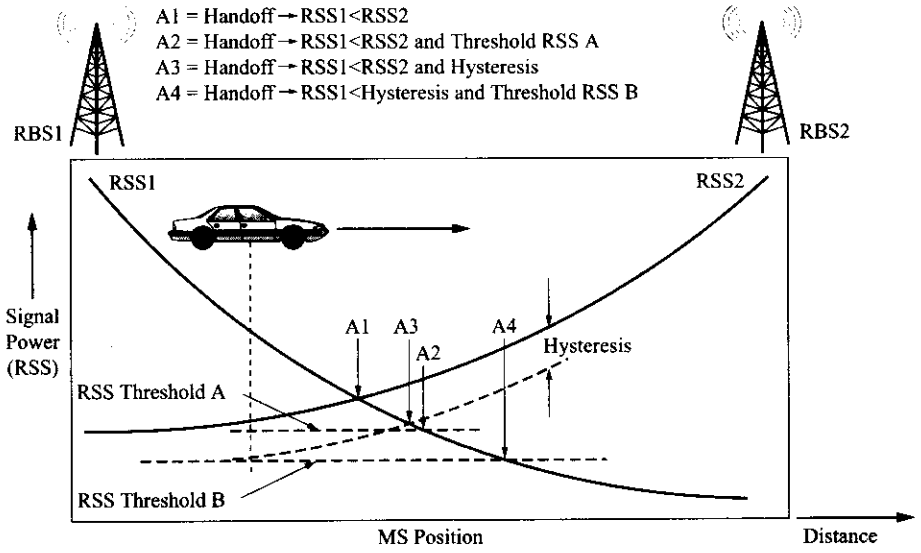


Figure 4–17 Typical handoff algorithms using RSS measurements.

Handoff Operations

Handoff management requires the transmission of messages between various network elements to facilitate the handoff process. As depicted in Figure 4–16, signal power levels being received by the current and handoff candidate radio base stations and the mobile station are first relayed to the radio base station and then to the base station controller (BSC). When these levels meet the criteria for a handoff, the process is initiated. A handoff message is sent to the mobile from the current radio base station that requests the mobile station register with a new radio base station that is also identified in the message. When the mobile performs this task, the MSC/VLR is updated to reflect the new mobile point of attachment (i.e., the new RBS) and any other changed system parameters. If the MSC/VLR most recently registered with is not the same as the last, then the new VLR must send an interrogation message to the home HLR to obtain the subscriber profile and authentication information. The HLR responds over the SS7 network with the authentication information. If the mobile is authenticated, then the new radio base station sends a message to the mobile assigning a new pair of traffic channels to the MS and the RBS for the continuation of a voice conversation. The HLR database is updated so that it knows where the mobile is and the new VLR database adds the new mobile to its list of subscriber terminals that are being serviced by the particular MSC/VLR. As a last act, the HLR sends a message to the old MSC/VLR to purge the mobile from its list of actively attached subscriber terminals. More detail about handover operations will be given in Chapter 5. Additionally, any data packets that were intended for delivery to the MS from the old MSC/VLR that may have been placed in a temporary network storage area should be either deleted or redirected to the new MS access point.

As one can see, there are many necessary message transfers occurring between wireless network elements and subsequent operations to be performed by these same elements for a successful mobile station handoff. There are also other types of possible handoffs that have not been addressed here such as the various types of intracell or intra-BSC handoffs. Since the exact details of mobility management procedures for different cellular systems are specific to those systems, more details will be provided about these topics in later chapters when individual systems (GSM, TDMA, and CDMA) are covered.

4.6 RADIO RESOURCES AND POWER MANAGEMENT

The efficient use of radio resources and the need for power management has already been mentioned several times in other sections of this text. At this time, some details pertaining to this topic will be offered to the reader. Recall that in a cellular system the use of many closely spaced low-power RBSs allows for frequency reuse and hence increased system capacity. At the same time, the closer the spacing of the RBSs the greater the interference produced by both the subscriber MSs and the RBSs with other MSs and RBSs in both adjacent cells and cells using the same channels.

The use of power control algorithms for the adjustment of both the MS output power and RBS output power allow for nearly constant received signal strength at both the MS and RBS receivers. This use of power control provides several system advantages: the amount of cochannel interference is reduced, the risk of signal coupler saturation is reduced at the RBS, and the power consumption of the MS is reduced. This last advantage has additional ramifications in the reduction of battery requirements, which translates to longer time between charging and lighter and smaller mobile terminals.

Additionally, other power saving schemes are also being employed by the MS to conserve battery life, new energy-efficient designs for both hardware and software are being implemented, and radio resource management is being used to enable an MS or wireless network to optimize the use of the available radio resources. These topics will be discussed further in the next sections.

Power Control

As stated previously, cochannel interference is the limiting factor for the reduction of cluster or frequency reuse size, N. The use of power control algorithms for the output power of the MS and the RBS allows the system to use the lowest possible output powers to achieve the minimum S/I ratio that can be tolerated and still provide good-quality communications. This means that for an MS close to the RBS both devices may lower their output power and as the mobile moves farther from the RBS both devices will in all likelihood have to increase their output power. Any reduction in output power from the nominal design power for the RBS or MS will produce a reduced amount of cochannel and adjacent channel interference for other cells using the same frequency channels.

Since the power output of both the RBS and the MS must be constantly adjusted due to the numerous changes in signal strength caused by fading and any motion of the mobile, several different methods of power control can be employed in a wireless network.

One typical system algorithm for power control usually consists of two phases. The first phase occurs when the MS initially registers with the system upon power-up. In this phase, the MS uses the nominal (maximum) power output allowed by the system. The first measurements of signal strength made by the RBS are used by the BSC to determine a valve of reduced MS output power. Power control messages are quickly sent to the MS to reduce its output power; however, this first power reduction is usually limited to avoid the possibility of a dropped call. In the second phase of this process, additional measurements are made and the MS power is adjusted as needed. The power output of the RBS is also adjusted on a case-by-case basis to yield the required signal strength at the MS. In this situation, whenever a new connection is made, the RBS initially transmits with its nominal or maximum output power. As done with the MS, the output power of the RBS is quickly reduced to a point where more stable measurements can be made and then the power control algorithm adjusts the output power as needed. If the mobile is operating in the discontinuous transmission mode, the algorithm must be modified to take this fact into account.

Another possible power control method employs a complex algorithm that uses information about all the active radio links in a system to adjust the output powers of all the RBSs and MSs to achieve maximum but also equal S/I ratios for all radio links. In each of these systems, output powers are usually adjustable in incremental steps of 2 dB or less.

Power Saving Schemes

In addition to the power saving schemes outlined in the previous section, there are several other ways in which MS battery power may be conserved. It is well known that the mobile consumes the greatest amount of power during the transmission of a signal to the RBS. Less power is consumed during the reception of a signal from the RBS. Another mode of MS operation can exist and that is known as "standby." In a standby mode, much less power is consumed by the mobile than in either the transmission or reception mode. There are several techniques used with mobile stations to achieve standby status.

Discontinuous Transmission

Using speech detection methods, a mobile may be programmed to only transmit when there is speech activity by the user. The radio base station sets a discontinuous transmission (DTX) bit to either permit or disallow this mode of operation and includes it in an overhead message to the mobile during initial registration by the mobile. Just using straight speech detection methods can cause problems due to the unnatural resulting sound of the system as perceived by the users. To compensate for this, a low-power background or comfort noise signal is generated by the mobile receiver during gaps of silence or no speech activity. This operation is also repeated at the base station controller or TRC for the benefit of the calling party.

Sleep Modes

Another common technique used to save MS battery power is to put the MS into a sleep mode when there are periods of no activity. For this scheme, the RF portion of the mobile's circuitry is powered off while waiting between messages. The mobile will periodically awaken and read control channel messages from the system so as to not miss a paging message but with much less overall power consumption.

Energy-Efficient Designs

The use of the most power-efficient semiconductor technologies is normally a given in the design of cellular mobile stations. Additional power saving can be achieved through the use of power-efficient modulation and coding schemes. However, another area that can provide power efficiencies is in the design of the protocols used in a wireless network and in the software design employed by the MS itself. As the cellular world evolves toward universal 3G deployment, system designers are implementing these protocol- and software-based power saving ideas and designs into new systems. As digital signal processor (DSP) technology advances, the eventual use of software radios will be another step in the evolution of lower-power, reconfigurable, advanced wireless radio systems that can last longer between battery recharges.

Radio Resource Management

Radio resource management is used to provide several functional improvements and necessary operations to permit the correct operation of a wireless network. The first and most important aspect of radio resource management is to implement system power control that reduces interference and therefore allows for system capacity to be maximized. As pointed out before, a side benefit of this control function is the increase in MS battery life. Another improvement afforded to the system is that the MS is directed toward the best radio channel connection available to it within the cell. This is made possible by the constant transmission of measurement information from the MS to the BS and then to the BSC. Finally, the use of a wireless network radio resource management scheme enables the handoff operation. Without this network management function, handoff could not operate as seamlessly and efficiently as it does in today's systems. More details of radio resource management functions and organization used by particular radio systems will be presented in the chapters devoted to the individual systems.

4.7 WIRELESS NETWORK SECURITY

Unlike wireline telecommunications systems that usually provide some modest amount of security through infrastructure design and physical installation, wireless technologies pose special security concerns. The unguided nature of wireless signals exposes them to the possibility of undesired interference and interception. This section will present some of the security requirements for both the air interface and the fixed infrastructure of the wireless network itself and conclude with a brief overview of present wireless security techniques.

Wireless Network Security Requirements

Just as the wireline telecommunications networks require increasingly more effective security in this post-9/11 world, the security requirements of wireless networks are very similar to their wired counterparts. The need for privacy in the transmission of a voice conversation is necessary regardless of the means used to deliver the signal. The ability for anyone to attain the unauthorized interception of a private conversation is not a desired feature of any telecommunications system. The newer digital cellular systems make any interception of voice conversations extremely difficult due to the conversion of the voice signals from analog to digital form and the ciphering of the transmitted digital information by the system. However, with the increasing use of wireless data services and e-commerce activities, the need for more secure wireless networks is becoming more important as more wireless cellular users avail themselves of these new data services and more sensitive economic information is transmitted over wireless networks.

In addition to the transmission of voice or data traffic over the air interface, a certain amount of sensitive control and identification information is transmitted over control channels to the fixed wireless network. There is certainly a potential for the misuse of this type of information and the possibility of someone obtaining telecommunication services (teleservices) fraudulently as happened with the first-generation analog cellular system.

Presently the GSM Association maintains a global central equipment identity register (CEIR) database in Dublin, Ireland, of all handsets that have been approved for use on GSM networks. The database categorizes these approved handsets as being on a White List. There is also a CEIR Black List of handsets that should be denied access to the network due to being reported as either lost or stolen or otherwise unsuitable for use. GSM cellular operators that employ an EIR in their network use it to keep track of handsets to be blocked. If they are also registered users of the CEIR, they call in daily to share their database with the CEIR, and each day the CEIR creates a master Black List that the operator can download the following day. In this way any stolen or lost handset is blocked by the next day after it has been reported missing.

Network Security Requirements

In addition to the privacy and fraud concerns for the air interface portion of the network, there are also security issues involving the fixed portion of the network. The fixed wireless network infrastructure includes numerous network elements that are involved in identification, authentication, billing functions, and so on. However, most of these network elements and the transmission facilities between them enjoy the same level of physical security (or lack thereof) as the traditional PSTN or PDN telecommunications systems. As the wireless cellular system transitions to an all-IP network there will be a need to employ increased security measures to prevent hacking of the system and the possible infection of system components by software viruses. After the events of 9/11, the threat of terrorism is all too real and one can no longer discount any type of infrastructure target as being unrealistic.

Network Security

There are several methods by which the security of air interface messages can be enhanced. The most viable method is to employ encryption techniques. These techniques rely on the scrambling of the message using a

particular key to perform the encryption. Various encryption techniques have been used since the need for confidentiality first arose. Most encryption techniques are known as secret-key algorithms since the key to the encryption is kept secret from everyone but the two end users of the communications channel.

However, as complex as one can make the encryption process it seems that it is always possible to break the code given enough computational power and time. The field of telecommunications infrastructure security is a very hot research topic right now with the reality of a proliferation of attacks on the Internet as well as the high threat of global terrorism. As with many of the topics discussed to this point, security details of particular wireless systems will be presented with the particular technology.

Security issues concerning wireless LANs will be presented in the chapters addressing IEEE 802.XX wireless technology.

QUESTIONS AND PROBLEMS

1. What factors determine frequency reuse distance?
2. What advantage does the use of a cellular architecture provide?
3. What factors limit cell size?
4. A cell tower located near an interstate highway would most likely provide service to what type (size) of cell?
5. Determine the frequency reuse distance for a cell radius of twenty kilometers and a cluster size of 7.
6. Determine the frequency reuse distance for a cell radius of two kilometers and a cluster size of 4.
7. Construct a chart that shows how a cellular system with a cluster size of 4 could have twenty-eight channels assigned to the system in such a manner as to maximize channel spacing.
8. For a particular radio transmission technology, a minimum S/I ratio of 15 dB is needed for proper operation. What is the minimum required cluster size?
9. What will be the resulting (ideal) increase in cellular system capacity for a typical cell splitting scheme?
10. For a cell splitting scenario, why must the cell transmit power be reduced?
11. How is cell splitting different from cell sectoring?
12. What possible limitations can you conceive that would impose a practical limit on cell sectoring?
13. What is the driving force for the adoption of microwave cellular backhaul networks?
14. What has been the traditional method used to provide connectivity between the cellular network and the PSTN?
15. If and when the all-IP core network becomes a reality, how will voice traffic be carried to the cellular network?
16. Mobility management consists of several basic functions. What are they?
17. When does the location updating function occur?
18. What two basic operations occur during the handoff process?
19. Why is power management so important for cellular wireless systems?
20. Describe the process of power control used by cellular systems.
21. What is meant by the term *discontinuous transmission* in the context of wireless cellular systems?
22. What is meant by the term *sleep mode* in the context of wireless cellular systems?
23. Describe how the GSM Association provides a form of security to its members.
24. What is the basic form of security employed by cellular wireless systems?
25. Describe secret-key encryption.

CHAPTER 5

GSM and TDMA Technology

Upon completion of this chapter, the student should be able to:

- Discuss the basic services offered by GSM cellular and the frequency bands of operation.
- Discuss the network components of a GSM system and the basic functions of the mobile station, base station system, and network switching system.
- Explain the concept of GSM network interfaces and protocols, and their relationship to the OSI model.
- Explain the GSM channel concept.
- Discuss the functions of the GSM logical channels.
- Explain the TDMA concept and how it is implemented in GSM.
- Explain the mapping of logical channels on to the GSM physical channels.
- Discuss the various GSM identities.
- Explain the GSM operations of call setup, location updating, and handover.
- Discuss the GSM operations that occur over the Um interface.

This chapter provides a detailed description of the GSM wireless cellular telephone system and the time division multiple access (TDMA) technology used to implement the air interface portion of the system. GSM cellular is by far the most popular wireless system in the world with over one billion subscribers. Because of this popularity, this chapter presents an in-depth explanation of the architecture of this system and the access technology used to implement it. Because of the amount of detail included in this chapter, the chapter has been organized into three parts: an overview of GSM, GSM network operations, and other TDMA systems.

Part I coverage starts with a short prologue to the evolution of GSM and the rationale behind its introduction. The GSM frequency bands are introduced and the channel numbering system is explained. Next, the network components that compose a GSM system are introduced and their functions are described in detail. How these subsystem components are interconnected and the messages that are sent between them are looked at from several different viewpoints. The GSM standards specify various system interfaces that are introduced to the reader along with the protocols used by the subsystems to deliver the messages and commands needed for overall system operation. The OSI model is used extensively to frame the theory of GSM operation.

Next, the GSM channel concept is introduced with descriptions of the various logical channels and their function. How the system uses time division multiplexing to provide a means by which system commands, messages, and traffic can be transmitted over the air interface during selected timeslots is examined in

detail. Several examples of possible TDMA frame timing schemes are presented to give the reader a feel for the complexity of the system and the number of operations needed to make the system functional.

Part II of the chapter reviews GSM system identifiers before a detailed coverage of GSM traffic cases is started. The three basic operations needed to support a subscriber's mobility within a wireless cellular system—call setup, location updating, and call handover—are now treated from the viewpoint of the interactions between subsystem components through command and message transmissions over the interfaces specified in the GSM standards. The last portion of this section takes the reader a step closer to the networking aspect of GSM system operation by examining typical system management functions in the context of the OSI model.

In Part III, the last section of this chapter introduces NA-TDMA, a cellular technology very similar to GSM but not compatible with it. Because of the amount of detail already provided about GSM, this topic is dealt with in a fairly superficial manner by simply indicating the system similarities and differences. This chapter does not cover the operations needed for high-speed packet data transmission over a GSM or NA-TDMA network. Discussion of that topic is delayed until Chapter 7.

PART I GSM SYSTEM OVERVIEW

5.1 INTRODUCTION TO GSM AND TDMA

As discussed in prior chapters, the GSM system evolved due to a desire by the European countries to develop a pan-European system that would allow roaming on an international basis. At the time, digital technology and microelectronics had advanced sufficiently to allow for the development of an entirely digital second-generation cellular system. Other TDMA digital cellular standards such as North American IS-136 are very similar to GSM. The GSM standards, as published by the ETSI, includes specifications for the air interface portion of the system as well as the fixed network infrastructure used to support the services offered over the wireless network. The GSM standards may be downloaded from www.etsi.org.

In 1982, the frequency bands of 890–915 MHz and 935–960 MHz were allocated for a pan-European second-generation digital cellular system (GSM 900) that would replace the incompatible first-generation systems that were already in existence in different countries. The allocation of the frequency bands was only the first step in this process. An international task force was also assembled during 1982 and by 1987 GSM was formally adopted by the European Commission. The ETSI took over development in 1989 and published the standards for the first phase of GSM in 1990. The development process continued, resulting in the deployment of a functional system in 1992. A new frequency band in the 1800-MHz range was added worldwide for what was originally named digital cellular system (DCS 1800). This upbanded version of GSM 900 was renamed GSM 1800 in 1997. GSM service in the 1900-MHz range (GSM 1900) using the PCS bands in the United States has been deployed recently. Also, the implementation of additional GSM services offered under Phase 2 and Phase 2+ of GSM has been an ongoing process and continues today under the direction of the ETSI. Today, the GSM system is by far the most popular cellular wireless system in the world.

GSM Services

The first-generation analog cellular systems were designed for basic voice service. Data services for fax or circuit-switched data transmission using a voiceband modem were classified as "overlay" services that run on top of the voice service. The second-generation GSM cellular system was designed to be an integrated wireless voice-data service network that offered several other services beyond just voice telephone service. The types of services to be offered over the GSM network were classified into two categories: teleservices and bearer services (see Figure 5–1). In addition, there are supplementary services that can be added to the teleservices.

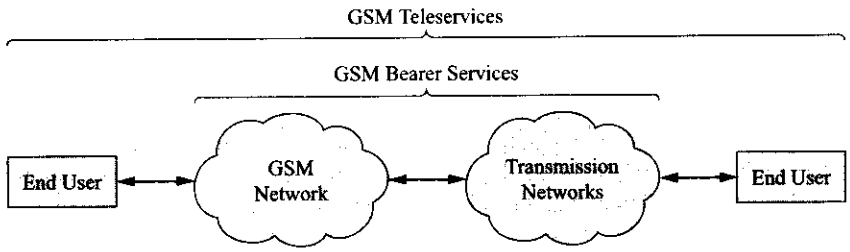


Figure 5–1 Relationship of teleservices and bearer services to the GSM system (Courtesy of ETSI).

Teleservices provide standard voice communications between two end users and additional communications between two end user applications according to some standard protocol. **Bearer services** provide the user with the ability to transmit data between user network interfaces. Supplementary services are services that enhance or support a teleservice provided by the network.

The planning of GSM system development and deployment called for the implementation of system services to be carried out in two phases. In the first phase, the GSM services offered were as shown in Table 5–1. In the second phase of GSM implementation, the service offerings would be expanded to include those shown in Table 5–2. Presently, the development of GSM system services has evolved into Phase 2+. Phase 2+ is primarily focused on the addition of high-speed packet data services to GSM. This initiative is embodied in general packet radio service (GPRS) and enhanced data rates for global evolution (EDGE). These topics will be discussed in more detail in Chapter 7.

Table 5–1 Phase 1 GSM services (Courtesy of ETSI).

Service Category	*Service*	*Additional Details*
GSM Teleservices	Telephony Emergency calls Short Message Service Videotext access Teletex, FAX, etc.	Full rate at 13 kbps voice "112" is GSM-wide emergency number Point-to-point (between two users) and cell broadcast types
GSM Bearer Services	Asynchronous data Synchronous data Synchronous packet data Others	300–9600 bps (transparent/nontransparent) 2400–9600 bps transparent
Supplementary Services	Call forwarding Call barring	All calls, when the subscriber is not available Outgoing calls with specifications

GSM Radio Frequency Carriers

For GSM cellular systems the air interface consists of channels that have a frequency separation of 200 kHz. For the three most widely used frequency bands devoted to GSM system operation this channel spacing yields a different total number of carrier frequencies per band. The GSM 900 band has 124 carrier frequencies, the GSM 1800 band has 374 carrier frequencies, and the GSM 1900 band has 299 carrier frequencies. Since each carrier can be shared by up to eight users, the total number of channels for each system is:

$$124 \times 8 = 992 \text{ channels for GSM 900}$$

Table 5–2 Phase 2 GSM services (Courtesy of ETSI).

Service Category	*Service*	*Additional Details*
GSM Teleservices	Half-rate speech coder Enhanced full rate	Optional implementation
Supplementary Services	Calling line identification Connected line identification Call waiting Call hold Multiparty communications Closed user group Advice of charge Operator determined call barring	Presentation or restriction of displaying the caller's ID Presentation or restriction of displaying the called ID Incoming call during current conversation Put current call on hold to answer another Up to five ongoing calls can be included in one conversation Restriction of certain features from individual subscribers by operator

$374 \times 8 = 2992$ channels for GSM 1800

$299 \times 8 = 2392$ channels for GSM 1900/PCS 1900

The frequency bands allocated to the five present GSM system implementations are shown in Table 5–3. The channels have absolute radio frequency channel numbers (ARFCNs) associated with them and are numbered as 1–124, 259–293, 306–340, 512–885, and 512–810 for Primary GSM 900 (P-GSM 900), GSM 450, GSM 480, GSM 1800, and GSM 1900/PCS 1900, respectively. Also note that Extended GSM 900 (E-GSM 900) and Railways GSM 900 (R-GSM 900) have added channels 975–1023 and 955–1023,

Table 5–3 GSM frequency bands and channel numbers (Courtesy of 3GPP).

GSM Band	*Uplink Frequency*	*Downlink Frequency*
P-GSM 900 ARFCN=1...124	890 - 915 MHz (ARFCN-1) × 0.2 MHz + 890.2 MHz	935 - 960 MHz Uplink frequency + 45 MHz
E-GSM 900 ARFCN=975...1023	880 - 890 MHz (ARFCN = 0 = 890 MHz) (ARFCN-975) × 0.2 MHz + 890 MHz	925 - 935 MHz Uplink frequency + 45 MHz
R-GSM 900 ARFCN=955...1023	876 - 890 MHz (ARFCN-1023) × 0.2 MHz + 890 MHz	921 - 935 MHz Uplink frequency + 45 MHz
GSM 1800 ARFCN=512...885	1710 - 1785 MHz (ARFCN-512) × 0.2 MHz + 1710.2 MHz	1805 - 1880 MHz Uplink frequency + 95 MHz
GSM 1900 ARFCN=512...810	1850 - 1910 MHz (ARFCN-512) × 0.2 MHz + 1850.2 MHz	1930 - 1990 MHz Uplink frequency + 90 MHz
GSM 450 ARFCN=259...293	450.4 - 457.6 MHz (ARFCN-259) × 0.2 MHz + 450.6 MHz	460.4 - 467.6 MHz Uplink frequency + 10 MHz
GSM 480 ARFCN=306...340	478.8 - 486 MHz (ARFCN-306) × 0.2 MHz + 478.8 MHz	488.8 - 496 MHz Uplink frequency + 10 MHz

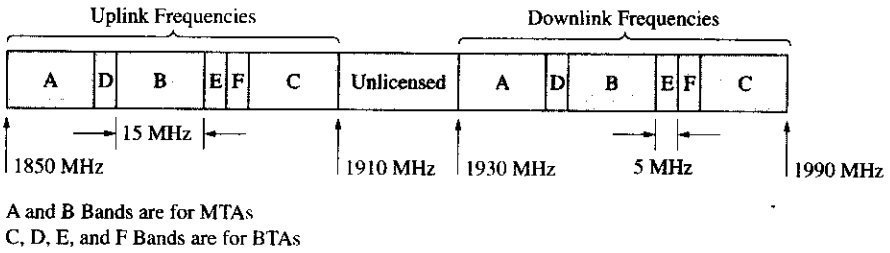


Figure 5–2 GSM frequency allocations in the 1900-MHz PCS bands.

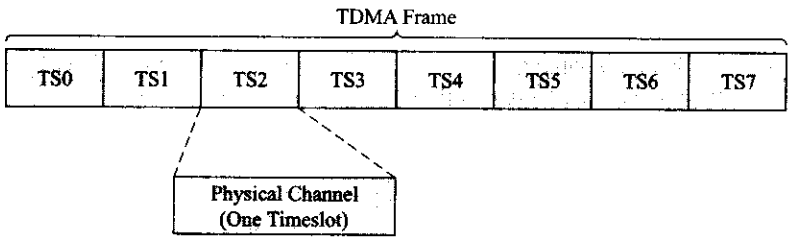


Figure 5–3 GSM timeslot in a TDMA frame.

respectively. Figure 5–2 shows some additional details about the bands within the PCS spectrum allocation that are used by the GSM 1900 system in the United States. As shown by Figure 5–2, the various bands are allocated for use in either major or basic trading areas (MTA and BTA). The A, B, and C bands are each 15-MHz wide and the D, E, and F bands are each 5-MHz wide. The reader should note that there is also limited usage of other bands for GSM at 450 and 850 MHz.

For a particular carrier frequency, a channel consists of a single timeslot that occurs during a TDMA frame of eight timeslots (see Figure 5–3). Each of these timeslots represents a physical channel. Therefore, each GSM TDMA frame represents eight physical channels. Furthermore, besides voice and data traffic there are a host of different system messages and other overhead information constantly being transmitted between the base transceiver station (BTS) and the MS.

5.2 GSM NETWORK AND SYSTEM ARCHITECTURE

Figure 5–4 shows the basic system architecture for a GSM wireless cellular network. As can be seen from the figure, the major GSM subsystems are the network switching system (NSS), the base station system (BSS), and the mobile station (MS). Most of these wireless network subsystems and their components have been discussed previously in Chapter 3 as the common components of cellular systems. Contained within the description of these components was a brief overview of their function and relationship to the other components in the wireless system. This section will provide a brief review of the description of the common components and their system functions. Components that are specific to GSM systems or not previously introduced to the reader will receive more complete coverage.

Mobile Station

The mobile station (MS) is the device that provides the radio link between the GSM subscriber and the wireless mobile network. In the GSM system, the MS provides subscribers the means to control their access to the PSTN and PDN and also to facilitate their mobility once connected to the network. The MS is a multifunctional system with a fairly large amount of signal and data processing power. It is constantly monitoring messages being broadcast from the base transceiver system (BTS) to support the setup and clearing of radio

channels used for the transmission of various forms of subscriber traffic. In addition, the MS is constantly performing power and bit error rate (BER) measurements on signals being received from the BTS that it is attached to and the neighboring BTSs in the MS's general vicinity. These measurements, in conjunction with the handover (handover is the term used by the GSM standard) algorithms performed by the BSS, support the MS's mobility as the subscriber moves about the GSM network.

The GSM system also makes use of a **subscriber identity module** or SIM card that when inserted into the MS makes it functional (the MS can only make emergency calls without the SIM card). The SIM is a smart card that is issued to the subscriber when the subscriber signs up for service with the wireless network operator. Besides containing information about the types of service available to the subscriber, the card contains the subscriber's IMSI number, the mobile MSISDN number, a SIM personal identification number (PIN), security/authentication parameters, and address book contact information (i.e., names and numbers) stored by the subscriber. The SIM card also stores SMS messages that the subscriber receives and saves. The SIM card allows for some unique possibilities for GSM subscribers. A single GSM phone can be shared by several users with different SIM cards or a subscriber could visit other countries and purchase a country-specific SIM card for use with a single GSM mobile that was carried by the subscriber.

In the GSM standard, the MS consists of two elements: the mobile equipment (ME), which is the physical phone itself, and the SIM card. The mobile is constantly being redesigned to incorporate new features and different form factors (mobile size, screen size, etc.) that the public is perceived to desire. Today, the newest mobile phones contain several video cameras with which the subscribers can use to send pictures or short video clips to each other or use as a videophone. Traditionally, the service providers have subsidized the cost of the rather expensive electronics incorporated into the mobiles to encourage more users to subscribe to the wireless services that they offer.

Base Station System

The base station system (BSS) is the link between the MS and the GSM mobile-services switching center (MSC). The BSS consists of two elements: a base transceiver system (BTS) and the base station controller (BSC). The BTS communicates with the MS over the air interface using various protocols designed for the wireless channel. The BSC communicates with the MSC through the use of standard wireline protocols. The BSC and BTS communicate with each other using **LAPD protocol,** which is a data link protocol used in ISDN. In essence, the BSS provides a translation mechanism between the wireline protocols used in the fixed portion of the wireless network and the radio link protocols used for the wireless portion of the network.

Today, the two elements of the BSS may be physically implemented by either two or three hardware systems depending upon the GSM hardware vendor. The BTS (often called a radio base station or RBS) is the BSS air interface device that corresponds to the subscriber's MS. It provides the radio link to the MS over the air interface. The usual basic components of the BTS are radio transceiver units, a switching and distribution unit, RF power combining and distribution units, an environmental control unit, a power system, and a processing and database storage unit. The BTS is physically located near the antenna for the cell site. Radio base station is the term usually used to describe the cellular radio transmitting and receiving equipment located at the cell site. Typically, an RBS may consist of three BTSs that service a standard sectorized cell site.

The functional elements needed by a base station controller to implement its operations may be all located in a single physical unit or split out into several separate units. The basic BSC components are input and output interface multiplexers, a timeslot interchange group switch, a subrate switch, speech coder/decoders, transcoders and rate adaptors, SS7 signaling points, environment control units, power supply and power distribution units, and various signal and control processors. As mentioned, the transcoder and rate adaptation unit is sometimes split out from the BSC to be a stand-alone unit that is known as a transcoder controller (TRC). Some system economies for suburban and rural areas can be gained through the use of separate BSCs and a shared transcoder controller. Urban and heavy-traffic areas are best served

by a combined BSC/TRC. Chapter 8 will provide more details about these BSS hardware elements and their operation.

Network Switching System

The wireless cellular network switching system (NSS) provides the necessary interface for the connection of the wireless network to other networks (i.e., the PSTN, PDN, and other wireless PLMNs). Additionally, it provides support for the mobility of the GSM subscriber within the GSM network. The switching system maintains databases that are used to store information about the system's subscribers and facilitate the connection of a mobile to the system as long as it has connection privileges. The GSM switching system was designed to communicate with the PSTN through ISDN protocols. The basic components of the network switching system include at least one mobile-services switching center (MSC), a gateway MSC, the visitor and home location registers, the equipment identity register, and the authentication center. In addition to these basic components, the switching system may also have a flexible numbering register and an interworking location register to provide more system functionality.

To handle short message service (SMS) the wireless switching system will need to have an SMS gateway MSC (SMS-GMSC) and an SMS-interworking MSC (SMS-IWMSC). The implementation of general packet radio service (GPRS) for high-speed data transmission and reception requires the use of two additional switching system elements: a serving GPRS support node (SGSN) and a gateway GPRS support node (GGSN). These last two units connect to IP networks and will be discussed along with the SMS elements in more detail in Chapter 7.

The MSC, in conjunction with several of the databases listed previously, performs the necessary telephony switching functions required to route incoming mobile-terminated telephone calls to the correct cell site and connect mobile-originated calls to the correct network (i.e., PSTN or PLMN). The MSC communicates with the PSTN and other MSCs using the SS7 protocol. The MSC that is connected to the PSTN is commonly referred to as the gateway MSC (GMSC). Additionally, the MSC is instrumental in the supervision and administration of mobility and connection management and authentication and encryption.

The GSM network switching system databases provide the wireless network with the necessary information to facilitate subscriber mobility. The visitor location register (VLR) is a temporary database used to hold information about mobile subscribers within the coverage area of a particular MSC. The temporary subscriber information contained in the VLR allows the MSC to provide service to the visiting mobile subscriber. Commonly, the MSC will be integrated with the VLR to create a combined MSC/VLR and hence reduce system signaling operations. For security reasons the VLR will assign a temporary mobile subscriber number (TMSI) to the visiting MS so as to avoid using the IMSI over the air interface. The home location register (HLR) database contains information about the subscriber's account. Commonly stored information will include such items as the MSISDN and IMSI numbers and types of services that have been subscribed to. Also included in the HLR database will be dynamic data such as the subscriber's current location (i.e., VLR address) and presently activated services. The HLR together with the VLR and the MSC provide support for the connection and mobility management of mobile stations either in their home location area or roaming within the GSM system. The authentication center (AUC) and the equipment identity register (EIR) in conjunction with the MSC/VLR and HLR provide additional GSM network security and help facilitate international roaming within the GSM network. The **flexible numbering register** (FNR) is used by the GSM system to provide number portability to a subscriber. With this feature a subscriber may change GSM operators and still maintain the same MSISDN number. The network switching system will use the FNR to redirect messages sent by a GMSC toward a particular HLR to the correct HLR. The interworking location register (ILR) is used to allow intersystem roaming. In the United States, this operation supports roaming between the legacy AMPS system and GSM 1900 system.

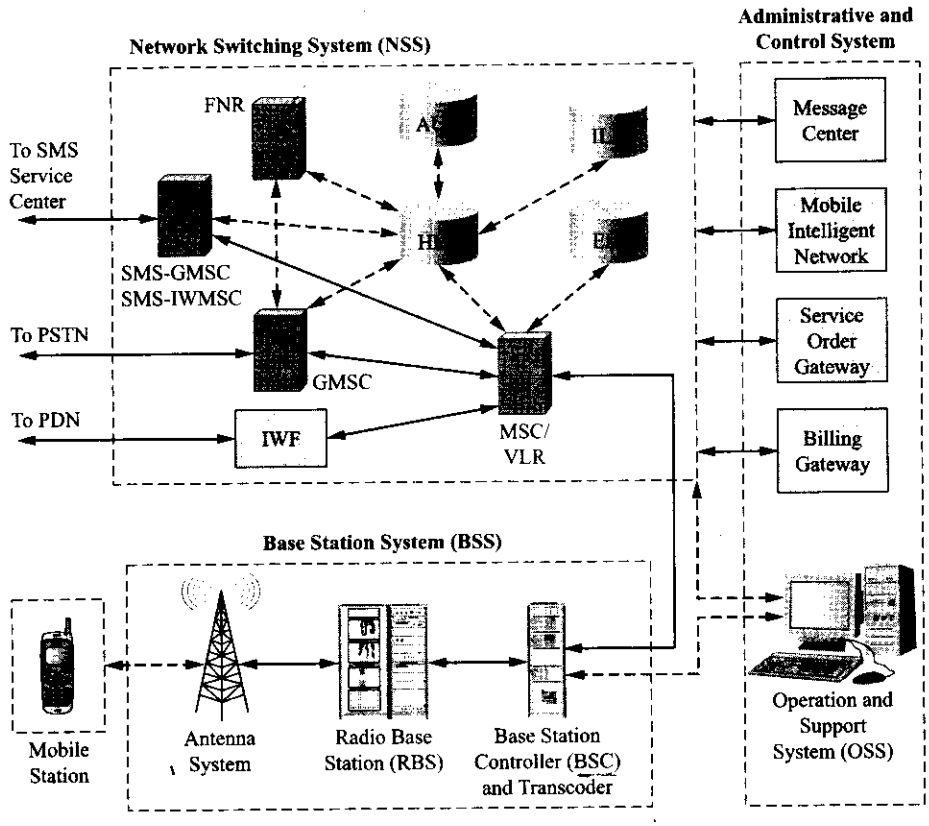


Figure 5–4 GSM network architecture.

Operation and Support System and Other Nodes

As shown by Figure 5–4, the entire GSM wireless network is monitored and controlled by an **operation and support system** (OSS) (the GSM standard refers to this functional entity as a operation and maintenance center). This centralized system can be used to provide surveillance of the complete network and thus provide the operator a means to support operation and maintenance of the entire network. Usually, there are several sublevels to the management functions that cover the circuit, packet, and radio network portions of the GSM network. The OSS software usually provides the system operator with the ability to perform configuration, performance evaluation, and security management of each portion of the wireless network along with the traditional display of alarms or fault indicators for specific system elements.

The other nodes shown in Figure 5–4 are used to interface the wireless network switching system with the operator's administrative computer systems and software. The titles of billing gateway and service order gateway are descriptive of the functions performed by these elements. The reader may refer back to Chapter 3 to review additional detail about these nodes.

GSM Network Interfaces and Protocols

The seven-layer OSI model was introduced in Chapter 1. At that time, the basics of electronic telecommunication protocols were also introduced in the context of the OSI model. Recall that a network protocol is an agreement on how to communicate between network elements or nodes. At this time, it will be instructive to take a brief look at the interfaces and protocols specified for use in the GSM system.

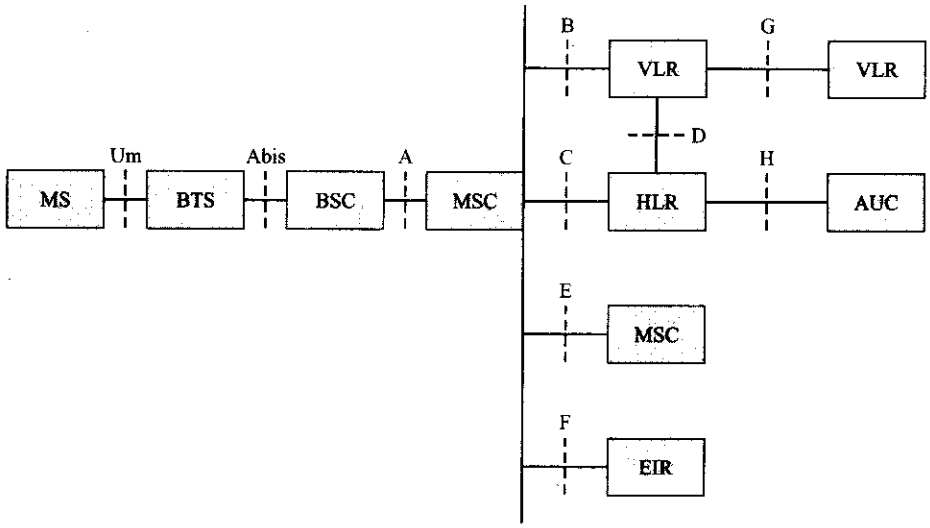


Figure 5–5 GSM network interfaces.

GSM Interfaces

The GSM standard specifies the various interfaces between the GSM elements. Figure 5–5 shows these GSM interfaces. As shown in the figure, the air interface between the MS and the BTS is the Um interface. The physical interface between the BTS and the BSC is known as the **Abis interface,** and the interface between the BSC and the MSC, is known as the A interface. The MSC has various interfaces between it and the other network switching system elements or other MSCs. Note that the interfaces for SMS and GPRS nodes will be discussed in Chapter 7.

Layered Structure/OSI Model Also, recall from Chapter 1 the layered structure of the OSI model (refer back to Figure 1–10). The OSI model views the communications between user application processes as being partitioned into self-contained layers that contain tasks that can be implemented independently of tasks in other layers. A message sent between two network nodes travels downward in the protocol stack of the sending node. As the message propagates through the layers, information is added to the original message at each layer. After transmission to the receiving network node, the message propagates upward through the receiving node protocol stack. At each layer the information added by the sending node is stripped off the message and analyzed by the corresponding peer layer (refer back to Figure 1–12) in the receiving node. The receiving layer is then able to offer various services to the higher layers within the receiving node. This model will be used to illustrate the operation and structure of the GSM system.

GSM Protocols and Signaling Model

Figure 5–6 shows a signaling model for the GSM system. As shown by the figure, the MS communicates with the MSC to provide system connection, mobility, and radio resource management by sending messages back and forth over the air interface from the MS to the BTS, between the BTS and the BSC, and between the BSC and the MSC. The figure indicates the various protocols that are used between the different GSM interfaces and at the different OSI layer levels. Additionally, the MSC communicates with the various networks that it is connected to (PSTN, PLMN, etc.) by using the various protocols shown in the figure. These operations will be briefly summarized in the next several sections and then explained in more detail in Section 5.6 of this chapter.

Um Interface The Layer 1, Um, air interface specifications will be detailed more extensively in Section 5.6 of this chapter and in Chapter 8. The Layer 2 protocol used on the Um interface is LAPDm, a modified

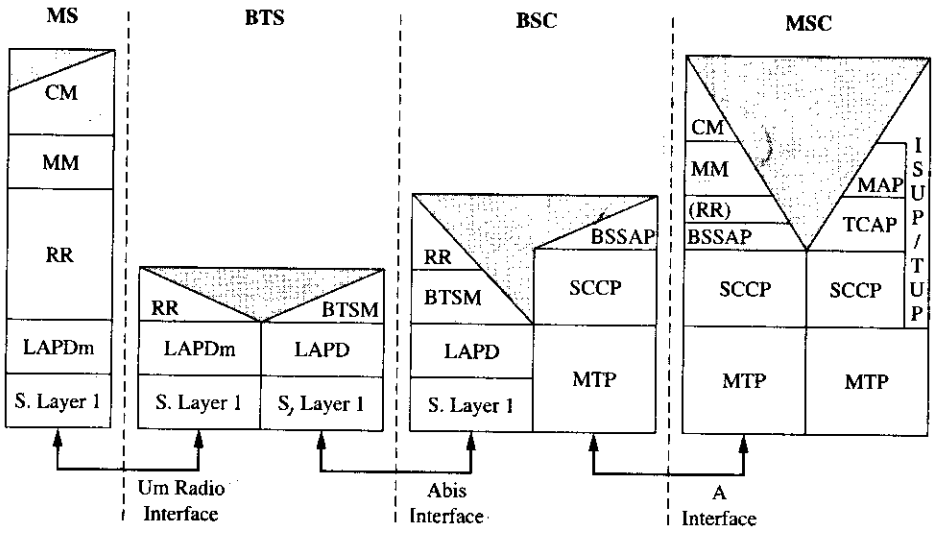


Figure 5–6 GSM signaling model.

version of the ISDN protocol LAPD. The major differences between LAPD and LAPDm protocol are the following: for LAPDm no error detection is employed since it has been built into Layer 1 signaling and LAPDm messages are segmented into shorter messages than LAPD to be compatible with the TDMA frame length used in GSM.

Abis Interface The Abis interface exists between the BSC and the BTS. The Layer 2 protocol used on the Abis interface is LAPD. At the Layer 3 level, most messages just past through the BTS transparently. However, there are some radio resource management messages that are closely linked to the system radio hardware that must be handled by the BTS. The BTS management (BTSM) entities manage these messages. An example of this type of radio resource message involves encryption. The ciphering message sends the cipher key, K_c, to the BTS and then the BTS sends the ciphering mode command to the MS. Abis Layer 1 signaling details will also be discussed further in Chapter 8.

A Interface The A interface exists between the BSC and the MSC. Signaling over the A interface is done according to base station signaling application part (BSSAP) using the network service part of SS7. In the MSC, in the direction of the MS, Layer 3 is subdivided into three parts: radio resource management (RR), mobility management (MM), and connection management (CM). More will be said about these sublayers in Section 5.6 of this chapter. As mentioned, the protocol used to transfer the CM and MM messages is BBSAP. The BBSAP protocol has two subparts: direct transfer application part (DTAP) and base station system management application part (BSSAMP). DTAP is used to send CM and MM messages between the MSC and the MS transparently through the BSS. BSSAMP is used to send messages between the MSC and the BSC. This operation is detailed in Figure 5–7.

Ater Interface The Ater interface only exists in GSM systems that have separate units for the transcoder controller and BSC (this is typical of some vendors' GSM equipment). Signaling between the BSC and the TRC is performed by the use of BSC/TRC application part (BTAP) protocol (BTAP is a vendor- [Ericsson] specific protocol) over the Ater interface. Figure 5–8 shows this type of operation. The figure indicates how BSSAP signaling is sent transparently through the TRC node. Ater Layer 1 signaling details will also be discussed further in Chapter 8.

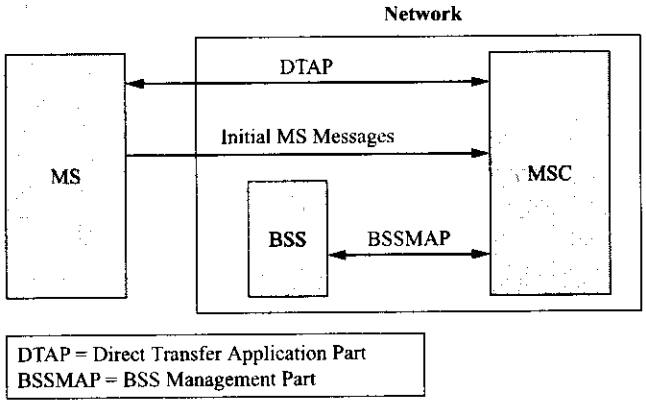


Figure 5–7 Signaling between the MSC, BSS, and MS in a GSM system.

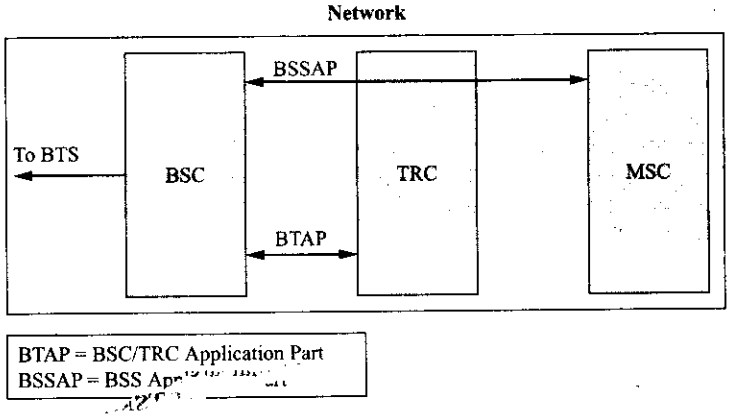


Figure 5–8 Signaling over the GSM Ater interface.

MSC Interfaces The GSM signaling model (Figure 5–6) shows two protocol stacks within the MSC node. The protocol stack on the left-hand side is associated with the A interface and has been discussed earlier. The right-hand protocol stack corresponds to the MSC network interfaces to the VLR, HLR, GMSC, and the PSTN or other PLMNs. Within the network interface stack are the following protocols: MTP, SCCP, TCAP, MAP, and ISUP/TUP. Message transfer part (MTP) is used to transport messages and for routing and addressing. MTP corresponds to OSI Layers 1, 2, and parts of 3. Signaling connection control part (SCCP) adds functions to SS7 signaling to provide for more extensive addressing and routing. Together, MTP and SCCP form the network service part (NSP) and correspond to Layers 1–3 in the OSI model. Transfer capabilities application part (TCAP) and mobile application part (MAP) are Layer 7 protocols. TCAP provides services based on connectionless network services. MAP is a protocol specifically designed for mobile communications. It is used for the signaling between databases (HLR, VLR, EIR, AUC, etc.) and is further designated as MAP-n where n is given as shown by Figure 5–5. ISDN-user part (ISDN-UP) and temporary user part (TUP) are used from Layer 3 up to Layer 7 and are used between the MSC and the ISDN/PSTN for call setup and supervision. More detail about these protocols and operations will be given later in this chapter.

5.3 GSM CHANNEL CONCEPT

As discussed in previous chapters, cellular telephone networks use various control and traffic channels to carry out the operations necessary to allow for the setup of a subscriber radio link for the transmission of

either a voice conversation or data and the subsequent system support for the subscriber's mobility. The GSM cellular system is based on the use of **time division multiple access** (TDMA) to provide additional user capacity over a limited amount of radio frequency spectrum. This is accomplished by dividing the air interface connection period into timeslots that can be used by different subscribers for voice or data traffic and also for the transmission of the required system signaling and control information. In essence, this process provides additional channels to the system over the same physical radio link.

As shown by Figure 5–9, the GSM system divides the radio link connection time into eight equal and repeating timeslots known as **frames** for both uplink and downlink transmissions. The timeslots can be considered logical channels. That is, from a system point of view, each timeslot may carry either subscriber traffic or signaling and control information required for the management of the radio link and other system resources. The system can use several different types of repeating frame structures known as **multiframes** depending upon the type of information being transmitted. The next several sections will provide more detail about the timeslots and the frame structure and the operations and the various functions performed by the signaling and control channels.

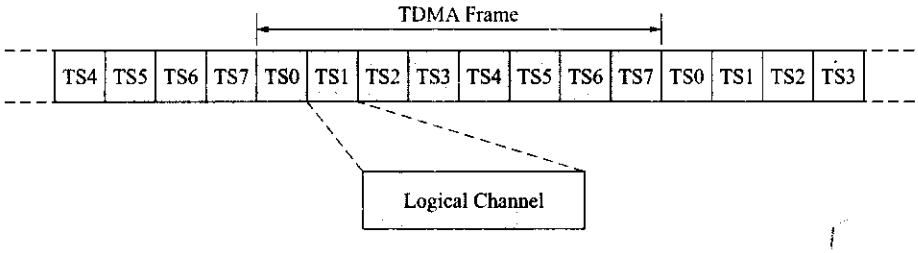


Figure 5–9 GSM TDMA frame.

Logical Channels

As previously mentioned, the **logical channels** may carry either subscriber traffic or signaling and control information to facilitate subscriber mobility. Presently, there are three types of traffic channels (TCHs). The full-rate traffic channel (TCH/F or Bm) carries one conversation by using one timeslot. The transmitted voice signal is encoded at a 13-kbps rate, but it is sent with additional overhead bits. This information plus additional channel overhead bits yields a final channel data rate of 22.8 kbps. The full-rate traffic channel may also carry data at rates of 14.4, 9.6, 4.8, and 2.4 kbps. The half-rate traffic channel (TCH/H or Lm) carries voice encoded at 6.5 kbps or data at rates of 4.8 or 2.4 kbps. With additional overhead bits, the total data rate for TCH/H becomes 11.4 kbps. Therefore, two conversations or a conversation and a data transfer or two data transfers may be transmitted over one channel at the same time. Enhanced full-rate (EFR) traffic encodes voice at a 12.2-kbps rate and like TCH/F adds overhead bits to yield a 22.8 kbps channel data rate. The EFR channel may also transmit data at the TCH/F rates. More will be said about these channels later.

The signaling and control channels consist of three channel sub-categories: broadcast channels, common control channels, and dedicated control channels. The function of these channels will be explained in more detail next. Later, the timing scheme used to transmit the signaling and control channels within the TDMA frame structure will be examined.

Broadcast Channels

The GSM cellular system uses broadcast channels (BCHs) to provide information to the mobile station about various system parameters and also information about the location area identity (LAI). The three types of BCHs are broadcast control channel, frequency correction channel, and synchronization channel. Using the information transmitted over these three BCHs, the MS can tune to a particular base transceiver system

(BTS) and synchronize its timing with the frame structure and timing in that cell. Each time the MS attaches to a new BTS it must listen to these three BCHs.

At present, the timing of different GSM cells is not synchronized. However, there are several emerging technologies that may be adopted in the near future that may alter this fact. The use of single-antenna interference cancellation (SAIC) algorithms to increase GSM system capacity is being investigated by the GSM industry. This noise cancellation technique is enhanced for synchronous networks. Therefore, eventually GSM cells may all be aligned to some master clock like the Global Positioning System (GPS).

Broadcast Control Channel The broadcast control channel (BCCH) contains information that is needed by the MS concerning the cell that it is attached to in order for the MS to be able to start making or receiving calls, or to start roaming. The type of information broadcast on the BCCH includes the LAI, the maximum output power allowed in the cell, and the BCCH carrier frequencies for the neighboring cells. This last information is used by the MS to allow it to monitor the neighboring cells in anticipation of a possible handover operation that might be needed as the MS moves about. The BCCH is only transmitted on the downlink from BTS to MS.

Frequency Correction Channel The frequency correction channel (FCCH) transmits bursts of zeros (this is an unmodulated carrier signal) to the MS. This signaling is done for two reasons: the MS can use this signal to synchronize itself to the correct frequency and the MS can verify that this is the BCCH carrier. Again, the FCCH is only broadcast on the downlink.

Synchronization Channel The synchronization channel (SCH) is used to transmit the required information for the MS to synchronize itself with the timing within a particular cell. By listening to the SCH, the MS can learn about the frame number in this cell and about the BSIC of the BTS it is attached to. The BSIC can only be decoded if the BTS belongs to the GSM network. Again, SCH is only transmitted in the downlink direction.

Common Control Channels

The common control channels (CCCHs) provide paging messages to the MS and a means by which the mobile can request a signaling channel that it can use to contact the network. The three CCCHs are the paging channel, random access channel, and the access grant channel.

Paging Channel The paging channel (PCH) is used by the system to send paging messages to the mobiles attached to the cell. The MS listens to the PCH at certain time intervals to learn if the network wants to make contact with it. The mobile will be paged whenever the network has an incoming call ready for the mobile or some type of message (e.g., short message or multimedia message) to deliver to the mobile. The information transmitted on the PCH will consist of a paging message and the mobile's identity number (e.g., ISMI or TMSI). The PCH is transmitted in the downlink direction only.

Random Access Channel The random access channel (RACH) is used by the mobile to respond to a paging message. If the mobile receives a page on the PCH, it will reply on the RACH with a request for a signaling channel. The RACH can also be used by the mobile if it wants to set up a mobile-originated call. The RACH is only transmitted in the uplink direction. For this last operation, the RACH also plays an important role in the determination of the required timing advance needed by the MS and the subsequent assignment of this parameter to the mobile by the network.

The format of the signal sent on the RACH provides enough information to the wireless network (i.e., the BSC) to allow it to calculate the distance of the mobile from the BTS. This measured time delay is then translated into a timing advance (TA) that is sent to the MS. The use of a TA allows any mobile within the cell to transmit information that will arrive at the BTS in correct synchronization with the start of the TDMA frame. In the GSM system, the structure of the RACH signal allows for a maximum cell radius of 35 km except when extended range cells are defined by the system.

Access Grant Channel The access grant channel (AGCH) is used by the network to assign a signaling channel to the MS. After the mobile requests a signaling channel over the RACH the network will assign a channel to the mobile by transmitting this information over the AGCH. The AGCH is only transmitted in the downlink direction.

Dedicated Control Channels

The last group of broadcast channels is known as the dedicated control channels (DCCHs). These dedicated channels are used for specific call setup, handover, measurement, and short message delivery functions. The four DCCHs are the stand-alone dedicated control channel (SDCCH), the slow associated control channel (SACCH), the fast associated control channel (FACCH), and the cell broadcast channel (CBCH).

Stand-alone Dedicated Control Channel Both the mobile station and the BTS switch over to the network-assigned stand-alone dedicated control channel (SDCCH) that is assigned over the access grant channel in response to the mobile's request that has been transmitted over the random access channel. The call setup procedure (i.e., the initial steps required to set up a radio link) is performed on the SDCCH. The SDCCH is transmitted in both the uplink and downlink directions. When the call setup procedure is complete, both the mobile and the BTS switch to a preassigned available traffic channel.

Slow Associated Control Channel The slow associated control channel (SACCH) is used to transmit information about measurements made by the MS or instructions from the BTS about the mobile's parameters of operation. In the uplink direction the mobile sends measurements of the received signal strength from its own BTS and those of neighboring BTSs. In the downlink direction, the MS receives information from the BTS about the mobile's output power level and the timing advance that the mobile needs to use. The SACCH is transmitted in both the uplink and downlink directions over the same physical channels as the SDCCH or the TCH.

Fast Associated Control Channel The fast associated control channel (FACCH) is used to facilitate the handover operation in a GSM system. If handover is required, the necessary handover signaling information is transmitted instead of a 20-ms segment of speech over the TCH. This operation is known as "stealing mode" since the time allotted for the voice conversation is stolen from the system for a short period. The subscriber is usually not aware of this loss of speech since the speech coder in the mobile simply repeats the last received voice block during this process.

Cell Broadcast Channel The cell broadcast channel (CBCH) is used to deliver short message service in the downlink direction. It uses the same physical channel as the SDCCH.

Speech Processing

Before examining the structure of a timeslot, it will be instructive to take a brief look at how speech is processed in a GSM system. Figure 5–10 depicts this process. In the mobile, speech is digitized and broken up into 20-ms segments. It is then coded to reduce the bit rate and to control errors. This process produces 8000 samples of 13 bits per sample per second or 160 samples of 13 bits per sample per 20 ms. The speech coder yields 260 bits per 20 ms or 13 kbps whereas channel coding yields 456 bits per 20 ms or a 22.8-kbps data rate. Interleaving, ciphering, and burst formatting yields 156.25 bits per timeslot. This yields an overall data transfer rate of 270.8 kbps over a GSM channel.

The receiver works in the following manner: signal bursts are received and used to create a channel model. The channel model is created in the equalizer where an estimated bit sequence is calculated for a received signal. After all of the bursts containing information about a 20-ms segment of speech have been received and deciphered, they are reassembled into the 456-bit message. This sequence is then decoded to detect and correct any errors that occurred during transmission. More details about the signal bursts will be forthcoming shortly and more information about the interleaving and ciphering operations will be presented in Chapter 8.

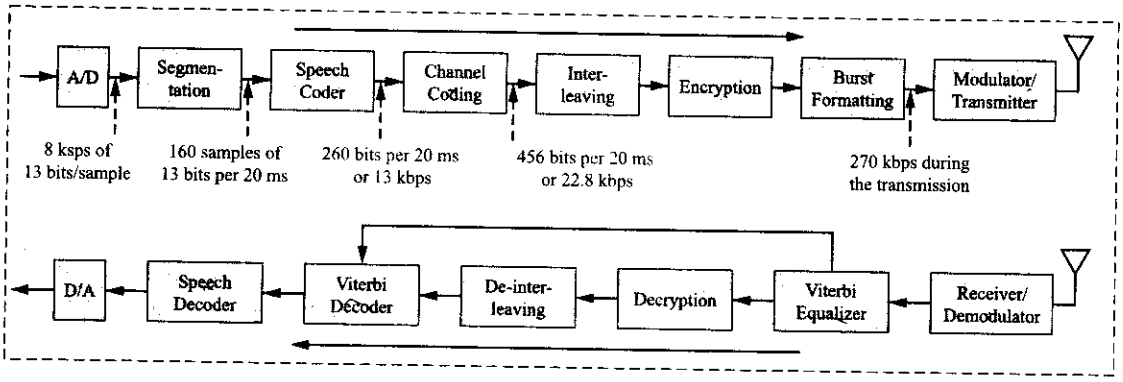


Figure 5–10 GSM speech processing.

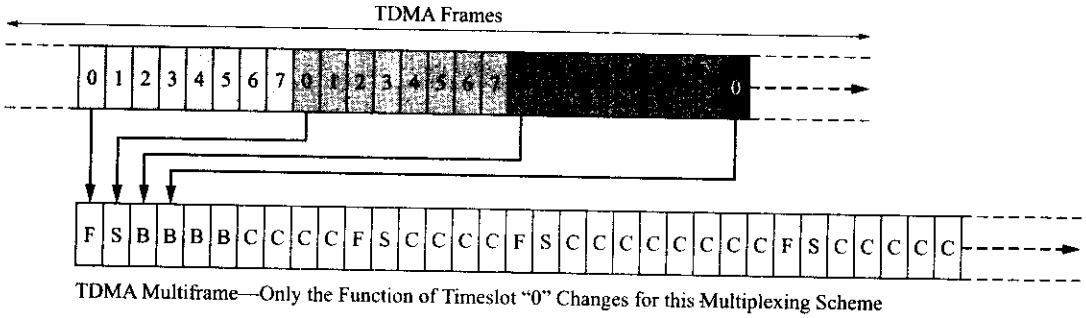


Figure 5–11 Relationship between timeslots and TDMA multiframes.

Timeslots and TDMA Frames

In a GSM system, both traffic and signaling and control information are transmitted over the same physical frequency channel. To accomplish this, time division multiplexing is used. The physical channels of the system used for the transmission of traffic are distinguished by virtue of their particular timeslot within a TDMA frame and the system signaling and control information is organized in terms of both the specific timeslot within the TDMA frame and the particular frame within a larger organization of TDMA frames (multiframes). The relationship between timeslots and TDMA multiframes is depicted in Figure 5–11. The next several sections will examine the concepts of timeslots and TDMA frames in more detail.

TDMA Frames

In the GSM system, eight timeslots constitute a TDMA frame. The system assigns numbers to the frames sequentially from 0 to 2,715,648 and then the process repeats itself. Our description of GSM timing will start with the largest system time period. This grouping of successive TDMA frames is known as a hyperframe. The hyperframe (as shown in Figure 5–12) consists of 2,048 superframes (2,715,648 frames) and takes 3 hours 28 minutes 53 seconds and 760 milliseconds to complete. Each superframe consists of 1,326 TDMA frames that take approximately 6.12 seconds to complete. These superframes may take on one of two possible formats. An explanation of why this is the case will be forthcoming shortly. One form of a superframe consists of 51 (26 frame) multiframes (i.e., each multiframe consists of 26 TDMA frames that take 120 ms to complete). The other superframe format consists of 26 (51 frame) multiframes (i.e., each multiframe consists of 51 TDMA frames that take about 235 ms to complete). Finally, as previously mentioned, within a TDMA frame there are eight timeslots that take approximately 4.615 ms to complete.

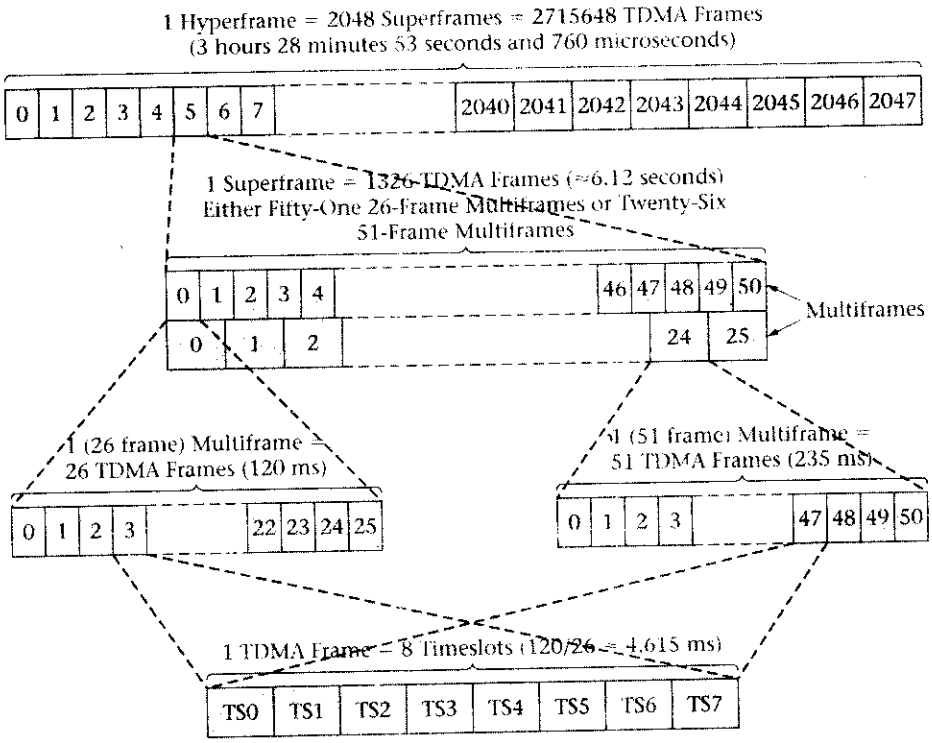


Figure 5–12 A GSM hyperframe (Courtesy of ETSI).

Timeslots

The organization of the transmitted digital bits within the air **timeslot** itself can take on several different formats depending upon the type of information being transmitted (i.e., voice traffic, data, or signaling and control messages). As shown in Figure 5–13, the air interface timeslot has a duration of 3/5200 seconds or approximately 577 μs (or 156.25 bit periods) whereas the typical transmitted burst is approximately 546 μs (or 148 bit periods). A bit time is 48/13 μs or approximately 3.69 μs. The overall bit rate over the air interface is approximately 270.8 kbps.

The start of a TDMA frame on the uplink is delayed by three timeslot periods from the downlink frame as shown in Figure 5–14. The purpose of this delay is so that the same timeslot may be used on both the downlink and uplink radio paths without the need for the MS to receive and transmit at the same time. This extends mobile battery life and makes it easier for the mobile's hardware to implement the RF operations needed for proper system functioning.

Timeslot Bursts The transmission of a normal (traffic and control channels) burst and the other types of burst signals are shown in Figure 5–15. In the case of a normal **burst,** two groups of 57 encrypted bits are transmitted on either side of a training sequence of bits. This **training sequence** consists of alternating 0s

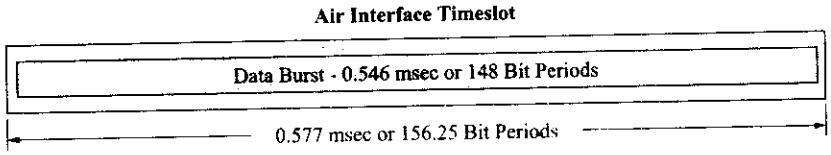


Figure 5–13 The GSM air interface timeslot.

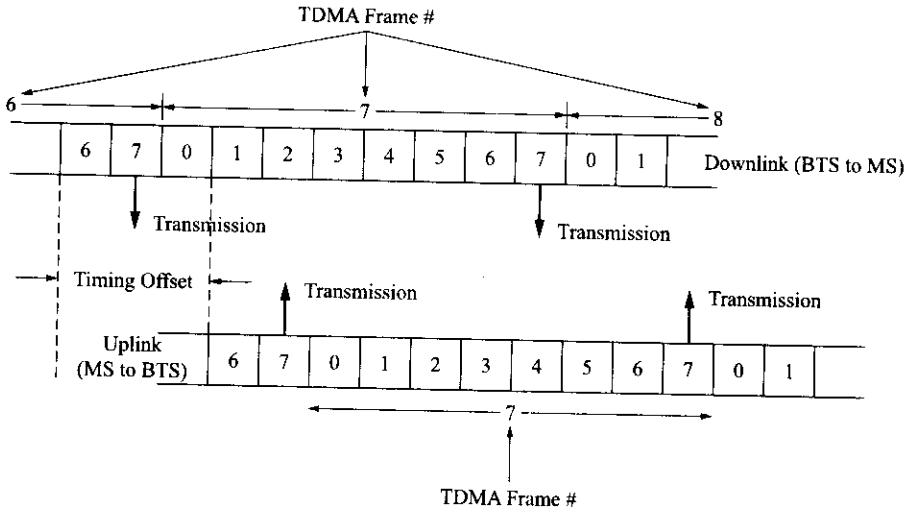


Figure 5–14 TDMA timing offset between uplink and downlink.

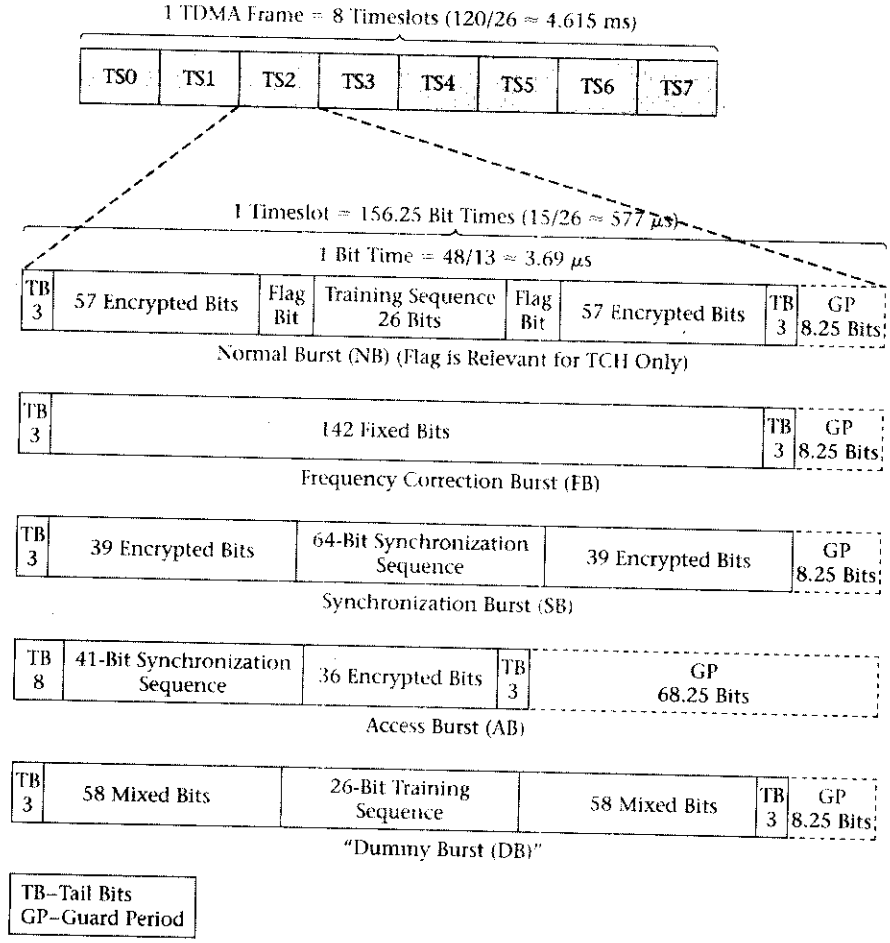


Figure 5–15 GSM traffic and control signal bursts (Courtesy of ETSI).

and 1s and is used to train the adaptive equalizer incorporated into the GSM mobile receiver. Three (3) tail bits precede the first group of traffic bits and 3 tail bits trail the last group of traffic bits. These tail bits consist of three zeros (unmodulated carrier) that provide time for the digital radio circuitry to initialize itself. Two single flag bits separate the training bit sequence from the encrypted bit groups. The flag bits are used to indicate whether the encrypted bits contain traffic or control information. The normal burst has an 8.25-bit long guard period at the end of the burst where no transmission activity takes place. When used as a traffic channel, a total of 114 encrypted bits are delivered per timeslot. Details of the encryption process will be presented later.

The frequency correction burst is used by the mobile to obtain frequency synchronization. It consists of 142 fixed bits (binary 0s or an unmodulated carrier) proceeded by 3 tail bits and followed by 3 tail bits. It also has the same 8.25-bit long guard period after it. The repetition of the frequency correction burst by the BTS within the GSM frame structure becomes the frequency correction channel (FCCH).

The synchronization burst is used by the mobile to obtain timing synchronization. It consists of 3 tail bits, followed by 39 encrypted bits, a 64-bit synchronization sequence, 39 more encrypted bits, 3 tail bits, and the same 8.25-bit long guard period. The encrypted bits contain information about the frame number (FN) and the base station identity code (BSIC). The repetition of the synchronizing sequence burst by the BTS within the GSM frame structure becomes the synchronizing channel (SCH).

The access burst is used by the mobile to facilitate random access requests by the mobile and handover operations. It consists of 8 tail bits followed by a 41-bit synchronization sequence, then 36 encrypted bits, and 3 tail bits. In this case, the length of the guard bit time period is equal to 252 μs or 68.25 bits. The reason for the long guard time is so a mobile that has just become active or has just been handed off and does not know the system timing advance can be accommodated. The value chosen allows for a cell radius of 35 km. The access burst is used on both the random access channel (RACH) and on the fast associated control channel (FACCH) during handover.

The dummy burst is transmitted on the radio frequency designated as c_0 when no other type of burst signal is being transmitted. It consists of 3 tail bits, 58 mixed bits, a 26-bit training sequence, 58 more mixed bits, 3 tail bits, and the same 8.25-bit long guard period. The purpose of the dummy burst is to ensure that the base station is always transmitting on the frequency carrying the system information. This affords the mobile the ability to make power measurements on the strongest BTS in its location and thus determine which BTS to attach to when first turned on. Furthermore, the mobile can also make measurements of other BTSs and therefore provide information to the system if handover is needed.

Mapping of Logical Channels to Physical Channels

Now that the reader has some sense of the structure of a timeslot and the overall TDMA frame and frame hierarchy, it is time to take a look at the transmission of information within the multiframe structure of GSM.

The system needs to be able to transmit both traffic (voice or data) and signaling and control information to the subscriber. The subscriber needs to be able to access the system and request radio resources to set up a call or to send data. The operations involving data transmission will be discussed more fully in Chapter 7.

Only certain combinations of logical channels are permitted within the GSM standard. This section will not attempt to treat all the various possibilities and mapping combinations but will attempt to provide a broad overview with several specific examples to give the reader an appreciation for the basic concept of how traffic and signaling and control information is passed over the radio link within the GSM multiframe structure.

As a first example to consider, for proper system operation, there is a standard combination of logical channels that must be transmitted during Timeslot 0 of the designated downlink radio frequency channel (known as c_0) within a cell. It is: FCCH + SCH + BCCH + CCCH. Also permitted within Timeslot 0 of c_0 is another similar combination of channels: FCCH + SCH + BCCH + CCCH + SDCCH + SACCH. In both

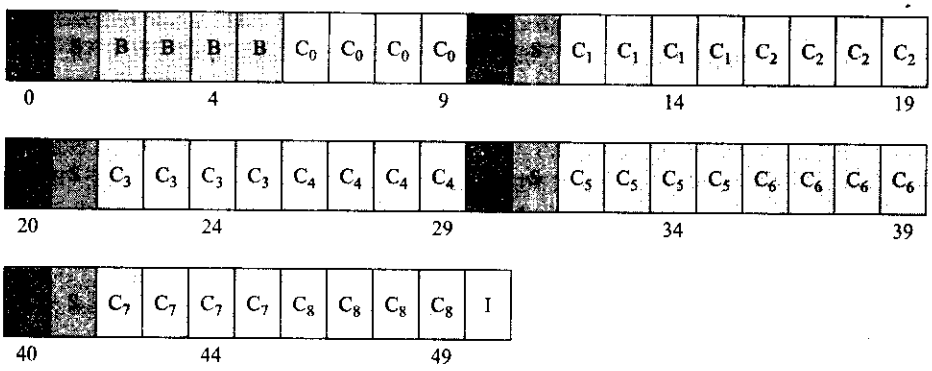


Figure 5–16 The multiplexing of GSM logical channels.

cases, the transmission of the three broadcast channels (BCHs) and the common control channels (CCCHs) provides both the information needed by the mobile to determine Timeslot 0 on c_0 and to synchronize with the frame structure of the cell and the means by which the mobile can access the network. How the first combination is multiplexed within the GSM multiframe structure is shown in Figure 5–16.

As shown in Figure 5–16, the sequence of FCCH, SCH, BCCH, and CCCH repeats every fifty-one TDMA frames (a multiframe). The last frame of the sequence (Frame #50) is an idle frame and carries no information. The nine groups of four frames carrying CCCH information are called paging blocks and the one group of four frames that carry BCCH information is needed due to the large amount of overhead information transmitted by the BTS over the BCCH. In the uplink direction, Timeslot 0 is reserved for use by the mobile for access to the GSM system (over the random access channel or RACH).

The second combination of channels that includes the SDCCH and SACCH channels along with the BCHs and CCCHs (known as a combined control channel) is implemented in the GSM multiframe structure as shown in Figure 5–17. In this case, one can see that only three paging blocks are present but four SDCCH and two SACCH channels are available. This type of channel combination (known as SDCCH/4) is effective in a rural cell where little traffic is expected to be generated. In this configuration, the physical

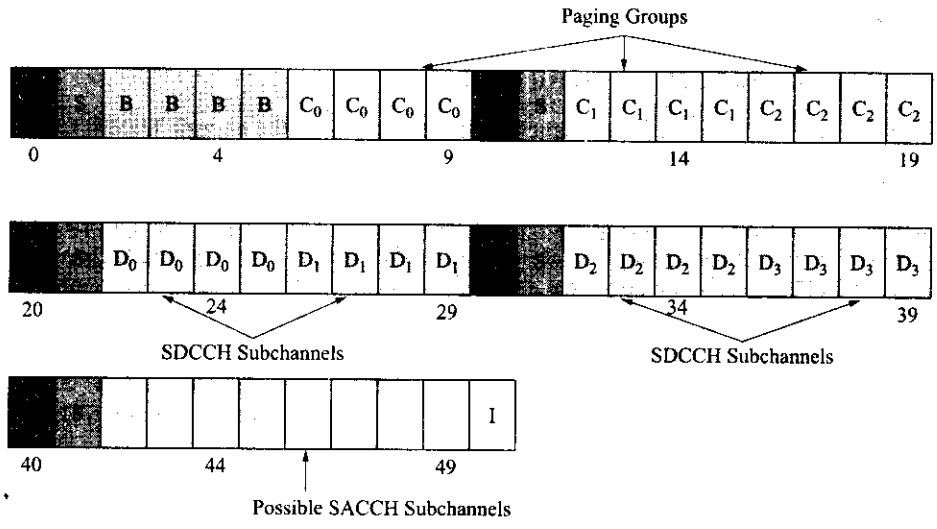


Figure 5–17 Another GSM multiframe configuration.

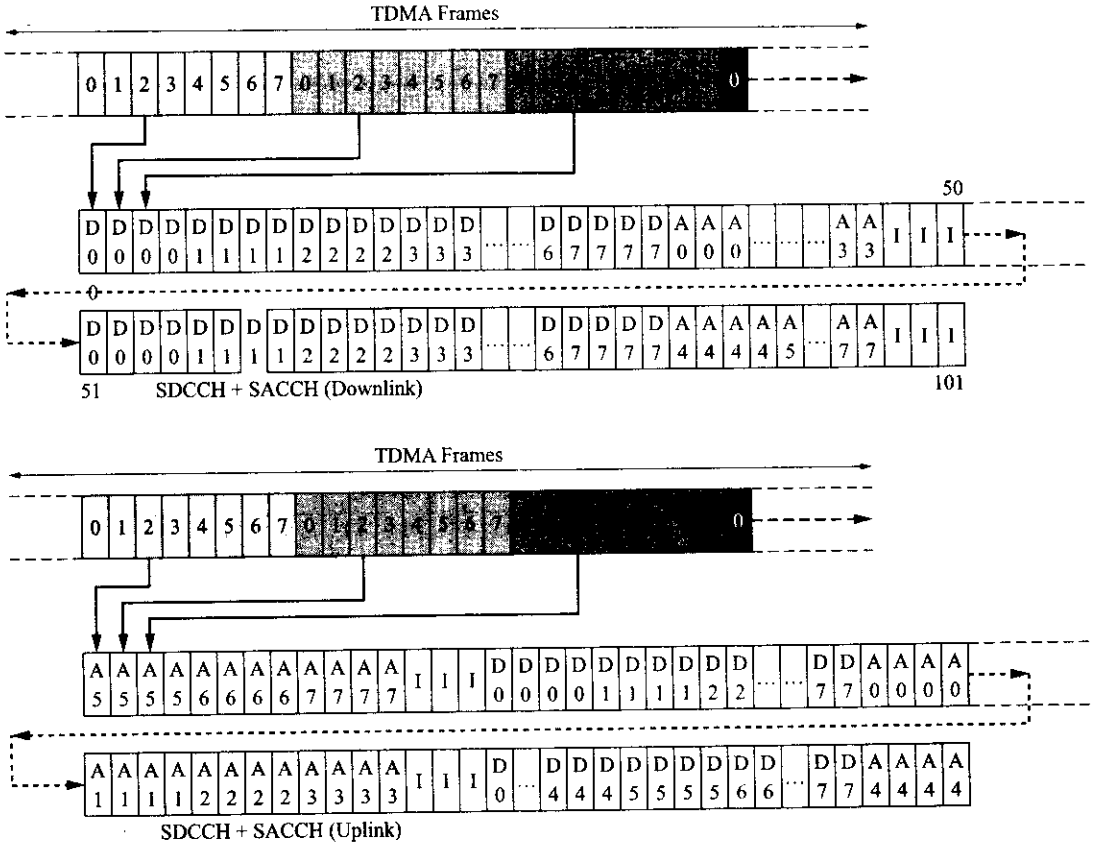


Figure 5–18 High-traffic GSM multiframe.

channels that would normally be assigned exclusively to SDCCH and SACCH (to be explained next) can be used as traffic channels and call setup and other system operations may all be carried out over one physical channel (Timeslot 0).

For high-traffic cells, another combination of channels, SDCCH and SACCH, can be transmitted on any timeslot and any carrier frequency except for Timeslot 0 on c_0. The repeating sequence of SDCCH and SACCH channels takes place over 102 TDMA frames. Figure 5–18 shows the multiplexing of the SDCCH and SACCH combination on both the downlink and uplink. As can be seen from the figure, this combination of SDCCH and SACCH channels can support up to eight mobile stations simultaneously and is known as channel combination SDCCH/8. The figure also shows how the timing between the channels on uplink and downlink are shifted by sixteen timeslots to accommodate the delay needed by the mobile for the processing of information received from the BTS. A cell may support up to four SDCCH/8 channels in addition to a broadcast channel on Timeslot 0 or up to three SDCCH/8 channels with a combined control channel on Timeslot 0.

Transmission of Short Messages

A cell broadcast channel (CBCH) is required for the transmission of short message service in the downlink direction. One of the SDCCH subchannels will be assigned for this purpose. Only one CBCH can be supported within a cell.

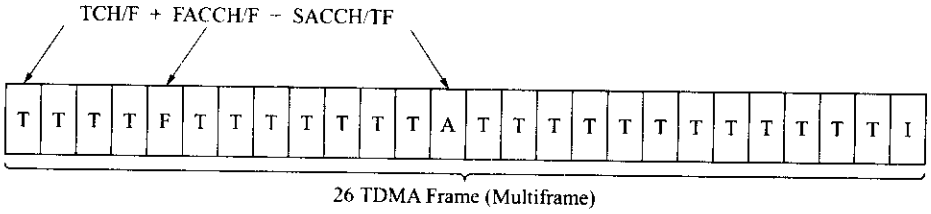


Figure 5–19 GSM traffic channel frames.

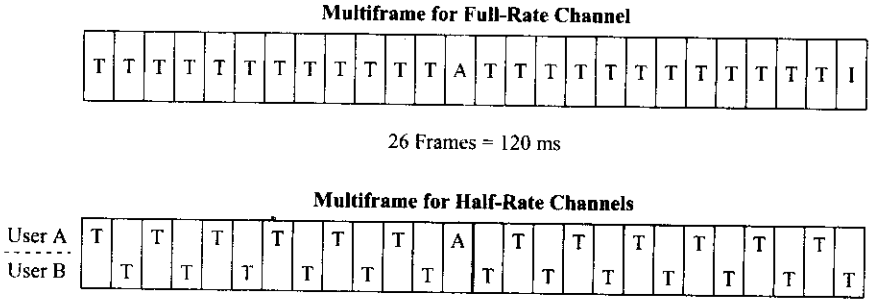


Figure 5–20 GSM full-rate and half-rate traffic frame structures.

Traffic Channels

With the channel combinations already mentioned, typically, Timeslot 0 and another timeslot (Timeslot 2 was used in Figure 5–18) are used by broadcast and control channels and the dedicated control channels. This leaves six timeslots (TS1 and TS3-TS7) free for use by traffic channels (TCH). Traffic channels are mapped onto physical channels (timeslots) along with a SACCH channel. The repetition of the TCHs and the SACCH occurs over a sequence of twenty-six TDMA frames. Figure 5–19 shows this situation. The FACCH channel may also be used in this sequence by stealing timeslots from bursts of speech as shown.

Half-Rate Traffic Channels The GSM system can use half-rate channels to double system capacity since two users share the same physical channel. Figure 5–20 shows both full-rate and half-rate frame structures. Using half-rate channels, the idle frame used in full rate will be used for the SACCH signaling for the second MS. Since a mobile uses only every other timeslot for a call, the multiframe will contain thirteen idle frames for each mobile. Therefore, the mobile could be allocated two traffic channels or a speech and a data channel.

Paging Groups

The mobile will be assigned by the network to a particular **paging group** through an algorithm (refer to GSM TS 05.02) that uses the mobile's IMSI number and other system information.

From system information messages, the MSs attached to a cell will receive information about the type of combined mapping of logical channels that the cell supports and the type of multiframe structure used between transmissions of paging messages to the same paging group. Using this information and its own IMSI number, the mobile will calculate which CCCH and which paging group that it belongs to. From its calculations, the MS will only listen for pages and make random accesses on a specific CCCH. The number of paging groups using noncombined mapping is greater than with combined mapping of the logical channels.

PART II GSM SYSTEM OPERATIONS

5.4 GSM IDENTITIES

Chapter 3 introduced the reader to the idea of wireless network element identities. The GSM standards use the same numbering systems as described previously. These numbering plans will be briefly described here again for continuity.

Mobile Station Associated Numbers

The MS has a **mobile station ISDN number** (MSISDN) that uniquely identifies a mobile telephone subscription in the PSTN numbering plan. It is therefore a dialable number and is linked to one HLR. The **international mobile subscriber identity** (IMSI) is a unique identity allocated to each subscriber by the wireless service operator and stored in the subscriber's SIM. All network-related subscriber information is linked to the IMSI. Besides being stored in the subscriber's SIM, the IMSI number is also stored in the HLR and VLR databases. The **temporary mobile subscriber identity** (TMSI) number is used by the GSM network to protect the subscriber's privacy over the air interface. The wireless network assigns a TMSI to the MS, and the TMSI number only has local significance within the particular MSC/VLR coverage area during MS attachment. The international mobile equipment identity (IMEI) number and the international mobile equipment identity and software version (IMEISV) number are used by the GSM network for equipment identification and to uniquely identify an MS as a piece of equipment.

Network Numbering Plans

The GSM system uses both location area identity (LAI) numbers and cell global identity (CGI) numbers. The LAI is used for MS paging and location updating. The CGI is used for cell identification within a location area (LA). Within the wireless network itself, the network elements will have identity numbers or addresses that are necessary to facilitate the correct operation of the system. Examples of this identification scheme are base station identity codes (BSICs) and mobile global titles (MGTs) that uniquely identify network switching system elements within a particular operator's network. The MGT concept was already covered in greater detail in Section 3.4.

Mobile Station Roaming Number

The **mobile station roaming number** (MSRN) was introduced in Chapter 3 during a discussion of the basic functions of the various network switching system databases during a mobile-terminated call. The MSRN is used by the GSM system during the call setup operation. This operation is shown in Figure 5–21 and is known as the interrogation phase.

As shown in the figure, several operations must be performed before the call setup operation can be completed. In Step #1, the GMSC receives a signaling message, "initial address message," from the PSTN about the incoming call for a particular MSISDN number. In Step #2, the GMSC sends a signaling message, "send routing information," to the HLR where the subscriber data for the particular MSISDN is stored. In Step #3, the HLR uses MSISDN to find the subscriber data in the database. The VLR address that corresponds to the subscriber location and the IMSI for the subscriber is retrieved from the HLR database in this step. In Step #4, the HLR sends a "provide roaming number" message to the MSC/VLR using the VLR address as the destination and the IMSI to identify the mobile subscriber. In Step #5, the VLR asks the MSC to seize an idle MSRN (this corresponds to a signaling path) from its available pool of numbers and to also associate it with the IMSI number received from the HLR. In Step #6, the MSC/VLR sends the MSRN back to the HLR. In Step #7, the HLR sends the MSRN back to the GMSC. In Step #8, the GMSC

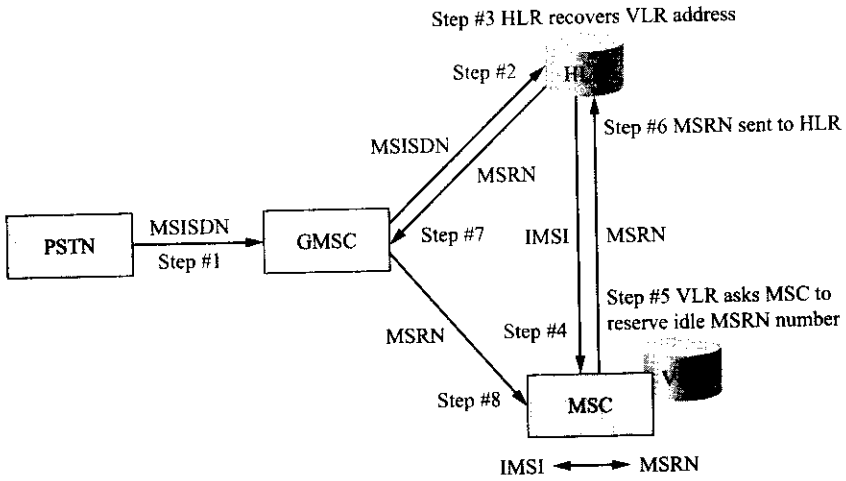


Figure 5–21 GSM call setup using the MSRN (Courtesy of Ericsson).

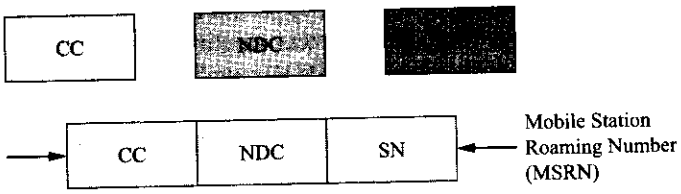


Figure 5–22 Formulation of the GSM MSRN.

uses the MSRN to route the call to the correct MSC. Now the serving MSC receives a signaling message, "initial address message," for the incoming call identified by the MSRN value. The MSC analyzes the incoming digits and associates them with the IMSI that corresponds to the subscriber. The MSRN number is released and made available for other calls. The IMSI is used by the MSC for final establishment of the call.

The MSRN follows the E.164 numbering plan. It consists of three parts as shown in Figure 5–22.

$$\text{MSRN} = \text{CC} + \text{NDC} + \text{SN}$$

Where CC = Country Code

NDC = National Destination Code

SN = Subscriber Number (This is the number of the serving MSC)

This last example pulls together some of the concepts presented earlier. The next several sections will also provide additional examples of overall system operation.

5.5 GSM SYSTEM OPERATIONS (TRAFFIC CASES)

The reader has already been introduced to the typical wireless network operations of call setup, location updating, and handoff in Chapters 2–4 as the common tasks and operations performed by the various elements of a wireless network system. The purpose of this section is to show the reader further detail about how the various typical traffic cases are handled within the GSM system. These examples will indicate the different types of system signaling that occur, the nodes of the GSM system involved in the assorted operations, and the functions that the nodes perform during these operations. The traffic cases considered in this

section will include calls and the operations that support a subscriber's mobility: location updating operations and handover cases. For a description of all of the possible GSM traffic cases, the reader will have to refer to the GSM standards. For the sake of continuity, the reader may want to review Section 3.5 for an overview of call establishment. Again, the details of data calls and short message service will be covered in Chapter 7.

Registration, Call Setup, and Location Updating

Before describing the call setup operations, one needs to consider the various states that the MS can be in. The MS can be powered off, or the SIM card can be removed from the mobile, or the mobile can be on but located in an area without service. In all these cases, the MS is considered to be in the detached condition. Otherwise, the MS can be powered on within the GSM system and will subsequently enter into an attached relationship with the system. The mobile can be in either of two states when attached: (1) the idle state in which the MS has no dedicated channel allocated to it and it just listens to the broadcast control channels (BCCH) and the paging channels (PCH) or (2) the active or dedicated state in which the MS has a dedicated connection to the GSM network. While in the attached mode, the MS may change from the idle to the active mode as the result of call setup, short message service transfers, location updating, or supplementary service procedures. Also, if the MS is in the active mode and changes cells, this operation is referred to as GSM handover.

Call Setup

Call setup within a GSM system consists of quite a few necessary operations. For either a mobile-originating call or a mobile-terminating call the following ten operations need to be performed. For a mobile-terminating call it is necessary to perform an initial additional operation as shown:

- Interrogation (only for a mobile-terminating call)
- Radio resource connection establishment
- Service request
- Authentication
- Ciphering mode setting
- IMEI number check
- TMSI allocation
- Call initiation
- Assignment of a traffic channel
- User alerting signaling
- Call accepted signaling

Interrogation Phase The interrogation phase has been described previously in this chapter in Section 4 under the heading Mobile Station Roaming Number. Figure 5–23 graphically illustrates the interrogation phase in a timeline/flowchart form. This will be the format used in this section to illustrate the various operations and signaling occurring between the different nodes of the GSM network. For the interrogation operation, one notes that the initial address message (IAM) comes from outside the GSM network using ISUP/TUP protocols. In some vendors' systems, the GMSC can send a request to the flexible numbering register (FNR) system node before being sent to the HLR. Also, for security reasons, the subscriber data can be simultaneously stored and updated in two HLRs. This built-in system redundancy assures successful operation in all but the most catastrophic disasters. In one final note about this operation, one observes that in the last operation performed, the two GSM system nodes (the MSC/VLR and the GMSC) use a non-MAP protocol to communicate with each other (i.e., the IAM message).

Radio Resource Connection Establishment Figure 5–24 shows a graphic of the radio resource connection establishment process. Figure 5–25 shows the detailed steps required for radio resource connection estab-

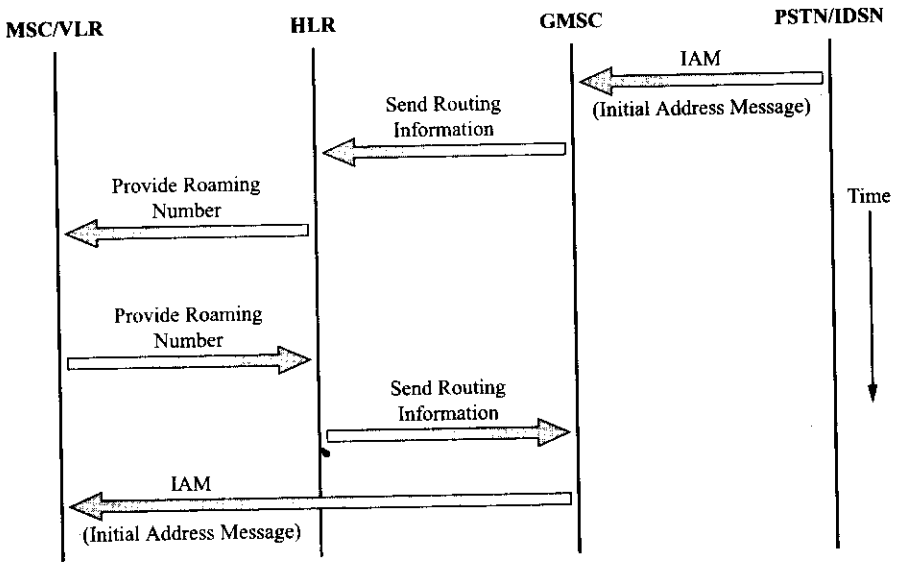


Figure 5–23 GSM interrogation phase of call setup.

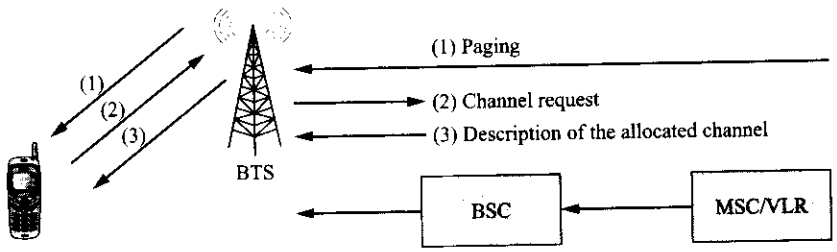


Figure 5–24 GSM radio resource connection establishment.

lishment. The MSC/VLR initiates the call setup process by sending a Layer 3 paging message to the appropriate BSC. The paging message will contain the subscriber's IMSI number so that the BSC can calculate the correct paging group to use. Recall that the MS can be paged in all the cells of a particular location area or even globally in all the cells of a MSC/VLR serving area. In most cases, the LAI is provided by the MSC to the BSC. The BSC receives the paging message and typically translates the LAI to a cell global identity (CGI) number if this information was not provided in the paging message.

The BSC sends the paging command message to the appropriate BTSs. This message will contain the following information: the IMSI or TMSI, the paging group, and the channel number. The channel number will contain enough information to indicate the channel type and the timeslot number. For this case, the channel type is a downlink common control channel (CCCH) (i.e., a paging channel [PCH]). For the GSM system, the paging group is determined by the subscriber's IMSI and other information defined in the BSC. When the MS has received the system information and knows its paging group, it will calculate when this paging group will be broadcast and thereafter will only listen for pages during the time they are expected to be sent.

Finally, the BTS sends a paging request message to the MS. This message is sent on the PCH. There are several different types of paging requests possible depending upon the use of IMSI or TMSI numbers. If TMSI numbers are used instead of IMSI numbers, up to four MSs may be paged in one paging message. The MS responds to the paging request message by sending a channel request message to the BTS. This message is transmitted on the random access channel (RACH) and contains information about the type of

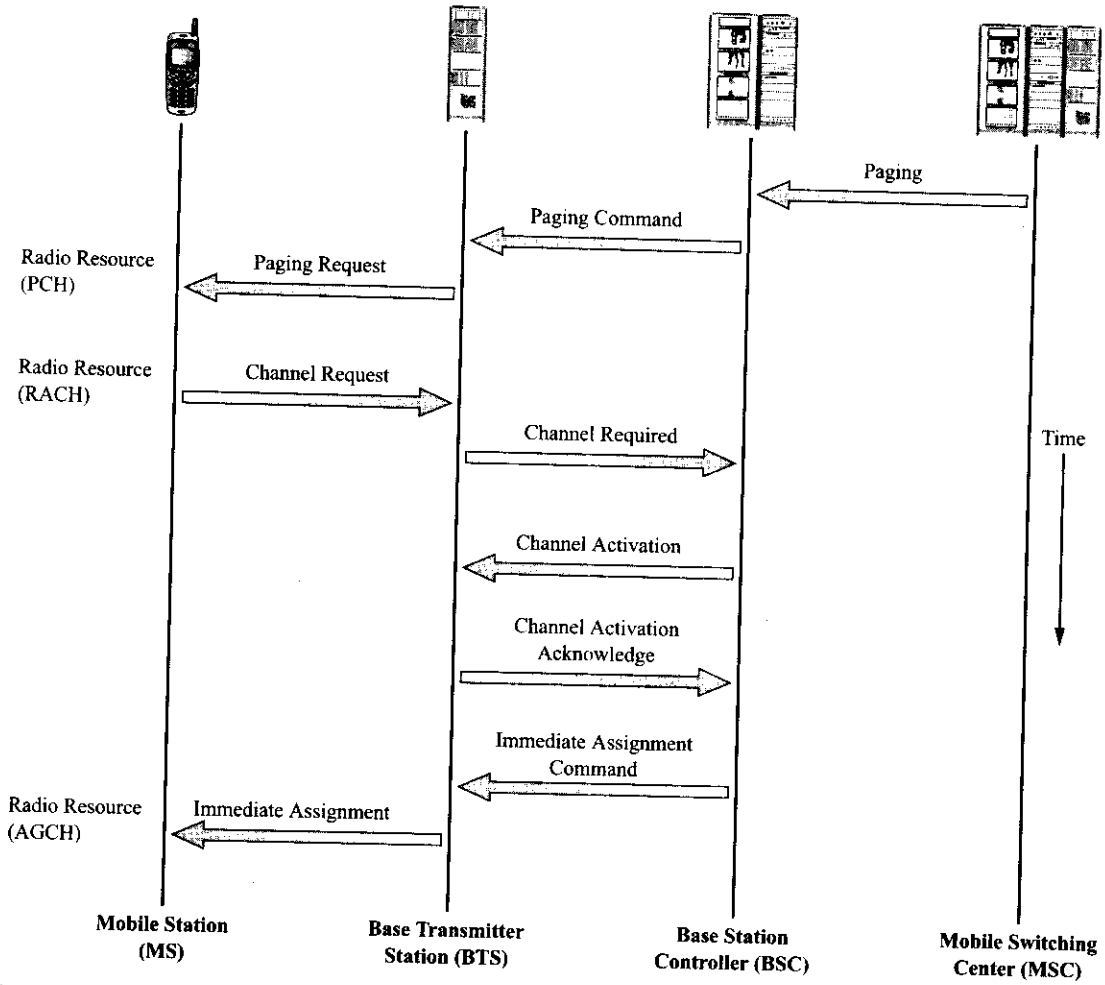


Figure 5–25 Detailed messaging during GSM radio resource connection establishment.

request (i.e., answer to page, originating call, location updating, emergency call, or other operations) to set priority if the system is experiencing heavy call volume and the radio resources are low. When the BTS detects an access burst, it sends a channel required message to the BSC. The BSC examines the information contained within the channel required message (access delay of the access burst, type of request, and TDMA frame number when the access burst was detected, etc.) and determines whether the MS is within the allowed range of the cell. The BSC determines what channel to use and sends a channel activation message to the BTS that contains the following: MS and BS power, timing advance (TA), DTX status, the reason for the allocation, and a complete description of the channel as shown in Figure 5–26. Figure 5–26 indicates that there are two possible modes of system operation: single carrier or multiple carrier (known as frequency hopping). In Figure 5–26, the value of the mobile allocation index offset (MAIO) is a number between 0–63 used to identify the hopping sequence of the mobile and the hopping sequence number (HSN) identifies the pseudo-random generator employed by the MS and the network to generate the frequency hopping sequence to be used.

The BTS activates this channel and then sends a channel activation acknowledge message back to the BSC. The BSC then sends an immediate assignment command message back to the BTS that includes ar immediate assign message for the MS. This immediate assign message is sent by the BTS to the MS o

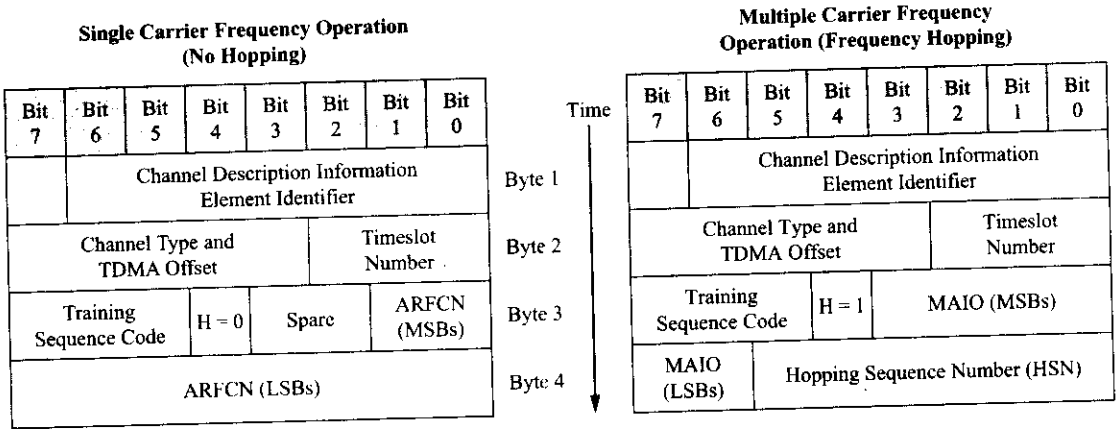


Figure 5–26 GSM channel description messages (Courtesy of ETSI).

the access grant channel (AGCH) and instructs the MS to switch to the allocated signaling channel and contains the channel description information element shown in Figure 5–26, the TA for the MS, some of the original information from the access burst, and in the case of frequency hopping a list of frequencies for the MS to hop between. If the information sent back to the MS from the original access burst agrees with the values stored by the MS, the mobile enters a new phase to be described next.

The GSM specifications allow for a modification of the just described procedure. If need be, the BSC may send an immediate assignment on TCH command to the MS. This allows the call setup signaling to be performed directly over the TCH. When the call setup procedure is complete, a channel mode modify command message can used to initiate a procedure that will return the TCH to the traffic mode. This strategy might be employed if there is congestion on the available system SDCCHs.

Service Request The service request phase occurs as soon as the MS has tuned to the new channel assigned to it by the immediate assignment message sent during the radio resource connection phase. Figure 5–27

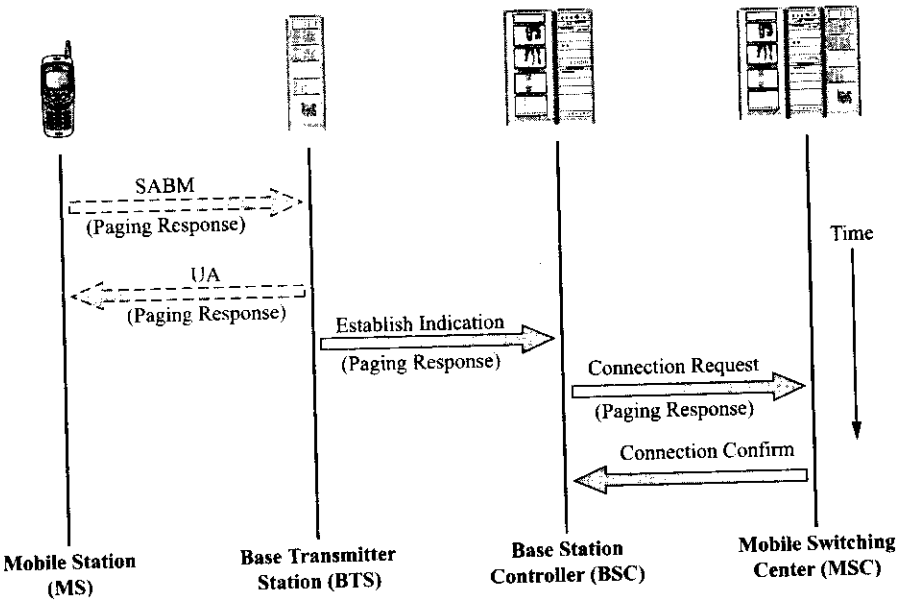


Figure 5–27 GSM service request operations.

shows these operations. At this time, a Layer 2 message known as set asynchronous balanced mode (SABM) is sent from the MS to the BTS. This Layer 2 message contains a Layer 3 message (i.e., the information field of the Layer 2 message contains the paging response message). Shortly thereafter, the BTS sends back to the MS a Layer 2 message in an unnumbered acknowledgement (UA) frame that contains the original paging response message. This operation prevents the chance occurrence of two MS accessing the same channel simultaneously.

The paging response message from the MS contains information about the MS identity, the ciphering key sequence number, and the MS class mark. When the paging response arrives at the BTS it is forwarded to the BSC in an establish indication message. This message causes the BSC to activate radio connection quality supervision and initiates power control algorithms for the dynamic control of the MS output power level. The paging response message from the MS is to be eventually delivered to the MSC and therefore the BSC sends it on to the MSC as a connection request message after it adds the CGI number to the Layer 3 information contained in the original paging response message. Finally, the MSC sends a connection confirm message back to the BSC. This means that the circuit-switched connection is established on the A interface.

Authentication The next step in the call setup procedure is authentication. The authentication process in shown in Figure 5–28. Depending upon the exchange properties stored in the MSC/VLR, as set up by the GSM operator, authentication is either activated or not activated. If authentication is activated, an authentication request message is sent transparently to the MS. The message containing a 128-bit random number (RAND) and the ciphering key sequence number (CKSN) is sent to the MS over the stand-alone dedicated control channel (SDCCH) from the BTS. The MS stores the CKSN and then calculates the value of a signed response (SRES) by using the RAND, the value of k_I (the subscriber authentication key that is stored in the SIM card), and K_C in several authentication algorithms (known as A3 and A8). The value of SRES is returned to the MSC/VLR as a transparent authentication response message. Between the BSC and the BTS a data request frame and a data indication frame are used to pass the Layer 3 message as shown. A timer is set in the MSC/VLR when the first authentication request message is sent. If the timer expires, the request is sent again. If the timer expires a second time, the radio resources (the channel) are released.

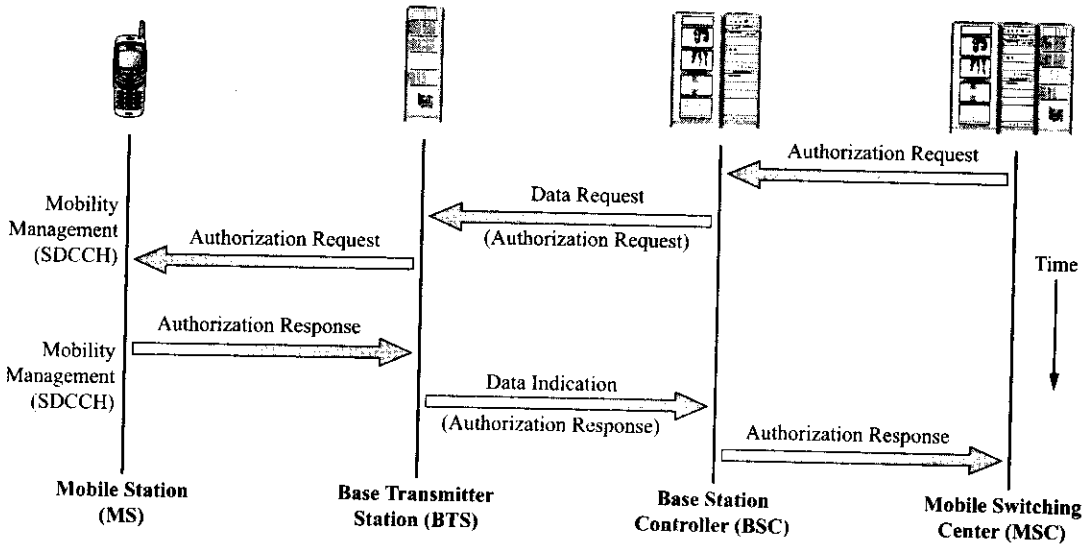


Figure 5–28 GSM authentication operations.

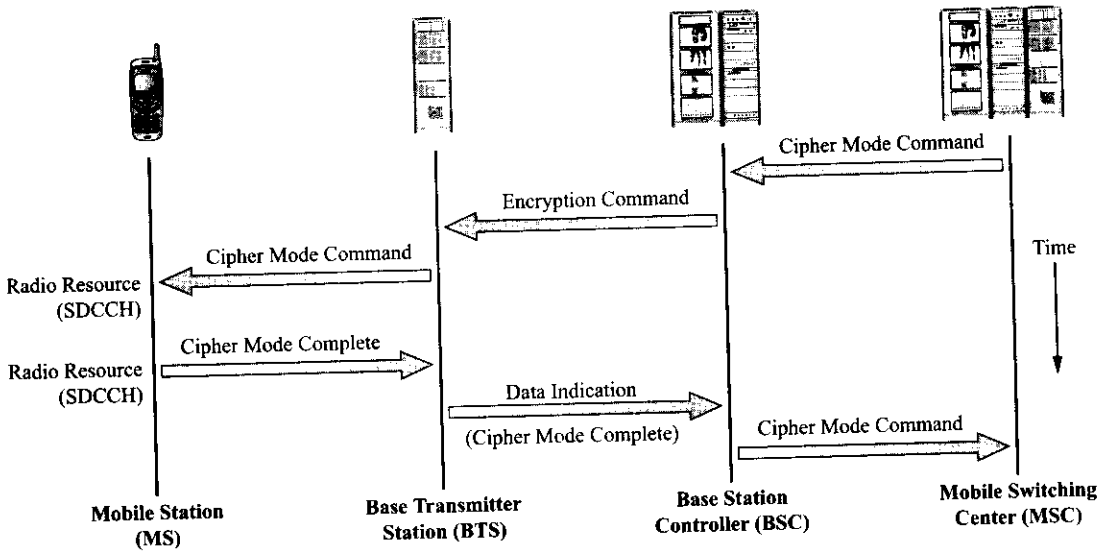


Figure 5–29 GSM ciphering mode setting operations.

If authentication is unsuccessful, the GSM system may initiate a procedure to identify the MS. Depending upon the results of this procedure the MS may be barred from the system or sent a message indicating that the "IMSI is unknown in VLR" or "PLMN not allowed."

Ciphering Mode Setting If the authentication process is successful, the next step in the call setup process is initiated. The process of ciphering mode setting is shown in Figure 5–29. The MSC/VLR sends the ciphering mode command to the BSC. This is a BSSMAP message that contains the value of K_C. This value is forwarded to the BTS within an encryption command message. The BTS stores the value of K_C and sends a nonciphered ciphering mode command message to the MS. The MS inserts K_C and the TDMA frame number into another authentication algorithm (A5). This creates a ciphering sequence that is added to the message that is to be sent. This ciphering mode complete message is sent to the BTS. The BTS upon receipt and correct deciphering of this message sends it transparently to the MSC via a data indication frame from BTS to BSC.

The ciphering key sequence number (CKSN) is used by the GSM system to reduce the number of steps required for call setup. Recall that the value of CKSN has been stored in the SIM card. If the MS makes another call without first detaching and reattaching to the network, the service request message from the MS to the MSC will include the CKSN. The system checks to see if the CKSN value is stored with the MS's IMSI in the VLR. If so, the MS may start ciphering immediately without first performing authentication. Obviously, this will ease the network signaling load. This process of selective authentication can be controlled by exchange properties set in the MSC/VLR by the system operator.

IMEI Check Again, the exchange properties set in the MSC/VLR determine whether an IMEI check is performed. If the IMEI number is to be checked, the MSC/VLR sends an identity request message to the MS as shown by Figure 5–30. As shown by the figure, this mobility management message and the MS identity response message are sent transparently between BTS and BSC. The value of IMEI sent by the mobile is checked against the equipment identity register (EIR) database. The EIR can return three status modes for the MS back to the network. The MS can be "white listed" and allowed to use the network, the MS can be "black listed" and not allowed to use the network, or the MS can be "grey listed." It is then up to the network operator to decide if the MS can use the network or not.

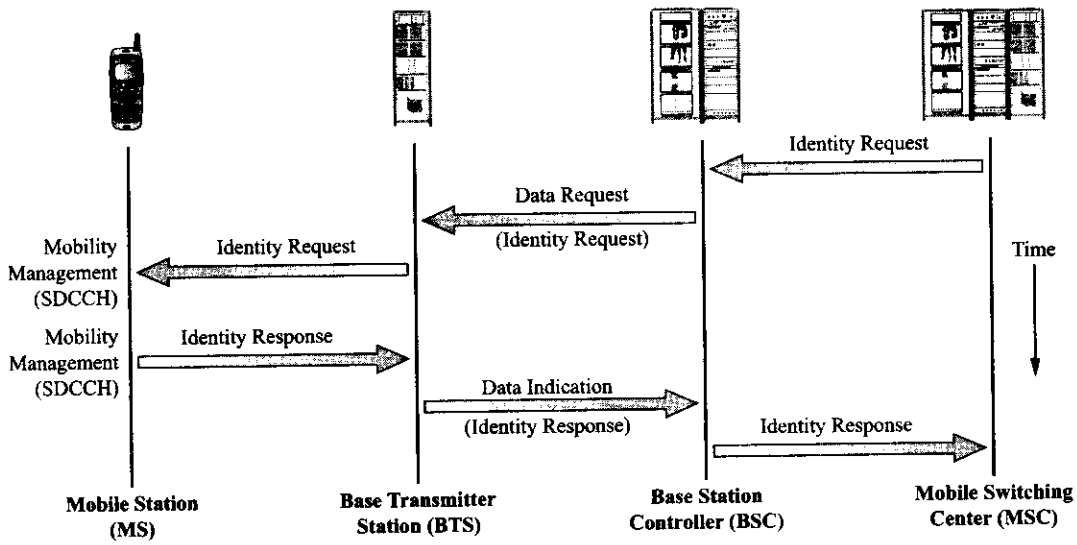


Figure 5–30 GSM IMEI check.

TMSI Reallocation The value of the TMSI number to be used for a particular traffic case or if one will be used at all is determined by the MSC/VLR software program. If a TMSI number is to be used, it is sent transparently to the MS from the MSC/VLR via the TMSI reallocation command as shown in Figure 5–31. This mobility management message is transmitted over the SDCCH from the BTS to the MS. The value of the TMSI number is stored in the SIM card and a TMSI reallocation complete message is sent transparently from the MS to the MSC/VLR over an uplink SDCCH.

Call Initiation Procedure The next step in the call setup process is the transmission of the setup message transparently from the MSC to the MS. As shown in Figure 5–32, this connection management message is sent over the downlink SDCCH from BTS to MS. This message contains a request for GSM bearer services (speech, data, fax, etc.). The MS will send a call confirmed message on the uplink SDCCH if it can handle

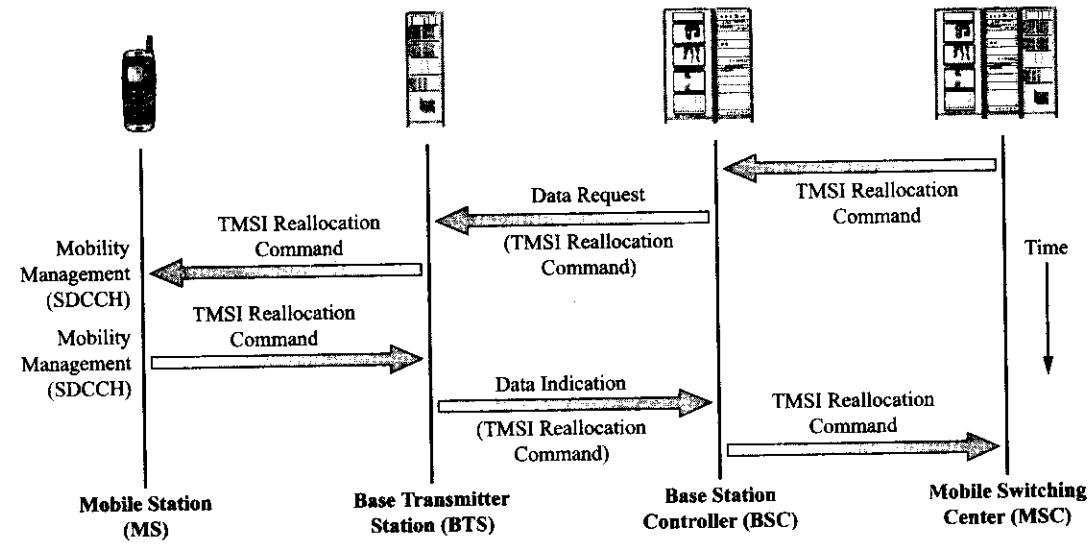


Figure 5–31 GSM TMSI reallocation operations.

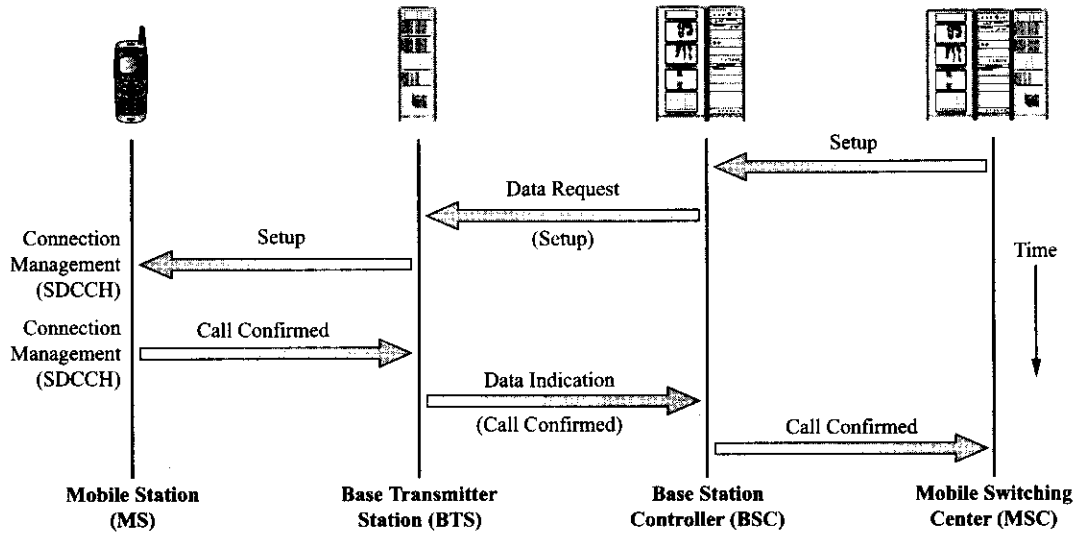


Figure 5–32 GSM call initialization operations.

the requested service. Today one can imagine many instances of incompatible mobiles unable to handle the newest multimedia data formats. This message is also sent transparently from MS to MSC. A timer is started in the MSC/VLR once the setup message is sent. If the timer expires before the call confirm message is received, the connections to the calling subscriber and the mobile subscriber are released.

Assignment of a Traffic Channel The traffic channel assignment is initiated by the MSC. As shown in Figure 5–33, the MSC sends an assignment request message to the BSC. This message contains information about the call priority, the status of DTX on the downlink, a circuit identity code (CIC) to indicate the transmission path for the speech or data traffic between the MSC and the BSC, and possibly a particular radio channel to facilitate some type of operations and maintenance function. The BSC could at this time assign the MS to the serving cell, another cell in the BSC serving area, or an external cell depending upon the status of the system and the available radio resources at the time.

If the assignment is to the serving cell, the BSC must obtain the timing advance information, calculate the MS output power level, select an idle traffic channel, and send a channel activation message to the BTS. This is the same message described in the section on Radio Resource Connection Establishment. However, instead of assigning SDCCH + SACCH as done in the RR connection establishment, this time the channel type is set to Bm + ACCH, which means a full-rate TCH + SACCH + FACCH. The BTS sends an acknowledgement back via a channel activation acknowledgement message to the BSC. The BSC sets up a path through its group switch for the traffic. The BSC sends an assignment command message to the MS that contains the information about the new channel assignment (i.e., TCH + SACCH + FACCH). This radio resource message is sent over the SDCCH. It consists of a complete channel description as was shown in Figure 5–26.

At this point, the MS tunes to the new channel and sends a SABM message over the FACCH to indicate successful seizure of the channel. As the BTS receives this message it sends a UA message to the MS and an establish indication message to the BSC. The UA message is sent back to the MS for the same reason as explained previously. The MS then sends an assignment complete message to the MSC to indicate that the traffic channel is working. Finally, the BSC sends a message to the BTS that the signaling channel is no longer needed in the form of a RF channel release message. The BTS sends an RF channel release acknowledgement message back to the BSC.

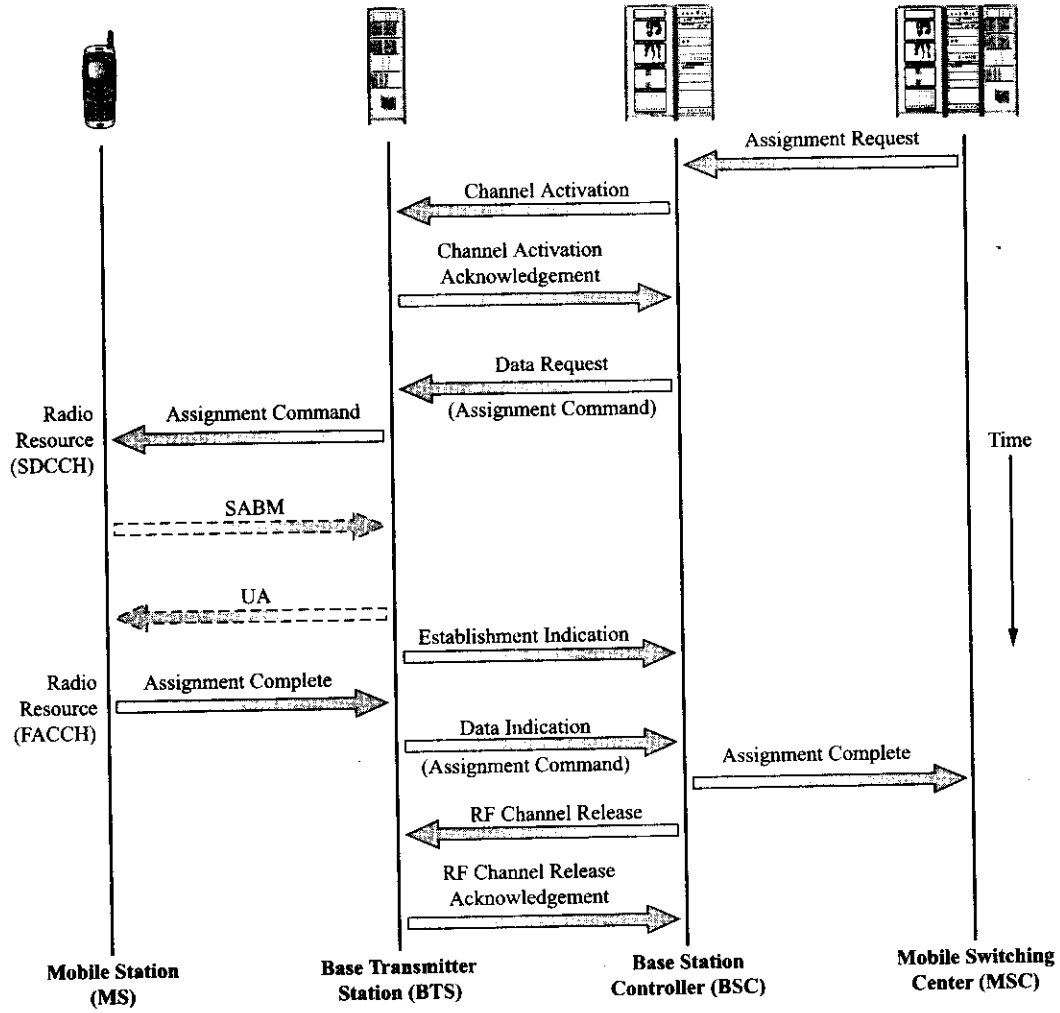


Figure 5–33 GSM traffic channel assignment.

Call Confirmation, Call Accepted, and Call Release The operations performed for call confirmation and call accepted are shown in Figure 5–34. The call confirmation procedure starts when the MS sends a transparent alerting message to the MSC. This message indicates that a ringing tone has been generated in the mobile and that it can be used for user-to-user signaling. When the alerting message is received the MSC/VLR sends the TUP address complete message to the calling subscriber who can now hear the ringing tone generated in the MSC. When the MS user answers, the connect message is sent to the MSC. This message, when received by the MSC, prompts a connect acknowledgement message to be sent back transparently to the MS. These are all connection management messages.

The system messages that occur at the end of a call have been already introduced in detail in Section 3.5. The reader may refer back to this section to review the details of the **call release** operation (refer to Figure 3–19) if so desired.

Other Aspects of Call Establishment

In early GSM systems, international calls to a GSM mobile were routed through each country's international exchanges. A call from one country to another required extensive signaling over each country's

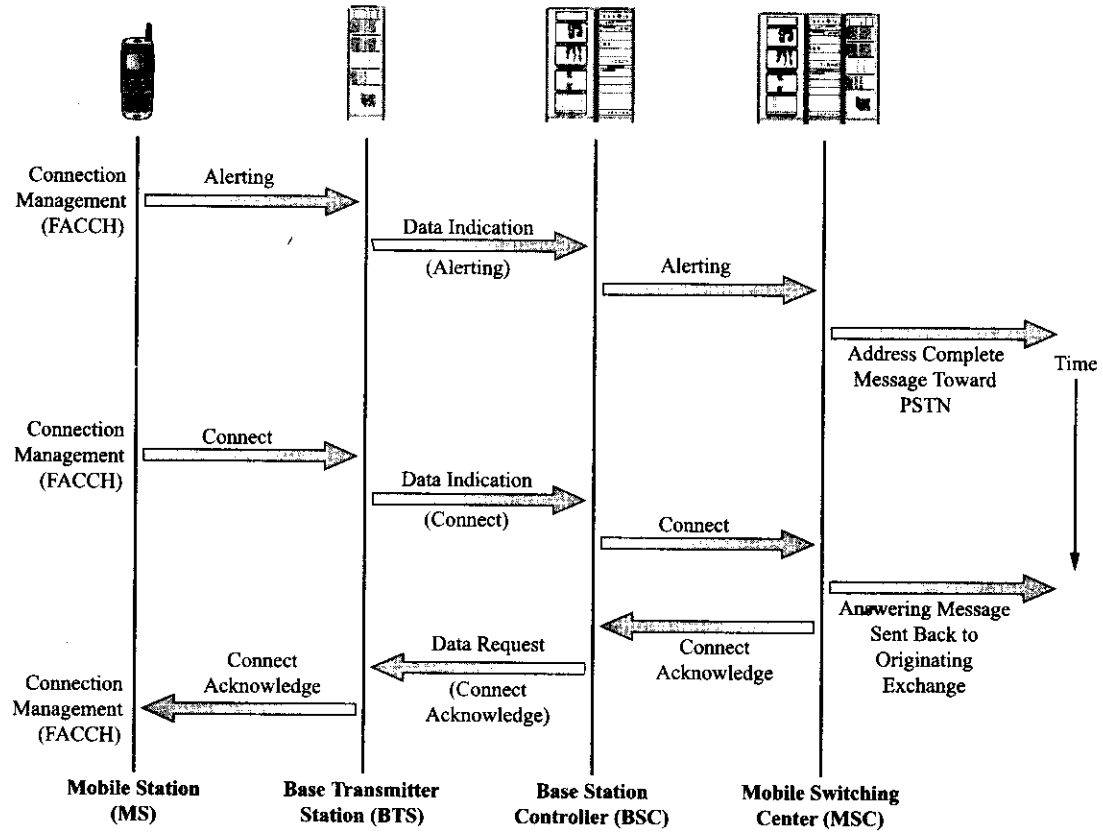


Figure 5–34 GSM call conformation and call accepted.

PSTN and the country's international facilities in order retrieve information about the location of the mobile from its HLR. One can conceive of various scenarios where a GSM mobile has registered in a MSC/VLR in a foreign country and a call is made to that mobile from the same country that it is now registered in. In the old system, information from the home HLR about the present whereabouts of the mobile would direct the call back to the serving MSC/VLR. Therefore, it was possible for a call to a GSM mobile to be sent back and forth through the international exchanges of different countries when the mobile was actually within the same country as the originating call.

Now, the local exchange where the call is being placed through has the ability to detect a GSM number and directly interrogate the proper HLR for the information needed to locate the mobile subscriber. This process saves a great deal of signaling and unnecessary routing.

Location Updating

The operation used to support the subscriber's mobility within the GSM network is known as location updating. At any given time, the subscriber may receive or initiate a call since the cellular system knows where the MS is located within the network. There are three different types of location updating used in the GSM system. The type of location updating used depends upon the status of the MS. These three location updating operations will be explained here. The three schemes are normal or forced registration, periodic, and ISMI attach. In addition to the location updating function, the MS will also inform the network when it is about to switch to a detached mode.

Normal Location Updating (Idle Mode) The basic steps involved with location updating look very similar to those used for call setup. The steps are radio resource connection establishment, service request, authentication (except for the case of periodic registration), cipher mode setting (depending upon the circumstances), location updating, and then radio resource connection release.

Recall that a location area is defined as a group of cells that is controlled by one or more BSCs but only one MSC. When an MS is in the idle mode, it listens to system information sent over the BCCH. This information includes the location area identity (LAI) of the serving cell. If the MS detects an LAI different from that stored in the SIM card (the value stored at the most recent attachment time), the MS must perform a normal location update. As shown in Figure 5–35, the first step of the process is to perform a radio resource connection establishment operation. Since this radio resource management operation is initiated by the MS, it will be shown here (see Figure 5–36). As shown by Figure 5–36, the MS sends a channel

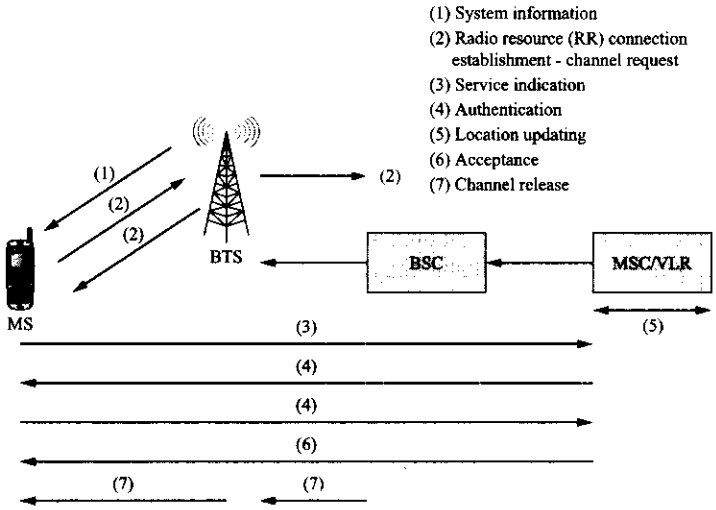


Figure 5–35 GSM location updating (Courtesy of Ericsson).

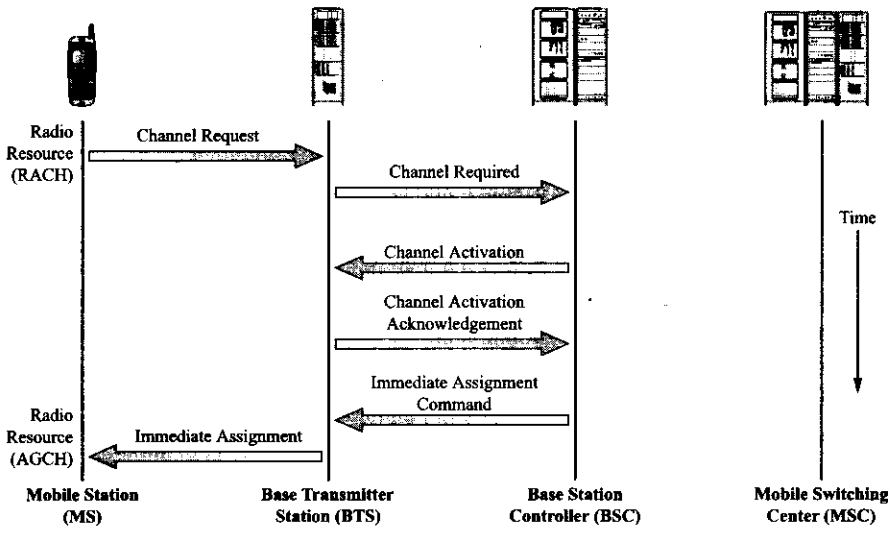


Figure 5–36 GSM location updating.

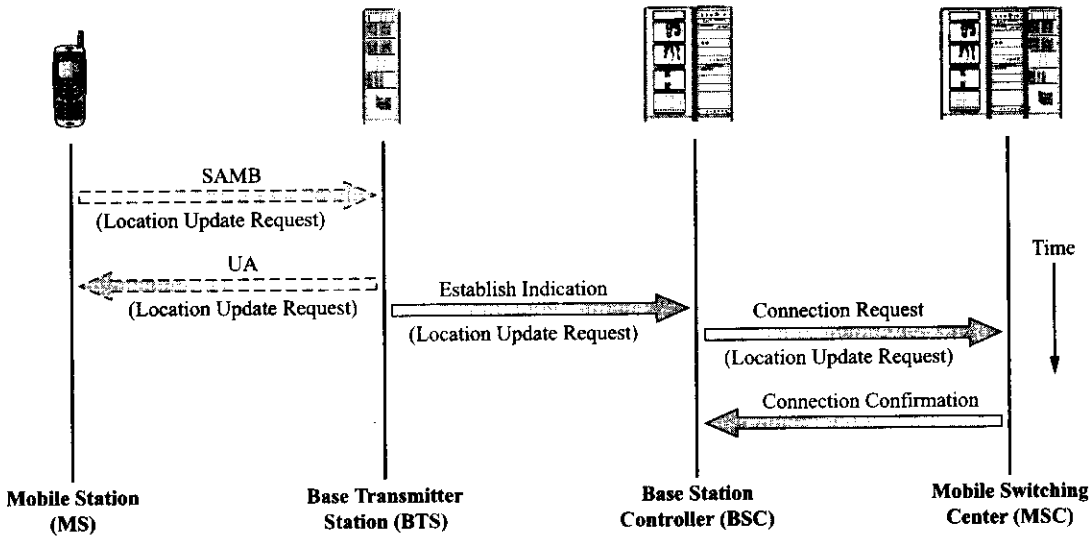


Figure 5–37 GSM location updating service request.

request message over the RACH. The BTS, in turn, sends a channel required message to the BSC. If a free SDCCH is available, the BSC sends a channel activation message to the BSC. Once a channel has been activated, the BSC send an immediate assignment message to the MS and starts a system timer. The reader may want to compare this process to that describer earlier under call setup.

When the MS receives the immediate assignment message, it switches to the ordered channel and sends a service request via an SABM message that contains a location updating request message to the BTS (see Figure 5–37). The message is looped back to the MS via a UA message for reasons mentioned previously and also forwarded to the BSC within an establish indication message. When this message arrives at the BSC the timer is disabled and the message is forwarded to the MSC within a connection request message. The location updating request message will include the old MS location and the new cell location (via the CGI number). The MSC acknowledges this Layer 3 information by sending a connection confirmed message back to the BSC.

Authentication and ciphering mode settings operations are similar to those performed during call setup described previously. Authentication is normally performed for new visitors to a MSC/VLR. Since selective authentication is normally employed by the system, the MSC/VLR will perform a check of the exchange properties to determine if authentication must take place. If the MSC/VLR needs to contact the HLR, this process may be delayed. Ciphering mode setting may or may not be activated depending upon the status of the MS. If a periodic updating is being performed, ciphering mode setting need not be activated since it has already been performed by the system.

The next step in the location update process is shown in Figure 5–38. If the MSC/VLR accepts the location updating, the MSC/VLR sends the location updating accepted message transparently through the BSC and BTS to the MS over a SDCCH. The message sent by the MSC/VLR may contain a new TMSI number. If this is the case, the MS responds by sending a TMSI reallocation complete message transparently back to the MSC/VLR. When a new TMSI is sent to the MS, a timer is enabled in the MSC. When the MS sends its acknowledgement of the new TMSI back to the MSC, the timer is disabled when the MSC receives the message. If the location updating request is rejected for whatever reason (IMSI unknown in HLR, "black listed" MS, etc.), the MSC sends a location updating reject message to the MS. The radio resource connection is released and the MS may be put into an idle state with only emergency call functionality.

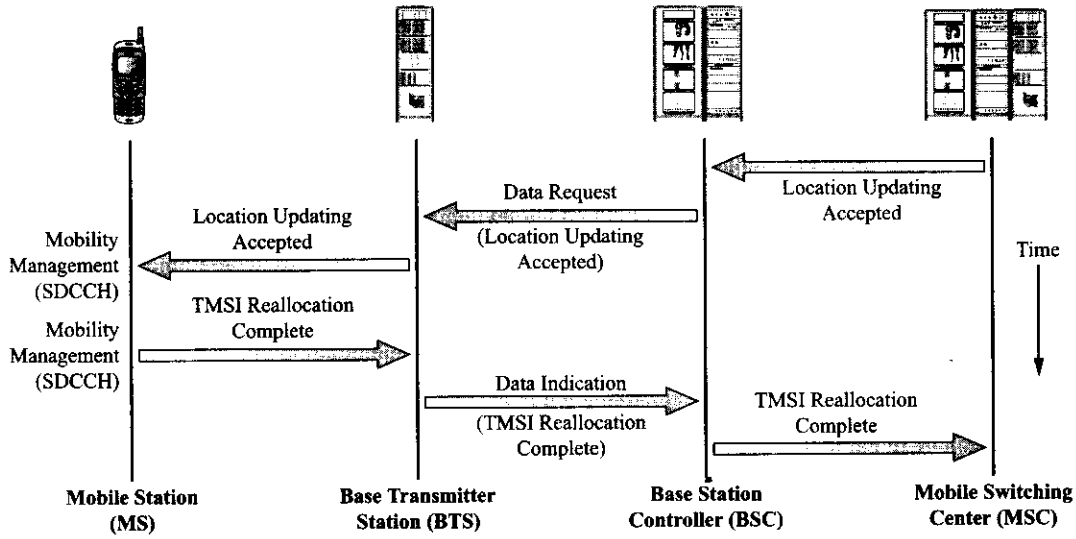


Figure 5–38 GSM location updating accepted.

The last step in the location updating process occurs when the radio resource connection is released. This process is identical to the call release operation already discussed. Figure 5–39, which shows this operation, is included here for purposes of continuity.

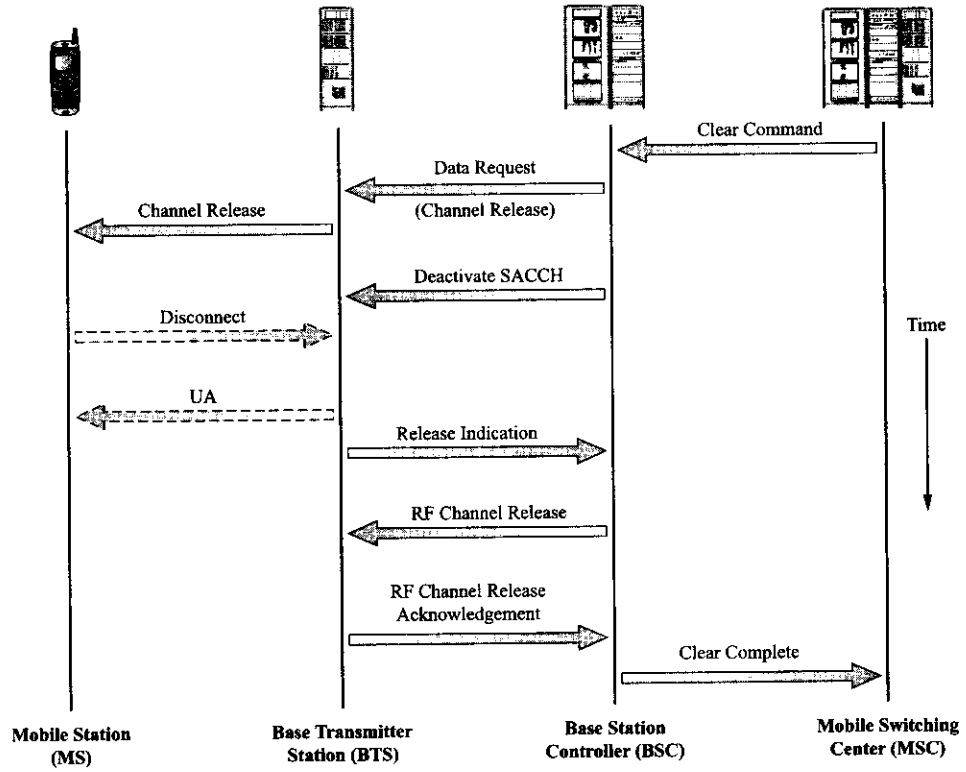


Figure 5–39 GSM connection release.

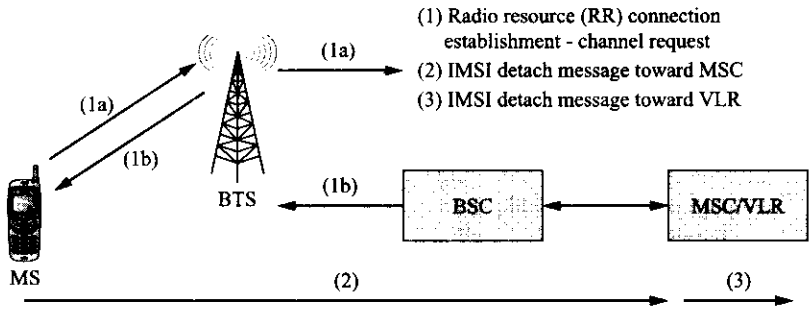


Figure 5–40 GSM IMSI detach (Courtesy of Ericsson).

If the mobile is in an active mode as it changes location area, the process just described must be delayed until after the call is released and it returns to the idle state. In this case, the mobile will have received the new LAI number over a SACCH.

IMSI Detach/Attach Location Updating Depending upon the GSM system, the MS may use the IMSI detach procedure when powering off. This process is shown by Figure 5–40. When the MS power is being turned off, the mobile requests an SDCCH. Over the SDCCH, the MS sends a message to the network that it is about to enter the detached state. The MSC denotes the MS status (IMSI detached) in the VLR. The VLR will reject incoming calls for the MS sending a voice message back to the caller that the subscriber is currently unavailable. Alternately or additionally, the VLR can send a message to the HLR indicating the detached condition of the subscriber. If the subscriber has voice mail, the caller will be directed to leave a message for the subscriber.

The IMSI attach procedure is the complementary operation to IMSI detach. If the MS is powered on in the same location area where it performed an IMSI detach, then the following operations take place as shown in Figure 5–41. The MS requests a SDCCH, the system receives the IMSI attach message from the MS, the MSC passes the attach message on to the VLR. The VLR returns the MS to active status and resumes normal call handling for the MS. The MSC/VLR returns an IMSI attach acknowledgement message to the MS. If the IMSI detach process caused the HLR to be updated with the MS's detached status, the normal location updating will have to be performed by the mobile.

If the mobile has changed location area while in the detached mode, it will also have to perform a normal location updating when it is switched on again. The signaling used to perform an IMSI attach is basically identical to that of a normal location updating.

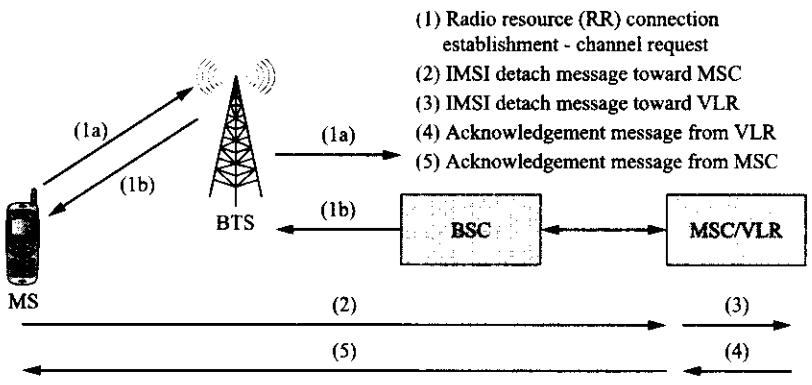


Figure 5–41 GSM IMSI attach (Courtesy of Ericsson).

Periodic Location Updating Periodic location updating is used to prevent unnecessary use of network resources such as the paging of a detached MS. If the system uses periodic registration, the mobile is informed how often it must register. Timers in both the MS and MSC control this operation. When the MS timer expires, the MS performs a location updating that does not require all of the steps involving authentication and ciphering mode setting operations. If the timer in the MSC expires before the MS performs a location updating, the MSC denotes the MS as detached. If the updating operation is performed on time, the MSC sends an acknowledgement to the MS. Any time the MS is activated within the system the periodic updating timers are reset.

In a related operation known as VLR purge, a subscriber that was registered in an MSC/VLR may be deactivated by the MSC/VLR due to lack of activity. For example, a particular subscriber has been detached for more than twenty-four hours, in order to avoid unnecessary signaling within the network; the VLR informs the HLR that the subscriber is no longer available. The HLR denotes the subscriber as unavailable and returns an acknowledgement (purge MS) message to the VLR. At this point the VLR deletes the subscriber from its database.

Call Handoff

The ability of the GSM cellular wireless system to support a subscriber's roaming mobility is made possible through the operations of location updating and handover. Together, these operations allow the wireless network the capability to locate the mobile outside of its home location and to maintain a connection to the mobile even if the subscriber is rapidly moving about the system. Previous parts of this section have already described the location updating operations of a GSM cellular system.

Handover occurs when an active mobile station changes cells. The BSC for the location area where the mobile is attached makes the decision to have the mobile change cells based on a locating algorithm. This algorithm takes into account the signal strength of the serving cell and the strongest nearby cells. When the calculated results indicate a need to change cells, the BSC initiates a handover. There are several different types of handover scenarios that are possible. The most common handover operations will be described next. For these descriptions, a graphical illustration will be used.

Intra-BSC Handover

The intra-BSC handover is shown graphically in Figure 5–42. In this case the handover is going to occur between two cells that are both controlled by the same BSC. During an ongoing call, the MS makes measurements of the received signal strength (RSS) of its own traffic channel (TCH) and the RSS of the neighboring cells. In Step #1, the MS sends a measurement report about the RSS levels to the BTS at a rate of about two times per second. The BTS also makes measurements of the TCH uplink signal strength and adds these to the measurement report from the MS. In Step #2, the combined report is forwarded to the BSC. The BSC uses a locating function to determine the necessity of handing over the call to another cell because of either poor quality or low signal strength in the cell that the MS is attached to. In Step #3, if handover is deemed necessary, the BSC sends a command to the BTS in the new cell to activate a TCH. In Step #4, when the new BTS acknowledges the activation of the new TCH, the BSC sends a message to the MS via the old BTS with information about the new TCH (i.e., frequency, timeslot, and mobile output power). In Step #5, the mobile tunes to the new TCH and sends short handover access bursts on the appropriate FACCH. At this time, the MS does not use any timing advance. In Step #6, when the BTS detects the handover access bursts, it sends timing advance information to the MS over the FACCH. The BTS also sends a handover detection message to the BSC. The BSC reconfigures the group switch to deliver the traffic to the new BTS. In Step #7, the MS sends a handover complete message to the BSC. In Step #8, the BSC sends a message to the old BTS to deactivate the old TCH and its associated signaling channel (SACCH). In the intra-BSC handover the MSC is not involved with the operations. The BSC would however send a record of the handover to the MSC for the generation of system statistics.

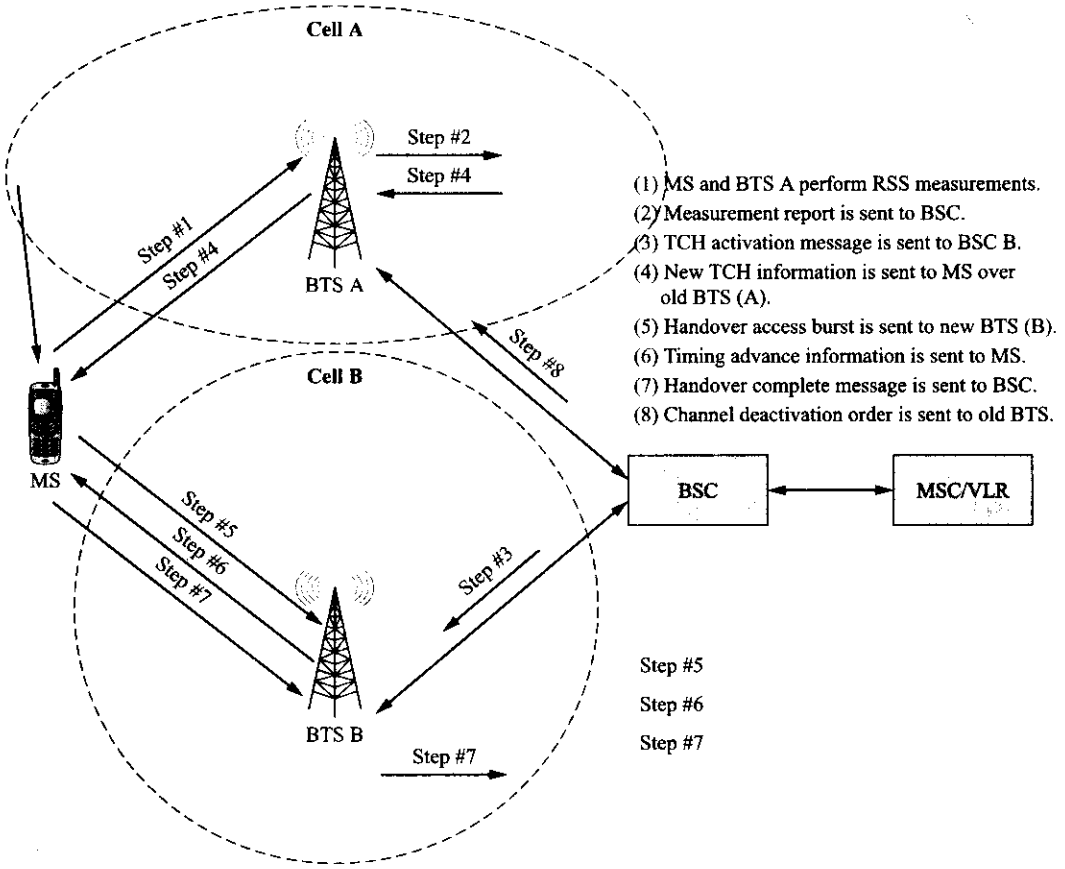
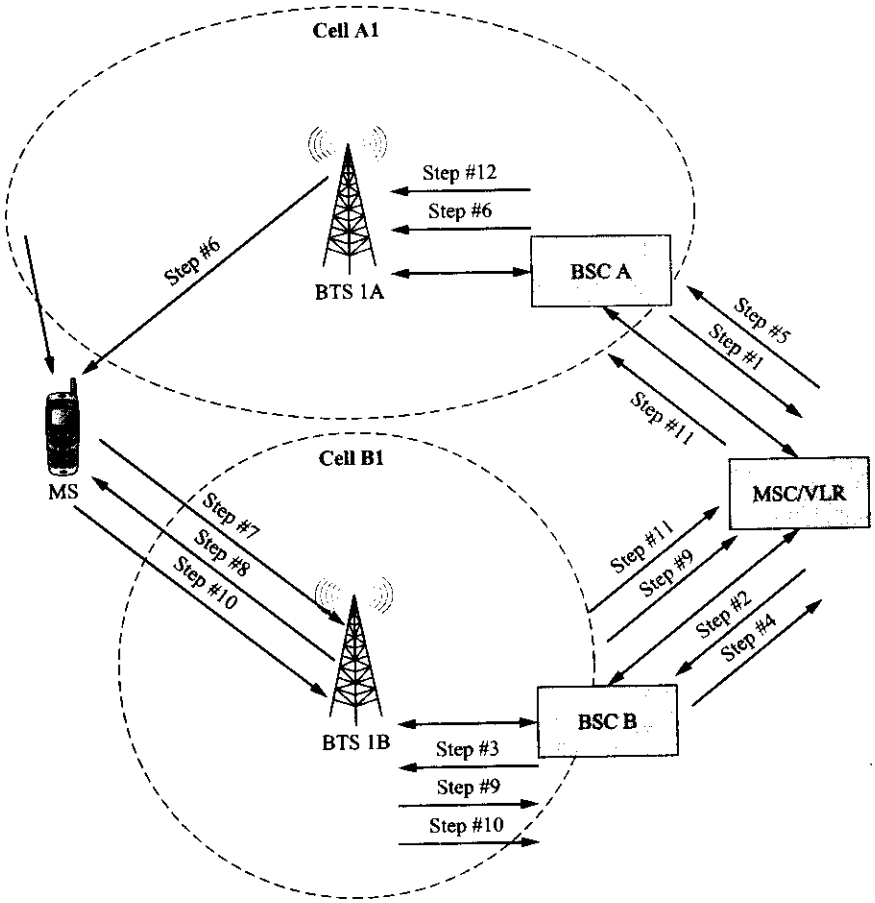


Figure 5–42 GSM Intra-BSC handover (Courtesy of Ericsson).

Inter-BSC Handover

In the inter-BSC handover the mobile has moved to a cell that is in a different location area and therefore has a different BSC. Figure 5–43 shows this situation. Again, the serving BSC decides that the call must be handed over to a new cell that belongs to a new BSC. In Step #1, the serving BSC sends a handover required message to the MSC with the identity of the new cell. In Step #2, the MSC determines the serving BSC for the new cell and sends a handover request to the new BSC. In Step #3, the new BSC sends an order to the new BTS to activate a TCH. In Step #4, when the new BTS activates the TCH, the BSC sends channel information (i.e., frequency, timeslot, MS output power) and a handover reference to the MSC. In Step #5, the MSC passes the channel information to the old BSC. In Step #6, the MS is instructed to change to the new TCH and it also gets the handover reference information contained in a handover command message. In Step #7, the MS tunes to the new TCH and sends handover access bursts containing the handover reference on the new FACCH. In Step #8, the new BTS detects the handover access bursts and sends timing advance information to the MS on the FACCH. In Step #9, the new BTS sends a handover detection message to the new BSC. The new BSC sends a message to the MSC informing it of the handover. The MSC changes the traffic path through the group switch in order to send it to the new BSC. In Step #10, when the MS receives the timing advance information it sends a handover complete message to the BSC. In Step #11, the new BSC sends a handover complete message to the old BSC via the MSC. In Step #12, the old TCH and SACCH are deactivated by the old BTS. The MS gets information about the new cell on the SACCH associated with the new TCH. If the cell is in a new location area, the MS performs a normal location updating after the call has been released.

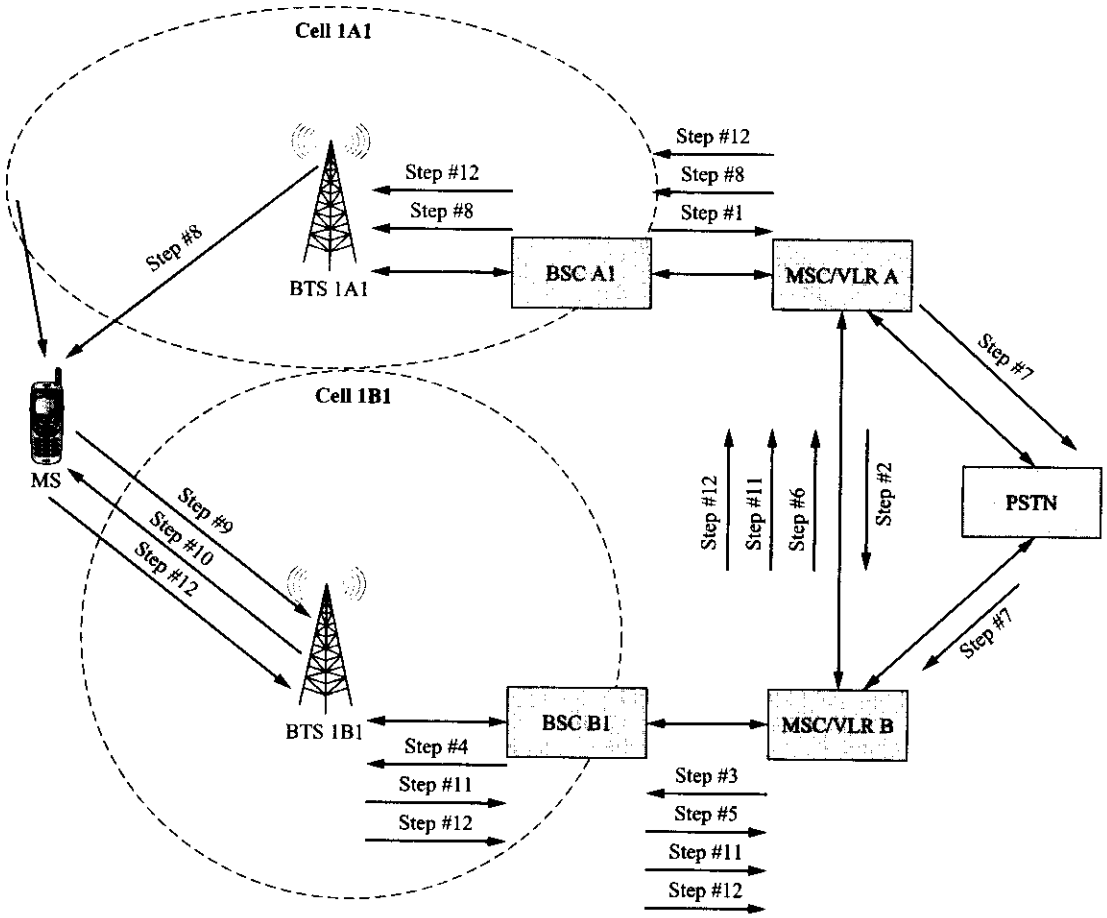


(1) Handover request is sent by serving BSC to MSC.
(2) Handover request is sent by MSC to new BSC (B).
(3) BSC B sends activation order to BTS 1B.
(4) BSC B sends handover information to MSC.
(5) MSC sends handover information to BSC A.
(6) BSC A sends MS new TCH information.
(7) MS sends handover access burst to new BTS (1B).
(8) Timing advance information is sent to the MS.
(9) BTS 1B sends handover detection message to BSC B.
(10) MS sends handover complete message to BSC B.
(11) BSC B sends handover complete message to the old BSC (A).
(12) Old BSC (A) sends channel deactivation message to old BTS (1A).

Figure 5–43 GSM Inter-BSC handover (Courtesy of Ericsson).

Inter-MSC Handover

Another possible handover that can occur is when the BSC decides that a handover should occur and the new cell belongs to another MSC. This type of handover is known as an inter-MSC and is shown by Figure 5–44. For this handover to be performed, Step #1, has the BSC sending a handover required message to the serving MSC as was the case for the inter-BSC handover. In Step #2, the serving MSC asks the new MSC for help. In Step #3, the new MSC allocates a "handover number" in order to reroute the call to



(1) Handover request is sent by serving BSC (A1) to MSC A.
(2) MSC A requests assistance from MSC B.
(3) MSC B provides MSC A with handover number and sends new BSC (B1) a handover request.
(4) New BSC (B1) sends handover activation order to new BTS (1B1).
(5) BSC sends handover information to new MSC.
(6) Handover information is send to old MSC.
(7) A signaling/traffic link is set up between the two MSCs.
(8) Handover message is sent to MS.
(9) MS sends handover access burst to new BTS.
(10) New BTS sends timing advance information to MS.
(11) Old MSC is sent handover detected message.
(12) MS sends handover complete message to new BSC.
BSC sends handover complete message to the old BSC.
Old BSC sends channel deactivation message to old BTS (1A1).

Figure 5–44 GSM Inter-MSC handover (Courtesy of Ericsson).

the new MSC. Also, a handover request is sent to the new BSC. In Step #4, the new BSC sends a command to the new BTS to activate an idle TCH. In Step #5, the new MSC receives the information about the new TCH and handover reference. In Step #6, the TCH description and the handover reference is passed on to the old MSC with the handover number. In Step #7, a signaling/traffic link is set up from the serving MSC to the new MSC. In Step #8, a handover command message is sent to the MS with the necessary information about channel and timeslot to be used in the new cell and the handover reference to use in the handover access burst. In Step #9, the MS tunes to the new TCH and sends handover access bursts on the FACCH. In Step #10, the new BTS detects the handover access bursts and then sends timing advance information to the

MS on the FACCH. In Step #11, the old MSC is informed about the handover access bursts (this info comes from the new BSC and MSC). In Step #12, a handover complete message is sent from the MS. The new BSC and MSC inform the old MSC. The old MSC informs the old BSC and the old BSC sends a message to the old BTS to release the old TCH. In this procedure the old MSC maintains control of the call until it is cleared. In this process, the old MSC is called the anchor MSC.

Since the call entered a new location area, the MS is required to perform a location updating as soon as the call is released. During this operation, the HLR is updated as to the whereabouts of the MS. Also, the HLR will send a cancel location message to the old VLR telling it to delete all stored information about the MS (again, this operation is known as a VLR purge).

Other Handover Operations

There is the possibility of an intercell handover. This can occur when the channel quality is worse than that expected from the RSS measured values. This would entail a change to a new TCH from an old TCH within the same cell. The handover of SMS occurs on the SDCCH. The procedure is identical to that used for the TCH. Also, there is the possible need to hand over the SDCCH during the call setup operation.

5.6 GSM INFRASTRUCTURE COMMUNICATIONS (UM INTERFACE)

The previous sections of this chapter have presented a considerable amount of detail pertaining to the network components and infrastructure, system timing, air interface signal formats, and the operations necessary to provide mobility to the subscribers of GSM wireless cellular networks. Earlier in Section 5.2, a brief overview of the GSM signaling model was presented. This OSI-based signaling model indicated the various interfaces between the GSM network elements and the protocol stacks that serviced these nodes as defined by the technical specifications of GSM. This section will supply some additional details about the type of communications and messages that are sent across the radio link or Um interface and the role of the various protocols in the processing of these messages. In particular, the signaling between peer network layers will be examined starting with Layer 3 of the protocol stack and working downward.

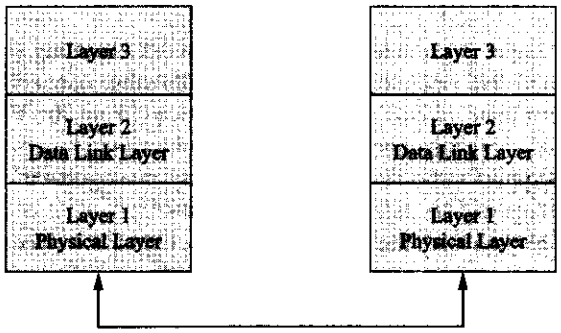


Figure 5–45 Information flow between two nodes in a network.

Additional detail will be provided about the physical layer (Layer 1) signaling across all of the GSM interfaces in Chapter 8 under the general topic of GSM hardware.

Review of GSM Protocol Architecture

Before considering specific examples of GSM peer-to-peer signaling, the reader is referred to Figure 5–45 and Figure 5–46 that illustrate the flow of information between two nodes in a network (e.g., the MS and the BTS across the Um interface). As previously described in Chapter 1, the information from the particular

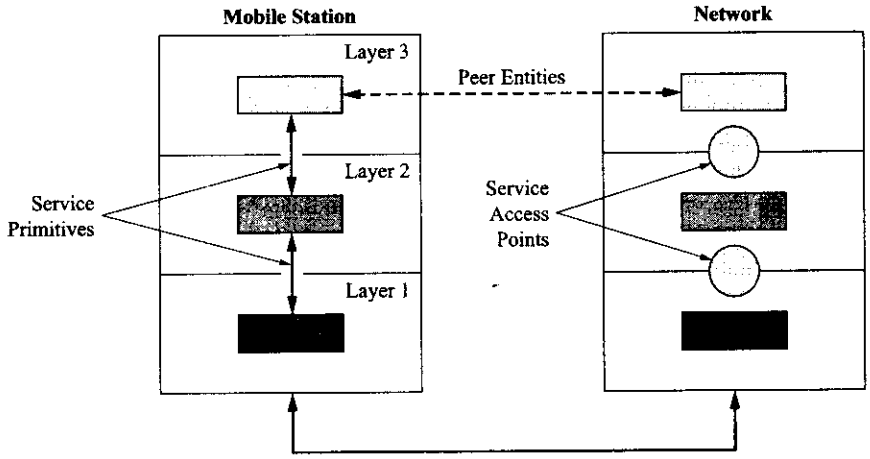


Figure 5–46 Information flow between two nodes in a GSM network (Courtesy of ETSI).

user application is sent down through the protocol stack across the physical interface and up through the protocol stack of the receiving node. In each layer, there exist protocol entities that are responsible for the specific signaling operations and procedures required to complete the transfer of information between nodes. Within the same layers in different nodes, there are peer entities that communicate with each other through the use of a specific protocol. Between adjacent layers, so-called service primitives are used for communication between the different protocol entities. These service primitives provide a means by which the information is carried over the boundary common to the adjacent layers. This information transfer occurs at a **service access point** (SAP); a logical concept defined by the OSI model. The SAP is identified by its service access point identifier (SAPI) value. A familiarity with these concepts will be useful to the reader while learning about the various peer-to-peer operations that will be described next.

Layer 3: Networking Layer Operations

Within the GSM network, Layer 3 provides the mobile network signaling (MNS) service for the mobile subscriber's application. The MNS operations include the following: connection management functions to establish, maintain, and terminate circuit-switched connections from the PSTN to a GSM mobile subscriber; functions to support short message service to the subscriber; functions to support supplementary services; and functions to support radio resource and mobility management operations. The discussion of wireless data service operations will be deferred until Chapter 7.

Within Layer 3, three sublayers with the appropriate protocol control entities must exist to provide these functions. These sublayers are connection management (CM), mobility management (MM), and radio resource management (RR).

Figure 5–47 shows the allocation of the signaling functions at Layer 3 for the Um interface. As the figure indicates, the MS contains all three sublayers and their respective protocol control entities. On the network side of the air interface, the CM and MM protocol entities only reside within the MSC. The RR entity resides primarily within the BSC; however, some RR functions may reside in the BTS and the MSC (hence the parenthesis in the figure). In most cases, RR messages are handled transparently by the BTS and MSC.

Connection Management

The CM sublayer contains functions for call control, call-related supplementary services management, non-call related supplementary service, and short message service. All MSs must support the call control

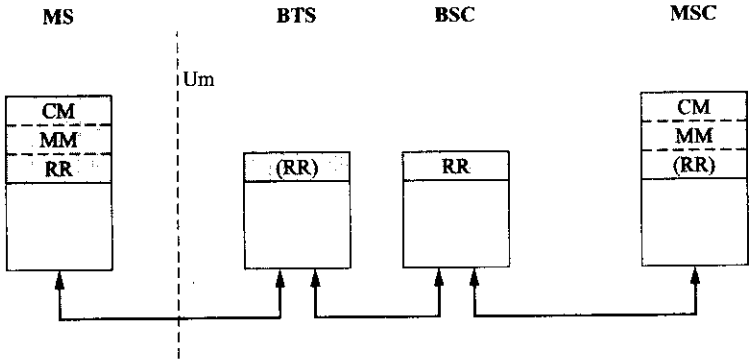


Figure 5–47 Distribution of Layer 3 signaling functions (Courtesy of ETSI).

protocol. The GSM specifications categorize the call control signaling procedures into four subsections: call establishment and clearing procedures, active state procedures, and miscellaneous procedures.

Call Control Call control (CC) procedures are used during call establishment. For a mobile-originated call, the mobile subscriber starts the call establishment procedure by dialing the digits and pressing the send button on the MS keypad. This process is known as a man-machine interface (MMI) procedure. These procedures are mapped onto call control procedures through an exchange of service primitives over the mobile network CC service access point (MNCC-SAP) as shown in Figure 5–48. When a request is made to establish a call, a free or idle CC entity is used to establish a CC connection between the MS and the GSM network. The CC entity initiates the call establishment by requesting the MM sublayer to establish a MM connection. When the MM sublayer confirms the establishment of a MM connection, the CC entity sends a setup message to its peer entity in the MSC. Other CC messages are then exchanged as needed. After connect and connect acknowledgement messages have been exchanged, the two peer sublayers enter an active state and the call establishment signaling phase is complete. When a mobile-terminating call occurs, the CC entity used to establish a connection between the network and the MS is located in the MSC. Call clearing procedures are initiated through the sending of a disconnect message by the CC entity. After the exchange of release/release complete messages, the MM connection is released and the CC entities return to an idle or null state. Detailed descriptions of the messages exchanged during these call establishment and call clearing procedures have been given in the previous section and in Chapter 3. During the active state, a CC entity may send a message to inform its peer entity of some type of call-related event or call rearrangement via a notify or modify message. Miscellaneous other CC procedures such as congestion control, call status, and DTMF are also possible.

It is possible to have parallel CC transactions taking place through the existence of more than one CC entity. This is shown in Figure 5–49. The CC entities are independent of one another and communicate with their peer entities through separate MM connections. The figure shows four CC entities: two call control entities, one SMS entity, and one supplementary service entity. The different CC entities use different transaction identifiers.

CC entity messages can be grouped into call establishment, call information, call clearing, messages for supplementary service control, and miscellaneous messages.

Short Message Service Support Short message service entities known as short message control (SMC) use short message control protocol (SM-CP). These entities are used to transfer short messages between the MS and the MSC. As shown in Figure 5–48, the SMC entities provide service to the SMS application through the mobile network SMS service access point (MNSMS-SAP).

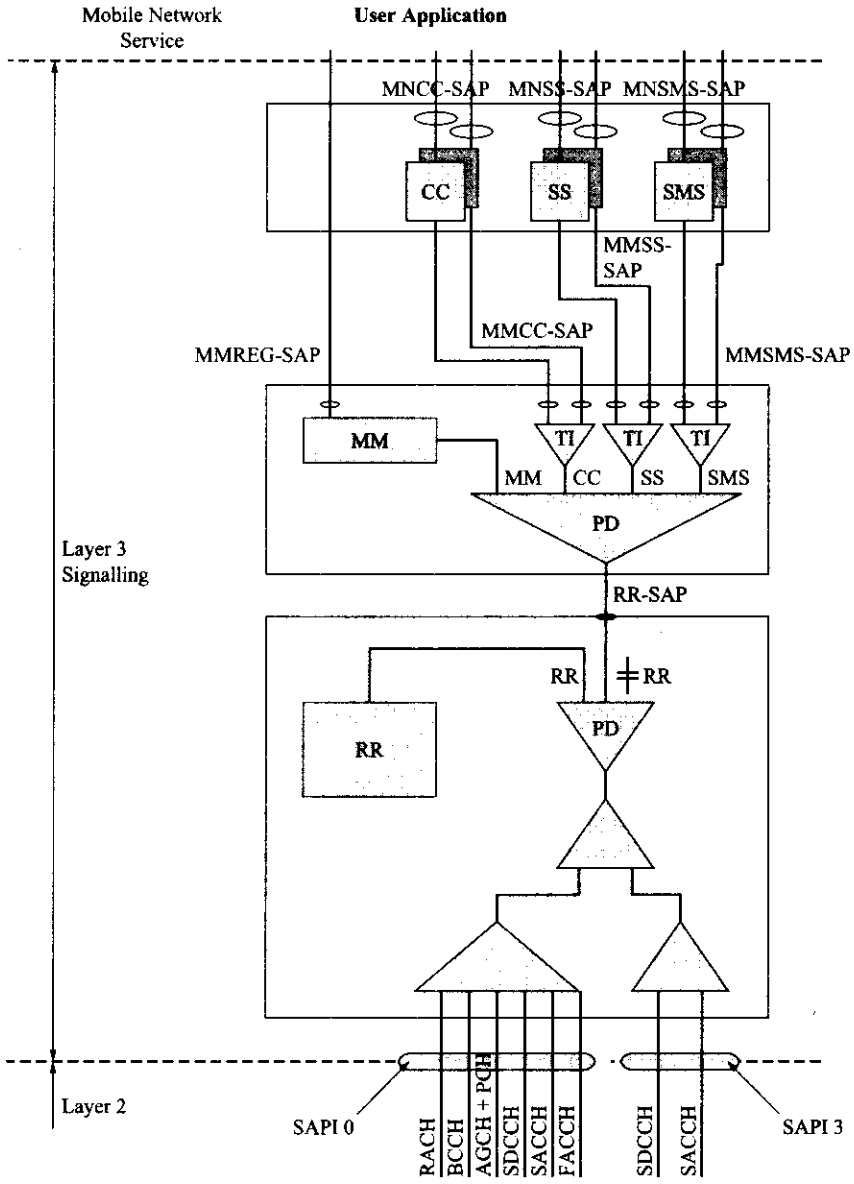


Figure 5–48 Call control procedures (Courtesy of Ericsson).

Supplementary Services Support Supplementary Services (SS) handle services that are not related to a specific call. Examples are call forwarding and call waiting. This information is transferred to the HLR through messages related to the appropriate service. The SS entities provide service through the mobile network SS service access point (MNSS-SAP).

Mobility Management

The mobility management sublayer performs three types of procedures that are related to mobility support, subscriber confidentiality, and service of the CM entity. The mobility support procedures are known as MM specific procedures. They include the different location updating types discussed in the prior section

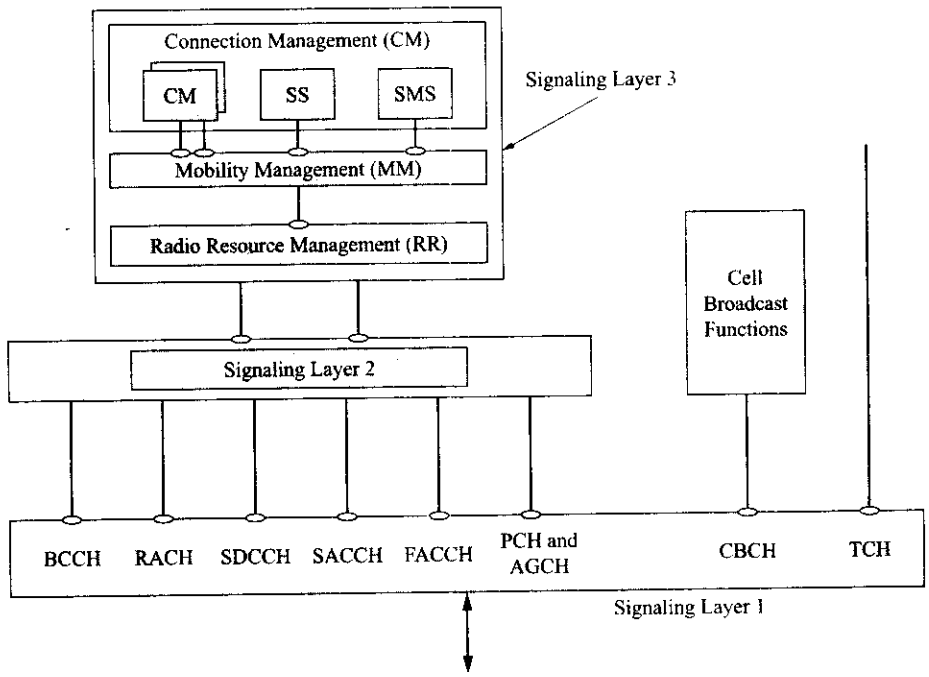


Figure 5–49 Parallel call control operations (Courtesy of Ericsson).

of this chapter. The MM procedures used to provide subscriber confidentiality include authentication, TMSI reallocation, and MS identification through IMSI or IMEI. The MM connection management procedures also provide service to the different entities in the CM sublayer. These services allow the CM entity the ability to use an MM connection for communication with its peer entity either in the MS or the MSC. When the MM sublayer receives a request for a MM connection from a CM entity, the MM sublayer sends a request message for the establishment of a RR connection to the RR sublayer. After the RR connection is established, the network may start the MM procedure of authentication and TMSI reallocation and the network may also ask the RR sublayer to perform ciphering mode setting. After the successful completion of these MM and RR procedures, the MM connection establishment is finished and the CM entity that requested the MM connection is informed that it exists.

Some of the types of messages used by MM are registration messages, security messages, connection management messages, and miscellaneous messages.

Radio Resource Management

The radio resource sublayer receives service from Layer 2 and provides service to the MM sublayer. Additionally, the RR sublayer communicates directly with Layer 1 for the exchange of information related to measurement control and channel management. The primary function of the RR procedures is to establish, maintain, and, when no longer needed, release a dedicated connection between the MS and the BTS (i.e., the wireless network). To achieve this end, the RR procedures include cell selection at power on, handover, cell reselection during idle mode, and recovery from lack of service during idle mode. The cell selection procedures are performed in conjunction with Layer 1 in fulfillment of GSM Phase 2 recommendations for PLMN selection. The MS RR functions include the procedures for the reception of BCCH and CCCH when in idle mode, and the network RR functions include the broadcasting of system information and the continuous transmission of paging information on all paging subchannels to MSs in the idle mode.

The establishment of an RR connection may be initiated by the MS or the network. On the MS side, the MM sublayer requests the establishment of an RR connection or on the network side a RR entity transmits a

paging message to the MS. In either case, the MS's RR entity transmits a channel request message that asks for a signaling channel. The network responds by allocating a dedicated channel to the MS by sending an immediate assignment message. There is an exchange of Layer 2, SABM and UA frames, and the RR connection is established. The MM sublayers in both the MS and in the network side are informed that an RR connection exists.

While in the RR connected mode, many operations can take place. Some of these operations are as follows: entities in upper layers can send messages to their peer entities, the RR sublayer on the network side sends system information on the downlink radio channel, the RR sublayer in the MS sends RSS measurement reports on the uplink radio channel, the network may use the RR ciphering mode setting procedure for setting the ciphering mode, the network side RR sublayer may request an intercell change of channel or change in channel mode (i.e., coding, decoding, and transcoding modes), and MS classmark information may be exchanged.

Only one RR connection can be established for one MS at a time. To release the RR connection, a normal release procedure may be initiated or a radio link time-out procedure may take place. Some of the types of RR connection messages are change establishment messages, ciphering messages, handover messages, channel release messages, paging messages, system information messages, and miscellaneous messages.

Message Format for Layer 3

The format of a GSM Layer 3 message is shown in Figure 5–50. The message consists of a sequence of a number of 8-bit bytes of information. As shown in the figure, the first field or header, consisting of 4 bits of the first byte, is a protocol discriminator (PD) code that indicates the type of protocol the message belongs to (i.e., CC or call-related SS, MM, RR, SMS, or non-call related SS). The next field of 4 bits is a transaction identifier (TI) or skip indicator. For every CM message, the TI identifies the particular CC transaction that is taking place within the CM sublayer. For both MM and RR connection messages the field is set to all zeros (0000). The next byte of the message indicates the function of the message within the specified protocol and for a given direction (i.e., the message meaning changes depending upon its direction; MS to network or network to MS). As shown in Figure 5–50, additional bytes of information are provided as necessary.

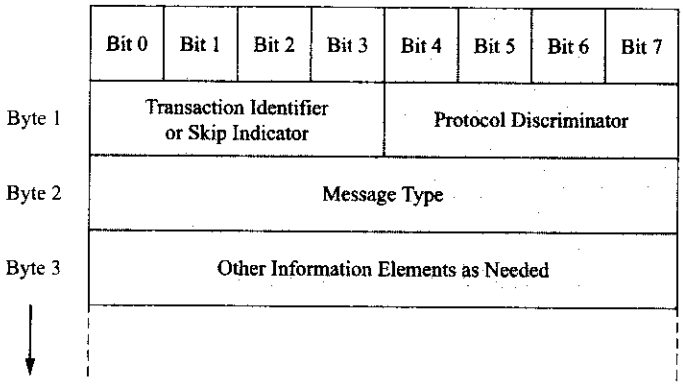


Figure 5–50 Format of a GSM Layer 3 message.

In the RR and MM connection sublayers, functions related to the transport of messages are defined. The task of RR and MM sublayers is to examine the message header to determine the correct routing of the message either to or from the correct protocol entity by virtue of its PD and TI codes. Refer to Figure 5–48 again.

Layer 2: Data Link Layer Operations

As discussed previously, link access procedures on the Dm channel (LAPDm) is the Layer 2 protocol used to carry signaling information between Layer 3 entities over the air interface. The designation of the Dm channel refers to any of the control channels discussed in Section 5.3. Each logical channel is allocated a separate protocol entity as shown in Figure 5–51. Only the RACH control channel does not use LAPDm. For RACH, LAPDm serves as an interface between Layer 3 entities and the physical layer (Layer 1). LAPDm is a protocol that is used at the data link layer of the OSI model. The purpose of this layer is to provide a reliable signaling link. Layer 2 receives services from the physical layer and provides services to Layer 3.

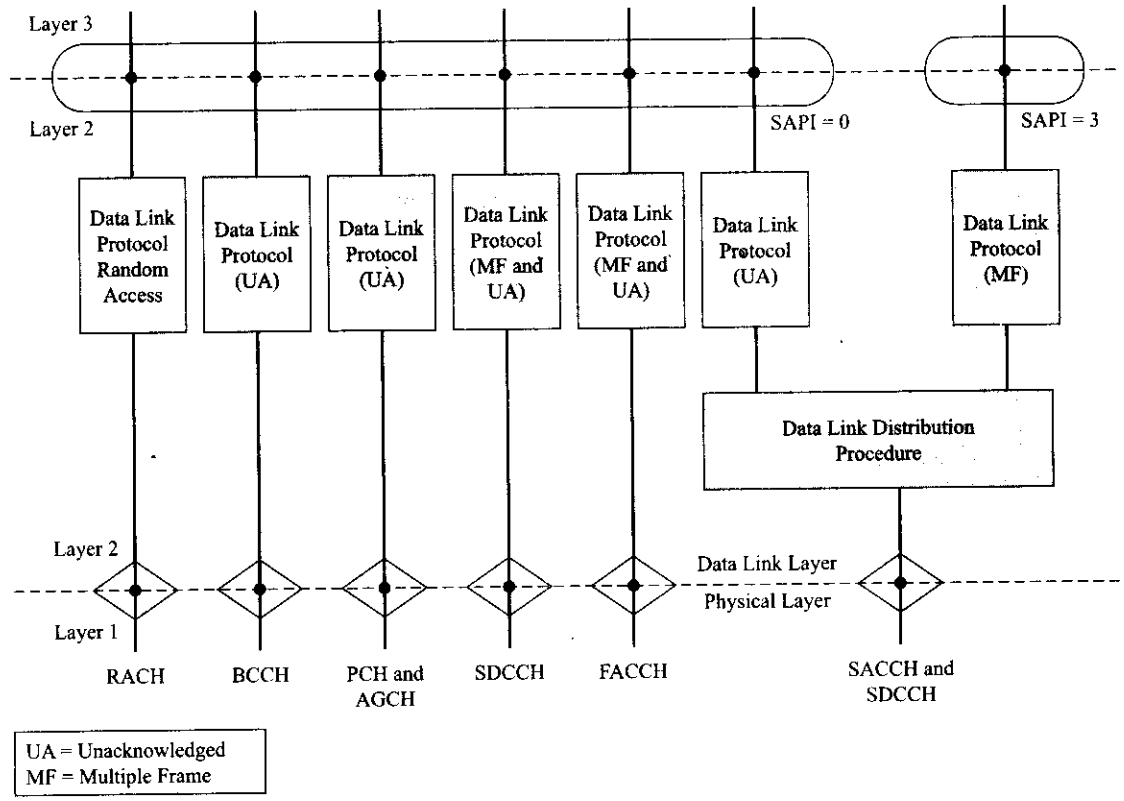


Figure 5–51 GSM protocol entities (Courtesy of Ericsson).

LAPDm used over the Um interface is a modified version of LAPD, the ISDN protocol. LAPD is used on the GSM Abis interface between the BSC and the BTS. LAPD messages can have up to a maximum of 260 bytes per frame. The LAPD message consists of a header and a Layer 3 message. For a transparent message from the network to the MS, the BTS removes the header information and the remaining bytes of data (251 maximum) are sent to the MS. However, the message is often too long for a single frame on the air interface. Therefore, LAPDm segments the message into a number of smaller messages and these messages are sent over either four bursts or eight half-bursts after undergoing convolutional coding to provide error correction capabilities. More detail will be provided about this process in Chapter 8.

LAPDm Operations

LAPDm supports two types of operation on the data link: unacknowledged and acknowledged. Messages that do not need to be acknowledged are sent via unnumbered information (UI) frames. This implies that

there is no error recovery or flow control operation in place during information transmission. Acknowledged operation occurs when information is sent within multiple frames. Layer 3 messages are sent in numbered I (information) frames. Each I frame must be acknowledged before the next frame may be sent. The multiple frame mode is initiated with the set asynchronous balanced mode (SABM) command.

LAPDm adds overhead information depending upon the type of frame to be sent. The different control channels use different frame formats. Some of the types of information contained within the frame fields are as follows: address, information format type, and length indicator. The details of the various frame fields will not be discussed here. There are numerous references about LAPD protocol available if the reader desires more information about this topic.

Service Access Points

The **service access points** (SAPs) of a layer are defined as the gateways through which services are offered to adjacent higher layers. The SAP identifier (SAPI) between Layer 3 and Layer 2 has a specific value for each of the functions on the Dm channel. As shown in Figure 5–51, SAPI = 0 for CC, SS, MM and RR signaling and SAPI = 3 for SMS. Between Layer 2 and Layer 1 there are SAPs defined for each control channel. In the GSM specification, the RR sublayer of Layer 3 controls the establishment and release of the SAPs between Layer 1 and Layer 2. This procedure differs from the OSI reference model where this function is performed at the data link layer.

Data Link Procedures

Figure 5–51 shows a functional block diagram of the data link layer in the MS. As shown by the diagram only the data link connections for SDCCH and SACCH can terminate at SAPI = 3; all other control channels terminate at SAPI = 0. The diagram also shows the three types of procedures that can be supported for the control channels. They are as follows: data link procedure, data link distribution procedure, and random access procedure. The data link procedure is performed once on each type of physical channel that is supported by the SAPI. The procedure examines the frame for the control field and the length indicator field. The procedure performs segmentation and reassembly of the Layer 3 message. The data link distribution procedure is invoked whenever there is more than one SAPI on a physical channel. The procedure examines the address field of the frame and the type of physical channel to determine the correct data link block to deliver the information to. The procedure also provides contention resolution for various data link procedure blocks on the same physical channel. The random access procedure is used for data links on the RACH. The procedure in the MS formats the random access frames and initiates the transmission of these frames. The BTS receives the frames and provides the appropriate indication to Layer 3.

Physical Services Required by the Data Link Layer

The data link layer requires the following services from the physical layer: frame synchronization, error protection and correction to ensure a low BER in the data link layer, transmission and reception by the MS and BTS, respectively, of random access bursts, and a physical layer connection that provides for the arrival of bits and frames in the same order as they were transmitted to the peer entity on the receiving side.

Data Link Timers

There are several system timers and counters used to keep track of the waiting time for the acknowledgement of a previously transmitted message and the number of times that retransmission may take place. The functions and names of these elements can be found in the LAPD specifications.

Layer 1: Physical Layer Operations

The physical layer or signaling Layer 1 is the actual physical hardware, modulation schemes, channel coding, and so forth used to send the bits over the physical channels on the air interface. The physical layer

interfaces with the data link layer (Layer 2) through the various control channels. Additionally, the physical layer interfaces with other physical units such as speech coders and terminal adaptors for the support of traffic channels (refer back to Figure 5–49). Furthermore, the physical layer provides services to the radio resource management sublayer through the assignment of channels (i.e., mapping of logical channels on to physical channels) and monitoring and the measurement reports that are sent to the RR sublayer about channel quality and RSS.

The GSM physical layer operations include various channel coding techniques, bit and frame interleaving of both traffic and control channels, ciphering, and burst formatting and modulation for the transmission of information and the complementary functions for the reception of the transmitted information. The details of these operations will be covered in Chapter 8 under the topic of GSM hardware.

Other Layer 1 operations include the setting of the timing advance as ordered by the network, power control functions, synchronization of the mobile receiver, cell selection strategy, and handover functions. These topics also will be given further coverage in Chapter 8.

PART III OTHER TDMA SYSTEMS

5.7 NORTH AMERICAN TDMA

At this time, North American TDMA (NA-TDMA) is deployed mainly in the Americas (North and South America). An interactive map of TDMA coverage is available at the following Web site: www.3Gamericas.org. Presently, there are over 110 million NA-TDMA subscribers, which represent slightly fewer than 9% of the total worldwide cellular users. A large portion of the recent growth in NA-TDMA has occurred in the Latin American countries. There are predictions that NA-TDMA will experience continued growth during this decade and despite the continued evolution of the cellular industry toward 3G networks, NA-TDMA is thought to be a technology that will be viable for another decade—however, just not as a 3G technology. In the United States, AT&T has announced that it will maintain its TDMA service indefinitely; however, AT&T is building a new GSM/GPRS network at 800/1900 MHz. Cingular is also converting to GSM/GPRS and EDGE technology.

As discussed earlier, NA-TDMA was developed as a true second-generation cellular system (recall D-AMPS discussed in Chapter 2) for use on the 800-MHz band and then on the 1900-MHz PCS band. The first implementation of NA-TDMA (i.e., 2G) does not support packet data transfer. NA-TDMA technology is very similar to GSM but it is not compatible with it since it uses a different timeslot and frame structure over the air interface. The specifications for the 3G version of NA-TDMA were published in August of 2001 and are described by the standard known as TIA/EIA-136-440-1. The 3G version of NA-TDMA has added an additional air interface standard that is GSM compatible to achieve the packet data transfer rates called for by the 3G standard. Therefore, the NA-TDMA mobile stations will have to be able to handle both air interface standards (i.e., dual band and dual mode) to be able to receive high-speed packet data. The use of GPRS for high-speed data over this second air interface is a step toward the eventual adoption of one universal 3G standard.

TIA/EIA-136 Basics

As already mentioned, the NA-TDMA system is very similar in its operation to GSM. Therefore, the treatment of this wireless technology will be brief and only highlight the major differences between it and GSM. Like GSM, the information transmitted over the air interface undergoes convolutional coding, interleaving, and so on. However, the air interface for TIA/EIA-136 consists of six timeslots per frame instead of the eight timeslots used by GSM. TIA/EIA-136 uses the same identification numbers as AMPS, the ESN, SID, and MIN, but also uses the GSM identifiers introduced earlier in this chapter. The frequency

Table 5–4 NA-TDMA channel allocations in the PCS bands.

Band	*Bandwidth (MHz)*	*Number of Channels*	*Boundary Channel Number*	*Frequency MS (MHz)*	*Frequency BS (MHz)*
A	15	499	1 499	1864.980 1850.040	1944.960 1930.020
D Not used	5	165 1	501 665 500	1865.040 1869.960 1865.010	1945.020 1949.940 1944.990
B Not used Not used	15	498 1 1	668 1165 666 667	1870.050 1884.960 1869.990 1870.020	1950.030 1964.940 1949.970 1950.000
E Not used Not used	5	165 1 1	1168 1332 1166 1167	1885.050 1889.970 1884.990 1885.020	1965.030 1969.950 1964.970 1965.000
F Not used Not used	5	165 1 1	1335 1499 1333 1334	1890.060 1894.980 1890.000 1890.030	1970.040 1974.960 1969.980 1970.010
C Not used	15	499 1	1501 1999 1500	1895.040 1909.980 1895.010	1975.020 1989.960 1974.990

allocations for TIA/EIA-136 in the 800-MHz band are the same as the original AMPS channels shown in Chapter 2 with the same 30-kHz channel bandwidths. For TIA/EIA-136 systems that operate in the 1900-MHz PCS band, the channel allocations are as shown in Table 5–4. The channel spacing is 30 kHz with approximately 80-MHz spacing between base and mobile transmitting frequency. The relationship between channel number and the transmitter center frequency is given by the following equations:

$$\text{TIA/EIA-136 Mobile} \quad \text{Transmit Frequency} = 0.030N + 1850.010 \text{ MHz}$$

$$\text{TIA/EIA-136 Base} \quad \text{Transmit Frequency} = 0.030N + 1929.990 \text{ MHz}$$

TIA/EIA-136 Channel Concept

As discussed in Chapter 2, the D-AMPS system afforded the wireless cellular service providers an option to use already allocated AMPS traffic channels as conventional AMPS channels or as TDMA channels. However, the control channels retained their analog nature. The NA-TDMA system replaces the analog control channels with digital TDMA control channels that are, not surprisingly, known as digital control channels (DCCHs). Figure 5–52 shows the TIA/EIA-136 channel organization. The digital traffic channels (DTCs) are divided into two groups: the forward (downlink) DTCs and the reverse (uplink) DTCs. The digital control channels are also divided into two groups: the forward and reverse digital control channels. The

Digital Traffic Channels (DTC)		**Digital Control Channels (DCCH)**					
Reverse DTC	Forward DTC	Uplink (Reverse DCCH)		Downlink (Forward DCCH)			
Fast ACCH	Fast ACCH	RACH		SPACH	BCCH	SCF	Reserved
Slow ACCH	Slow ACCH			PCH	Fast BCCH		
TRAFFIC	TRAFFIC			ARCH	Extended BCCH		
				SMS Channel	SMS BCCH		

ACCH—Associated Control Channel
ARCH—Access Response Channel
BCCH—Broadcast Control Channel
PCH—Paging Channel
RACH—Random Access Channel
SCF—Shared Channel Feedback
SPACH—SMS Point-to-Point, Shared, ACKed Channel

Figure 5–52 NA-TDMA channel organization.

forward digital control channels are further divided into three groups: the broadcast control channels, the shared channel feedback, and the SMS, point-to-point, paging, and ACKed channels. Table 5–5 shows a listing of these channels with a short summary of the channel function and direction of transmission. As the reader can see, this channel organization is similar to the GSM system channel organization with similar channel names and functionality.

Upon powering up, the TIA/EIA-136 system requires the MS to scan the DCCHs within a cell and, through the use of RSS measurements and a hashing algorithm, select a suitable DCCH. This hashing process selects suitable candidate DCCHs for the MS through a specific identifier known as the paging channel ID (PAID) and other cell characteristics. In an effort to reduce scanning time, the NA-TDMA system standard provides recommended DCCH channel allocations for both the 800- and 1900-MHz bands. As with the GSM system, once the MS locks on to a suitable DCCH it will use broadcast information transmitted by the BTS to become timeslot and frame synchronized.

TIA/EIA-136 Timeslot and Frame Details

The organization of NA-TDMA channels is similar conceptually to that used in GSM, but the implementation of the two schemes is fairly different and therefore will be discussed briefly here.

Figures 5–53 to 5–56 show the hierarchy of the NA-TDMA timeslots and frames. Like GSM, before a typical Layer 3 network control message is sent over a forward DCCH (FDCCH), it is first encapsulated within a Layer 2 LAPDm frame. The frame undergoes channel coding and interleaving operations before it is sent as data during a timeslot burst. As shown in Figure 5–53, the timeslot format for information sent from the BTS to the MS on a DCCH contains a synchronization (SYNC) field used for timeslot synchronization, equalizer training, and timeslot identification; two shared-channel feedback (SCF) fields that carry information about the status of RACH and RDCCH channels; a coded superframe phase (CSFP) field used

Table 5–5 Summary of NA-TDMA channels.

Channel	*Channel Acronym*	*Function*	*Direction of Transfer*
Digital Traffic Channel (DTC)	FDTC RDTC FACCH SACCH	User information and signaling User information and signaling Burst signaling Continuous signaling	BS-to-MS MS-to-BS Up- and downlink Up- and downlink
Digital Control Channel (DCCH)			
Reverse Digital Control Channel (RDCCH)	RACH	Used to gain access to system	MS-to-BS
Forward Digital Control Channel (FDCCH)			
SMS, Point-to-point, Paging, ACKed Channel (SPACH)	PCH ARCH SMSCH	Page MS moves to ARCH after RACH operation Short message channel	BS-to-MS BS-to-MS BS-to-MS
Broadcast Control Channel (BCCH)	F-BCCH E-BCCH S-BCCH	Initialization, exchange IDs, etc. Less time-critical information SMS broadcasts	BS-to-MS BS-to-MS BS-to-MS
Shared Channel Feedback (SCF)		Controls RACH access	BS-to-MS

by the MS to find the start of the superframe; two data fields (data) that compose the message; and a field, RSVD, that is reserved for future system use. As in GSM, there are other timeslot formats depending upon the usage (type of channel, direction, etc.). Several additional examples are shown in Figure 5–54.

One timeslot has a duration of 6.67 ms. There are three timeslots to a TDMA block and two TDMA blocks or six timeslots to a TDMA frame. Each TDMA frame has a duration of 40 ms or a transfer rate of twenty-five frames per second. A **superframe** consists of thirty-two TDMA blocks (sixteen TDMA frames) and is therefore 640 ms long. A hyperframe consists of two superframes. The first superframe is known as the primary superframe and the second is known as the secondary superframe. The information contained in the primary superframe can be repeated in the secondary superframe or the SPACH and E-BCCH information may be different in each superframe (See Figure 5–55). The number of data bits per burst or timeslot varies depending upon the format used within the timeslot. In the downlink direction the data field is 260 bits in length whereas in the uplink direction, the data field is either 200 or 244 bits in length, again depending upon the format of the timeslot.

In GSM, a TDMA frame (equivalent to one physical radio frequency channel) can support up to eight users simultaneously. In NA-TDMA, a TDMA block supports up to three users at a time (i.e., a user gets a timeslot every three timeslots). Therefore, within a frame, the user receives two timeslots (see Figure 5–56). Over a traffic channel, the final data rate for compressed speech traffic is 7950 bps.

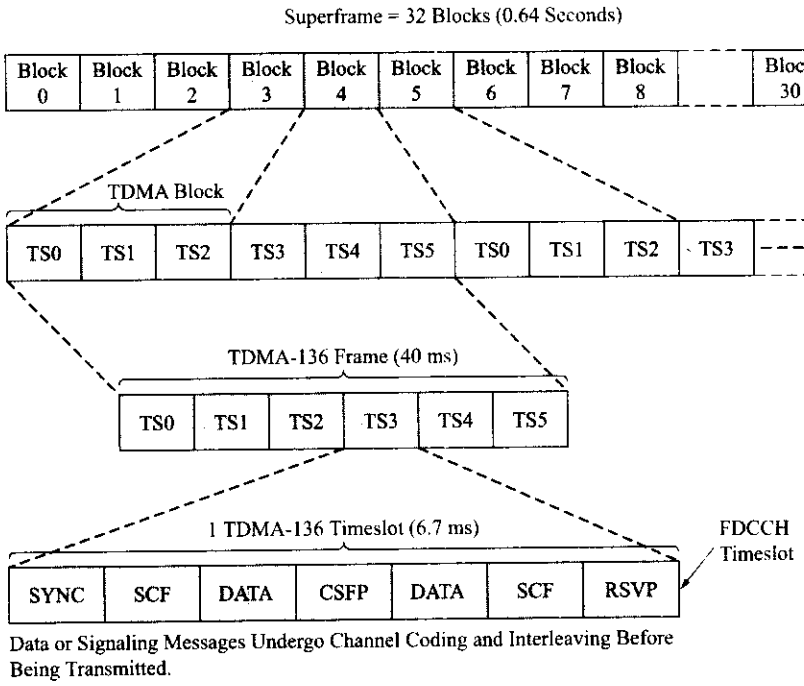


Figure 5–53 Hierarchy of NA-TDMA timeslots.

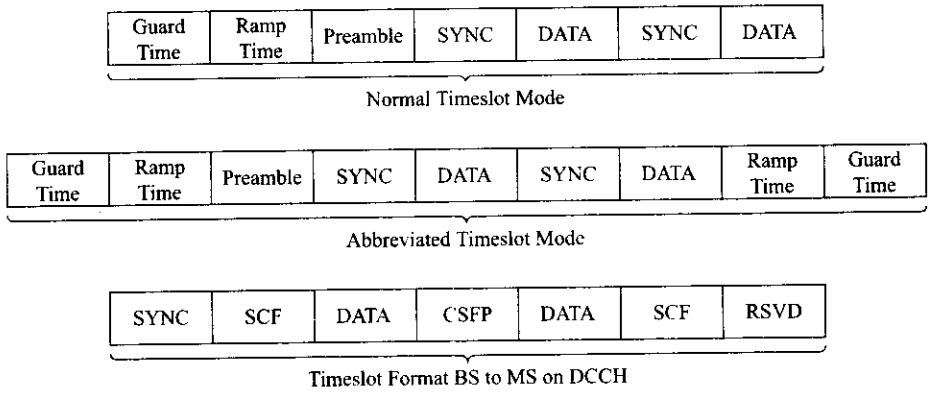


Figure 5–54 NA-TDMA hyperframe structure.

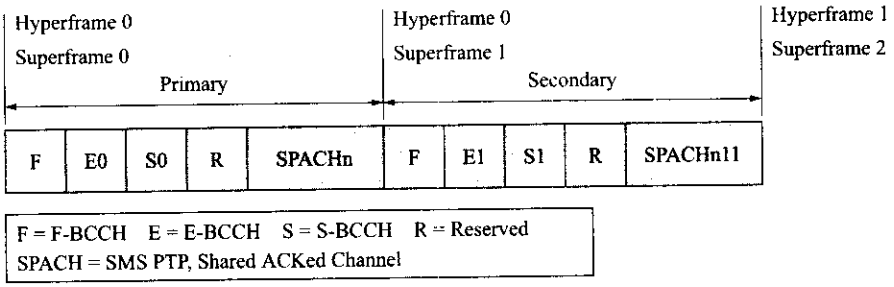


Figure 5–55 NA-TDMA superframe structure.

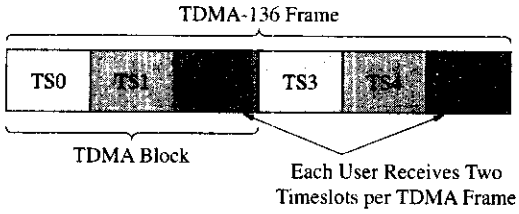


Figure 5–56 NA-TDMA block structure.

This completes our brief coverage of the NA-TDMA system. As is the case with GSM, there are literally hundreds of pages of system specifications and recommendations for NA-TDMA. For further details of NA-TDMA, the reader should consult the most recent TIA/EIA-136 standards available from www.tiaonline.org.

QUESTIONS AND PROBLEMS

1. Describe the TDMA frame structure used by GSM cellular.
2. What is the standard bandwidth of a GSM channel?
3. Name the three major subsystems of a GSM wireless cellular network.
4. What is a GSM SIM card? What purpose does it serve?
5. Describe the Um interface.
6. Why are there two protocol stacks (refer back to Figure 5–6) within the MSC node for a GSM system?
7. Name the subcategories of GSM signaling and control channels.
8. Contrast the digital encoding of voice by a typical vocoder and a PCM telecommunications system.
9. Describe the GSM TDMA timeslot.
10. Contrast the GSM hyperframe, superframe, multiframe, and TDMA frame.
11. Describe a typical normal GSM "burst."
12. What is the purpose of the GSM burst training sequence?
13. What is the purpose of the GSM synchronization burst?
14. What is the function of the GSM access burst?
15. What is the significance of Timeslot 0 on channel c_0 for a GSM system?
16. What is the purpose of the GSM dedicated control channels?
17. What GSM control channel is specifically tasked with the facilitating of the handover operation?
18. How does the GSM mobile station know what paging group it belongs to?
19. What advantages does a GSM half-rate channel offer?
20. Why are there several types of GSM multiframes?
21. The GSM MS roaming number is constructed according to what numbering plan?
22. What purpose does the TMSI number have?
23. Define the "attached" condition for a GSM mobile.
24. Define the "detached" condition for a GSM mobile.
25. What is the purpose of periodic location updating?
26. What is the basic difference between intra-BSC handover and inter-BSC handover?
27. What is the basic difference between inter-BSC and inter-MSC handover?
28. What basic functions are located within the connection management sublayer?
29. What basic functions are located within the mobility management sublayer?
30. What basic functions are located within the radio resource sublayer?
31. What is the function of the various GSM system timers and counters?
32. Why is a modified version of LAPD necessary for the Um interface?
33. What is the fundamental difference between GSM and NA-TDMA in the context of access technology?
34. Contrast the required bandwidth requirements of AMPS, GSM, and NA-TDMA
35. What is the first operation performed by a NA-TDMA mobile upon powering up?