

Cellular Networks

Jingyuan Zhang, *University of Alabama*
Ivan Stojmenovic, *University of Ottawa, Canada*

Introduction	654	Alternatives to Cellular Networks	661
Basic Concepts	654	Summary	661
Multiple Access Methods	656	Acknowledgments	662
Location Management	657	Glossary	662
Hand-Off Strategies and Channel Assignment	658	Cross References	662
Authentication and Encryption	659	References	662
Evolution of Cellular Networks	660		

INTRODUCTION

Cellular communications has experienced explosive growth in the past two decades. Today millions of people around the world use cellular phones. Cellular phones allow a person to make or receive a call from almost anywhere. Likewise, a person is allowed to continue the phone conversation while on the move. Cellular communications is supported by an infrastructure called a cellular network, which integrates cellular phones into the public switched telephone network.

The cellular network has gone through three generations. The first generation of cellular networks is analog in nature. To accommodate more cellular phone subscribers, digital TDMA (time division multiple access) and CDMA (code division multiple access) technologies are used in the second generation (2G) to increase the network capacity. With digital technologies, digitized voice can be coded and encrypted. Therefore, the 2G cellular network is also more secure. The third generation (3G) integrates cellular phones into the Internet world by providing high-speed packet-switching data transmission in addition to circuit-switching voice transmission. The 3G cellular networks have been deployed in some parts of Asia, Europe, and the United States since 2002 and will be widely deployed in the coming years.

This chapter gives an introduction to cellular networks. The rest of this chapter is organized as follows: the second section introduces basic cellular concepts, and the third section describes how the air interface is shared by multiple users. The fourth and fifth sections discuss how to track mobile users and how to assign channels to hand-off calls and new calls, respectively. The sixth section discusses the relevant security issues. The seventh section describes the evolution of cellular networks. The eighth section presents alternatives to cellular networks. The final section gives a summary.

BASIC CONCEPTS

A cellular network provides cell phones or mobile stations (MSs), to use a more general term, with wireless access to the public switched telephone network (PSTN). The service coverage area of a cellular network is divided into

many smaller areas, referred to as cells, each of which is served by a base station (BS). The BS is fixed, and it is connected to the mobile telephone switching office (MTSO), also known as the mobile switching center. An MTSO is in charge of a cluster of BSs and it is, in turn, connected to the PSTN. With the wireless link between the BS and MS, MSs such as cell phones are able to communicate with wireline phones in the PSTN. Both BSs and MSs are equipped with a transceiver. Figure 1 illustrates a typical cellular network, in which a cell is represented by a hexagon and a BS is represented by a triangle.

The frequency spectrum allocated for cellular communications is very limited. The success of today's cellular network is mainly due to the frequency reuse concept. This is why the coverage area is divided into cells, each of which is served by a BS. Each BS (or cell) is assigned a group of frequency bands or channels. To avoid radio co-channel interference, the group of channels assigned to one cell must be different from the group of channels assigned to its neighboring cells. However, the same group of channels can be assigned to the two cells that are far enough apart such that the radio co-channel interference between them is within a tolerable limit. Typically, seven neighboring cells are grouped together to form a cluster, as shown in Figure 2. The total available channels are divided into seven groups, each of which is assigned to a cell. In Figure 2, the cells marked with the same number have the same group of channels assigned to them. Furthermore, the cells marked with different numbers must be assigned different groups of channels.

If there are a total of M channels allocated for cellular communications and if the coverage area consists of N cells, there are a total of $MN/7$ channels available in the coverage area for concurrent use based on the seven-cell reuse pattern. That is the network capacity of this coverage area. Because of explosive growth of mobile phone subscribers, the current network capacity may not be enough. Cell splitting is one technique that used to increase the network capacity without new frequency spectrum allocation (Black, 1996; Rappaport, 2002). In this technique, the cell size is reduced by lowering antenna

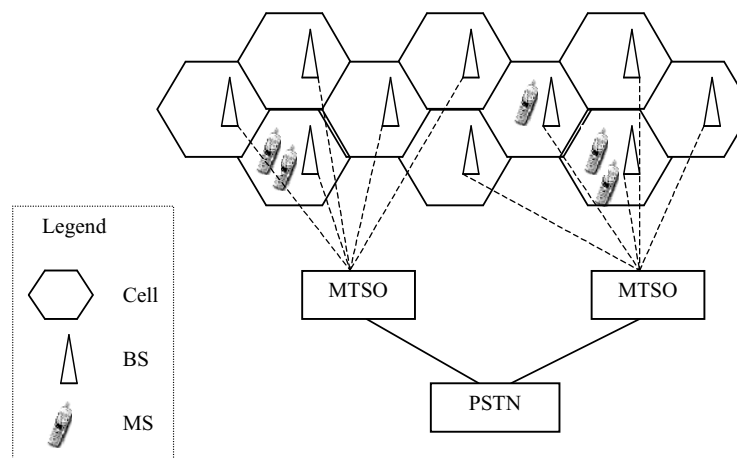


Figure 1: Typical cellular network.

height and transmitter power. Specifically, an original cell is divided into four smaller cells. After cell-splitting, the coverage area with N cells originally will be covered by $4N$ smaller cells. Therefore, the new network capacity is $4MN/7$, which is four times the original network capacity. In reality, bigger cells are not completely replaced by smaller cells. Therefore, cells of different sizes (e.g., pico, micro, and macro cells) may coexist in one area. This allows high-speed subscribers to use bigger cells, which reduces the number of hand-offs (to be explained later).

Sectoring is another technique to increase the network capacity (Black, 1996; Rappaport, 2002). In sectoring, the cell size remains the same, but a cell is divided into several sectors by using several directional antennas at the BS instead of a single omnidirectional antenna. Typically, a cell is divided into three 120° sectors or six 60° sectors. The radio co-channel interference will be reduced by dividing a cell into sectors, which reduces the number of cells in a cluster. Therefore, the network capacity is increased.

Digital technology can also be used to increase the network capacity. Transmission of digitized voice goes through three steps before the actual transmission: speech

coding, channel coding, and modulation. Speech coding is to compress voice. For example, a short voice segment can be analyzed and represented by a few parameter values. These values cannot be transmitted directly because wireless transmission is error prone, and a small change in these values may translate into a big change in voice. Therefore, data representing compressed voice should be arranged carefully, and redundancy should be introduced such that a transmission error can be corrected or at least detected. This process is called channel coding. Finally, the output data from channel coding are modulated for transmission. The detailed information on speech coding, channel coding, and modulation can be found in three corresponding chapters in Rappaport (2002). A good speech-coding scheme combined with a good channel-coding scheme will greatly reduce the amount of bandwidth needed by each phone user and therefore increase the network capacity while keeping the quality of voice unchanged.

The channels assigned to a cell are used either for voice or for control. A voice channel is used for an actual conversation, whereas a control channel is used to help set up conversations. Both voice and control channels are further divided into forward (or downlink) and reverse (or uplink). A forward channel is used to carry traffic from the BS to the MS, and a reverse channel is used to carry traffic from the MS to the BS. The channels assigned to a cell are shared by MSs located in the cell. Multiple access methods are used to share the channels in a cell.

Each MS has a home MTSO, which is the MTSO where the mobile user originally subscribed for wireless services. If an MS moves out of the home MTSO area, it is roaming. A roaming MS needs to register in the visited MTSO. An MS needs to be authenticated against the information kept in its home MTSO before any service can be rendered by the network. The services include making a call, receiving a call, registering the location, and so forth. These services are possible because of a widely used global, common channel-signaling standard named SS7 (Signaling System No. 7) (Modarressi & Skoog, 1992; Rappaport, 2002).

To make a call from an MS, the MS first needs to make a request using a reverse control channel in the current cell.

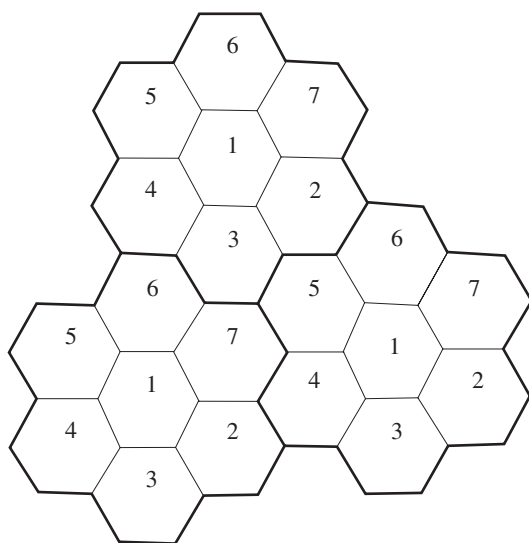


Figure 2: Frequency reuse.

If the request is granted by MTSO, a pair of voice channels (one for transmitting and the other for receiving) is assigned for the call. Making a call to an MS is more complicated. The call is first routed to its home MTSO or its visited MTSO if it is roaming. The MTSO needs to know the cell in which the MS is currently located. Finding the residing cell of an MS is the subject of location management. Once the MTSO knows the residing cell of the MS, a pair of voice channels is assigned in the cell for the call.

If a call is in progress when the MS moves into a neighboring cell, the MS needs to get a new pair of voice channels from the BS of the neighboring cell so the call can continue. This process is called hand-off. A BS usually adopts a channel assignment strategy that prioritizes hand-off calls from neighboring cells over the new calls initiated in the current cell.

MULTIPLE ACCESS METHODS

Within a cell covered by a BS, there are multiple MSs that need to communicate with the BS. Those mobile stations must share the air interface in an orderly manner so that no MSs within the cell interfere with each other. The methods for MSs to share the air interface in an orderly manner are referred to as multiple access methods. The popular multiple access methods include (frequency division multiple access) FDMA, TDMA, and CDMA.

FDMA divides the frequency spectrum assigned to the BS into several frequency bands, as known as channels, as shown in Figure 3. These channels are well separated and do not interfere with each other. An MS can use the assigned channel(s) exclusively. FDMA is used in the Advanced Mobile Phone System (AMPS) (Black, 1996, 1999). AMPS uses a total of 40 MHz in the 800-MHz spectrum, 825–845 MHz and 870–890 MHz, to be exact. (For ease of clarification, the additional 10 MHz added later is not considered here.) In AMPS, each channel has a bandwidth of 30 kHz, and the 40-MHz bandwidth translates into about 1332 channels. In the United States, it is required that two cellular communication providers be present in every market to encourage competition. Therefore, each

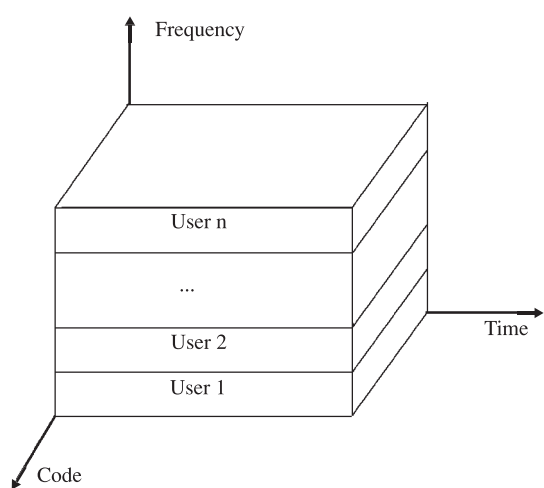


Figure 3: FDMA (frequency division multiple access).

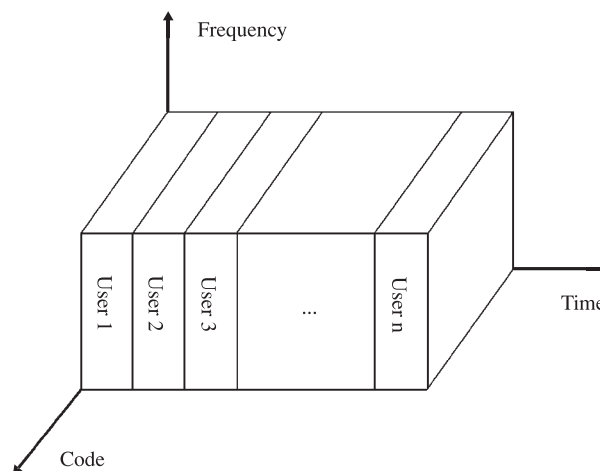


Figure 4: TDMA (time division multiple access).

cellular communication provider has 666 channels. AMPS uses FDD (frequency division multiplexing). That is, 333 channels are for communication from the BS to MSs and the other 333 channels are for communication from mobile stations to the base station. Among these 333 channels, only 312 channels are for voice traffic because 21 of them need to be used for control. Based on the seven-cell reuse pattern, only about 45 MSs within a cell can communicate with the BS simultaneously.

TDMA usually builds on FDMA and allows multiple MSs to share the same channel. In TDMA, time is slotted. In each time slot, only one MS is allowed to use the shared channel to transmit or receive. MSs take their turn transmitting or receiving in their allocated slots in a round-robin fashion. Although the channel is shared, no interference can arise among those sharing MSs because only one MS can use the channel at one time. Figure 4 illustrates the concept of TDMA.

Because an MS is not able to use the channel all the time, it is challenging to deliver voice, which is supposed to be continuous. Fortunately, an ordinary human can tolerate a delay of 20 milliseconds (ms). In D-AMPS (D for digital), a speech segment consists of 20-ms durations of speech. The speech segment is first digitized and then compared with the VSELP (vector sum excited linear predictive) Cookbook (Black, 1999). The index to the entry in the VSELP Cookbook that is closest to the digitized voice is transmitted instead of the digitized voice. The index is 159 bits long. At the receiving end, the digitized voice that is very close to the original voice can be retrieved based on the 159-bit index. In D-AMPS, which uses the same 30-kHz channel as AMPS, 159 bits (along with overhead bits for a total of 260 bits) can be transmitted in two of six time slots in a frame. At 25 frames per second, D-AMPS has three times the capacity of AMPS using the same number of channels. TDMA can operate in either the 800-MHz cellular spectrum (IS-54/D-AMPS; EIA/TIA, 1990) or the 1900-MHz PCS spectrum (IS-136; EIA/TIA, 1995).

CDMA takes an entirely different approach from TDMA. In CDMA, multiple MSs share the same wideband of spectrum. Instead of being assigned to time slots as in TDMA, each MS is assigned a unique sequence code. Each

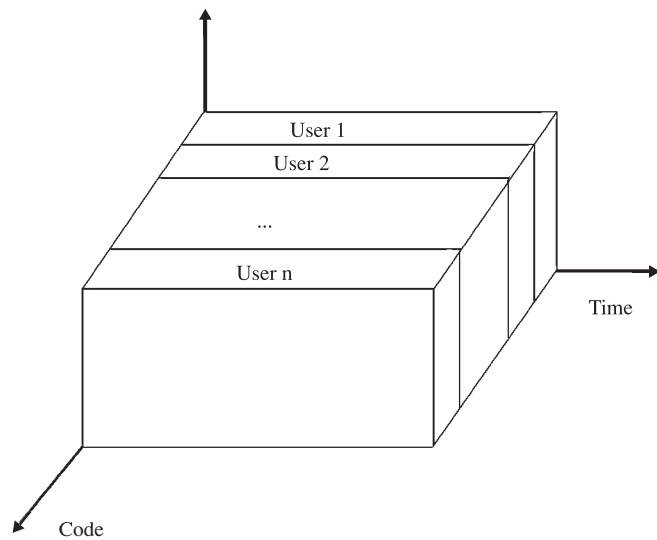


Figure 5: CDMA (code division multiple access).

MS's signal is spread over the entire bandwidth by the unique sequence code. At the receiver, that same unique code is used to recover the signal. Although the radio spectrum is shared, no interference can arise because the sequence codes used by the sharing MSs are basically orthogonal. Figure 5 illustrates the concept of CDMA. An excellent illustration of how CDMA encodes and decodes can be found in Stallings' text (2002). In principle, for CDMA to work, the signals received by the BS from different MSs must be at the same power level. To achieve this, a few bits in the forward control channel are set aside for power control. Specifically, the BS uses these bits to instruct each MS to increase or decrease its output power level such that all signals received at the BS have the same strength.

LOCATION MANAGEMENT

This section describes how to track an MS. How to keep track of an MS is the subject of location management in cellular networks. Because the exact location of an MS must be known to the cellular network when a call is in progress, location management tracks an active MS that is not in a call. (An MS is active when it is powered on.) Specifically, the cellular network needs to find out the exact cell in which an MS is located when an incoming call arrives for the MS.

An extreme case, known as "Never-Update" in Bar-Noy, Kessler, and Sidi (1995), is that an MS never tells the cellular network its location when it moves around. When an incoming phone call arrives for the MS, the cellular network needs to page all cells in the service area to find out the cell in which the MS is currently located so the incoming call can be routed to the BS of that cell. It will cost a great deal to page all cells in the service area. The other extreme case is known as "Always-Update," in which an MS needs to update its location whenever it moves into a new cell. When an incoming phone call arrives for the MS, the cellular network can just route the incoming call to the cell that is last reported by the MS. Obviously, there

is no paging cost involved. However, the MS needs to tell the cellular network its location when it moves from cell to cell, which can also be very expensive.

There are two basic operations involved with location management: location update and paging. The paging operation is performed by the cellular network. When an incoming call arrives for an MS, the cellular network will page the MS in all possible cells to find out the cell in which the MS is located so the incoming call can be routed to the corresponding BS. The number of all possible cells to be paged is dependent on how frequent the location update operation is performed. The location update operation is performed by an MS. Both operations consume wireless bandwidth: location update uses reverse control channels, whereas paging utilizes forward control channels. Although the cost of location management involves both the wireline portion and the wireless portion of the cellular network, only the wireless portion is usually considered. This is mainly because the radio-frequency bandwidth is limited, whereas the bandwidth of the wireline network is easily expandable. Therefore, the location management cost is measured by the total wireless bandwidth consumed by both location update and paging operations, that is, the location update cost and the paging cost. There is a trade-off between the location update cost and the paging cost. For example, if an MS updates its location more frequently, the network knows the location of the MS better. It incurs a higher location update cost, but the paging cost will be lower when an incoming call arrives for the MS.

The location-areas scheme is used in current cellular networks. In the location-areas scheme, the whole service coverage area is partitioned into location areas, each of which consists of several contiguous cells. The BS of each cell broadcasts the ID of the location area to which the cell belongs. An MS knows the location area it is in by listening to the broadcasting from the BS. An MS updates its location (i.e., location area) whenever it moves into a cell that belongs to a new location area. When an incoming call arrives for an MS, the cellular network pages all the cells of the location area that was last reported by the MS. Figure 6 illustrates a coverage area with three location

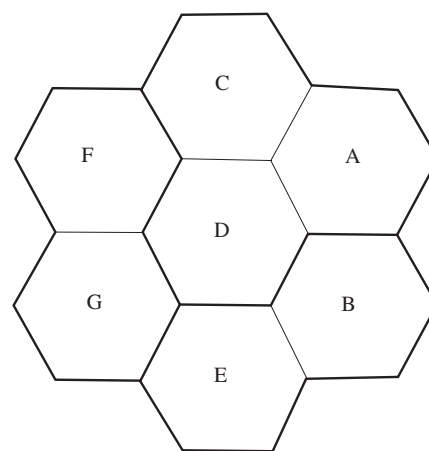


Figure 6: A service area with three location areas.

areas separated by wide lines. When an MS moves from cell A to cell B in the figure, it will report its new location area because cell A and cell B are in different location areas. No location update is necessary if the MS moves from cell A to cell C because cell A and cell C are in the same location area.

The location-areas scheme is a generalization of both Never-Update and Always-Update extreme cases. In Never-Update, all the cells in the coverage area belong to the same location area, and in Always-Update, each cell forms a location area. How to divide the whole coverage area into locations greatly affects the location management cost. In reality, a location area consists of all the cells under the control of an MTSO. Each MTSO consists of two main databases: HLR (Home Location Register) and VLR (Visitor Location Register). HLR contains information for each MS that considers the location area covered by the MTSO as its home. Each MS can subscribe to one MTSO as its home. VLR records a list of the MSs that are currently located in the location area covered by the MTSO but do not consider the MTSO as their home. When an MS moves out of its home MTSO and visits another MTSO, the MS needs to register with the visited MTSO. The visited MTSO contacts the home MTSO about the MS. The home MTSO records the visited MTSO of the MS in its HLR. The visited MTSO updates its VLR about the MS. When a call is made to the MS, it is first addressed to the home MTSO of the MS. By checking the information stored in HLR about the MS, the call can eventually be routed to the visited MTSO. Finally, the visited MTSO pages the MS in the cells of its covered location area.

The location-areas scheme is global in the sense that all MSs transmit their location updates in the same set of cells, and it is static in the sense that the location areas are fixed. A location-update scheme can be classified as either global or local (Bar-Noy, Kessler, & Sidi, 1995; Ramanathan & Steenstrup, 1996). A location-update scheme is global if all MSs update their locations at the same set of cells, and a scheme is local (or individualized) if an individual MS is allowed to decide where to perform location update. From another point of view, a location-update scheme can be classified as either static or dynamic (Bar-Noy, Kessler, & Sidi, 1995; Ramanathan & Steenstrup, 1996). A location-update scheme is static if there is a predetermined set of cells at which location updates must be generated by an MS regardless of its mobility. A scheme is dynamic if a location update can be generated by an MS in any cell depending on its mobility. Recently there was a swarm of research on location management, especially how to reduce the cost of location management. For detailed information on recent research on location management, please refer to the survey in Zhang (2002).

HAND-OFF STRATEGIES AND CHANNEL ASSIGNMENT

The previous section described how to track the movement of an MS when it is not in a call. This section deals with the movement of an MS when it is in a call. When an MS is in a call, it has acquired two channels (one for transmitting and the other for receiving) from the current

cell for communication with the BS. When the MS moves out of the current cell and enters a neighboring cell, the MS needs to acquire two channels from the neighboring cell in order for the call to continue. The process of transferring a call from the current cell to a neighboring cell is called hand-off. To a cell, a hand-off call is a call that is in progress in a neighboring cell and needs to be continued in the cell because of the movement of an MS. In contrast, a new call is a call that is started in the cell. As mentioned earlier, the number of channels assigned to each cell is limited. Channel assignment deals with how to assign available channels to new calls as well as hand-off calls.

The simplest channel assignment scheme is the fully shared scheme (FSS), in which all available channels are shared by hand-off calls and new calls. No distinction is made between a hand-off call and a new call. FSS is widely used in the current cellular networks because of its simplicity. In addition, FSS has the advantage of maximizing the utilization of wireless channels. The disadvantage is the increased dropping rate of hand-off calls. In general, it is less desirable to drop a hand-off call than to block a new call. The dropping probability of hand-off calls is considered as one of major metrics that measure the quality of service of calls.

Recently intensive research on channel-assignment schemes has been conducted to decrease the dropping probability of hand-off calls. One such scheme is the hand-off queuing scheme (HQS; Tekinay & Jabbari, 1991). When an MS detects that the received signal strength from the current BS is below a certain level, called the hand-off threshold, a hand-off operation is initiated. The hand-off operation first identifies the new BS into which the call is moving. If the new BS has unused channels, the call will be transferred to the new BS. If there is no unused channel available, the hand-off call will be queued until a channel is released by another call. The HQS scheme is feasible because there is a difference between the signal strength at the hand-off threshold and the minimum acceptable signal strength for voice communication. This gives an MS some time to wait for a channel at the new BS to become available. A new call will be blocked in the new cell until all the hand-off calls in the queue are served. Therefore, the HQS scheme decreases the dropping probability of hand-off calls while increasing the blocking probability of new calls because the scheme gives higher priority to hand-off calls.

The scheme proposed by Li, Shroff, and Chong (1999) goes one step further. When a hand-off call is not able to acquire the necessary channels in the new cell, the call is allowed to carry the channels in the current cell to the new cell, a concept the authors called channel carrying. However, carrying the channels from the current cell to the new cell may reduce the reuse distance of the channels and violate the minimum reuse distance requirement. To ensure the minimum reuse distance requirement is not violated, an $(r + 1)$ -channel assignment scheme is used. That is, the same channels are reused exactly $(r + 1)$ cells apart. Here r is the minimum reuse distance. In this scheme, a channel can only be carried from the assigned cell to a neighboring cell, and the carried channel will be returned as soon as a local channel is available. By using channel

carrying, a hand-off call can continue, even if there is no channel available in the new cell. Therefore, the dropping probability of hand-off calls will be reduced. However, the capacity of the cellular network is reduced because the $(r + 1)$ -channel assignment scheme is used instead of r -channel assignment.

Another channel assignment scheme that prioritizes the hand-off call is referred to as the guard channel scheme (GCS; Tekinay & Jabbari, 1991). In GCS, each BS reserves a fraction of wireless channels exclusively for hand-off calls, and the remaining channels, called normal channels, are shared between hand-off calls and new calls. Both hand-off calls and new calls use the normal channels first. When the normal channels are used up, a new call will be blocked, but a hand-off call can still use the reserved channels. In this way, the dropping probability of hand-off calls will be reduced. The improvement in the dropping probability of hand-off calls is dependent on the number of channels reserved. However, a new call is blocked if there are only reserved channels left even though no hand-off calls exist. Therefore, the total utilization of wireless channels is decreased. There is a trade-off between decreasing the blocking probability of hand-off calls and increasing the total channel utilization. The number of channels that should be set aside for hand-off calls depends on a lot of factors such as the mobility of MSs, the call duration, and so forth.

Kim et al. (1999) propose a dynamic channel reservation scheme (DCRS) based on mobility. Their goal is to guarantee the required dropping probability of hand-off calls while minimizing the blocking probability of new calls. As in GCS, the normal channels of DCRS are shared by both hand-off calls and new calls, and the guard channels are reserved for hand-off calls. However, the guard channels can also be used by new calls whose request probability is a function of the mobility of calls. The mobility of calls in a cell is defined as the ratio of the hand-off call arrival rate to the new call arrival rate. If there is no arrival of hand-off calls, the request probability will be one, and the guard channels will be used by new calls. If there is no arrival of new calls, the request probability will be zero, and the guard channels will be used by hand-off calls. When the arrival rate of hand-off calls is larger than that of new calls, the request probability is decreased quickly so the hand-off calls can use the guard channels. In this way, the dropping probability of hand-off calls is guaranteed. When the arrival rate of hand-off calls is lower than that of new calls, the request probability is decreased slowly so new calls have a chance to use available guard channels. This decreases the blocking rate of new calls without sacrificing the dropping probability of hand-off calls, which increases the total channel utilization.

AUTHENTICATION AND ENCRYPTION

Because cellular communication is carried out over the air interface, it is more vulnerable to fraud and eavesdropping. It is essential to have cell phone users authenticated and voice or data traffic encrypted. Authentication verifies whether a user is who he or she claims to be. Encryption is to use a key to scramble a message such that the

message cannot be read by anyone but the holder of the correct key. There are two popular encryption methods, private key encryption and public key encryption. Both encryption methods can also be used for authentication. For more information on how these two encryption methods work and how they can be used for authentication, the reader is referred to Stallings' text (2003) or the corresponding chapters in this book. This section discusses how authentication and encryption are performed in cellular networks, including global system for mobile communications (GSM) and IS-41.

Every GSM mobile subscriber has an SIM (subscriber identity module) card that can be inserted into a mobile phone (Black, 1999; Mouly & Pautet, 1992). A phone with an SIM card inserted can be activated by entering the four-digit PIN (personal identification number) correctly. The SIM stores identity- and security-related information such as the IMSI (international mobile subscriber identity), the authentication key (K_i), the A_3 authentication algorithm, and the A_8 ciphering key generating algorithm. The authentication in GSM checks if the SIM is genuine, and it is based on a challenge-and-response method. A 128-bit random number is sent by the network to the MS as the challenge. The MS inputs the challenge and the authentication key (K_i) to the A_3 algorithm to generate the signed response (SRES), which is sent back to the network. The network compares the received SRES with the SRES provided by the authentication center. If they match, the mobile user is authenticated; otherwise, the user is not. Once the mobile user is authenticated, the A_8 algorithm uses the random number and the authentication key (K_i) to generate a ciphering key (K_c) at both ends of the air interface. The ciphering key (K_c) is used by the A_5 ciphering algorithm of the MS to encrypt and decrypt data.

To make GSM secure, all security specifications such as the A_3 and A_8 algorithms are kept secret. However, it is claimed a secret encryption algorithm prevents the brightest minds in the world from identifying flaws in them. Recently a group of researchers discovered that the method used for authentication in GSM was not strong enough to resist attack (GSM Cloning, n.d.). An attacker can repeatedly send the challenges and collect the responses. By analyzing the responses, the secret key can be determined. However, despite its security vulnerability, GSM is still the most secure public wireless system in the world.

For the cellular systems in North America including IS-54, IS-136, and IS-95, IS-41 is used for security (Black, 1999; EIA/TIA, 1993; Mohan, 1996). There are five authentication and privacy operations defined in IS-41: authentication of mobile station registration, authentication of mobile station originating a call, authentication of a call to a terminating mobile station, unique challenge-response procedure, and updating the shared secret data. The idea used for authentication in IS-41 is similar to that used in GSM. Both the authentication center and the MS share a secret key, referred to as SSD (shared secret data). SSD consists of two parts: SSD-A and SSD-B, each of 64 bits. SSD-A is used for authentication, and SSD-B is used for encryption. A BS periodically broadcasts a 32-bit random number, that is, the authentication challenge. At request of the network, an MS inputs the received authentication challenge, SSD-A, its ESN (electronic serial number), and

its last six digits of the phone number to the CAVE (cellular authentication and voice encryption) algorithm to obtain the 18-bit response to the authentication challenge, which is sent back to the network. If the response matches the one calculated by the authentication center, the MS is authenticated. Once authenticated, the MS will receive two values, the signaling message encryption key for message encryption and the voice privacy mask for voice encryption.

EVOLUTION OF CELLULAR NETWORKS

Cellular systems became popular because of radio-frequency reuse, which allows more cell phone users to be supported. The cellular concept was first used in the AMPS in the United States. As a first generation of cellular systems, AMPS is a FDMA-based analog system. The 2G of cellular systems uses digital technologies. Two interim standards, IS-95 (CDMA-based) and IS-136 (TDMA-based), are used in the United States, and TDMA-based GSM is used in European countries. It is clear that the 3G of cellular systems will be CDMA-based. However, the GSM community is developing WCDMA to be backward compatible with GSM while the CDMA community tries to evolve CDMA into CDMA2000. Currently researchers are studying technologies for Beyond 3G (B3G) or fourth-generation (4G) networks.

In the late 1970s and early 1980s, AT & T Bell Laboratories developed the AMPS, which was the first-generation cellular system used in the United States (Rappaport, 2002; Young, 1979). It was first deployed in Chicago and Washington, DC, then in all the major U.S. cities. Currently it is still used in many rural areas. The Federal Communications Commission (FCC) initially allocated a 40-MHz spectrum in the 800-MHz band for the AMPS in 1983, and later in 1989 added an additional 10 MHz to accommodate the increasing demand for cellular phone services. AMPS is FDMA-based, with each channel occupying a narrow band of 30 kHz. AMPS is an analog system. It transmits 3-kHz voice signal over the 30-kHz channel using frequency modulation. IS-41 was originally developed to support the operations with AMPS. However, as an analog system, AMPS does not support voice encryption.

To overcome the limited capacity of AMPS, especially in large cities, D-AMPS (IS-54) was developed in the early 1990s (EIA/TIA, 1990). D-AMPS inherited a lot of features from AMPS. Specifically, in D-AMPS, the same AMPS allocation of frequency spectrum is used, and each channel is still 30 kHz wide. However, in D-AMPS, a 30-kHz channel can be shared by three users through the 2G TDMA digital technology. In a typical D-AMPS cell, some of the 30-kHz channels are assigned for analog AMPS traffic, whereas the others are for digital TDMA traffic. It means that D-AMPS allows a service provider to migrate from the first-generation analog technology to the 2G digital technology on a gradual basis. In a less densely populated area, all the channels can be assigned for AMPS traffic. When the demand increases, some channels will be converted from analog to digital. In a densely populated area, all the channels need to be assigned for TDMA traffic to meet the demand.

IS-136, another prominent TDMA-based cellular system in the United States, is built on D-AMPS. Whereas D-AMPS provides dual-mode operations (both analog and digital), IS-136 provide pure digital operations. All the 30-kHz channels are shared by three users via TDMA digital technology. In addition, unlike D-AMPS, IS-136 also uses the digital control channels. IS-136 was initially developed on the 800-MHz cellular spectrum. It can be adopted onto the 1900-MHz PCS spectrum.

GSM is a 2G system developed to solve the incompatibility problem of different first-generation systems in Europe (Black, 1999; Mouly & Pautet, 1992; Rahnema, 1993). It is now widely deployed around the world including in the United States. GSM was first developed for Europe in the 900-MHz band (GSM 900), then expanded to the 1800-MHz band (1710–1880 MHz), which is named DCS 1800, and later renamed to GSM 1800. The North America version of GSM is called PCS 1900 because of its use of the 1900-MHz PCS spectrum. GSM uses the TDMA digital technology. The allocated spectrum is divided into multiple channels of 200 kHz using FDMA, and each 200-kHz channel is shared by as many as eight users using TDMA. One feature of GSM worth mentioning is the SIM card that can be inserted into a cellular phone to provide the owner's identity information. A cell phone without a SIM card inserted does not work. A SIM card can be inserted into any cell phone to make the phone usable.

Whereas IS-54, IS-136, and GSM are all TDMA-based, IS-95 is CDMA-based (EIA/TIA, 1995). As mentioned earlier, each user in CDMA is assigned a unique code to encode the data to be transmitted. Knowing the code of the transmitter, the receiver is able to recover the original data from the received data. CDMA is a very new 2G digital technology. Since its first launch in 1995, CDMA quickly became one of the world's fastest-growing wireless technologies. CDMA uses channels that are 1.25 MHz wide, and it is able to support up to 64 users with orthogonal codes. With CDMA, the same channel can be reused in a neighboring cell. CDMA is superior to FDMA and TDMA. In fact, CDMA provides roughly 10 times more capacity than analog systems, whereas TDMA provides 3 to 4 times more capacity than analog systems. In 1999, CDMA was selected by the International Telecommunications Union as the industry standard for new 3G cellular systems.

The goal of a 3G cellular system is to provide all kinds of services: voice, high-speed data, audio/video, and so forth. The high-speed data transmission is the main development focus. The CDMA and GSM communities are two major players in this effort. The 3G path adopted by the GSM community is first to GPRS, then to EDGE, and ultimately to WCDMA (Qualcomm CDMA Technologies, n.d.). Currently GSM provides data services of 9.6 Kbps using a single TDMA channel. Although multiple TDMA channels can be combined to provide high-speed data service, it is circuit switched. A service called general packet radio service (GPRS) is first developed to allow users to connect to packet-switched data networks via a different connection from the voice network. With GPRS, the raw data rate increases to approximately 170 Kbps. The next step is enhanced data rates for global evolution (EDGE). EDGE provides practical raw data rates of up to 384 Kbps using a new high-speed physical layer. Finally, WCDMA

is used for the 3G version of the GSM community. In WCDMA, CDMA is used instead of TDMA, and the carrier bandwidth jumps from 200 kHz to 5 MHz to provide data rates of up to 2 Mbps. However, new frequency allocations, BSs, and MSs are required because of the change from TDMA to WCDMA.

The 3G path adopted by the CDMA community is first to CDMA2000 1x, then to 1xEV, and ultimately to 3x (Qualcomm CDMA Technologies, n.d.). The first step is to use one 1.25-MHz carrier to support packet data. In fact, CDMA phones and networks are already capable of handling packet data. The CDMA2000 1xEV consists of two phases. In Phase 1, one carrier (1.25 MHz) is dedicated to high-speed packet data, and one or more additional carriers are used for voice. In Phase 2, packet data and voice can be combined in the same carrier. Finally, CDMA2000 3x can use up to three 1.25-MHz carriers. When CDMA2000 uses three 1.25-MHz carriers, its total bandwidth approaches that of WCDMA. However, 3x is more flexible because three channels can be used independently or together as a single 3.75-MHz channel. The main advantage of the 3x approach is that it can be implemented in existing frequency allocations in CDMA, which also uses 1.25-MHz carriers.

Now researchers are developing technologies for B3G or 4G networks (Technologies Beyond 3G, n.d.). It is expected all the 4G network elements are digital and the entire network is packet-switched. 4G networks will integrate wireless local area networks (LANS) such as IEEE 802.11 and Bluetooth with wide area cellular networks. The data transmission rate of 4G communications will be much higher than 3G, at 20 to 100 Mbps in mobile mode.

ALTERNATIVES TO CELLULAR NETWORKS

This section describes alternatives to cellular networks. Although cellular networks use either the 800-MHz cellular spectrum or the 1900-MHz PCS spectrum, most of the alternatives use the license-free 2.4-GHz industrial, scientific, and medical (ISM) band. The alternatives introduced here include IEEE 802.11 wireless LAN, IEEE 802.16 wireless MAN (metropolitan area network), mobile ad hoc networks, and multihop cellular networks.

IEEE 802.11 wireless LAN is an infrastructure wireless network, similar to a cellular network (Crow, Widjaja, Kim, & Sakai, 1997; IEEE P802.11, n.d.). (IEEE 802.11 also supports peer-to-peer communication that is not frequently used.) Access points (APs) that are analogous to BSs in cellular networks are established to provide an extension to the wired network. The area covered by an AP is called the basic service area, which is analogous to a cell in cellular networks. The nodes in a basic service area form the basic service set. Any node in the basic service set can communicate with the AP directly. Their radio communication uses the license-free ISM band with direct sequence spread spectrum. Under 802.11b (a variation of the IEEE 802.11), the communication is kept at a maximum speed of 11 Mbps whenever possible. It drops back to 5.5 Mbps, then 2 Mbps, and finally down to 1 Mbps if signal strength or interference is corrupting data. New

IEEE 802.11g extends the data rate to 54 Mbps from 11 Mbps of IEEE 802.11b. With IEEE 802.11 wireless LAN, a node is connected to the Internet world, and it is able to get any service the Internet provides.

IEEE 802.11 works well for wireless access in a local area. To use it for wireless access in a metropolitan area, its bandwidth is often insufficient, and it can experience interference from competitors because they use the same license-free ISM band. IEEE 802.16 standards are for broadband wireless access in a metropolitan area, and they use licensed bands (IEEE 802.16 Task Group e, n.d.). IEEE 802.16a is just for fixed broadband wireless access. IEEE 802.16e supports both fixed and mobile broadband wireless access. Therefore, hand-off between two towers is allowed in IEEE 802.16e.

Unlike cellular networks, mobile ad hoc networks (MANET for short) are infrastructureless. All nodes within the network can be mobile, and there is no fixed BS or centralized control (Macker & Corson, 1998). A mobile node can communicate with another one directly if one is within the transmission range of the other. Because the network topology for a MANET can dynamically change as a result of node mobility, a MANET is usually modeled by a dynamically changing graph. A MANET finds its applications in an environment where it may not be economically practical or physically possible to provide the necessary infrastructure, or the expediency of the situation may not permit its installation. For example, in a battlefield, it is not possible to install an infrastructure wireless network in the enemy's territory. In such a situation, a MANET is a good solution. Unlike in the cellular network, the route for communication between two nodes is not fixed in MANET. To send a message from one mobile node (the source) to another mobile node (the destination), a route between the source and the destination must be found. Designing a good routing protocol in a dynamically changing network is very challenging. For more information on routing protocols, please refer to routing surveys in Royer and Toh (1999) and Stojmenovic (2002).

The traditional cellular network is called a single-hop cellular network (SCN) because an MS in a cell can reach the corresponding BS with a single wireless hop, and a MANET is considered a multihop wireless network because intermediate nodes are required for communication between two nodes that are not within the transmission range of each other. Lin and Hsu (2000) proposed a new architecture, referred to as multihop cellular network (MCN), which combines the features of SCN and MANETs. MCN has many advantages over SCN. In SCN, two MSs communicate only via a BS even though they are in the same cell and mutually reachable. In MCN, these two MSs can communicate directly. In MCN, if a MS is not reachable from a BS in one hop, the BS will seek intermediate nodes to forward, which is not feasible in SCN.

SUMMARY

Cellular networks are based on the frequency reuse concept. In a cellular network, a service coverage area is divided into cells, each of which is served by a BS. Each BS is assigned a group of channels. The same set of channels can be assigned to the two cells that are far apart such that

the radio co-channel interference between them is within a tolerable limit. Within a cell covered by a BS, multiple MSs want to communicate with the BS. Multiple access methods deal with how to share the channels assigned to a BS by the MSs located within the cell. FDMA, TDMA, and CDMA are three popular multiple access methods.

When an incoming call arrives for a MS, the cellular network needs to find out the exact cell in which a MS is located. How to track the movement of a MS is the subject of location management. Location management for cellular networks was discussed in this chapter. When a MS is in a call and moves from the current cell to a neighboring cell, the cellular network needs to transfer the call from the current cell to the neighboring cell. This process is called hand-off. A hand-off call is usually prioritized over a new call. Several channel assignment schemes that deal with new calls and hand-off calls were presented in this chapter.

Because wireless communication is carried out over the air interface, it is more vulnerable to fraud and eavesdropping. Therefore, authentication and encryption are two important issues in cellular networks. This chapter discussed authentication and encryption in GSM and IS-41. A cellular network integrates cell phones into the public switched telephone network. Whereas the first generation of cellular networks was analog, the 2G is digital. The 3G integrates cell phones into the Internet world by providing high-speed packet-switching data transmission. The evolution of cellular networks from the first generation to the 3G was presented. Technologies for B3G or 4G networks are being studied currently. Finally, this chapter discussed alternatives to cellular networks, including IEEE 802.11 wireless LAN, IEEE 802.16 wireless MAN, MANETs and MCNs. For further reading on cellular networks, please refer to these three excellent books: Agrawal and Zeng (2002), Black (1999), and Rappaport (2002).

ACKNOWLEDGMENTS

The authors thank Professor Hossein Bidgoli and five anonymous referees for their constructive comments and suggestions that greatly improved the quality of this chapter. The authors also thank Professor Phillip Bradford for many helpful discussions and his time to proofread the manuscript, and Mr. Zhijun Wang for drawing the figures in this chapter.

GLOSSARY

AMPS (Advanced Mobile Phone System) The first-generation analog cellular network developed in the United States.

BS (Base Station) A fixed station in the cellular network that connects mobile stations to the land telephone network by communicating with mobile stations via radio.

CDMA (Code Division Multiple Access) A digital cellular technology that allows multiple users to share a channel using different orthogonal codes.

Cell The area covered by a base station.

Cell Splitting A technique that divides a cell into multiple smaller cells to increase the capacity of a cellular network.

Cellular Network A network that divides a geographic area into cells such that the same radio frequency can be reused in two cells that are a certain distance apart.

FDMA (Frequency Division Multiple Access) An analog cellular technology that divides a wide band into multiple narrow bands to be used by multiple users.

GSM (Global System for Mobile Communications) A second-generation TDMA cellular network developed in Europe to achieve compatibility among European countries.

Hand-Off Channel switching for a call to continue when the mobile station involved in the call moves from one cell to another.

HLR (Home Location Register) An MSC database that stores information about the users who consider the area covered by the MSC as their home.

IS-54 A second-generation TDMA cellular network in the United States that uses the same cellular spectrum (800 MHz) as AMPS. It is also known as Digital-AMPS.

IS-95 A second-generation CDMA cellular network in the United States for both the cellular and PCS spectrums.

IS-136 A second-generation TDMA cellular network in the United States for both the cellular (800 MHz) and PCS (1900 MHz) spectrums.

Location Management A subject that deals with how to track an active mobile station.

MS (Mobile Station) A station in the cellular network that can communicate with base stations via radio even when it is in motion. Examples include cellular phones and wireless-capable laptops.

MTSO (Mobile Telephone Switching Office) The switching center that interconnects cellular phones with the land telephone network. It is also known as MSC (mobile switching center).

Multiple Access Method A method that allows multiple users in a cell to share the air interface in an orderly manner. Multiple access methods include FDMA, TDMA, and CDMA.

PSTN (Public Switched Telephone Network) The regular wireline telephone network.

TDMA (Time Division Multiple Access) A digital cellular technology that allows multiple users to share a channel using different time slots.

VLR (Visitor Location Register) An MSC database to temporarily store information about the users who are visiting the area covered by the MSC.

CROSS REFERENCES

See *Computer and Network Authentication; Encryption Basics; Radio Frequency and Wireless Communications Security; Wireless Channels*.

REFERENCES

- Agrawal, D. P., & Zeng, Q.-A. (2002). *Introduction to wireless and mobile systems*. United States: Brooks/Cole.
- Bar-Noy, A., Kessler, I., & Sidi, M. (1995). Mobile users: To update or not to update? *Wireless Networks*, 1(2), 175–185.

REFERENCES

663

- Black, U. (1996). *Mobile and wireless networks*. Upper Saddle River, NJ: Prentice Hall.
- Black, U. (1999). *Second generation mobile and wireless networks*. Upper Saddle River, NJ: Prentice Hall.
- Crow, B. P., Widjaja, I., Kim, J. G., & Sakai, P. T. (1997). IEEE 802.11 wireless local area networks. *IEEE Communications Magazine*, 35(9), 116–126.
- EIA/TIA. (1990, May). Cellular system dual-mode mobile station–base station compatibility standard (IS-54). United States: EIA/TIA
- EIA/TIA. (1993, May). Cellular radio telecommunications intersystem operations: Authentication, signaling message encryption and voice privacy (TSB-51). United States: EIA/TIA.
- EIA/TIA. (1995, May). Mobile station–base station compatibility standard for dual-mode wideband spread spectrum cellular system (IS-95). United States: EIA/TIA.
- GSM cloning. (n.d.). Retrieved April 17, 2005, from <http://www.isaac.cs.berkeley.edu/isaac/gsm-faq.html>
- IEEE 802.16 task group e (mobile wireless MAN). (n.d.). Retrieved April 17, 2005, from <http://www.ieee802.org/16/tge/>
- IEEE P802.11, the working group for wireless LANs. (n.d.). Retrieved April 17, 2005, from <http://grouper.ieee.org/groups/802/11/>
- Kim, Y. C., Lee, D. E., Lee, B. J., Kim, Y. S., & Mukherjee, B. (1999). Dynamic channel reservation based on mobility in wireless ATM networks. *IEEE Communications Magazine*, 37(11), 47–51.
- Li, J., Shroff, N. B., & Chong, E. K. P. (1999). Channel carrying: A novel handoff scheme for mobile cellular networks. *IEEE/ACM Transactions on Networking*, 7(1), 38–50.
- Lin, Y.-D., & Hsu, Y.-C. (2000). Multihop cellular: A new architecture for wireless communications. *Proceedings of IEEE INFOCOM*.
- Macker, J. P., & Corson, M. S. (1998). Mobile ad hoc networking and the IETF. *Mobile Computing and Communications Review*, 2(1), 9–14.
- Modarressi, A. R., & Skoog, R. A. (1992). An overview of signaling system no. 7. *Proceedings of the IEEE*, 80(4), 590–606.
- Mohan, S. (1996). Privacy and authentication protocols for PCS. *IEEE Personal Communications*, 3(5), 34–38.
- Mouly, M., & Pautet, M. B. (1992). *The GSM system for mobile communications*. [location not available]: Telecom.
- Qualcomm CDMA Technologies. (n.d.). Retrieved April 17, 2005, from http://www.cdmatech.com/resources/glossary_full.jsp
- Rahnema, M. (1993). Overview of the GSM system and protocol architecture. *IEEE Communications Magazine*, 31(4) 92–100.
- Ramanathan, S., & Steenstrup, M. (1996). A survey of routing techniques for mobile communication networks. *Mobile Networks and Applications*, 1(2), 89–104.
- Rappaport, T. S. (2002). *Wireless communications—Principles and practice*. Upper Saddle River, NJ: Prentice Hall.
- Royer, E. M., & Toh, C. K. (1999). A review of current routing protocols for ad hoc mobile wireless networks. *IEEE Personal Communications*, 6(2), 46–55.
- Stallings, W. (2002). *Wireless communications and networks*. Upper Saddle River, NJ: Prentice Hall.
- Stallings, W. (2003). *Cryptography and network security* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.
- Stojmenovic, I. (2002). Position-based routing in ad hoc networks. *IEEE Communications Magazine*, 40(7), 128–134.
- Technologies beyond 3G—4G vision, concepts & efforts. (n.d.). Retrieved April 17, 2005, from <http://www.mobileinfo.com/3G/4GVision&Technologies.htm>
- Tekinay, S., & Jabbari, B. (1991). Handover and channel assignment in mobile cellular networks. *IEEE Communications Magazine*, 29(11), 42–46.
- Young, W. R. (1979). Advanced mobile phone service: Introduction, background, and objectives. *Bell Systems Technical Journal*, 58, 1–14.
- Zhang, J. (2002). Location management in cellular networks. In I. Stojmenovic (Ed.), *Handbook of wireless networks and mobile computing* (pp. 24–49). New York: John Wiley & Sons.