CHAPTER 1

Wireless Telecommunications Systems and Networks

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

- Discuss the general history and evolution of wireless technology from a North American viewpoint and explain the cellular radio concept.
- Discuss the evolution of modern telecommunications infrastructure.
- Discuss the structure and operation of the Public Switched Telephone Network, the Public Data Network, and the SS7 Network.
- Explain the basic structure of broadband cable TV systems.
- Explain the basic concept and structure of the Internet.
- Discuss the usage of the various telecommunications networks and their relationship to one another.
- Discuss the OSI model and how it relates to network communications.
- Discuss wireless network applications and the future of this technology.

Practical electrical communications began in the United States over 150 years ago with the invention of the telegraph by Morse. The invention of the telephone by Bell in 1876 brought with it the first manually switched wireline network. Radio or wireless was invented at the turn of the twentieth century, adding the convenience of mobile or untethered operation to electronic communications. For many years, wireless communications primarily provided entertainment and news to the masses through radio broadcasting services. Wireless mobility took the form of a car radio with simplex (one-way) operation. Two-way mobile wireless communications were limited to use by various public service departments, government agencies and the military, and for fleet communications of various industries. As technology decreased the size of the mobile unit, it became a handheld device known as a "walkie-talkie."

Further advances in integrated circuit technology or microelectronics gave us cordless telephones during the late 1970s that foreshadowed the next wireless advance. Starting in 1983, the public had the opportunity to subscribe to cellular telephone systems. These wireless systems, which provide mobile access to the public switched telephone network infrastructure, have become immensely popular and in many cases have even replaced subscribers' traditional fixed landlines. Technology advances and network build-outs have increased wireless system capacity and functionality.

Today's cellular networks provide access to the public telephone network from almost anywhere and provide access to the public data network or Internet. In two decades, cell phones have become indispensable communications devices and Internet appliances. During the same time period, wireless local area

network (LAN) technology has come of age and is gaining in acceptance by both the *Enterprise* (for profit and not-for profit business ventures) and the general public. Today, many homes and apartments have their own wireless LANs. This chapter will present a short history of wireless technology, a brief summary of the evolution and operation of the fixed public networks, a general idea of how these networks fit together, and an overview of how wireless systems connect to this modern infrastructure. Additional topics covered by the chapter are a review of the OSI model, an overview of wireless network applications, and a brief look at the future of wireless.

1.1 THE HISTORY AND EVOLUTION OF WIRELESS RADIO SYSTEMS

One can trace the evolution of wireless radio systems back to the late 1800s. In 1887, Heinrich Hertz performed laboratory experiments that proved the existence of electromagnetic waves, just as Maxwell predicted back in 1865. An obscure inventor by the name of Mahlon Loomis was in fact issued a U.S. patent for a crude type of aerial wireless telegraph in 1872. Although several prominent inventors of the day (Lodge, Popoff, and Tesla) experimented with the transmission of wireless signals, Marconi seems to have received most of the credit for the invention of radio because he was first to use it in a commercial application. From 1895 to 1901 Marconi experimented with a wireless telegraph system. He initially started his experiments at his family's villa in Bologna, Italy. He then moved to England in 1896 to continue his work. He built several radio telegraph stations there and started commercial service between England and France during 1899. However, the defining moment for wireless is usually considered to have taken place on December 12, 1901, when Marconi sent a message (the signal was a repetitive letter "s" in Morse code) from Cornwall, England to Signal Hill, St. John's, Newfoundland—the first transmission across the Atlantic Ocean. This was accomplished without the aid of any modern "electronic devices"—vacuum tubes and transistors did not exist at the time.

Historical Note: For another view of what actually happened during those early days, read Dr. Jack Belrose's account of the development of wireless radio at www.radiocom.net/Fessenden/Belrose.pdf.

Early AM Wireless Systems

A typical early wireless transmitter is shown in Figure 1–1. Note the inductance and capacitance used to tune the output frequency of the spark-gap. The resonant frequency of these two components tended to maximize output power at that particular frequency.

Due to the nature of the spark-gap emission, maximum power output typically occurred at a very low frequency with its corresponding long wavelength. Although little was known about antenna theory at the time, it was discovered early on that for a conductor to effectively radiate long wavelength signals the antenna had to be oriented vertical to the earth's surface and physically had to be some appreciable fraction of a wavelength. It was common for early wireless experiments to use balloons and kites to support long lengths of wire that served as the antennas. Also at that time it was thought that transmitting distance would be limited by the curvature of the earth. This belief became an additional rationale for tall antennas.

The wireless transmitter shown in Figure 1–1 would emit a signal of either long or short duration depending on the length of time the telegraph key was closed. The transmitted signal was the electromagnetic noise produced by the spark-gap discharge. This signal propagated through the air to a receiver located at some distance from the transmitter. At the receiver, the detected signal was interpreted by an operator as either a dot or a dash depending upon its duration. Using Morse code, combinations of dots and dashes stood for various alphanumeric characters. This early wireless transmission form is now known as amplitude modulation (AM) and in particular, on-off keying (OOK).

The next generation of wireless transmitters used more stable radio-frequency (RF) alternators or high-powered Poulsen spark-gap transmitters for their signal sources. These RF alternator-based transmitters were used to transmit another form of AM that is sometimes referred to as binary amplitude-shift keying

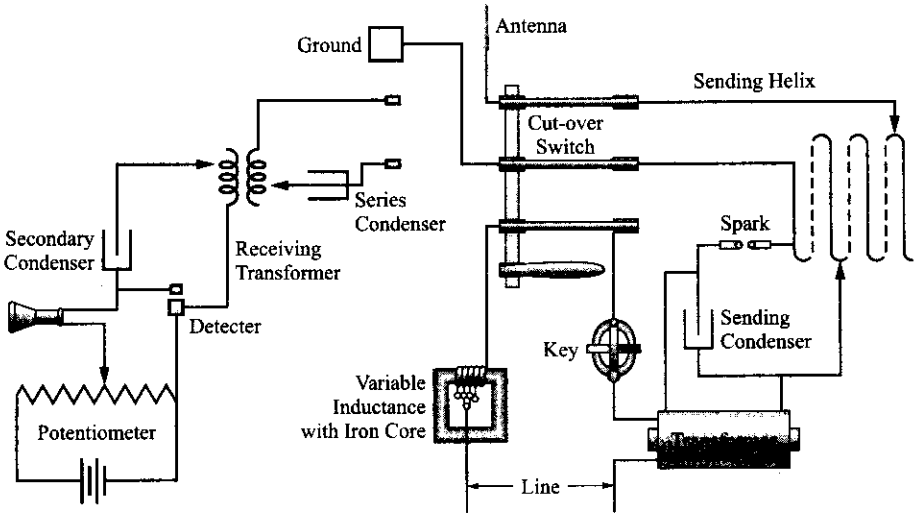

Figure 1–1 Typical early wireless transmitter.

(BASK), which is essentially the same as OOK. The Poulsen transmitters used a form of frequency-shift keying (FSK) to transmit a signal that was received and interpreted as a BASK signal by the detector of the radio receiver.

The First Broadcast

Beginning in 1905, Reginald Fessenden conducted experiments with continuous wave (CW) wireless transmissions at Brant Rock, Massachusetts, using 50-kHz high-frequency alternators built by General Electric. The output of this type of generator was much more stable than that of a spark-gap or Poulsen transmitter, allowing him to experiment with a continuous form of amplitude modulation. His experiments culminated on Christmas Eve of 1906, when he is credited with transmitting the first ever radio broadcast. This broadcast was repeated on New Year's Eve the following week. Prior to this time, it is reported that Fessenden, while experimenting with a wireless spark-gap transmitter at an experimental station on Cobb Island on the Maryland side of the Potomac River, had successfully broadcast a voice message over a distance of 1600 meters on December 23, 1900.

During the 1910s, the U.S. Navy led a major effort to develop wireless radio for ship-to-ship and ship-to-shore communications. Historical accounts of the sinking of the *Titanic* on the night of April 14, 1912, tell of the transmission of futile "SOS" distress messages by the ship's wireless operator. The start of World War I during the last part of the decade was also a major driver of the development of radio technology by the U.S. military.

The 1920s might well be characterized as the decade of high-frequency or short-wave radio development. During this era, Marconi's research on radio wave propagation revealed that transatlantic radio transmission was feasible at frequencies much higher than had previously been though possible. At the same time, vacuum-tube technology had improved to such an extent as to increase the upper-frequency limit of their operation. Radio wave propagation studies had demonstrated that ionospheric layers could be used to reflect high-frequency waves back and forth between the earth's surface and the ionosphere, hence allowing for the propagation of radio waves around the earth. Other technological advances in antennas and their application helped make transatlantic communications a practical reality. By 1926, transoceanic telephone calls were available via high-frequency radio transmission. The 1930s and 1940s saw more advancement in radio technology with the invention of television, radar, and vacuum tubes with the ability to generate "microwaves."

Modern AM

Amplitude modulation is now used for low-frequency legacy radio broadcasting, which had its corporate beginnings after World War I; short-wave broadcasting; low-definition (NTSC) television video-signal transmission; amateur and CB radio; and various other low-profile services. Newer uses of AM include quadrature amplitude modulation (QAM or *n*-QAM, where *n* is a power of 2). QAM is a hybrid form of amplitude and phase modulation (PM) used for high-speed data transmission at RF frequencies. QAM is considered a digital modulation technique. Today, QAM is used extensively by broadband cable and wireless systems to achieve bandwidth efficiency.

The Development of FM

Major Edwin Armstrong, a radio pioneer who invented first the regenerative and then the superheterodyne receiver in the 1910s, worked on the principles of frequency and phase modulation starting in the 1920s. It was not until the 1930s, however, that he finally completed work on a practical technique for wideband frequency modulation (FM) broadcasting. FM broadcasting became popular during the late 1960s and early 1970s when technological advances reduced the cost of consumer equipment and improved the quality of service. Many public safety departments were early adopters of FM for their fleet communications. AMPS cellular telephone service, an FM-based system, was introduced in the United States in 1983. Today FM is used for transmissions in the legacy FM broadcast band, standard over-the-air TV-broadcasting sound transmission, direct-satellite TV service, cordless telephones, and just about every type of business band and mobile radio service. FM is capable of much more noise immunity than AM, and is now the most popular form of analog modulation.

The Evolution of Digital Radio

AT&T built its original long-distance network from copper wires strung on poles. The first experimental broadband coaxial cable was tested in 1936, and the first operational L1 system that could handle 480 telephone calls was installed in 1941. Microwave radio relay systems developed in tandem with broadband coaxial cable systems. The first microwave relay system was installed between Boston and New York in 1947. AT&T's coast-to-coast microwave radio relay system was in place by 1951. Microwave relay systems had lower construction and maintenance costs than coaxial cable (especially in difficult terrain). By the 1970s, AT&T's microwave relay system carried 70% of its voice traffic and 95% of its broadband television traffic.

At the time, most of these systems used analog forms of modulation, although simple digital modulation forms like binary frequency shift keying (BFSK) existed. Advances in microwave digital radio technology and digital modulation techniques that provide increased data rates over the same radio channel caused these systems to gain in popularity during the 1970s and 1980s. However, fiber-optic cables and geosynchronous satellites proved to be disruptive technologies as far as the use of microwave digital radio by the Bell system was concerned. Many of the analog and digital microwave relay systems in use became backup systems to newly installed fiber-optic cables or they were removed from service entirely.

Today, microwave **digital radio** systems are enjoying a resurgence of sorts. Many service providers of point-to-point connectivity are employing microwave and millimeter-wave radio transmission systems that use the most modern digital modulation techniques to obtain high data rate links. Cellular operators are using economical point-to-point microwave radio systems to backhaul aggregated bandwidth signals to a common network interface point from both remote and not-so-remote cell sites. Wireless Internet service providers (WISPs) are using digital radio equipment designed for the Unlicensed National Information Infrastructure **(U-NII) bands** for point-to-point and point-to-multipoint systems that provide high bit-rate Internet connections to their customers. Cordless telephones adopted digital radio technology years ago and all the newest wireless systems and network technologies use modern digital modulation techniques to

achieve higher data rates and better noise immunity. Today, the television broadcasting industry is in the process of transitioning to a high-definition television (HDTV) standard for over-the-air broadcast that uses a digital transmission system. It is not too radical a concept to conjecture that FM broadcasting might follow suit in the not-too-distant future. All but the oldest analog cellular systems (these systems are in the process of being phased out) are digital, and all of the newest wireless LAN, MAN, and PAN technologies use complex digital modulation schemes. It may be that analog modulation schemes, with their inefficient use of radio spectrum, might disappear entirely for over-the-air applications in the not-too-distant future!

The Cellular Telephone Concept

The cellular telephone concept evolved from earlier mobile radio networks. The first mobile radios were used primarily by police departments or other law enforcement agencies. The Detroit Police Department is cited for its early use of mobile radios (beginning in 1921) by numerous references. These one-way mobile radio systems, operating at about 2 MHz, were basically used to page the police cars. They did not become operational two-way (duplex) systems until much later in 1933. There was no thought at the time for these systems to be connected to the public telephone system. It was not until after World War II that the Federal Communications Commission (FCC), at the request of AT&T, allocated a small number of frequencies for mobile telephone service. AT&T made a request for many mobile frequencies on behalf of the Bell telephone companies in 1947, but the FCC deferred any action on this request until 1949. At that time, the FCC only provided a limited number of frequencies that were to be split between the Bell companies and other non-Bell service providers. The FCC apparently felt that since these frequencies were used by the police and fire (public service) departments that the public interest would be best served by limiting the number of frequencies released to this new service. It should be observed that the state of the art of wireless technology at the time restricted the given spectrum available for any new wireless service that might be desired.

The mobile phone service that grew out of these new frequency allocations was very primitive by today's standards. It usually consisted of a single, tall, centrally located tower with a high-power transmitter that could only service one user at a time per channel over a particular metropolitan area. This also precluded the reuse of the same frequency within approximately a seventy-five-mile radius. Due to the limited number of frequency allocations, only several dozen simultaneous users were possible. The capacity of these systems was quickly exhausted in the major cities by the mid-1950s. Customers of the service paid extremely high monthly or yearly rates and it was perceived to be a service that only a business or the wealthy could afford. At the time, the available wireless transceiver technology (which usually had to be located in the car trunk due to its bulk) offered no way of reusing the frequencies within the same general area or any other way of increasing the capacity of the system.

The FCC in 1968, in response to the congestion of the presently deployed system, asked for technical proposals for a high-capacity, efficient mobile phone system. AT&T proposed a **cellular** system. In this cellular system there would be many towers, each low in height, and each with a relatively low-power transmitter. Each tower would cover a "cell" or small circular area several miles in diameter. Collectively, the towers would cover the entire metropolitan area. Each tower or cell site would use only a few of the total number of frequencies available to the entire system. Due to the small cell sizes, these same frequencies could then be reused (hence the term frequency reuse) by other cells at a much shorter spacing than previously possible thus increasing the total potential number of simultaneous users within the entire system. Additionally, as a mobile user (car) moved within the metropolitan area it would be "handed off" from cell to cell and to different frequencies as assigned to the different cells. All the cells would be connected to a central switch that would in turn connect them to the wireline telephone network. As more users signed up for the service and cells became too congested, the cells could be split into several smaller cells to increase their capacity. In theory, this process could be repeated many times yielding an almost infinite number of potential simultaneous users for a limited number of available frequencies. In 1970, the FCC released 75 MHz of additional spectrum for use by the current system and authorized AT&T's Bell Laboratories to test the cellular concept under urban traffic conditions. In 1971, Bell Labs reported that the test

had been successful. The cellular concept worked! In 1974 the FCC released 40 MHz more of spectrum for the development of cellular systems. In a far-reaching decision, the FCC also determined that both the incumbent Bell Telephone Company and other nonwireline entities would share the newly made available spectrum. By late 1982, the FCC started to award construction permits for cellular systems, and by late 1983 and early 1984 most major metropolitan areas had functioning systems that supported the user's ability to roam between systems. The rest is history. Twenty years later the cell phone has become a ubiquitous information appliance with well over one billion users worldwide.

1.2 THE DEVELOPMENT OF MODERN TELECOMMUNICATIONS INFRASTRUCTURE

The wireless networks and systems that have been rapidly evolving over roughly the past two decades have the basic function of connecting users to the public switched telephone network (PSTN) or, more recently, the public data network (PDN). Therefore, it is instructive to examine exactly what these two public networks are and how they have evolved over the course of time.

Over the last four decades, several other telecommunications networks have evolved. The SS7 network is a packet data network used in conjunction with the PSTN to establish, manage, and terminate interexchange telephone calls. Broadband cable television networks have been developed for the delivery of video services and more recently high-speed data services (Internet connectivity) and telephony service. The Internet, which is the world's largest computer network, has experienced phenomenal growth over the past two decades and continues to expand both its reach and high-speed data capacity. Finally, in the United States, cellular telephone networks have become nationwide providing subscribers access to both the PSTN and the PDN.

The Early Days

Morse invented the telegraph in 1837 and formed a telegraph company based on his new technology in the mid-1840s. The Western Union Telegraph Company was established in 1856 and within a decade bought out most of its competitors. Early long-distance telegraph systems required many relay points because signals had a limited maximum range. In 1867, an improved telegraph relay was invented by Elisha Gray. Gray's company was bought out by another company that in 1872 became the Western Electric Manufacturing Company. Alexander Granam Bell received a patent for the telephone in 1876 and formed the Bell Telephone Company in 1877. In 1882, Bell bought the Western Electric Company and in 1885 formed what was to become the American Telephone and Telegraph Company (AT&T).

By 1900, the Bell system served approximately 60% of telephone subscribers in the United States. During the next decade, AT&T bought out most of its competitors and in essence formed a telecommunications monopoly. Starting in the late 1940s the U.S. Department of Justice (DOJ) sued AT&T for violations of the Sherman Antitrust Act. This event signaled the start of deregulation of the existing telecommunications industry. Eventually, other FCC decisions and U.S. government lawsuits resulted in what is known as the "Modified Final Judgment," which took effect on January 1, 1984. In simple terms, AT&T was required to divest itself of all the Bell Operating Companies (BOCs) and the long-distance telecommunications market became deregulated, therefore allowing competition. These events, coupled with the more recent Telecommunications Reform Act of 1996, have helped shape our present-day telecommunications infrastructure.

The Public Switched Telephone Network

In the United States and most other industrialized nations, the present- day PSTN has evolved over time to become an almost entirely digital network. Deregulation has allowed other competitors to sell telephone service but they all essentially use the same technology. In an effort to explain the physical infrastructure of

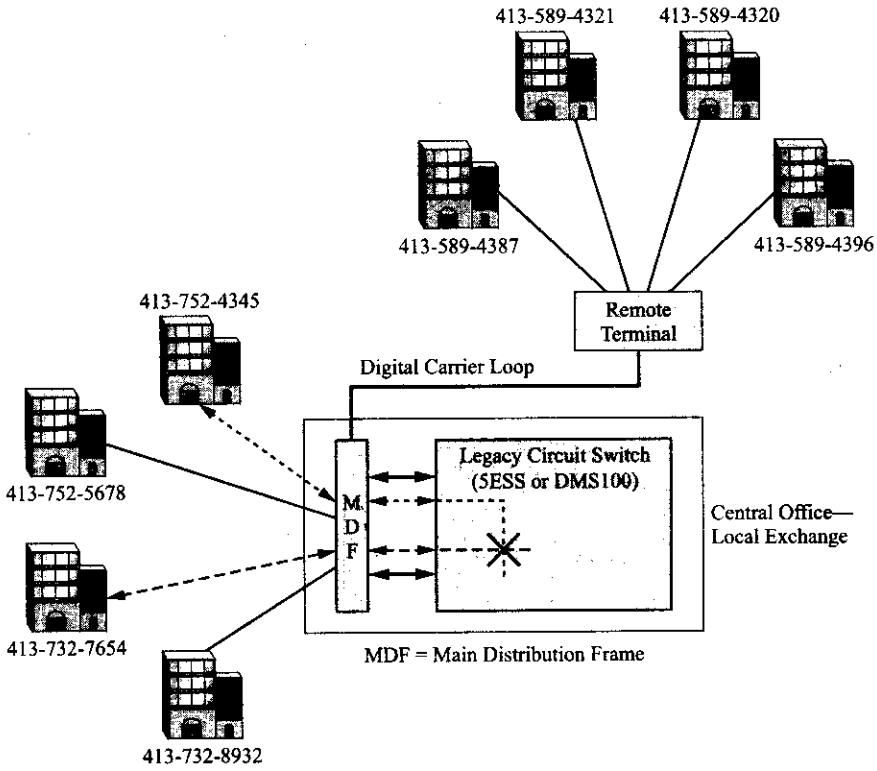

Figure 1–2 A PSTN intraoffice call through a local exchange.

the PSTN, it is instructive to consider the various pathways of communication available through the system.

Within a **local exchange** or company office (CO) a subscriber may be connected to the exchange in several different ways as shown in Figure 1–2. For plain-old telephone service (POTS) the subscriber may be connected through a local loop connection consisting of a pair of copper wires. In this case, dialing information (via dual-tone multifrequency [DTMF] or traditional rotary dialing [pulsing]) signals are interpreted by the local exchange switch to set up the correct pathway or connection through the switch to the desired called party. Call signaling information (dial tone, ringing, call-waiting tones, etc.) is sent to the called party and also sent back to the caller.

For an **intraoffice** call between two subscribers connected to the same switch, the analog voice signal from the subscriber's telephone propagates through the copper wire pair to a line card located at the switch. The line card converts the analog signal to a digital pulse code modulated (PCM) DS0 signal, which gets timed through the switch in such a way as to be connected to the corresponding line card of the called party. This counterpart line card performs the complementary conversion of the digital PCM signal from the switch into an analog signal that is sent to the called party over another pair of copper wires. A separate return path or connection is also created from the called party's line card through the switch to the calling party's line card. The line cards also provide the necessary opposite signal conversion functions for this return path and together the two paths through the switch provide for duplex operation for the duration of the telephone call. Since the call appears to be physically connected by a circuit and is using the resources of the switch, this type of operation is termed **"connection-oriented"** or in particular a **"circuit-switched"** connection. If the party to be called is connected to a different switch at another exchange within the same calling area (an **interoffice** call), the PCM signal from the calling subscriber's switch is timed through the switch in such a fashion that it is eventually forwarded to a multiplexer and then transmitted over a digital interoffice transmission facility (**trunk** line). Figure 1–3 shows this type of interoffice connection.

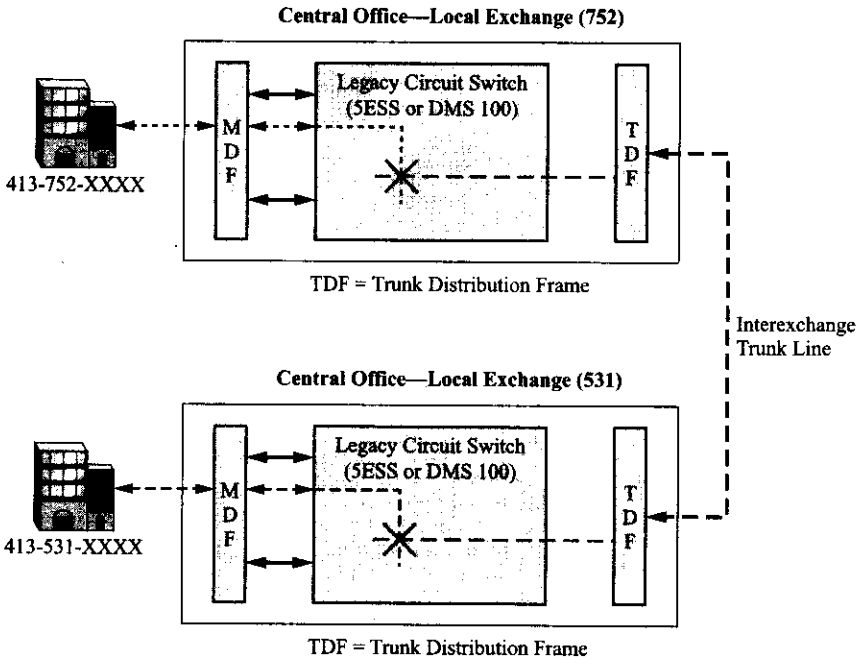

Figure 1–3 A PSTN interoffice call over an inter-exchange trunk line.

This interoffice connection might use some type of T-carrier transport technology (T-1, T-3, etc.) that might be carried over copper wires, but most likely it will be some form of fiber-optic connection that is transporting data at OC-1, OC-3, or OC-12 data rates using SONET transport technology. If the party to be called is in a different calling area (a long- distance call), the local switch will forward the caller's PCM packets to a long-distance carrier's multiplexed facilities using the area code of the called number to direct the call. The long-distance carrier's network will have switching centers located in different parts of the country typically connected by long-haul fiber facilities. Once the caller's signal is delivered by the long-distance carrier's network to the correct local end exchange it is demultiplexed back to a DS0 signal, the process that occurs to connect to the called party is the same as before. The signal is timed through the appropriate end switch to connect it to the called party's line card. Again, a completely separate circuit will be set up in the call's return direction.

Subscribers that live a substantial distance from the local exchange are usually connected by copper pairs to a remote terminal that provides an extended service area at some distance out from the local exchange. The remote terminal might use T-carriers or fiber-optic technology to connect to the local exchange, thus extending the digital network farther out from the switch and also providing the descriptive term "carrier service area" to describe the area served by the remote terminal. Refer to Figure 1–2 again.

To recap, the PSTN consists of copper pairs that transmit analog signals from the subscriber to a digital network that digitizes the signal and then processes it through a digital switch, at which point it might be converted back to an analog signal and delivered to another subscriber through another pair of copper wires. Or, after processing by the switch, the signal is forwarded to a number of digital facilities (multiplexers, demultiplexers, and various transmission media) that transmit the signal to other digital switches using any one of a variety of digital transport technologies. The digital switches use stored programs to control their operation and the sequence of operations needed for the appropriate transmission of calls between users connected to the switch or users connected to other switches.

Signaling System #7

The early PSTN used "in-band" signaling to set up and tear down interoffice and long-distance telephone calls. By this, we mean that the same facilities used to transport the call were first used to create an actual physical circuit for the call to be sent over. A big disadvantage of this type of system is that a voice trunk (an interoffice facility) or possibly many trunks had to be "seized" in order to do the signaling necessary to set up the call. If the call is nonchargeable (the called party is unavailable or the line is busy), the charges for the seizure of the trunk circuits must be paid for by the service provider that owns the local exchange. Furthermore, this type of system was very prone to fraud since the signaling was performed by sending easily reproducible audio tones over the trunking circuits. As the PSTN evolved into a digital network, for economic reasons and for both efficiency and security, an entirely separate network was created for the purpose of routing long-distance calls (calls between different exchanges or switches). This system of using a separate facility or channel to perform the call routing function is known as "out-of-band" signaling. AT&T's early out-of-band system was called Common Channel Interoffice Signaling (CCIS). With advances in technology, this common channel signaling system has been adopted by the international telecommunications community for use with both PSTNs and public land mobile networks (PLMNs). Today, it is known as CCIS #7 or simply Signaling System #7 (SS7).

The **SS7** system is a packet network that consists of **signal transfer points** (STP) and transmission facilities linking the signal transfer points as shown in Figure 1–4. The signal transfer points connect to **service switching points** (SSP) at the local exchange and interface with the local exchange switch or mobile switching center in the case of a PLMN. The service switching points convert signaling information from the exchange voice switch into SS7 signaling messages in the form of data packets that are sent over the

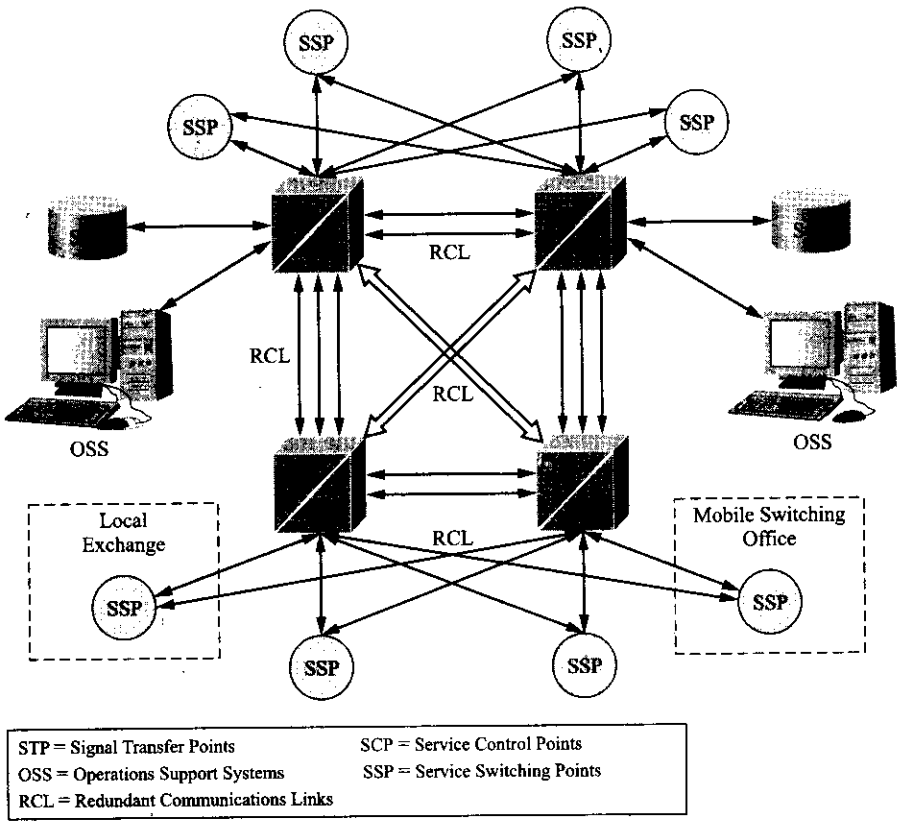

Figure 1–4 The network elements of the SS7 system.

SS7 network. All SS7 data packets travel from one service switching point to another through signal transfer points that serve the network as routers, directing the packets to their proper destination. There are several different types of redundant links between the signal transfer points to provide the SS7 network with a high degree of reliability.

The SS7 system provides two forms of services: circuit related—the setting up and tearing down of circuits, and non-circuit-related—the access of information from databases maintained by the network (e.g., 8XX number translation, prepaid calling plans, Home Location Register interrogation). **Service control points** act as the interface between the SS7 network and the various databases maintained by the telephone companies. The entire network is connected to a remote maintenance center for monitoring and management. This maintenance center is commonly referred to as a network operations center or NOC.

In the cellular telephone system, a service provider's cell sites in a metropolitan area are all connected to switching systems that are all tied to a common switch that is in turn connected to the fixed wireline network. This common gateway mobile switching center (GMSC) uses SS7 to signal between itself and the other switches and between itself and the fixed network. All PSTNs and PLMNs use SS7 for signaling operations within the network and between the network and other networks.

The Public Data Network

In the early days of data transmission, the PSTN was used to carry data and it is still used today for this purpose. This was accomplished both then and now through the use of modems. After performing handshaking functions to set up a circuit connection to another modem at a remote location, the modems perform the function of converting data from the host computers into digital signals (audio tones) that can be transmitted across the PSTN. Modem technology has gradually increased data rates close to the theoretical maximum possible through the PSTN switch or the digital network that extends outside the local exchange (recall the remote terminals used by the telephone companies to extend their service area to the suburbs).

The physical limitation to these modem data rates is in turn driving new technologies such as adaptive digital subscriber line (ADSL) and cable modems to provide high-speed data to the home or business. However, even though data can be sent through the circuit-switched voice network this does not mean that it serves the same purpose or is as efficient as the public data network.

The public data network (PDN) has been evolving for many years in response to the connectivity needs of business, industry, and government for the transport of data over wide area networks (WANs). The PDN is often depicted as a fuzzy "cloud" on diagrams that only show how the end users are connected to it. The reason for the use of a cloud is that the network uses many different types of SONET, transport technologies (T-carrier, xDSL, Ethernet, frame relay, ISDN, ATM, etc.) and physical media to transmit data within it and therefore from end point to end point. The connections to it might be through leased lines (copper pairs), fiber facilities, or wireless radio links using any one of the transport technologies mentioned earlier. Furthermore, the data network transports packets of data that, depending upon the type of transport protocol, can be configured in many different ways (size, overhead, etc.) and can take many different routes or paths through the network. See Figure 1–5 for one possible view of the PDN.

Additionally, the PDN can support many different types of service structures, including permanent virtual circuits, switched virtual circuits, and semipermanent virtual circuits. These different so-called connection-oriented service structures provide different levels of quality of service (QoS) to the customer. The PDN also consists of **"connectionless"** systems that use connectionless protocols to forward data packets through the network. This type of data transmission tends to reduce overhead requirements and therefore be faster. Finally, many modern networks use a combination of both connection-oriented and connectionless protocols to obtain the benefits of both technologies.

Private data networks use all the same technologies previously mentioned and may be constructed, owned, and maintained by the user or leased from some service provider. **Virtual private data networks**

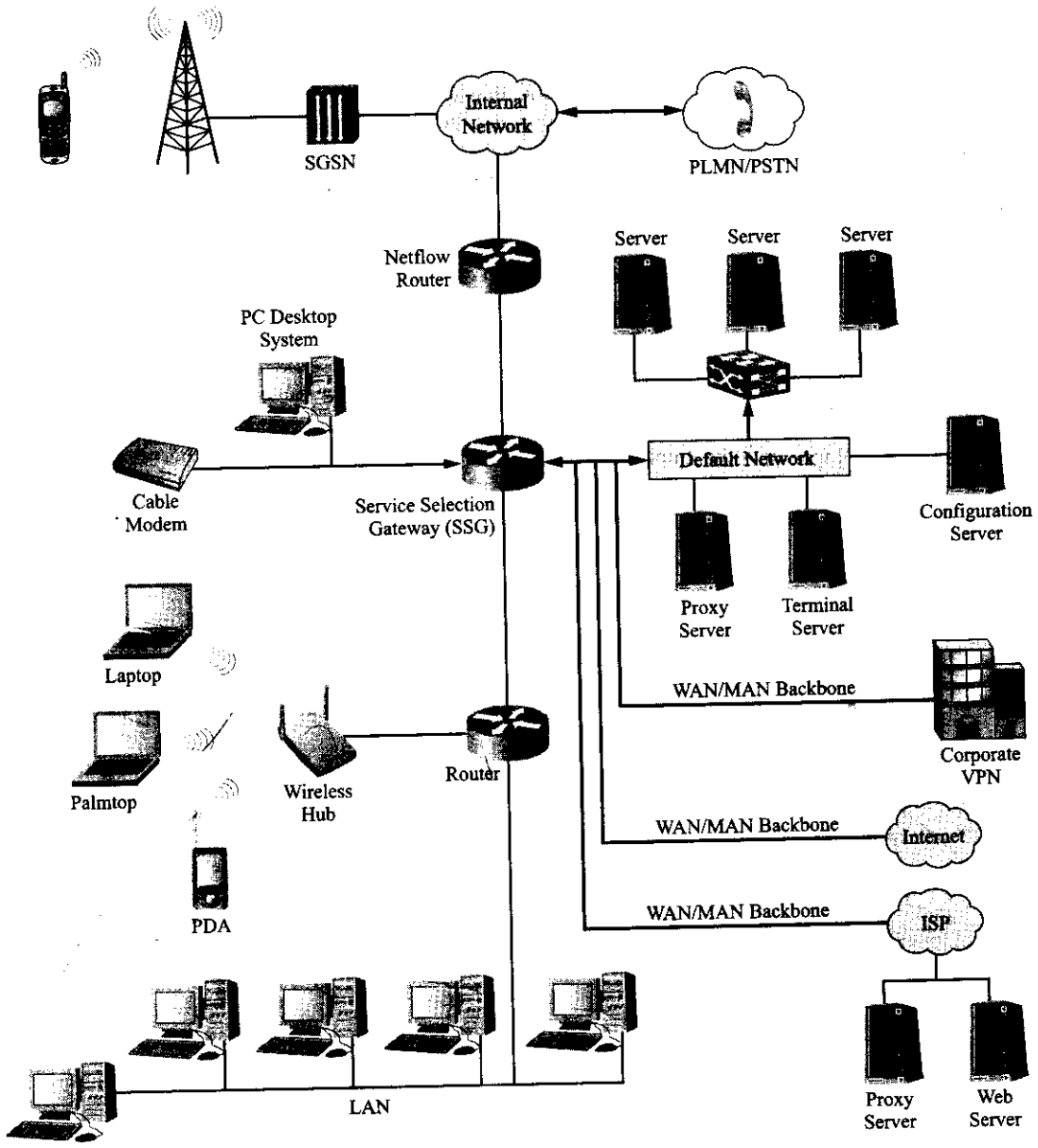

Figure 1–5 A depiction of the public data network.

use the public data network, maintaining privacy through the use of a **tunneling protocol** that effectively conceals the private network data and protocol information by encapsulating it within the public network transmission packets. Typically, further security is provided through the use of data encryption and decryption procedures.

Modern cellular telephone systems are currently in an evolutionary upgrade phase in an effort to provide mobile subscribers with high-speed connectivity to the PDN. Rapidly advancing technology for the implementation of wireless LANs and MANs is also providing this type of untethered connectivity at steadily increasing data rates and at an ever increasing number of geographical locations.

Broadband Cable Systems

Broadband cable systems have evolved from their early beginnings to become sophisticated and complex wideband networks designed to deliver analog and digital video signals (including HDTV), data, and plain-old telephone service to the subscriber. The video content can come from local off-air television stations, satellite feeds of network or distant-station program content, and local access facilities. The data service typically connects to an Internet service provider (ISP) and telephone service connects to the PSTN. The most important change in the legacy cable-TV plant is the migration to the two-way hybrid fiber-coaxial cable system shown in Figure 1–6. The bandwidth of cable systems has been expanded to 870 MHz, and the use of the frequency spectrum between 5 and 42 MHz now allows for upstream data transmission over the network. Another important aspect to the evolution of the cable system is the development and standardization of the cable modem (CM). The data-over-cable-service interface specification (**DOCSIS**) project has led to multiple-vendor interoperability of cable modems located at the subscriber premise and cable modem termination systems (CMTS) located at the cable service providers' network centers or "head ends." These systems allow for a shared high-speed data connection over the cable network to the Internet (access to the Internet is provided at the CMTS) that passes Ethernet packets to and from the subscriber's cable modem to the subscriber's PC. The modern broadband cable network has become just one more connection to the public data network.

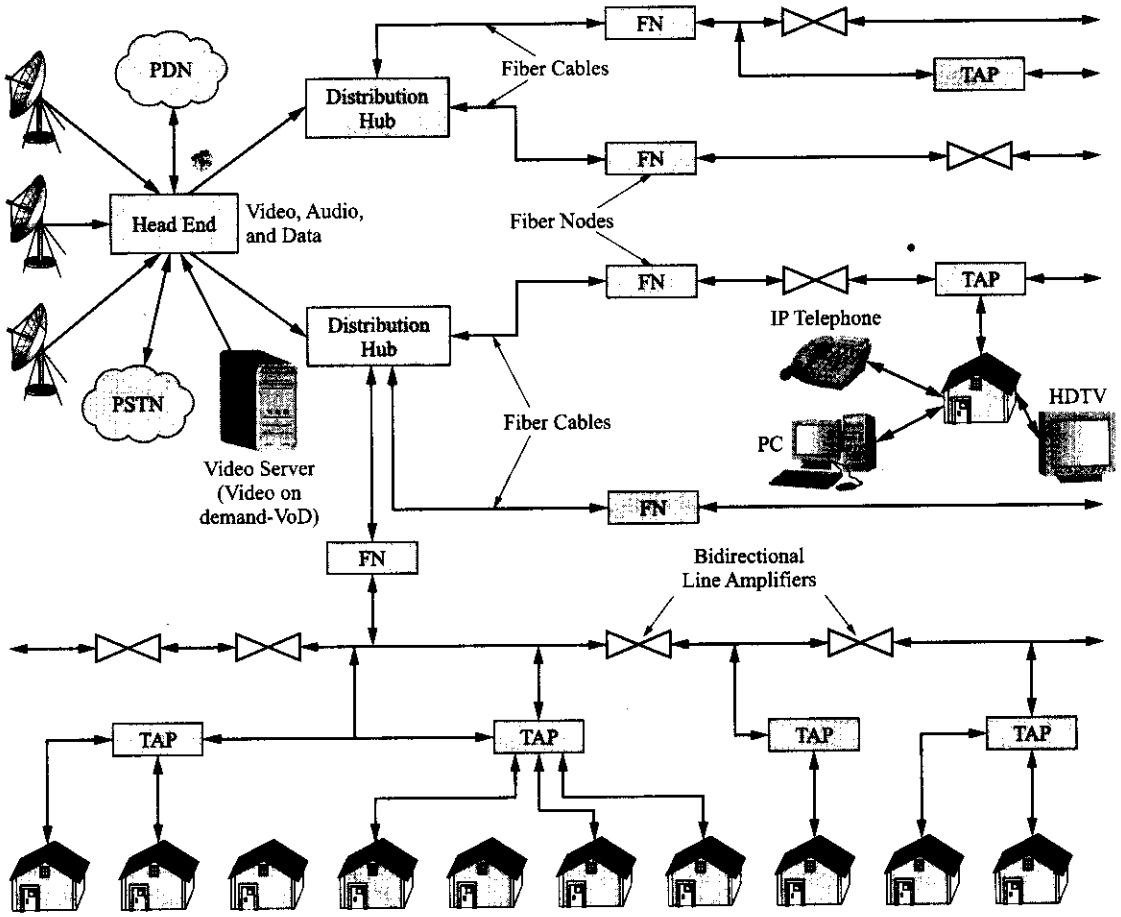

Figure 1–6 Modern two-way hybrid fiber-coaxial cable-TV system with fiber nodes.

The Internet

The Internet is the world's largest computer network. Over the Internet any computer or computer network may access any other computer or computer network. The structure of the Internet is shown conceptually in Figure 1–7. It consists of thousands of computer networks interconnected by dedicated special-purpose switches called routers. The routers are interconnected by a **wide area network** (WAN) backbone. This WAN backbone actually consists of several networks operated by national service providers (SprintLink, UUNet Technologies, internet MCI, etc.) These networks consist mainly of high-speed, fiber-optic, long-haul transport systems that are interconnected at a limited number of hubs that also allow for the connection of regional ISPs.

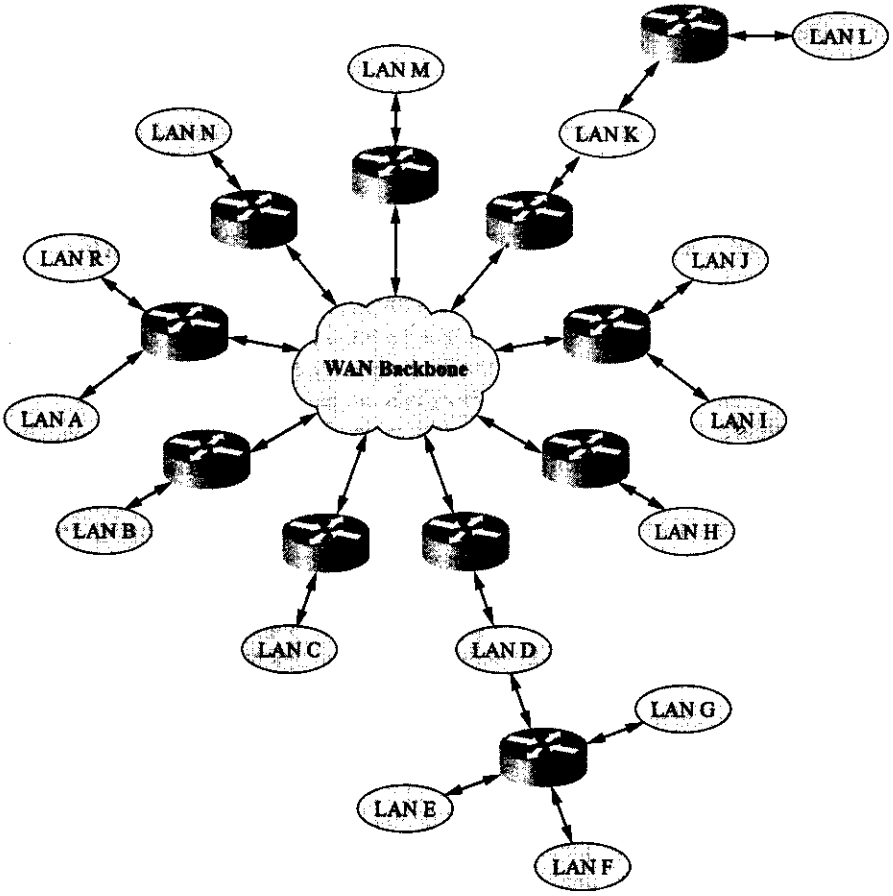

Figure 1–7 Conceptual structure of the Internet.

These national service provider (NSP) networks are interconnected to each other at switching centers known as network access points (NAPs). Regional ISPs may tap into the backbone at either the NSP hubs or the NAPs. If an individual wants to connect to the Internet, he or she must usually go through an ISP. The user might connect to the ISP through the PSTN over a low-speed dial-up connection using a modem that communicates with a "modem pool" at the ISP, or through high-speed cable-modem or ADSL (adaptive digital subscriber line) service. These services are usually connected through the PDN to the ISP. A **local area network** (LAN) at an Enterprise location will usually be connected to the ISP through some type of high-speed connection to the PDN (usually leased from a service provider) and then through the

ISP's high-speed connection to the PDN. The ISP will in turn be connected to the Internet through another high-speed network connection.

Today, one may be connected to the Internet by a wireless device while roaming or while connected to a LAN. Cellular telephones and personal digital assistants (PDAs) allow one to connect through the packet data network. The Web "browser" experience is not the same as with a desktop computer, but it is an Internet connection nevertheless.

Cellular Telephone Systems

Since their first deployment some two decades ago, cellular telephone systems have grown at a phenomenal rate. The public has been quick to adopt cellular technology, and the operator's networks have expanded to become national in scope. The technology used to implement cellular systems has also quickly evolved from analog (first generation or 1G), to digital (second generation or 2G), to systems with medium-speed data access (called 2.5G). High-speed data-access third-generation or 3G systems are already being deployed worldwide. Cellular operators have expanded coverage and capacity by using new frequency allocations, new air interface technologies, and cell splitting, and they have increased the functionality of their systems by expanding their scope to include access to the PDN, as well as the PSTN.

1.3 OVERVIEW OF EXISTING NETWORK INFRASTRUCTURE

Figure 1–8 shows an overview of the existing network infrastructure in place today. As mentioned previously, wireless telecommunication systems and networks perform the function of connecting to the existing network infrastructure. The three major types of traffic carried by the telecommunications network infrastructure are voice, video, and data (often known collectively as multimedia).

The PSTN was originally designed for voice transmission. To provide this function, the PSTN was structured in such a way as to provide a circuit-switched path for the conversation, which occurs in real time and therefore requires a certain Quality of Service (QoS). This physical path would be set up during the dialing of the call and torn down at the completion of the call. Supervisory, alerting, and progress tones and signals are generated by the system to facilitate creation of the connection and perform call handshaking functions. The network would take care of authentication and billing functions. Today the PSTN is an almost entirely digital system except for the analog signals that propagate over the copper wire pairs that provide a subscriber access to the network. The cellular telephone system gives a subscriber access to the PSTN.

The fixed network infrastructure developed to transport broadband analog video or television signals to the public is the hybrid fiber-coaxial cable (HFC) broadband cable television network. This system broadcasts the same video signals to all the subscribers connected to the network. A cable modem or set-top box allows the system to provide different levels of access to the entire suite of video signals transmitted over the system. If a subscriber has paid for premium services, these services are allowed to be passed to the subscriber's television tuner or are decrypted if they had previously been encrypted or "scrambled."

Through system upgrades and rebuilds and use of the DOCSIS standard, the modern cable television network has become bidirectional, allowing both downstream and upstream data transmission. Most cable operators now offer shared high-speed Internet access over their systems. Presently, cellular systems do not provide access to this network infrastructure. However, one may easily set up a wireless LAN to connect to this infrastructure in one's own home or apartment.

The data network was originally developed to carry bursty data traffic for business and industry. As technology has evolved the Enterprise data network has also evolved. Today's Enterprise data networks tend to have a wide area network (WAN) or **metropolitan area network** (MAN) high-speed backbone with a collection of local area networks (LANs) connected to it. This backbone network may be dedicated or switched and might use several different types of data transport technologies. Voice, data, and video can share these

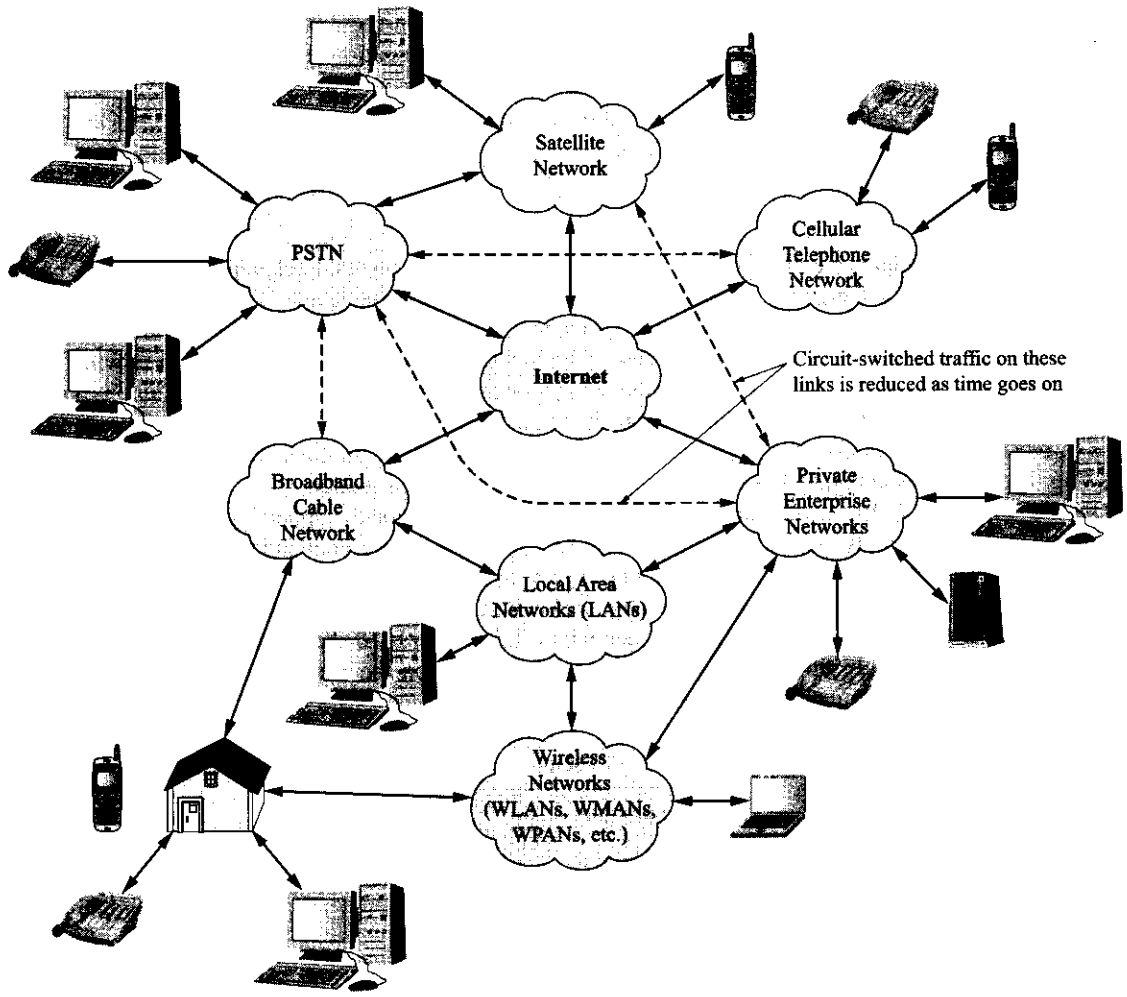

Figure 1–8 Today's existing network infrastructure.

transport facilities. Usually, the Enterprise private branch exchange (PBX) is also connected to the high-speed backbone as well as the PSTN. The Internet and the PDN are also interconnected. Some would argue that they are one and the same! As Voice over IP (VoIP) becomes more popular that line will blur even more. Wireless data networks tie into this infrastructure at the Enterprise level through wireless LANs and through cellular systems connected to the PDN.

1.4 REVIEW OF THE SEVEN-LAYER OSI MODEL

As we move toward a totally digital telecommunications infrastructure it is important to have some knowledge of the Open System Interconnection (OSI) reference model. This model describes how information moves from a software application in one computer to a software application in another computer either over a simple network or through a complex connection of networks or **internetwork.** An internetwork is usually defined as a collection of individual networks that are connected together by intermediate networking devices. To the user, an internetwork functions as a single large network.

Since one of the major functions of today's wireless systems and networks is to link the user to the installed wireline, wireless, coaxial cable and fiber-optic telecommunications infrastructure (i.e., PSTN and PDN), it is very often instructive to examine that particular interconnection and other wireless system interconnections through the OSI model. Therefore, this section will give a brief overview of the OSI model with emphasis on the lower layers of the model.

The OSI Model

The OSI reference model is a conceptual model that consists of seven layers. Each layer of the model specifies particular network functions. The model was developed by the International Organization for Standardization (ISO) in the 1980s, and it is now considered the major architectural model for network and internetwork data communications. The OSI model divides the tasks that are involved with moving information between networked computers into seven groups or layers of smaller, more manageable tasks. The tasks assigned to each layer are relatively autonomous and therefore can be implemented independently of tasks in other layers. This allows for changes or updates to be made to the functions of one layer without affecting the other layers. Figure 1–9 shows the seven layers of the OSI model.

Usually, the seven-layer model can also be divided into two categories: upper layers and lower layers. The upper layers are usually associated with application issues and are implemented in software. The lower layers handle data transport issues. The lowest two layers, data link and physical, are implemented in hardware and software. The lowest layer, the physical layer, is a description (electrical and physical specifications) of the actual hardware link between networks. Figure 1–10 shows the two sets of layers that make up the OSI model. In this text, emphasis generally will be on the data transport layers and on the physical layer when discussing particular air interfaces.

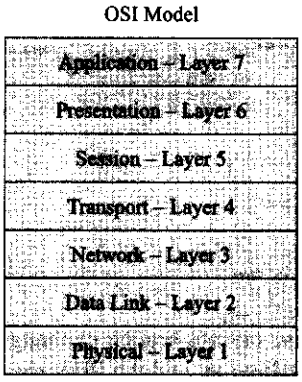

Figure 1–9 The seven-layer OSI model.

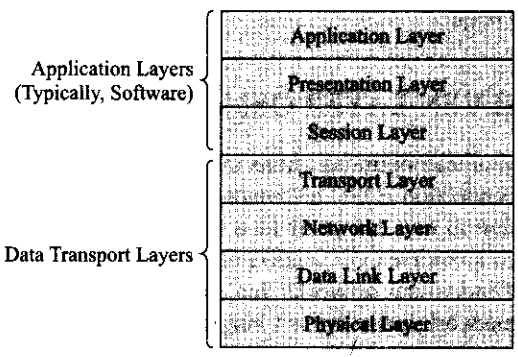

Figure 1–10 The two sets of layers that make up the OSI model.

Protocols

Although the OSI model provides one with a conceptual structure for the communication of information between computers, it itself is not a means of communication. The actual transmission of data is made possible through the use of communications **protocols.** A protocol is a formal set of rules and conventions that allows computers to exchange information over a particular network. A protocol is used to implement the tasks and operations of one or more of the OSI layers. There are many different communications protocols that are used today for different types of networks. LAN protocols define the transport of data for various different LAN media (CAT-n, fiber, wireless, etc.) and work at the physical and data link level. WAN protocols define the transport of data for various different wide area media and work at the network, data link, and physical level. Routing protocols work at the network layer level. These protocols are responsible for exchanging the required data between routers so that they can select the correct path for network traffic. Network protocols are the various upper-layer protocols that exist in any particular protocol suite. An example of a network protocol is AppleTalk Address Resolution Protocol (AARP).

Relation of OSI Model to Communications between Systems

Information that is being transferred from a software application in one computer through a computer network to a software application in another computer must pass through the OSI layers. Figure 1–11 depicts this process.

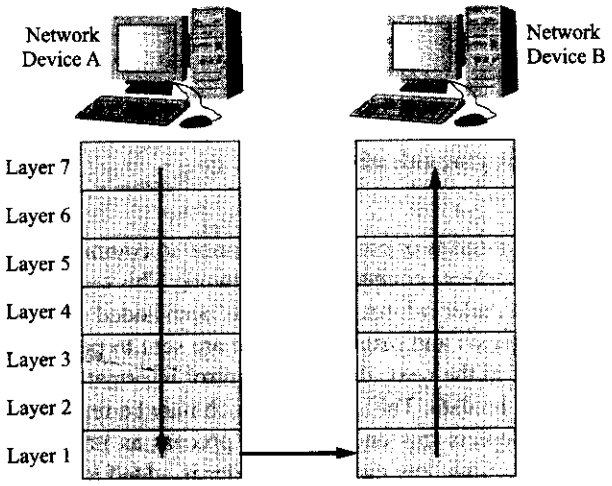

Figure 1–11 Information transfer between network devices as depicted by the OSI model.

Referring to Figure 1–11, data from an application in Device A is passed down through the OSI model layers, across the network media to the other OSI model, and up through the OSI layers to the application running on Device B. A closer look at this process might be instructive. The application program running on Device A will pass its information to the application layer (Layer 7) of Network Device A. The application layer of Device A will pass the information to the presentation layer (Layer 6) of Device A. The presentation layer then passes the information down to the session layer (Layer 5) of Device A and so on until the information is finally at the physical layer (Layer 1) of Device A. At the physical layer level, the information is placed on the physical network medium and transmitted (sent) to Device B. The process is reversed in Device B with the information being passed from layer to layer upward until it finally reaches the application layer (Layer 7) of Device B. At this point, the application layer of Device B passes the information on to the application program running on Network Device B.

More OSI Model Detail

In general, each OSI layer can communicate with three other layers: the layer directly below it, the layer directly above it, and the layer directly equivalent to it (its peer layer) in the other networked computer system (See Figure 1–12). The OSI layers communicate with their adjacent layers to make use of the services provided by these layers. The services provided by the adjacent layers in the OSI model are designed to allow a given OSI layer to communicate with its equivalent or peer layer in another computer system connected to the network.

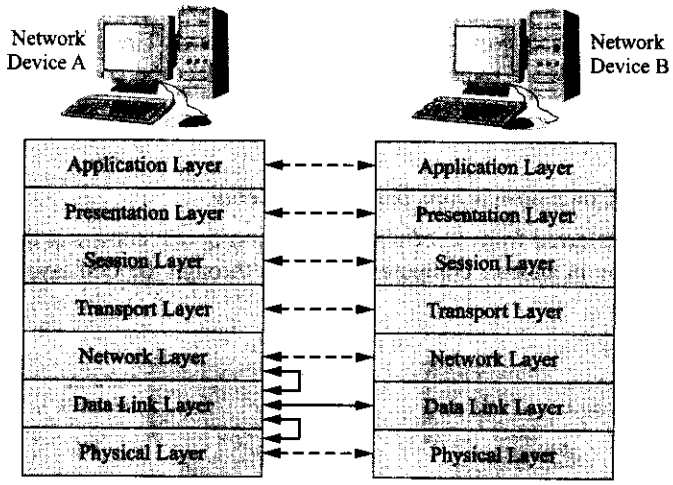

Figure 1–12 OSI model layers communicating with other layers.

The seven OSI layers use different forms of control information to communicate with their peer layers in other computer systems or system elements connected to a network. This control information (protocol specific) usually consists of header and trailer information that has been added to data passed down from upper layers, and it represents various requests and instructions that are sent to peer OSI layers. As the information is passed down from one layer to the next, the added control information from the prior layer is now considered by the next layer to also be data. This process, which may be repeated several times, is referred to as encapsulation. Figure 1–13 depicts this encapsulation process as information is passed downward through the seven-layer OSI model. Note how each new header is added to any previously added headers and the original data. Eventually, the entire assembled data unit is transmitted via the physical network connection.

Information Exchange

At this point it would be instructive to give an example of the entire process of information exchange between two computer systems that are both connected to a network. Computer System A has information from a software application to send to Computer System B. The information (data packet) is forwarded to the application layer. The System A application layer adds any control information to the data packet that will be needed by the application layer in Computer System B. The resulting information packet (control information and data) is forwarded to the presentation layer. The presentation layer now adds any required control information for the presentation layer of Computer System B. This process is repeated several more times as the information packet is forwarded to each succeeding lower layer. At each layer the information packet grows in size as the required control information is added to the packet. Finally, at the physical layer the entire data packet is placed onto the network medium. Refer to Figure 1–13 again. The physical layer of

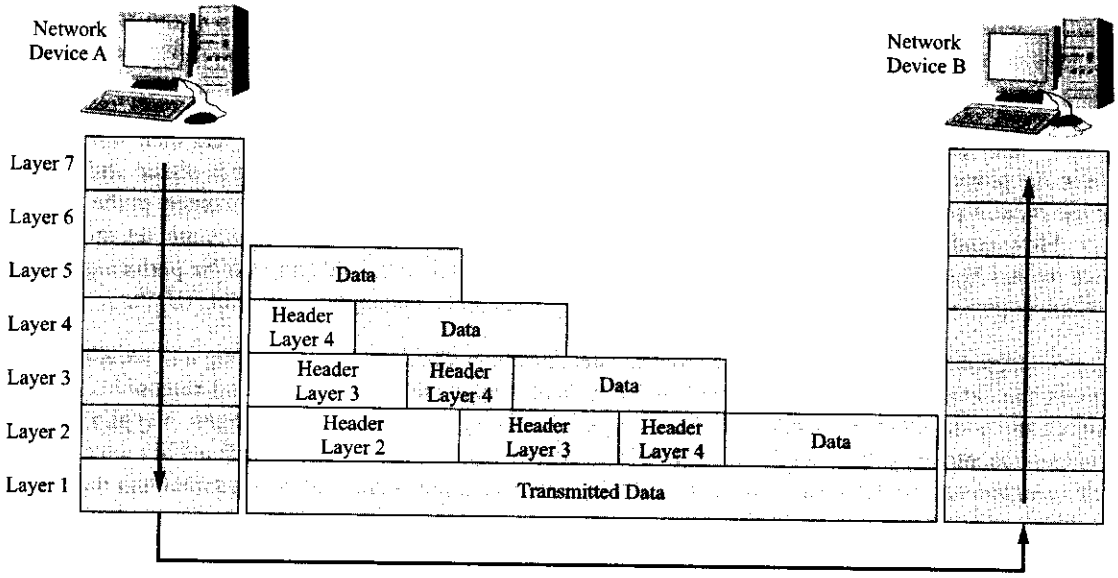

Figure 1–13 The process of data and header encapsulation during information exchange over a network.

Computer System B receives the data packet and forwards it to the data link layer of System B. The data link layer of System B reads the control information contained in the data packet, strips this information from the packet, and forwards the remaining data packet to the network layer. The network layer performs the same type of process, reading the control information meant for it, stripping this information from the data packet, and forwarding it upward to the next layer. This process is repeated by each layer until finally the application layer forwards the exact same data packet sent by Computer System A to the software application running on Computer System B.

Overview of the OSI Model Layers

As a final wrap-up to coverage of this topic, this section will present an overview of each OSI model layer. Where possible, examples of typical layer implementations will be presented.

Layer 7—Application Layer The application layer is closest to the end user. By that, we mean that the OSI application layer and the end user both interact with the software application that is running on the computer. Some examples of application layer implementations include File Transfer Protocol (FTP) for file transfer services, Simple Mail Transfer Protocol (SMTP) for electronic mail, Domain Name System (DNS) for name server operations, and Telnet for terminal services.

Layer 6—Presentation Layer The presentation layer provides a variety of conversion and coding functions that are applied to application layer information/data. This is done to assure that information sent from the application layer of one machine is compatible with the application layer of another machine (possibly a different type of computer system). Some of the types of coding and conversion that are performed are common data representation formats (standard multimedia formats), conversion of character representation formats (e.g., EBCDIC and ASCII converted to a syntax acceptable to both machines), common data compression schemes (e.g., GIF, JPEG, and TIFF), and common data encryption schemes. In each case, the use of these presentation layer coding and conversion functions allows data from the source system to be properly interpreted at the destination system.

Layer 5—Session Layer The session layer has the task of establishing, managing, and terminating communications sessions over the network. An example of a session layer implementation is the AppleTalk Session Protocol (ASP) from the AppleTalk protocol suite.

Layer 4—Transport Layer The transport layer accepts data from the session layer and then segments the data in the proper manner for transport across the network. Since the transport layer is tasked with delivering the data in proper sequence and in an error-free fashion, flow control is implemented at the transport layer. Flow control manages the data transmission between devices. Virtual circuits are set up and torn down by this layer, error checking and correction is executed, and multiplexing may be performed. Familiar transfer protocols are Transfer Control Protocol (TCP) and User Datagram Protocol (UDP).

Layer 3—Network Layer The network layer isolates the upper OSI layers from routing and switching functions in the network. The functions within the network layer create, maintain, and release connections between the nodes in the network and also manage addressing and routing of messages. A typical network layer implementation is Internet Protocol (IP). With IP, network addresses may be defined in such a way that route selection can be determined systematically by comparing the source network address and the destination address and applying the subnet mask.

Layer 2—Data Link Layer The data link layer provides for the reliable transmission of data across a physical network connection or link. Different data link layer specifications define different network and protocol characteristics. Some of these characteristics include physical addressing, error notification, network topology, sequencing of frames, and flow control. The Institute of Electrical and Electronics Engineers (IEEE) has subdivided the data link layer into two sublayers: Logical Link Control (LLC) and Media Access Control (MAC). The LLC sublayer of the data link layer manages communications over a single link of a network. The MAC sublayer of the data link layer manages protocol access to the physical network medium. The IEEE MAC specification defines MAC addresses. This specification allows multiple devices to be uniquely identifiable at the data link layer.

Layer 1—Physical Layer The physical layer defines the electrical and mechanical specifications for the physical network link. Characteristics such as voltage levels, timing, physical data rates, maximum distance of transmission, and physical connectors are all part of this specification for wireline media. For fiber-optic media similar specifications exist for the type of network transport technology employed (FDDI, ATM, SONET, etc.). Usually, physical layer implementations can be categorized as LAN, MAN, or WAN specifications. For cellular wireless networks, characteristics such as frequency of operation, channel format, modulation type, timing, hopping sequences, coding, and other technical specifications are grouped under the term "air interface" and depend upon the particular cellular system being discussed. Today the most prominent players in the development and publication of cellular specifications are the European Telecommunications Standards Institute (ETSI), the Telecommunications Industry Association (TIA), and the two Third-Generation Partnership Projects: 3GPP and 3GPP2.

For wireless LANs (local area networks), PANs (personal area networks), and MANs (metropolitan area networks) the IEEE produced the IEEE 802.11x, 802.15x, and 802.16x specifications (respectively) to cover the particular technology implementation.

The physical layer also defines the procedural and functional specifications for activating, maintaining, and deactivating the physical link between the devices communicating over the network systems.

During the discussion of various wireless systems and networks in other parts of this book, the operation of these systems will at times be illustrated through the relationship of the particular operations with the layers in the OSI model. For most of these illustrations, the lower transport layers will be of most importance and the physical layer will usually receive the most attention since it contains the details of the air interface that differentiate these various systems.

1.5 WIRELESS NETWORK APPLICATIONS: WIRELESS MARKETS

The markets for wireless services have evolved into two basic categories: the traditional voice-oriented market and the newer data-oriented market. The first market has enjoyed an amazing acceptance by the general public with an extremely fast take-up rate for the service offered—a connection to the PSTN via cellular telephone. With newer personal communications services (PCS) systems being built out on a different frequency band, there is the potential for a new generation of digital phones that offer voice, data, and fax services to the subscriber. However, there is actually little to differentiate between cellular and PCS service, and most network operators are offering both (subscribers have dual- and tri-mode phones) in an effort to achieve better coverage. Most operators are offering nationwide plans with an ever increasing number of minutes of system connection time per dollar. The newer data-oriented market has evolved around the Internet and computer LAN technology. More recently the cellular telephone data-oriented market has been driven by short messaging service (SMS), instant messaging (IM), and multimedia messaging service (MMS) applications and other novel entertainment- type applications. This shift has been noted by both service providers and content developers and is focusing applications development on entertainment- or "infotainment"-based uses of the cell phone. Wireless LANs have been steadily gaining in popularity and acceptance for both Enterprise and home use. The adoption of new standards that provide for the use of new unlicensed frequency bands and higher data rates have led to many predictions of a fast uptake rate for this technology and a potential threat to the cellular operators as they evolve their networks to offer faster data rates. In addition, wireless MANs and PANs are starting to be seen in the marketplace.

Voice Network Evolution

The development of voice-oriented wireless networks began in earnest during the early 1970s at AT&T's Bell Labs. The technology for frequency division multiple access (FDMA) analog cellular systems was developed but not deployed in the United States until much later in 1983 as the Advanced Mobile Phone System (AMPS). Similar systems were deployed in the Nordic countries a year earlier in 1982 as the Nordic Mobile Telephony (NMT) system. At about the same time, digital cellular networks were starting to be developed under the Groupe Spécial Mobile standardization group that eventually became the Global System for Mobile Communications or GSM. The **GSM** group was formed in an attempt to deal with international roaming, which was a serious problem for the European Union countries. This led to a new digital time division multiple access (TDMA) second-generation technology (i.e., GSM) that was deployed in many parts of the world beginning in late 1992. At present approximately 72% of cellular telephone users are serviced by GSM systems. In the United States, in an effort to increase capacity, a digital North American version of TDMA was introduced in the early 1990s. This new digital system was a hybrid air interface that used both first- and second-generation technology. At this time, its successor standard, TDMA IS-136, has a worldwide subscriber base of over 100 million users or approximately 9% of all cellular telephone users. The most recent entry into the cellular mix has been a **code division multiple access** (CDMA) technology-based system. First deployed in the United States in 1995, this standard now has a worldwide subscriber base of over 170 million users. An additional standard, personal digital cellular (PDC), is a Japanese TDMA-based standard. It also has a subscriber base of approximately 60 million users; however, some Japanese operators have already announced plans to phase out their PDC systems and shift to CDMA systems.

Although not as high profile as cellular technology, cordless telephones belong to the wireless voice network class of products also. First introduced in the late 1970s, these devices, which provided a wireless connection from the telephone handset to a fixed "base station," became an instant commercial success. Second-generation digital cordless telephones appeared in the early 1980s and the concept of the PCS device evolved in the early 1990s.

A PCS service was considered the next generation of residential cordless telephone. Although there have been some deployments of PCS systems worldwide, none of the PCS standards have become a major

commercial success or a competitor to the cellular telephone systems. In the United States the PCS bands have been put into use, but the PCS standards were not adopted for use in these bands. As mentioned before, most operators are using the PCS bands for cellular service and to fill gaps in their coverage area.

Data Network Evolution

The concept of data-oriented wireless networks started in the 1970s, but development did not start in earnest until the early 1980s. Amateur radio operators had built and operated simple wireless packet radio networks earlier, but commercial development of radio-based LANs did not begin until 1985 when the FCC opened up the industrial, scientific, and medical (ISM) bands located between 920 MHz and 5.85 GHz to the public. During the early 1990s the Institute of Electrical Engineers (IEEE) formed a "working group" to set standards for wireless LANs. The IEEE finalized the initial 802.11 standard in 1997 for operation at 2.4 GHz with data rates of 1 and 2 mbps.

Advances in wireless LAN technology have been occurring at a rapid pace. The IEEE 802.11x family of standards has been expanded to include operation in the 5-GHz U-NII bands with data rates of up to 54 mbps through the use of complex digital modulation schemes.

The IEEE has adopted two other families of standards, IEEE 802.15x for operation of wireless personal area networks (wireless PANs), also known as Bluetooth, and IEEE 802.16x for the operation of wireless metropolitan area networks (wireless MANs), also known as broadband wireless access. A great deal of promise lies in these new standards and the technologies that will be used to implement them, but only time will tell if they will enjoy widescale adoption.

There are parallel development activities occurring in Europe under the European Telecommunications Standards Institute (ETSI) for high-speed wireless LANs that appear to have characteristics similar to IEEE 802.11x-based products. The ETSI **HiperLAN/2** standard specifies operation in the 5-GHz band and also has the same maximum data rate of 54 mbps. More detailed coverage of these topics is given in Chapters 9 through 11.

Mobile data services were first introduced in North America with the ARDIS project sponsored by Motorola and IBM in 1983. Mobitex (an open version of ARDIS) was introduced in 1986 and then in 1993 **cellular digital packet data** (CDPD) service was introduced in the United States. Both ARDIS and Mobitex are based on proprietary packet-switched data networks. Mobitex service is still available from Cingular Wireless in the United States with data rates of about 8 kbps. CDPD service allows cellular systems to deliver packet data to subscriber phones albeit at low data rates (generally below 19.2 kbps). Currently there exist several data-only wireless operators. Data speeds over these networks range from less than 2.4 kbps for two-way paging applications to 19.2 kbps for the fastest systems. Note however that these peak data rates do not translate into real throughput rates. These values are typically 50% of the peak rate.

General packet radio service (GPRS) with its slightly higher user data rates of 20 to 50 kbps has been available over GSM systems since the early 2000s. GSM systems are implementing EDGE technology (2.5G to 3G) worldwide to achieve enhanced data rates. **EDGE** stands for Enhanced Data Rate for Global Evolution. The first operational CDMA systems offered data throughput rates to 14.4 kbps. The second phase of CDMA systems (IS-95-B) offer higher data rates (up to 56 kbps). The evolutionary pathway to third-generation (3G) CDMA systems includes a phasing in of greater packet data transfer rates. The first implementation phase (offered in 2002) of 3G is known as cdma2000 1xRTT and offers packet data rates of up to 144 kbps with real throughput rates of from 60 to 80 kbps.

Cellular service providers are spending a great deal of money to upgrade their systems to offer higher-speed data throughput rates to and from the PDN, as well as continued traditional access to the PSTN voice network. Future plans for all cellular technologies include upgrades to 3G technologies with even higher standard data rates and increased mobility. Will the public desire these digital services and subscribe to them? Only time will tell.

1.6 FUTURE WIRELESS NETWORKS

Present-day research efforts in the wireless field are geared toward the concept of seamless connectivity. It is conceived that in the not-too-distant future, an individual will be able to be connected to the installed telecommunication infrastructure in a seamless fashion. That is, the individual can be roaming throughout different service providers' networks that possibly use different delivery technologies and still be connected to the Internet without losing connectivity. Mobile IP will allow for both universal mobility and high data rate access either in a fixed location or while in motion. The user will not notice any loss of connectivity or change in service regardless of the conditions or the type of wireless network. Even before the installation of 3G cellular systems has become commonplace, 4G systems with ATM access speeds (over 100 mbps) are being discussed by the wireless research community. Many, including this author, believe that almost all access to the Internet will become wireless in the future. The future of wireless telecommunications technology appears to be unlimited!

QUESTIONS AND PROBLEMS

1. Do an Internet search for information about Mahlon Loomis. Write a short description of the theoretical operation of his patented aerial wireless telegraph system.
2. Do an Internet search for information about Marconi's first wireless experiments. What frequencies did Marconi first use for his wireless experiments?
3. Determine the length of a half-wave antenna for Fessenden's 50-kHz transmitter that he used at Brant Rock. What was the actual length of the antenna he used? Hint: Do an Internet search for information about Fessenden's early experiments. Hint: $\lambda = c/f$ where c is the speed of light and f is the frequency.
4. Use the Internet to research the deployment of over-the-air HDTV. By what date is HDTV broadcasting expected to be totally deployed?
5. What is the data rate of a DS0 signal?
6. In theory, how many DS0 calls can be transported by an OC-3 fiber-optic facility? After multiplexing to higher DSn rates, what is the practical capacity of DS0 calls that can be handled by an OC-3 facility?
7. What is the typical data transfer rate (in bps) over a SS7 transfer link?
8. Describe how a high-speed cable modem, xDSL service, or cellular telephone service extends the PDN.
9. In your own words, define the extent of a local area network (LAN).
10. In your own words, define the extent of a metropolitan area network (MAN).
11. In your own words, define the extent of a wide area network (WAN).
12. Describe the encapsulation process in the context of the OSI model.
13. At what OSI layer does flow control occur?
14. What is the function of the MAC sublayer?
15. Which OSI layer provides the specifications for the wireless air interface?
16. Go to an Internet Web site devoted to the GSM industry and determine the present total number of worldwide GSM subscribers.
17. Go to an Internet Web site devoted to the cellular telephone industry and determine the percentage of subscribers for the different major cellular telephone technologies (GSM, NA-TDMA, CDMA, PDC, etc.).
18. Go to the IEEE Wireless Standards Web site. Check the status of the IEEE 802.11 wireless LAN standard. Write a short one-paragraph report on the state of one of the IEEE 802.11 working group's amendments to 802.11.
19. Describe a cellular telephone use that would be considered an infotainment use.
20. Compare 3G cellular telephone data transfer rates with those available over wireless LANs. Comment on the difference. Is it important?

CHAPTER 2

Evolution and Deployment of Cellular Telephone Systems

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

- Discuss the concept of the different generations of wireless cellular systems.
- Discuss the evolution and deployment of wireless cellular systems on a worldwide basis.
- Explain the basic operations and structure of a 1G cellular system.
- Explain the difference between a 1G, 2G, and 2.5G cellular system.
- Discuss the different subscriber services available over 2G mobile systems.
- Discuss the characteristics of 3G wireless mobile systems.
- Explain the concept of 4G wireless.
- Explain the function of standards bodies.

The wireless mobile industry started many decades ago with systems that the prevailing technology of the day could support. These first rudimentary systems found restricted use in the fields of public safety and utilities, transportation, government agencies, and the like. Lack of accessible radio spectrum, inefficient transmission techniques, and immature technology made these systems expensive and not easily adaptable to mass markets. As time went on, technologic innovation allowed for the design of new mobile systems that were able to utilize heretofore unavailable radio spectrum and also allow for improved operation and reliability. Eventually, access to the public switched telephone system became available over limited-capacity mobile radio systems. New, more efficiently designed cellular telephone systems evolved from the earlier systems. Since the start of cellular service in the United States slightly over twenty years ago, there has been a tremendous expansion of cellular service that has resulted in nationwide networks that offer both voice and data services. These systems have undergone generational changes that have improved their performance and potential uses and in turn gained remarkable acceptance by the general public. Today, mobile cellular systems have well over a billion subscribers worldwide.

Recently, other wireless telecommunications systems and networks have become available that allow one to be connected to the Internet or some other data network through a wireless device. The market for wireless local area network equipment is growing at a very rapid pace. The vision of a wirelessly enabled mobile information society is becoming more widespread and more obtainable as technology keeps improving.

The wireless mobile industry is standards based. The need for a standards-based industry is dictated by today's global marketplace. In today's society, there is an inherent need for the interoperability of user

subscriber devices within different systems, in different locations, and increasingly in worldwide roaming situations.

This chapter will attempt to give the reader a feel for the fundamental nature of wireless cellular systems. This will be accomplished by providing an overview of the various generations of mobile cellular systems, the new types of technology involved, and a detailing of the steps entailed in the standardization, adoption, and deployment of these new generational systems. Some details of the evolution to third-generation systems and predictions about the fourth generation of wireless are given, but as always, no one can predict the future. Only time will tell how it all plays out.

2.1 DIFFERENT GENERATIONS OF WIRELESS CELLULAR NETWORKS

Aside from use by the military and transportation industries, some of the first strictly land-based two-way mobile radio systems commenced operation in the early 1930s in the United States. These systems were typically used for fleet communications by the public service sector (e.g., police and fire departments). Operating in a **time division duplex** mode, one mobile radio user at a time could talk and then use two-way radio jargon to indicate who should speak next or if the communication was over. Then in 1946, AT&T and Southwestern Bell commenced operation of a mobile radio-telephone service to private customers in St. Louis, Missouri. The system operated on a small number of channels licensed by the FCC in the 150-MHz band. It was not until 1947 that AT&T, on behalf of the Bell Operating Companies, petitioned the FCC for additional radio frequency spectrum meant for use by a mobile radio system that would connect the user to the public switched telephone network. The FCC granted the use of a limited number of channels for this application in 1949. At the time, the FCC felt that the interests of the public were better served by the legacy public land radio services than public use of this new service proposed by AT&T. As it turned out, this new technology, known as mobile telephone service (MTS), was extremely popular in large metropolitan areas and its capacity became totally exhausted by the mid-1950s. Technical improvements like automatic dialing were quickly added to the MTS system making the system easier to use and more transparent to the user.

The cellular telephone concept that had first been put forward in the late 1940s received minimal research and development effort during the 1950s, primarily due to the FCC's continued reluctance to increase the amount of available spectrum for MTS use. Even so, the cellular concept and possible models for its implementation were the subject of several internal Bell Laboratory technical memoranda in the late 1950s that later served as the basis for several publicly available journal articles published by the Institute of Radio Engineering (IRE) in 1960. Even though the Bell system petitioned the FCC for additional spectrum for the MTS system in 1958, the FCC did not act on the request. In 1964, the Bell system began to introduce Improved Mobile Telephone Service (IMTS). This new service allowed full duplex operation (i.e., both parties could talk at the same time) and provided for automatic channel selection, direct dialing, and more efficient use of the spectrum by reducing channel spacing. However, IMTS did not increase the capacity of the system enough to meet the public demand. It should be pointed out that the other providers of mobile radio services, the radio common carriers (RCCs), which owned half of the available mobile frequencies, constantly opposed the Bell system's requests to the FCC and in essence helped delay the implementation of any new high-capacity technology. Ultimately, ten years later, in 1968, in response to the backlog of requests for MTS and IMTS service, the FCC asked for technical proposals for a high-capacity and efficient mobile phone system to augment or replace the current system.

AT&T proposed a new mobile phone system using the cellular concept. Through the use of small coverage areas or **cell** sites, many low-power transmitters would be used to provide coverage to a metropolitan area. Furthermore, the use of low-power transmitters with their limited range would allow for the reuse of the scarce number of radio frequencies or channels available to the entire system on the basis of a much shorter spacing than previously feasible with earlier systems. Additionally, the system would implement a

process by which the mobile subscriber would be "handed off" as needed to a new cell site as the subscriber moved about the metropolitan area.

In 1970, Bell Laboratories, under authorization from the FCC, tested its cellular concept with prototype systems operating in the Newark, New Jersey, and Baltimore, Maryland, areas. In 1971, Bell Labs reported that its tests had proven that the cellular concept worked with cells as small as 2.8 miles in diameter. In 1974, the FCC released some 40 MHz more of frequency spectra for the development of early analog modulation-based cellular systems. There are reports that in 1976, 545 customers in New York City had Bell system mobile phones and a waiting list for this service had over 3700 names on it. After more delays, in 1978, a trial cellular telephone system, known as the Advanced Mobile Phone System (**AMPS**), was put into operation in the Chicago area by Illinois Bell and AT&T using the newly allocated 800-MHz band. Shortly thereafter, a service test with real customers was conducted and it proved that a large cellular system could work. Worldwide commercial AMPS deployment followed quickly but not in the United States! The impending breakup of the Bell system and the FCC's new competition requirement delayed commercial deployment. Finally, in 1983, commercial AMPS operation began in the United States.

In July of 1983, the FCC released Bulletin No. 53 from the Office of Science and Technology (OST). This bulletin, titled "Cellular System Mobile Station—Land Station Compatibility Specification," provided the core specifications and thus became the defining standard for the AMPS system. This document was developed to ensure the compatibility of mobile stations with any cellular system operating in the United States. To ensure compatibility, it was essential that both radio-system parameters and call-processing procedures be specified. Release of this document thus heralded the start of the use of technical specifications or standards-based technology for the development of modern cellular radio.

Earlier, an AMPS system with eighty-eight cells began operation in Tokyo in late 1979, and in the Nordic countries of Norway, Denmark, Finland, and Sweden a similar analog-based, voice-oriented, AMPS firstgeneration or 1G cellular system was put into operation in 1981 without the bureaucratic delays that were experienced in the United States. This first multinational cellular system, known as the Nordic Mobile Telephone (NMT) system, used the 450-MHz band and immediately became extremely popular.

The rest of the world was not waiting for the United States to create a universal standard and therefore several different systems evolved in different technologically advanced areas of the world. These systems used slightly different technical implementations and very often used different portions of the frequency spectrum for their first-generation systems. The use of different frequency bands is due to the fact that each country has its own frequency administration agency and had made previous frequency allocation decisions for various other legacy radio services or spectrum utilization schemes. These decisions were usually based on use of the particular service only within the individual country without regard for use in other countries. Only in recent years when the World Administrative Radio Conference has met has worldwide spectrum coordination started to result in the almost harmonious use of several different bands of radio spectrum for the same mobile service. In many instances, however, legacy uses of various radio services still preclude the universal use of much of the radio spectrum. Nevertheless, there are usually frequencies close to the desired bands that nations can free up for the same type of radio technology and service offerings as most other countries worldwide.

The legacy of this reality is still with us today as several prominent first- and second-generation cellular systems still maintain popularity around the globe. Third-generation or 3G systems are theoretically going to bring the world closer to a universal standard in the near future, but most feel that day is still many years away.

As new technology has been deployed to upgrade cellular system capacity and functionality, comprehensive technology changes have been designated as new generational systems. Digital modulation schemes are generally referred to as second-generation or 2G technology. The GSM system first deployed in the European countries and now worldwide is considered a second-generation technology as is North American TDMA or IS-136. Within a particular generational technology there are usually many updates and changes to the standard to reflect the use of newly evolving technology, new frequency spectra allocations, and new

functionality or applications built into the system. The ability to provide medium- to high-speed data access to and from the public data network over a cellular telephone system has resulted in a half-generational step that is presently referred to as 2.5G (halfway between second- and third-generation technology). The next generation of cellular telephones with functionality that meets the recently adopted IMT-2000 (International Mobile Telecommunications—2000) standards are referred to as third-generation or 3G technology.

Many are already referring to cellular telephone systems with more advanced functions and near asynchronous transfer mode (**ATM**) data transfer speeds as 4G technology!

2.2 1G CELLULAR SYSTEMS

As previously mentioned, the first analog-based, voice-oriented cellular telephone systems, which became available in other countries during the late 1970s to the early 1980s and in the United States during 1983, are now referred to as first-generation or 1G cellular technology. This chapter will provide an overview of the characteristics and operational aspects of these first-generation systems. While there are several other types of first-generation systems, most of this chapter's coverage will be devoted to the AMPS technology first deployed in the United States. Although the requirement to provide support for AMPS technology is now due to be phased out in the United States by 2007, it is instructive to look at the technical characteristics of AMPS because all succeeding generations of cellular telephone systems have evolved from this earlier technology.

Introduction

All first-generation cellular systems used analog frequency modulation schemes for the transmission of voice messages with two separate bands for **downlink** (from base station to mobile) and **uplink** (from mobile to base station) transmissions. This type of system is known as **frequency division duplex** (FDD). Also, within these two separate bands, frequency division multiplexing (FDM) is used to increase system capacity. The exact characteristics of the audio channel frequency response, other audio-processing details, and the allowed transmitter frequency deviation were defined by the particular system standard. The channel spacing was typically set by the appropriate regulatory agency (the FCC in the United States) as were the allowed frequency bands of operation (channels).

Identification (ID) numbers were assigned to the cellular system (SID) and the subscriber's device (mobile transmitter or handset). These ID numbers are used to determine mobile status (within home area or roaming), to perform authentication of the mobile, and to define the mobile's telephone number for correct operation of the network.

The system standard further defines physical layer technical parameters such as maximum permissible power levels, audio preemphasis standards, and maximum out-of-band emission levels. Most importantly, the standard sets the required procedures for the operations between the mobile subscriber's device and the cell site transmitter or base station. The standard also prescribes the required protocols and signals necessary for the successful exchange of messages between the mobile and the base station that will implement these operations.

AMPS Characteristics

The AMPS system began operation in the 800-MHz band with the eventual following frequency assignments. The downlink or forward band was from 824 to 849 MHz and the uplink or reverse band was from 869 to 894 MHz. The channel spacing was set at 30 kHz and each base station's transmit and receive frequency was separated by 45 MHz. The FCC introduced competition into the mobile phone arena by dividing the allotted frequency spectrum into "A" and "B" bands. The A band was allocated to one service provider and the B band was allocated to another service provider within a specific serving area. These

serving areas were created primarily from statistical data from the U.S. Office of Management and Budget and were known as cellular market areas (CMAs) similar to the concept of basic and major trading areas—BTAs and MTAs. In the vast majority of these market areas, one of these service providers was the incumbent Bell Telephone Company (assigned channels in the B band by default).

AMPS Channels

Initially, the A and B bands both consisted of 333 channels. Of these 333 channels, Channels 1–312 in the A band were **traffic channels** (TCHs) used for the subscriber's calls, and channels 313–333 in the A band were used for system control functions. These 21 **control channels** are used by the mobile and base station to set up and clear calls and other network operations such as handoff. The B band used channels 334–354 for control channels and channels 355–666 for traffic channels. An additional 5 MHz of spectrum was later added for use by the system with the channels again evenly split between the two operators. This yielded a total of 416 traffic and control channels per operator (see Table 2–1). The system operators are able to utilize the control channels in whatever way they deem most appropriate. Therefore, in most cases, operators group the voice channels and associate each group with a particular control channel to increase the effectiveness of the system.

Table 2–1 Table of AMPS channel numbers and frequencies.

System Band	*Bandwidth in MHz*	*Number of Channels*	*Boundary Channel #s*	*Transmitter Center Frequency in MHz*	
				MS	*BTS*
A	10	333	1 to 333	825.030 to 834.990	870.030 to 879.990
B	10	333	334 to 666	835.020 to 844.980	880.020 to 889.980
A[1]	1.5	50	667 to 717	845.010 to 846.480	890.010 to 891.480
B[1]	2.5	83	717 to 799	846.510 to 848.970	889.510 to 883.970
A[1]	1	33	991 to 1023	824.040 to 845.000	869.040 to 870.00
Not Used	N/A	1	990	824.010	869.010

[1]Additional frequency spectrum added later to system

AMPS System Components and Layout

As shown in Figure 2–1, the typical early AMPS cellular system consisted of the following components: several to many **base stations,** many **mobile stations,** and a **mobile telephone switching office** (MTSO). Today the MSTO has been replaced by the mobile switching center or MSC. The base stations (often

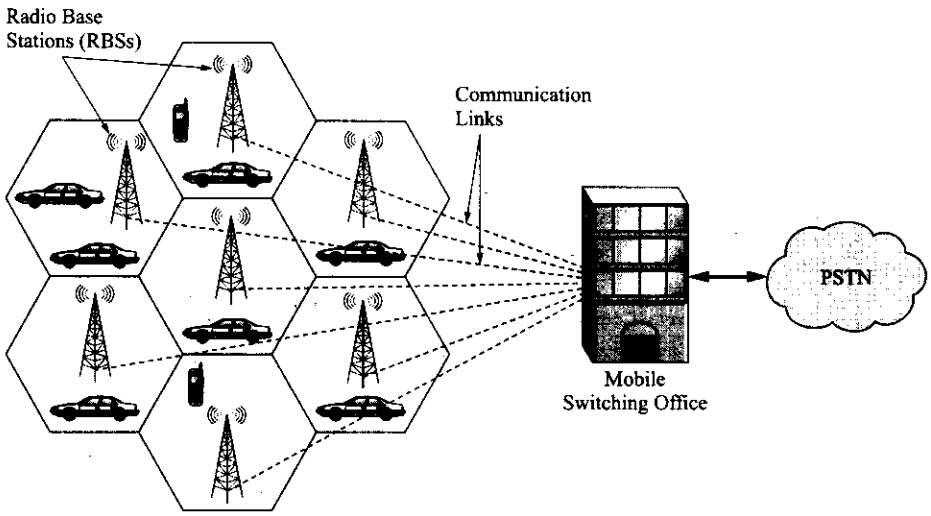

Figure 2–1 An early AMPS cellular system.

referred to as base transceiver stations) form cells that provide coverage to mobile subscribers over a particular geographic area. The base stations are connected to the MTSO that is in turn connected to the public switched telephone network (PSTN). Together, the base stations and the mobile stations provide the **air interface** that permits subscriber mobility while connected to the PSTN. The MSC performs system control by switching the calls to the correct cells, interfacing with the PSTN, monitoring system traffic for billing, performing various diagnostic services, and managing the operation of the entire network. The mobile unit is a frequency and output power agile radio transceiver that has the ability to change its operating frequencies to those designated by the MSC and its output power level if so instructed. The base station provides the interface between the MSC and the mobile subscriber. The base station receives both signals and instructions from the MSC that allow it to receive and send traffic to the mobile station.

Typical AMPS Operations

The first part of this section will provide an overview of the typical operations performed by the mobile station and the base station. The second part of the section will give a brief overview of the operations that occur between the base stations and the MTSO. The purpose of providing this information is to give the reader an insight into typical cellular system operation. Very little detail will be included about the exact nature of the signals or the formats of the data sent between cellular system components (the reader can obtain OST Bulletin No. 53 from the FCC Web site if more detail is desired). The details of more advanced cellular systems (second generation and up) will be covered in other chapters of this book.

The AMPS base station uses the dedicated control channels mentioned previously to send a variety of control information to idle (turned on but not being used) mobile stations within its cell, and the mobile stations use the corresponding reverse control channel to communicate with the base station while in the idle mode. When the mobile station is engaged in a voice call, control and signaling information may be also be transmitted over the traffic channel being used by the mobile and base station. Figure 2–2 depicts the flow of information over these channels. The need to transmit "radio link status" signaling information over active voice channels is facilitated by the use of supervisory audio tones (SATs), also known as **analog color codes.** Three SAT frequencies are used: 5970 Hz, 6000 Hz, and 6030 Hz. These tones give the base and mobile station the ability to keep informed about each other's transmitting capabilities and to confirm the success or failure of certain mobile operations. The base station periodically adds a SAT signal to the forward voice channel (FVC), thus transmitting it to the mobile station. The mobile station, acting like a **transponder,** transmits the same frequency tone on the reverse voice channel (RVC) back to the base

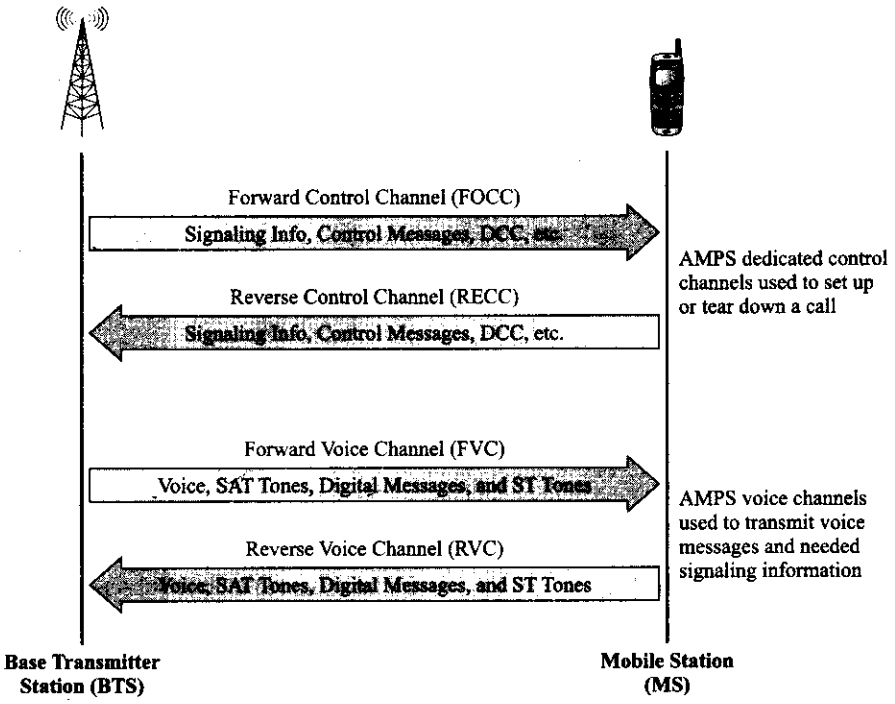

Figure 2–2 AMPS forward and reverse control and voice channels.

station. If for whatever reason a mobile station is captured by an interfering base station or a base station is captured by an interfering mobile station, this situation will be detected by the system due to the return of an incorrect SAT and the mobile receiver will be muted. A similar function is performed by the transmission of a **digital color code** (DCC) (an example of overhead information) over the forward control channel (FOCC) by the base station and returned over the reverse control channel (RECC) by the mobile station. A SAT color code (SCC) may also be transmitted to the mobile within certain mobile station control messages.

Additionally, a signaling tone (ST) of 10 kHz can be transmitted over a voice channel to confirm orders and to signal various requests. In some cases, signaling over active voice channels is accomplished through the use of changes in the SAT status or through the use of short bursts of the signaling tone or a combination of the two. For instance, the handoff operation makes use of both the SAT and ST signals to first initiate and then complete this process.

Additionally, both the forward voice channel and the reverse voice channel may be used to transmit digital messages from the base station to the mobile station and from the mobile station to the base station as needed. The base station may transmit Mobile Station Control messages that specify orders to the mobile over the forward voice channel, and the mobile station may transmit two types of messages over the reverse voice channel—an Order Confirmation message or a Called-Address message. The response to a digital message sent to the mobile station over the FVC will be either a digital message or a status change of SAT and/or ST signals transmitted back to the base station over the RVC. Voice signals are inhibited when digital messages are sent over these channels. The SAT and ST signals are filtered out of the audio delivered to the mobile user.

More Details When an AMPS mobile station is turned on but not connected to the PSTN for a telephone call, it tunes to the strongest control channel in its area and locks on to it. The mobile station will continuously monitor this control channel to receive control information from the base station. The base station

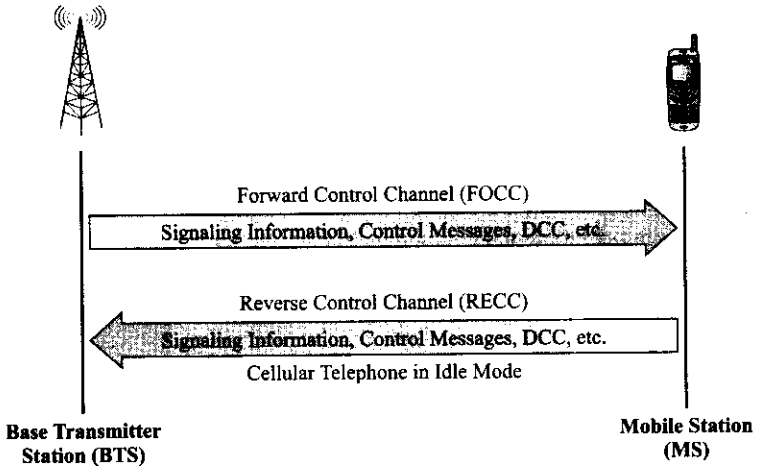

Figure 2–3 The transfer of control information over the AMPS forward and reverse control channels.

sends data over the forward control channel while the mobile station sends data over the reverse control channel as shown in Figure 2–3.

The FOCC transmits three data streams in a time division multiplexed (TDM) format as depicted by Figure 2–4. These three data streams are known as Stream A, Stream B, and the busy-idle stream. Messages to mobile phones with the least significant bit of the mobile's identification number (MIN) equal to "0" are sent on Stream A and messages to mobile phones that have a MIN with the least significant bit equal to "1" are sent on Stream B. The use of Streams A and B doubles the capacity of the control channel. The busy-idle stream indicates the current status of the reverse control channel. The reason for the busy-idle stream is to counteract the fact that the RECC can be shared by many mobile phones in a particular cell thus creating contention for its use. With the status of the RECC indicated by the busy-idle stream, message collisions can be minimized to some degree by software algorithms used by the mobiles within the cell. Both control channels operate at a 10 kbps data rate.

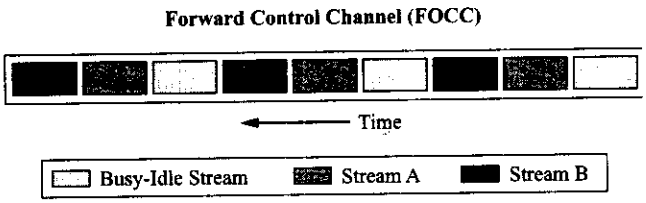

Figure 2–4 Data transfer over the AMPS forward control channel.

Each FOCC message can consist of one or more words. The types of messages to be transmitted over the FOCC are overhead messages, mobile station control messages, and control-filler messages.

Overhead message information is used to allow mobile stations to perform the Initialization Task, to update mobile stations that are monitoring a control channel by providing the latest system parameters, and to support system access by mobile stations. Two types of mobile station control messages can be sent by the base station. The base station may either page the mobile station or send it an **order message** that initiates a particular operation. The control-filler message consists of one "space filler" word that is sent whenever there is no other message to be sent on the FOCC. The control-filler message is also used to specify a control mobile attenuation code to adjust the output powers of mobile stations accessing the system on the RECC.

Typically, the base station in an AMPS system controls the mobile phone by sending order messages to the mobile. Some of these order messages are the *alert order message*—used to inform the mobile phone that there is an incoming phone call; the *audit order message*—used by the base station to determine if the mobile is still active in the system; the *change power order message*—used to alter the mobile's RF output power; the *intercept order message*—used to inform the user that a procedural error has been made in placing a call; the *maintenance order message*—used to check the operation of a mobile station; the *release order message*—used to disconnect a call; the *reorder order message*—used to indicate that all facilities are in use; the *send called-address order message*—used to inform the mobile station that it must send a message to the base station with dialed-digit information; and the *stop alert order message*—used to inform a mobile station that it must stop alerting (ringing) the user.

AMPS Security and Identification

Three identification numbers are used by the AMPS system: the mobile station's electronic serial number (ESN), the mobile service provider's system identification number (SID), and the mobile station's mobile identification number (MIN). The ESN is provided by the mobile phone's manufacturer and is not able to be easily altered. SIDs are 15-bit binary numbers that are uniquely assigned to cellular systems. These numbers are exchanged by the base and mobile station to determine the status of the mobile—at home or roaming. The MIN is a 34-bit binary number, derived from the mobile station's 10-digit telephone number—24 bits are derived from the 7-digit local number, and 10 bits are derived from the 3-digit area code. In the AMPS standards, these two groups of binary bits are referred to as MIN1 and MIN2, respectively.

Summary of Basic AMPS Operations

As one can see, the AMPS cellular telephone system uses several methods to provide control, signaling, and identification information between the base and mobile stations. Depending upon the control signals sent between the base and mobile stations, the mobile station might have to perform one or more complex sequences of steps to perform the required operation. Furthermore, this information may be sent over either control channels or traffic channels, a necessity dictated by the use of single transceivers in this first-generation cellular system. Finally, although the AMPS system uses analog FM voice transmission, it should be noted that the majority of control information is transmitted using a form of digital modulation—binary frequency shift keying (BFSK). The next several sections will provide some additional insight into common AMPS operations.

Initialization The process of AMPS mobile phone initialization is depicted in Figure 2–5. When the mobile phone is first powered up, it goes through an initialization process. This process allows the cellular phone to set itself to use either cellular provider A or B (designated as Task #1 in the figure). The second step in the process is the scanning of the twenty-one dedicated control channels of the selected service provider's system by the mobile phone (Task #2). At the completion of Task #2, the mobile station will select the strongest control channel to lock onto. This control channel will in all probability be associated with the cell in which the mobile is located. The third step in the process will be the updating of overhead information by the mobile station. The base station transmits a system parameter message that is used to update the data stored by the mobile station about the cellular system. If the mobile station cannot complete this task within three seconds, it will go to the next strongest control channel signal and attempt to complete the task within a second three-second time interval. If unable to complete this task, the mobile will now return to Task #1 and enable itself to use the other provider's system. If the mobile station can complete Tasks #1–3, it moves on to the next task. Task #4 requires the mobile station to scan the **paging channels** (a control channel) of the system and then lock onto the strongest paging channel. Within three seconds the mobile must receive an overhead message and verify certain overhead information. If this portion of the task cannot be completed, the mobile will go to the next strongest paging channel and attempt to complete the task within a second three-second time interval. During this task the mobile will compare its home system ID (SID) to

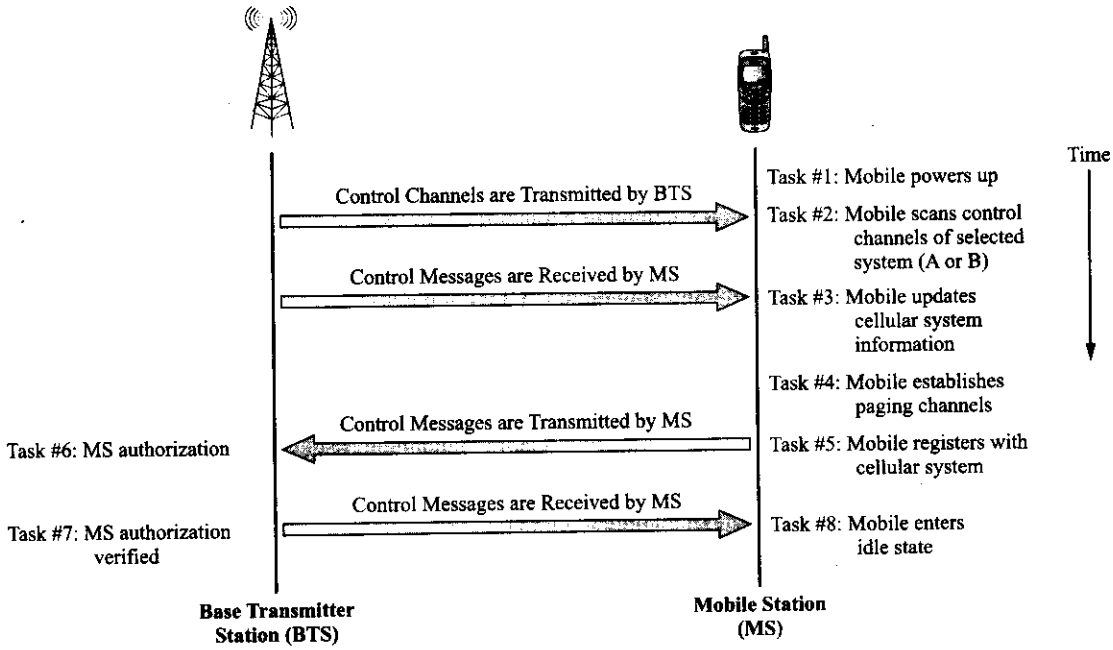

Figure 2–5 AMPS mobile phone initialization.

that of the system ID delivered to it in the overhead message. If the two system IDs are not the same, the mobile station knows that it is in a roaming status and sets parameters to allow roaming operations to take place between itself and the system that it is attached to. This action is necessary for the home system to be able to update the location of the mobile phone. If Task #4 cannot be completed successfully, the mobile returns to Task #1 and starts over. If Tasks #1–4 are complete, the mobile will identify or register itself with the network by sending its ESN, MIN, and SID numbers over the RECC (Task #5). These ID numbers will be compared against a database at the MSC to validate the mobile station's ability to have roaming status (Task #6). Finally, the base station sends a control message to the mobile to verify that the initialization process has been completed (Task #7). After Tasks #1–7 have been successfully executed the mobile goes into an idle mode (Task #8) during which it continually performs four ongoing tasks.

AMPS Ongoing Idle Mode Tasks While in the idle mode, the AMPS mobile phone will respond to continuous control messages from the base station. The mobile phone must execute each of the following four tasks every 46.3 milliseconds:

Idle Mode Task #1—Respond to overhead information. The mobile must continue to receive overhead messages and compare the received SID with the last received SID value. If the most recently received SID is different, the mobile station enters the initialization procedure again. If the SID value is the same, the mobile phone updates the received overhead information. Once the last task has been performed, the mobile responds to any messages received, if any, in the overhead message.

Idle Mode Task #2—Page match. The mobile station must monitor mobile station control messages for page messages. If paged, the mobile will enter the System Access Task with a page response.

Idle Mode Task #3—Order. The mobile station must monitor mobile station control messages for orders. If an order is received, the mobile must respond to it.

Idle Mode Task #4—Call initialization. When the mobile subscriber desires to initiate a call, the System Access Task must be entered with an origination indication.

This next section will provide several more examples of AMPS operations.

Mobile-to-Land Calls If the mobile subscriber wants to make a call, several handshaking messages must be exchanged between the mobile phone and the base station over the various control channels. Figure 2–6 shows the steps needed to complete this task. First, the mobile station enters the System Access Task mode and then attempts to seize the RECC once it becomes idle (Step #1). Once the mobile has seized the RECC, it starts to transmit a service request message to the base station over the RECC (Step #2). This message will include the mobile station's MIN, ESN, and the phone number of the dialed party. After transmitting a service request message to the base station the mobile station goes into an Await Message mode. If the base station grants the service request it will send an initial voice channel designation message (Step #3). The base station has also passed this info on to the network side (i.e., the MSC) usually through some proprietary vendor-specific messaging system (also, Step #2). Today, messaging between MSC and base stations has been standardized as TIA/EIA-634-B. The mobile will switch to the initial voice channel number provided by the base station. Other information is also included in the base station message—the power level for the mobile and an SCC that will designate what SAT tone to use on the traffic channel. At this point, both the base and mobile stations have switched their communications to the voice channels (Step #4). In the next step of the process, the base station sends a mobile control message over the FVC with the SAT

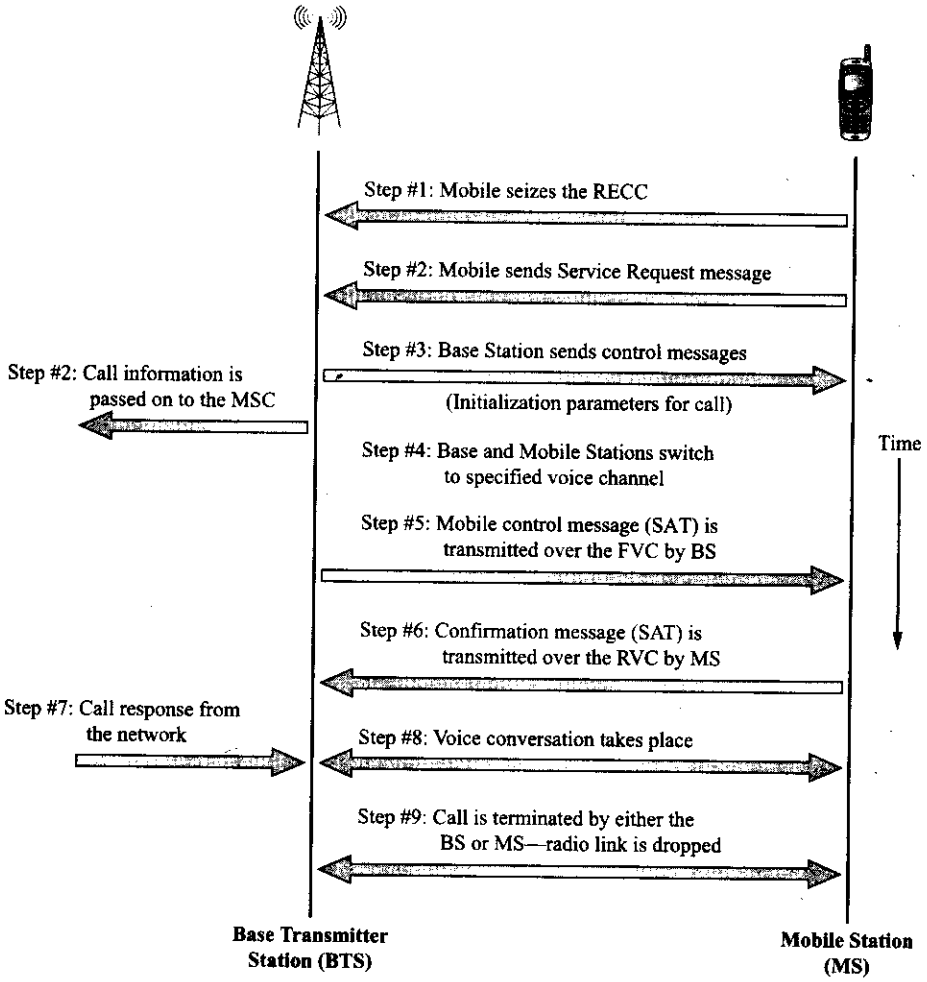

Figure 2–6 AMPS mobile-originated call.

signal (Step #5). As explained before, the mobile station responds to this message over the RVC with the SAT signal, which confirms the radio link (Step #6). The mobile station now awaits completion of the call with the resultant signal coming from the network (MSC) (Step #7). Finally, the conversation takes place (Step #8). To disconnect or complete the call, either the base station sends a release order message or the mobile sends a signaling tone (ST) for 1.8 seconds, at which point the base and mobile station drop the voice channel radio link (Step #9).

Land-to-Mobile and Mobile-to-Mobile Calls The mobile station can receive a call from another mobile station or from a telephone connected to the PSTN (a landline). For both cases, the needed handshaking steps are the same. As shown in Figure 2–7, the network (MSC) sends the ID of the mobile station to the base station (Step #1). The base station constructs a page control message. The ID information (ESN, MIN, and SID) is added to the message as is the initial voice channel information (Step #2). The mobile station responds to the page by returning identification information over the RECC in a page response message (Step #3). Another control message is sent over the FOCC by the base station that contains an SCC value to inform the mobile as to the correct SAT to be used on the voice channel (Step #4). The base and mobile station both switch to the voice channels (Step #5) and alternately use SAT tones to verify the radio link (Step #6 and #7). After this last handshake occurs, the traffic channel is then opened to conversation (Step #8).

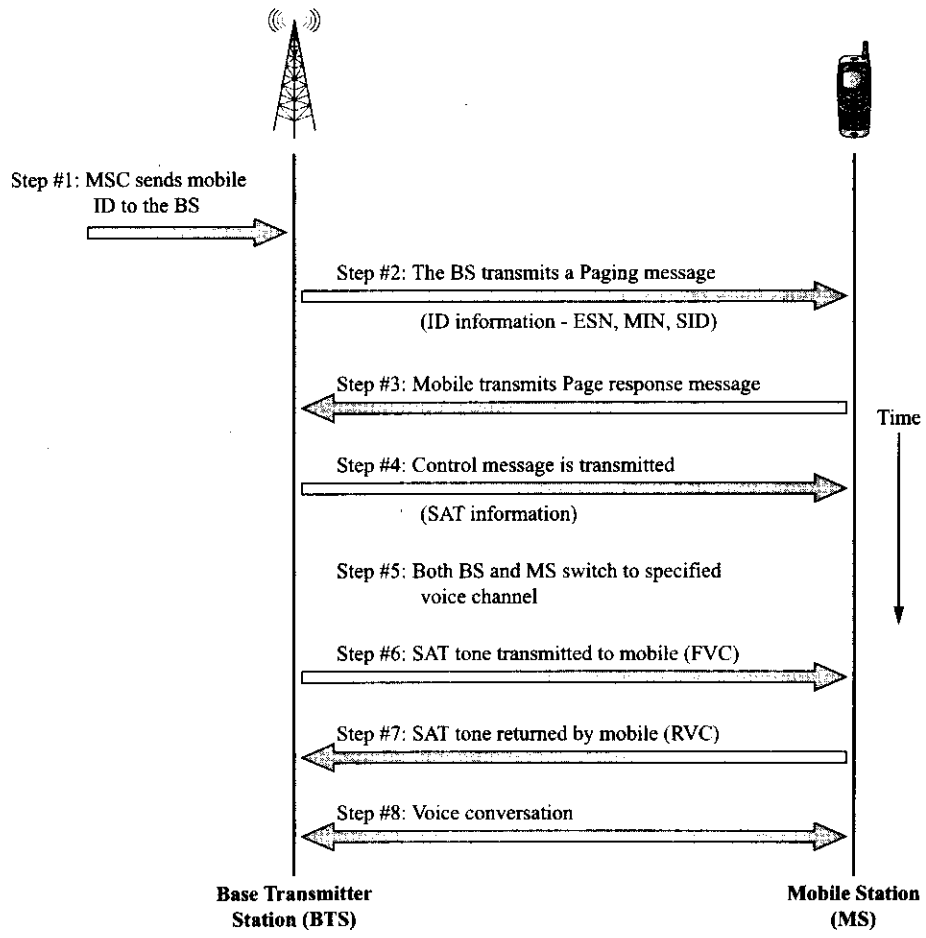

Figure 2–7 AMPS mobile-terminated call.

AMPs Network Operations At this time, it will be instructive to look at what is happening on the network side of the cellular system (base station to MSC and MSC to PSTN operations). Figure 2–8 shows some of the details of these operations. Consider a mobile-originated call like that shown before in Figure 2–6. On the network side of the cellular system, there are messages exchanged between the base station and the MSC and between the MSC and the PSTN. These messages are a combination of IS-41 and SS7 messages. Today, TIA/EIA-41-D is the intersystem standard and TIA/EIA-634-B is used between the mobile switching center and the base station. Notice that after the handshaking between the mobile station, base station, and MSC the PSTN is contacted. After the radio link between the mobile station and the base station is confirmed, the telephone call is put through to the called party over the PSTN. Several more operations are performed as handshaking between the called party and the mobile station. If the called party answers, the alert ring-back signal is removed and a conversation ensues on the forward and reverse voice channels. Either the called party or the mobile station may terminate the call.

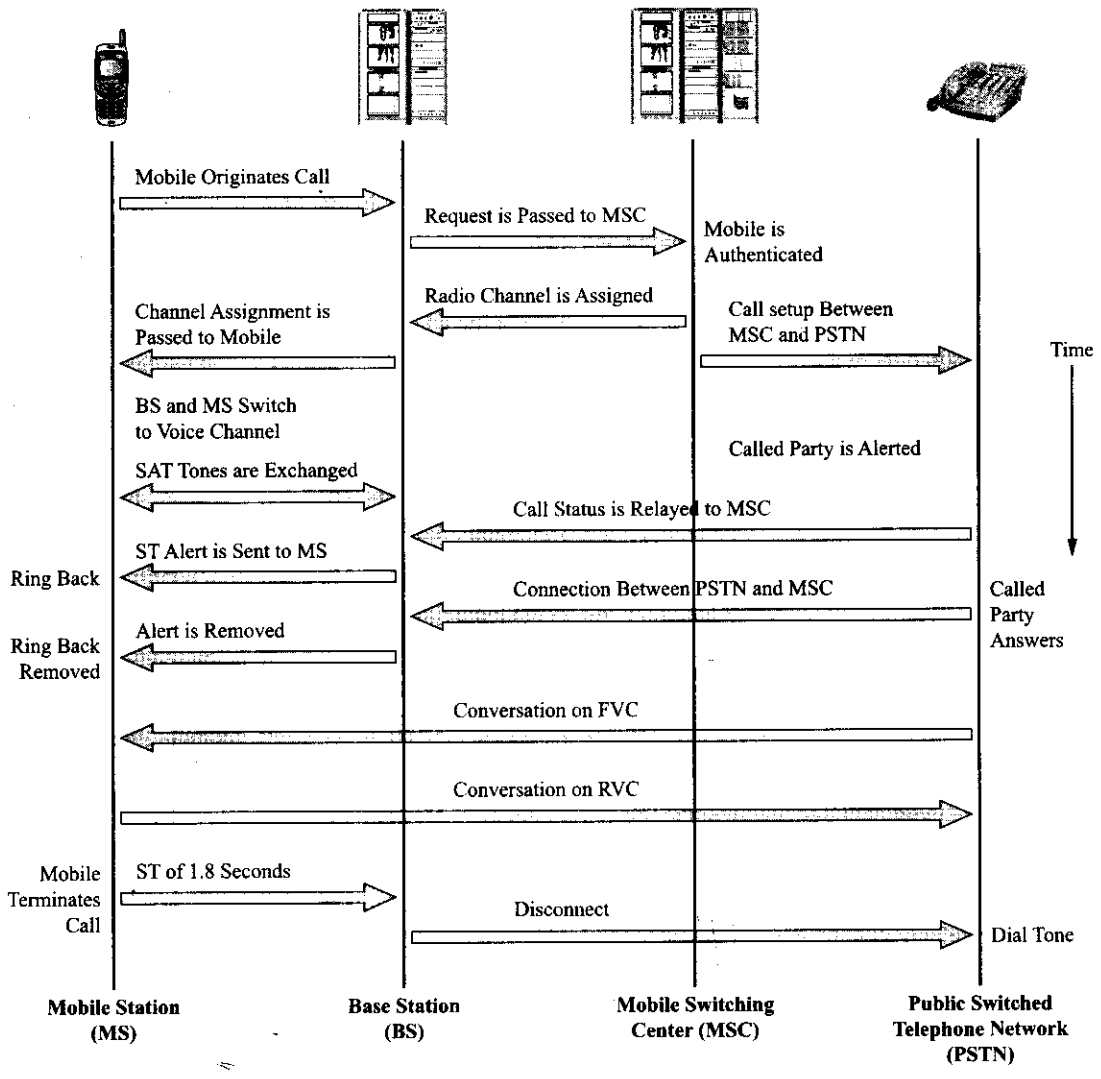

Figure 2–8 AMPS network operations for a mobile-originated call.

Handoff Operations A **handoff** operation occurs in a cellular system when a mobile station moves to another cell. Figure 2–9 details the handshaking operations that take place for handoff to occur. In this case, the figure depicts a mobile switching center connected to two or more base stations within some geographic area. Consider that Base Station A is handling an active call from a mobile station within its area of coverage.

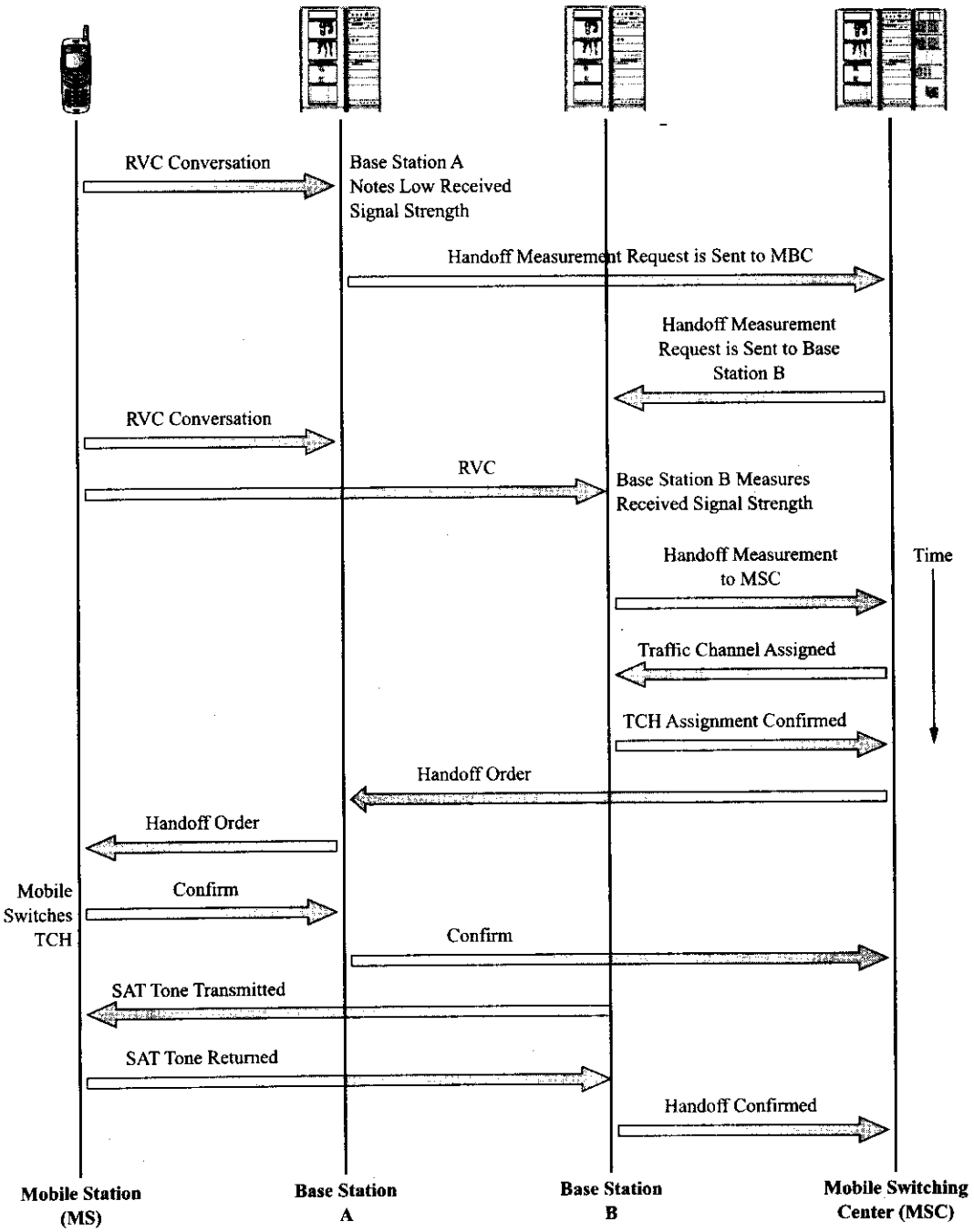

Figure 2–9 AMPS handoff operation.

However, the mobile station is in transit and is moving away from Base Station A and toward Base Station B's coverage area. Base Station A constantly monitors the received signal power from the mobile station. When the signal from the mobile station goes below a predetermined threshold level, Base Station A sends a handoff measurement request to the MSC. The MSC requests that all base stations that are able to receive the transmissions from the specified mobile station monitor its power level. It is determined that Base Station B is receiving the strongest signal from the mobile. The MSC assigns a traffic channel (TCH) to Base Station B. Base Station B responds and the handover order is sent from the MSC to Base Station A. Base Station A sends a handoff control signal to the mobile station with the necessary new channel information and then the mobile switches to the new voice channel with its newly prescribed output power and new SCC code. As before, the mobile receives Base Station B's SAT and returns it. If everything goes well, the handoff is successful.

These examples of AMPS operations should give the reader a feel for the handshaking that is necessary to perform the many operations needed to create a working functional cellular mobile system.

Other AMPS Details

One other type of information transmitted to the base station from the mobile station in the AMPS system is the **station class mark** (SCM). The SCM contains information about the mobile station's maximum output power (its class) and some additional details about the mobile station's ability to support various operations concerning output power changes.

Other 1G Systems

As mentioned before, numerous other analog first-generation cellular systems began to be deployed around the world starting in the early 1980s. Considering the recent rapid deployment of advanced digital cellular systems, the importance of these systems is limited at this time. However, it is possible that these first-generation systems will continue to be supported in some of the less developed countries in the world for a long time. Taking a brief look at some of these other systems now will help the reader develop a better appreciation and understanding of how the cellular industry has arrived at its current position. A listing of deployed cellular systems by country and type is available at www.cwt.vt.edu in the Wireless FAQ (frequently asked questions) section.

TACS Cellular

The TACS (Total Access Communications System) cellular system developed by Motorola began operation in the United Kingdom (UK) in 1985 and spread to other countries of the European Community (now the European Union) shortly thereafter. This system was a variation of the AMPS system and operated in the 800-MHz and 900-MHz bands (refer to Table 2–2). The system employed a reduced channel spacing of 25 kHz thus yielding a total of 1000 channels in the allotted spectrum. Two UK service provider networks evolved—Cellnet and Vodaphone. Due to a need for additional capacity, the networks persuaded the government to release additional frequency spectrum for TACS. TACS was upgraded shortly thereafter to Extended-TACS or E-TACS in the UK. TACS cellular systems or some variation of TACS systems are presently still employed in approximately twenty-five countries worldwide.

NMT Cellular

The NMT 450 cellular system was another variation of AMPS that was first deployed during 1981 in the Nordic countries of Denmark, Finland, Norway, and Sweden. The first NMT systems operated in the 450-MHz band with channel spacing of 25 kHz. An upbanded NMT cellular system operating in the 900-MHz band came online about five years later in 1986 with a narrower channel spacing of 12.5 kHz. NMT cellular systems have since been deployed in approximately fifty countries worldwide.

NTT Cellular

The NTT (Nippon Telegraph and Telephone) cellular system went into operation in Japan in December of 1979. A proprietary system, it used frequencies in both the 400-MHz and the 800-MHz band with channel spacing of 25 kHz. The system was not well received due to its high cost of use. Later on, during the late 1980s and early 1990s, and only after the Japanese government's Ministry of Posts and Telecommunications allowed competition in the mobile telephone market, several new first-generation systems were deployed. The JTACS/NTACS (Japanese TACS/Narrowband TACS) cellular systems operated in the 800-MHz and 900-MHz bands with 25-kHz and 12.5-kHz channel spacing, respectively. These systems, developed by Motorola, were derived from the original TACS system.

Other Analog Cellular Systems

Several other first-generation analog cellular systems (again, refer to Table 2–2) were placed in operation in different countries during the early days of cellular. NAMPS, a narrowband version of the AMPS cellular system that uses 10-kHz channel spacing (hence, triple the capacity) has been developed and introduced worldwide. One might note that Europe, with the introduction of the West German C-Netz, French Radiocom 2000, Swedish Comvik, and Italian RTMS systems, had many incompatible systems, which led the European community to become early adopters of the next generation of digital cellular technology in an effort to create a pan-European cellular system.

Table 2–2 Worldwide 1G analog cellular systems.

Cellular Standard	*Downlink Frequency Band*	*Uplink Frequency Band*	*Channel Spacing*	*Region*
AMPS	824–849 MHz	869–894 MHz	30 kHz	United States
TACS	890–915 MHz	935–960 MHz	25 kHz	European Union
E-TACS	872–905 MHz	917–950 MHz	25 kHz	United Kingdom
NMT 450	453–457.5 MHz	463–467.5 MHz	25 kHz	European Union
NMT 900	890–915 MHz	935–960 MHz	12.5 kHz	European Union
C-450	450–455.74 MHz	460–465.74 MHz	10 kHz	Germany & Portugal
RMTS	450–455 MHz	460–465 MHz	25 kHz	Italy
Radiocom 2000	165.2–168.4 MHz 192.5–199.5 MHz 215.5–233.5 MHz 414.8–418 MHz	169.8–173 MHz 200.5–207.5 MHz 207.5–215.5 MHz 424.8–428 MHz	12.5 kHz	France
NTT	915–918.5 MHz 922–925 MHz 925–940 MHz	860–863.5 MHz 867–870 MHz 870–885 MHz	6.25 kHz 6.25 kHz 6.25/25 kHz	Japan
JTACS/NTACS	898–901 MHz 915–925 MHz 918.5–922 MHz	843–846 MHz 860–870 MHz 863.5–867 MHz	12.5/25 kHz 12.5/25 kHz 12.5 kHz	Japan

Digital AMPS

Digital AMPS (D-AMPS) technology was introduced in North America during the early 1990s in an attempt to increase the capacity of the original AMPS cellular system. The AMPS system, as technologically advanced as it was at the time, had limited capacity and had become very bandwidth inefficient considering the rapid evolution of new highly efficient digital modulation techniques. The use of D-AMPS technology provided a desirable migration path that the cellular service providers could use to increase system capacity without having to totally change over their systems. As desirable as new technology is, economics must be considered when one considers changing over to a new system. With a large installed base of equipment and many mobile subscribers using AMPS mobile phones, the service providers welcomed the use of a hybrid system that used second-generation technology but was backward compatible with the installed base of first-generation equipment.

The D-AMPS system allows for the continued use of the AMPS bandwidth (channel spacing) and many of the AMPS procedures. The novel aspect of D-AMPS cellular systems is that this second-generation system using time division multiple access (TDMA) technology is able to use the same traffic channels as the first-generation AMPS system. This allows a D-AMPS system to be overlaid onto an existing AMPS system. In many cases, to upgrade to D-AMPS, the service provider could colocate D-AMPS equipment at the cell site in the same base station cabinet as the AMPS equipment. The use of D-AMPS therefore gave the service provider an evolutionary path to provide the capability of digital services to subscribers while maintaining traditional service.

Typically, in a D-AMPS/AMPS environment, a certain percentage of channels would be reserved for analog traffic and the rest allocated to TDMA traffic. As subscribers migrated to D-AMPS service the allocations could be adjusted as needed. Although the number of analog channels would be reduced, the total system capacity would increase because the D-AMPS system could support three users simultaneously in a single analog channel that formerly was capable of only supporting a single user.

The original specifications for D-AMPS were published as Interim Standard 54-B or simply IS-54-B. IS-54-B defined dual-mode operation within the same 800-MHz cellular network. All the frequency specifications remained identical to the AMPS specification in IS-54-B, as did the specifications for the analog control channels. However, both analog and digital traffic channels were defined by IS-54-B. IS-54-B was rescinded in September of 1996 and replaced by TIA/EIA-627. With the introduction of a true second-generation TDMA system developed for North America and published as IS-136, D-AMPS technology has been effectively superceded.

2.3 2G CELLULAR SYSTEMS

The first-generation cellular systems used the technology available at the time of their design. If one looks back at the technology of the early 1980s when these systems were first introduced, one realizes that this was the same era as the introduction of the IBM personal computer or PC. Those of us old enough to have lived through that time period and be involved with technology recall that the Intel microprocessors used in these devices could only process 16 bits at a time and access an extremely limited amount of memory. Most early PCs came without a hard drive! Mass storage was provided by floppy disks! Software applications had to be programmed using specialized programming languages and so on. Compare those early PCs with today's PCs and it is difficult to imagine that we were once "counting our blessings" just to be able to use a PC. If you are not old enough to relate to what I have just said, ask someone who is to tell you about these early days, or just imagine all the digital technology-based consumer electronics products removed from the shelf of your favorite technology store and the world without an Internet!

One more fact to consider is the following: over the last twenty years the number of transistors that may be put on an IC chip has increased by well over a factor of 1000, the cost of the same IC has gone down by a substantial factor, and the functionality of the IC has increased significantly. Today's mobile phones have the processing power of yesterday's supercomputers!

This continual and unrelenting onrush of technology has quickly brought us from the first generation of cellular telephone systems through the second generation to the half-generational step of 2.5G and beyond (2.5G+), with the promise and implementation of the third generation of wireless cellular service at our doorstep.

Introduction

There are several defining differences between first- and second-generation systems that will be outlined here. The most basic difference is that first-generation systems used analog modulation techniques for the transmission of the subscriber's voice over the traffic channel. All subsequent generations of cellular systems convert a user's voice from an analog signal to digital form and then use some form of digital modulation to transmit the digital encoding of the voice message. This conversion to a digital format usually results in the ability of a communications link (in this case, a traffic channel) to accommodate more than one user at a time. This attribute is usually referred to as multiplexing. The two most popular forms of multiplexing used by second-generation cellular systems are time division multiple access (TDMA) and code division multiple access (CDMA).

The control signals for first-generation systems used digital modulation to send digital control messages over the dedicated control channels and over the forward and reverse voice channels when the mobile station was in the conversation mode and thus using a traffic channel. First-generation systems also relied on supervisory audio tones and signaling tones to facilitate system operations. Second-generation systems also use digital modulation techniques to send digital control messages but have no need for analog supervisory or signaling tones.

As a further consequence of using digital encoding for the user traffic, digital encryption may be employed that provides both security and privacy for the mobile network subscriber. This was not possible in first-generation cellular systems and it led to the use of scanners that could be used to listen to private conversations as well as numerous cases of the fraudulent use of a subscriber's intercepted identification numbers (ESN, MIN, and SID). Furthermore, the use of digital encoding and modulation allows for the use of error detection and correction codes, the use of which, to some extent, combats the type of fading and noise effects peculiar to the radio channel (more about this topic in Chapter 8).

The AMPS system worked remarkably well when it was first deployed in the United States. Subscribers could move country-wide between different service provider systems and as long as they were in a coverage area they could receive service. Roaming was not a problem within the United States since all systems had to be compatible. This was not so in other areas of the world. As just outlined, many different systems were deployed in different regions of the world. This situation was nowhere more troublesome than in the European countries.

Therefore, in the early 1980s, the European countries began working together to develop a pan-European cellular system. This process was set in motion when, in 1982, the Conference of European Posts and Telegraphs (CEPT) formed a Groupe Spéciale Mobile study group to research and then develop this new system. The study group proposed that the new system meet certain operational criteria and in 1987 the Global System for Mobile Communications was formally initiated by the European Commission in the form of a directive. In 1989, responsibility for the continuing development of the new system was transferred to the European Telecommunication Standards Institute (ETSI). In 1990, the first phase of the GSM standards were published and commercial operation commenced soon afterwards in late 1992. The system chosen used digital technology and became known as the GSM cellular system.

General Characteristics of 2G Systems

Only a brief overview of the general characteristics of second-generation systems will be provided here since these systems and their succeeding implementations will be covered in much greater detail in subsequent chapters.

The ability of these cellular systems to support more than one user per radio channel is through the use of advanced digital multiplexing techniques. TDMA systems (GSM, North American TDMA, and PDC) all use **timeslots** to allocate a fixed periodic time when a subscriber has exclusive use of a particular channel (frequency). The GSM system uses a transmission format with eight timeslots and therefore the system can support eight users per radio channel simultaneously. CDMA cellular systems use a digital modulation technique known as **spread spectrum.** In this system, at the transmitter, each user's digitally encoded signal is further encoded by a special code that converts each bit of the original digital message into many bits. At the receiver, the same special code is used to decode or recover the original bit stream. The special codes used to perform this encoding/decoding function have the unique property that each received signal looks like noise to a receiver that does not share the same code as the transmitter of the signal. Therefore, in a CDMA system many radio signals may be simultaneously transmitted on the same radio channel without interfering with each other. The only detracting aspect of this technology is that CDMA signals have a very broadband spectrum compared to other digital modulation schemes.

For either TDMA or CDMA cellular systems, both control information and traffic share the same radio channel. For TDMA systems, since both forms of information are in a digital format they can be intermingled within a data stream and transmitted by a single transmitter over a radio link. For CDMA systems, control information is carried by dedicated channel elements and traffic is placed on any available traffic channel element. **Channel elements** (CEs) are individual transmitters that are all transmitting on the same frequency simultaneously.

Finally, although mobile data services were available over first-generation cellular systems, these early proprietary systems were generally not used by the general public. In 1993, cellular digital packet data (CDPD) service was introduced in the United States. The ability to connect to the public data network by any cellular subscriber is a distinguishing feature of all second- and succeeding generation cellular systems.

Note: First-generation cellular systems were capable of transmitting data over the PSTN (circuit switched) the old-fashioned way by using a modem.

GSM

The first GSM systems, originally scheduled to be deployed in 1991, began operation in late 1992 when GSM handsets first became available. Before the end of 1993 over one million customers had signed up for service. GSM technology has become the most popular cellular telephone technology with approximately 72% of the world's cellular customers subscribing to the service. At this point, there are over 500 GSM networks in operation in 174 countries worldwide with an estimated one billion users as of early 2004. GSM technology uses TDMA to allow up to eight users per channel. Channels are spaced 200 kHz apart. The basic system uses frequencies in the 900-MHz band (GSM 900), but later an upbanded version was added at 1800 MHz (GSM 1800) and the 1900-MHz band was added in the United States for PCS service (GSM 1900). There are current plans to expand into the 850-MHz and 450-MHz bands (GSM 850 and GSM 450). GSM service when first introduced supported circuit-switched data rates of up to 9.6 kbps.

CDMA

In the early 1990s, in response to the Cellular Telecommunications Industry Association's (CTIA) user performance requirements for the next generation of wireless service, a totally new digital technology known as Code Division Multiple Access or CDMA was developed by Qualcomm Corporation. In 1993, the CDMA air interface standard, IS-95, was adopted and the first CDMA commercial network began operation in Hong Kong in 1995. Since that time, CDMA systems have been used in both the cellular and PCS bands extensively in the United States and throughout the rest of the world. CDMA has experienced very rapid growth and presently 13% of the world's cellular telephones use this technology.

TDMA

In the United States a true second-generation TDMA system was developed for use at the 800-MHz and then the 1900-MHz PCS bands. This TDMA system is published as IS-136 and it has many similarities to GSM. Today it is known as North American TDMA (NA-TDMA). Currently, only 10% of the world's cellular subscribers use this technology.

PDC

In 1989, the Japanese Ministry of Post and Telegraph began a development study with the ultimate goal of creating a digital cellular system with a common air interface. From this study came the Japanese Personal Digital Communications (PDC) system in 1991. Using TDMA technology similar to IS-54 in both the 800-MHz and 1500-MHz bands, PDC systems supplied by Motorola were deployed starting in 1993. Currently, only 5% of the world's cellular subscribers use PDC technology.

PCS Systems

During the mid-1990s, in response to the Omnibus Budget Act of 1993, the FCC auctioned off portions of the electromagnetic spectrum in the United States for use by commercial mobile radio service providers. The FCC had allocated 153 MHz of spectrum for Personal Communication Services (**PCS**) and took the stance that the marketplace should dictate the use of this spectrum. Many cellular service providers bid on the two frequency blocks available in the fifty-one **major trading areas** (MTAs) and the 453 frequency blocks available for **basic trading areas** (BTAs). A large number of these licenses have been used to extend cellular coverage by the cellular service providers. In only a limited number of cases, service providers have deployed pure PCS networks (e.g., Sprint PCS and T-Mobile). Typically, CDMA, GSM 1900, and NA-TDMA technology have been used to provide service in these PCS bands.

2.4 2.5G CELLULAR SYSTEMS

After second-generation cellular systems began operation there was an increasing desire for mobile data delivery. During the 1990s, the PC had been in existence for over a decade and the Internet was starting its explosive growth. Worldwide, more and more telecommunications was becoming computer-to-computer oriented and society had become extremely mobile through the growth and efficiency of modern transportation systems. Several proprietary systems had been developed early in the life cycle of 1G systems, but in 1993, IBM and several mobile carriers published a specification for a system called cellular digital packet data (CDPD) that could be overlaid on the AMPS system. Although an improvement that allowed users wireless e-mail access, file transfer capabilities, and the like, CDPD service could only deliver data at very limited transfer rates (typically, 9.6 kbps).

Evolution of Mobile Data Services

With the advent of all digital second-generation cellular networks came the very real likelihood of increased data transfer rates over cellular systems. It was not long before the service providers and the cellular standards organizations set their sights on third-generation cellular systems that would offer high-speed data rates and many more features tied to the access of the PDN by their subscribers. However, before the appropriate technology and sufficient frequency spectrum exists to build these systems, an evolutionary approach to upgrading the existing cellular systems was outlined by the interested parties. A broad framework of 3G specifications has already been laid out, but for most systems in operation, however, we must pass through 2.5G and 2.5+G first!

Today, the most important cellular systems are GSM, CDMA, and NA-TDMA. Together these systems represent approximately 95% of the world's cellular subscribers. The next few sections will give a brief description of the technologies used to provide access to the PDN over these systems. A more detailed discussion of these technologies will be given in Chapter 7.

CDPD

CDPD was originally designed to provide mobile packet data services as an overlay system for the now legacy AMPS cellular system. It can be extended to CDMA service but CDMA is following a different path. CDPD service may continue as a viable alternative for the delivery of low-speed bursty packet data in the near term but will most likely fade away as time goes on.

HSCSD

Although **HSCSD** (High-Speed Circuit-Switched Data) is not a packet-switched data service, it should be included here because it was the first planned enhancement for increasing circuit-switched data rates on GSM networks. This enhancement takes place in two steps. Phase one, deployed in 2001, yields data transfer rates up to 43.2 kbps, and then a follow-up enhancement, phase two, will allow transfers to 64 kbps. This technology works by giving a mobile subscriber multiple timeslots out of the standard GSM TDMA frame with its eight timeslots. Since this technology deals with circuit-switched data, its importance to enhanced data services is not as great now as when it was first proposed. Effectively, HSCSD service has been superceded by GPRS technology.

GPRS

General Packet Radio Service (GPRS) was defined by the European Telecommunication Standards Institute as a means of providing packet-switched data service that allows full mobility and wide area coverage on GSM networks. The standards were published in the late 1990s and the service was introduced at the beginning of the new millennium. GSM GPRS service is designed to ultimately provide data transfer rates up to 160 kbps. This technology is also being deployed by NA-TDMA systems with data rates up to 45 kbps. Interestingly, it is felt that the use of GPRS technology for packet-switched data services for both GSM and NA-TDMA will eventually drive these two similar technologies toward a converged system as 3G is approached.

Packet Data over CDMA

The CDMA system used an **InterWorking Function** (IWF) component that is necessary for both circuit and packet data (see Figure 2–10). For circuit-switched data, the IWF supplies a modem connection to the PSTN and the modem function is built into the mobile subscriber's CDMA telephone. For first-generation CDMA systems (IS-95A), the maximum possible data rate for circuit-switched data is 14.4 kbps. For packet data, the IWF provides the interface between the wireless system and the external packet network with a maximum data rate of 14.4 kbps also.

For 2.5G CDMA systems (IS-95B revision) higher data rates of 115.2 kbps are possible. However, the real data throughput of the system is more in the range of 60 to 80 kbps. Note that both IS-95A and IS-95B systems are now referred to as **cdmaOne** cellular systems.

2.5 3G CELLULAR SYSTEMS

The term "third-generation mobile systems" or 3G is used to represent a number of cellular systems and their associated standards that have the ability to support high data rate services, advanced multimedia

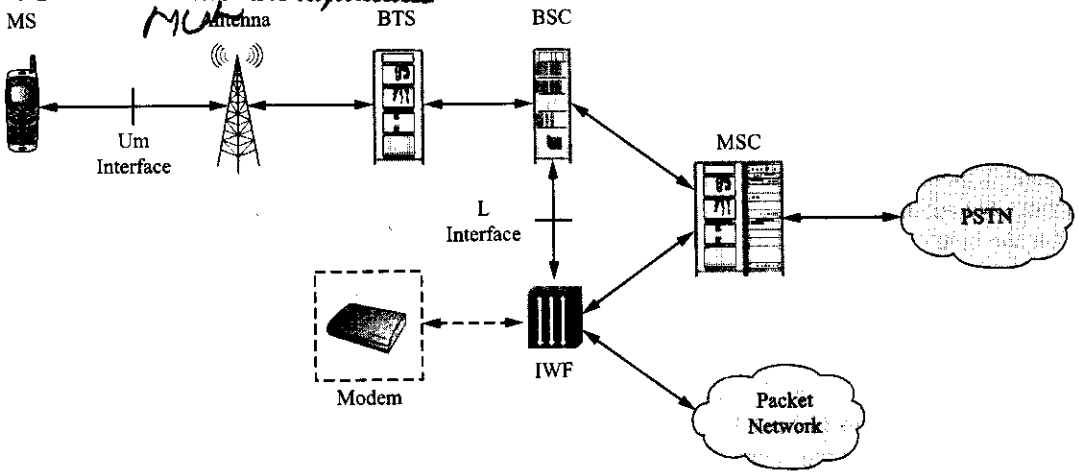

Figure 2–10 The CDMA interworking function node.

services (e.g., voice, data, and video), and global roaming. These standards are being facilitated by the International Telecommunications Union (ITU) and other regional bodies around the world (see Figure 2–11). In the late 1990s, the ITU formed the International Mobile Telecommunication- 2000 (IMT-2000) forum to address the mobile telecommunications needs of the twenty-first century (see www.ITU.org). Worldwide deployment of new 3G cellular systems has started and will be ongoing as new evolutionary phases of the 3G standard are used to build out the systems. Presently, the 3G Partnership Project (3GPP) group and the 3GPP2 group are overseeing these efforts on behalf of the GSM and CDMA mobile systems stakeholders, respectively.

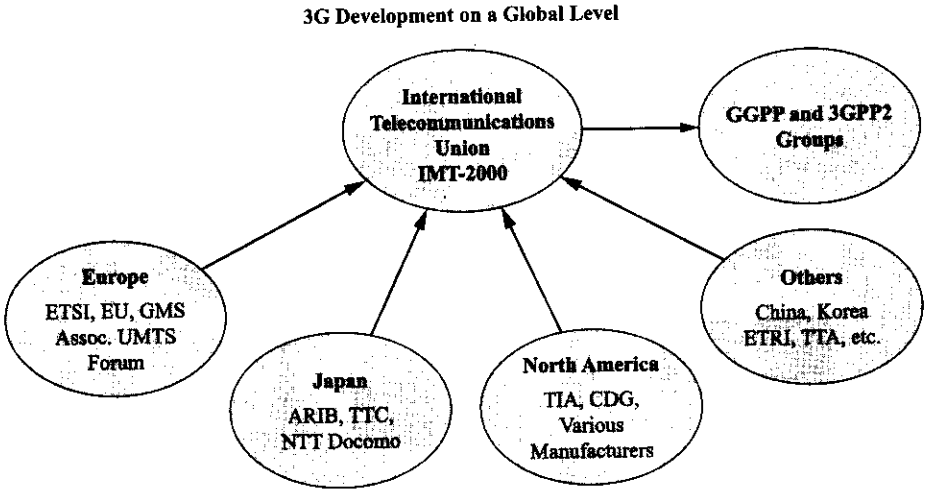

Figure 2–11 Organizations involved with the development of the 3G cellular standards.

Introduction

The deployment of second-generation cellular systems only served to fuel the fire of an increasing demand for more system capacity and new services. 2G systems primarily provided voice service even though they could also support low data rate services. The arrival of the Internet revolutionized the data market. The

demand for data over the PSTN increased drastically and spawned the development of new wireline technologies like broadband cable modems and digital subscriber line (DSL) for high-speed Internet access. At the same time, the wireless subscriber's demand for Internet access and data services grew but 2G could not satisfy the demand.

2G cellular systems are limited by bandwidth and roaming capability. Since the first several generations of cellular systems were designed primarily for voice service and not data, they do not have sufficient bandwidth for the high data transfer rates desired today. Also, since multiple air interface standards are used for the many different cellular systems already deployed around the world, the systems are not compatible with each other and therefore prevent global roaming. Additionally, 2G systems have data services limitations, lack of support for packet data networks, and lack of support for multimedia services.

3G Characteristics

3G mobile networks need to be able to provide high-speed data transfer from packet networks and to be able to permit global roaming. Furthermore, they need to support advanced digital services (i.e., multimedia) and to be able to work in various different operating environments (low through high mobility, urban to suburban to global locations, etc.). In other words, as shown in Figure 2–12, anywhere a mobile subscriber might be located (except for the most severe radio environments) should be supported by 3G networks. The IMT-2000 has defined these various hierarchical cell structures, their corresponding size, allowed subscriber mobility rate, and minimum supported data rate as shown in Table 2–3.

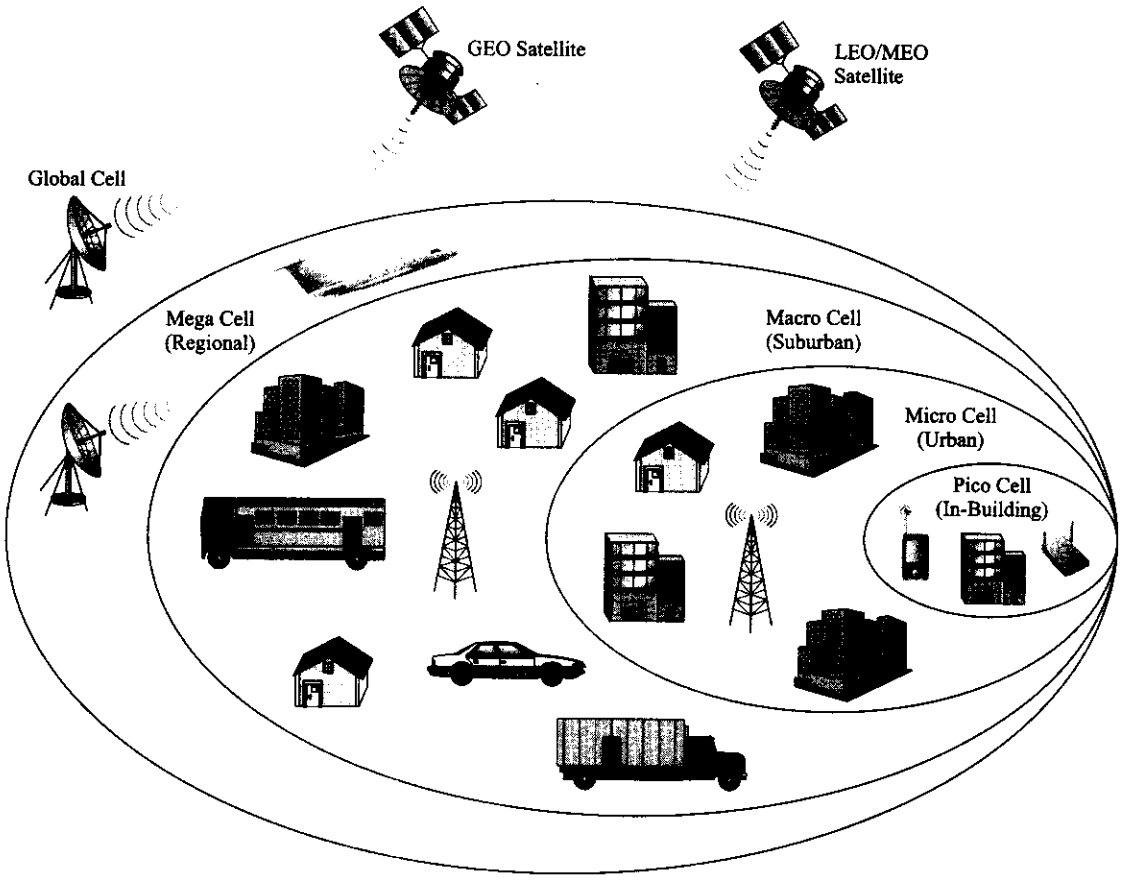

Figure 2–12 3G operating environments.

Table 2–3 3G characteristics by cell size and mobile speed.

Cell Type	*Global Cell*	*Mega Cell*	*Macro Cell*	*Micro Cell*	*Pico Cell*
Maximum Cell Radius	1000's of km	100–500 km	35 km	1 km	50 m
Operating Environment	Global	Regional	Suburban (low user density)	Urban (high user density)	In-building
Installation Type	Satellite GEO, MEO, LEO	Satellites LEO	Tower or building mounted	Building facade or lamp-post	Inside of a building
Data Rate	100's of kbps to several mbps*	100's of kbps to several mbps*	144 kbps	384 kbps	2 mbps
Maximum Mobile Speed (Km/h)	N/A	N/A	500 km/h	100 km/h	10 km/h

*Not part of the 3G standard

3G systems must be able to support varying data rates by providing bandwidth on demand to the subscriber. 3G subscriber devices (SDs) or end terminals (ETs) will be required to support multiple technologies and frequency bands and have the ability to be reprogrammed by their home cellular system. Today's mobile phones have dual-band and tri-mode capabilities, can provide limited video multimedia support, and have limited reprogramming features. Advanced, reconfigurable, multimedia mobile phones or subscriber devices based on software radios are under development now. Additionally, 3G systems must be able to support multiple simultaneous connections, IP addressing, and be backward compatible with 2G networks.

3G Radio Interfaces

The IMT-2000 requirements for radio transmission technology (RTT) are driven by the basic 3G requirements. Therefore, the radio technology used to implement a 3G system must have the ability to support all of the features referred to in the previous section. As mentioned before, the vast majority of cellular subscribers use either GSM, CDMA, or NA-TDMA technology. Many different proposals were submitted to IMT-2000, but only five were accepted by the International Telecommunications Union (ITU). Presently there are only two major 3G cellular technology proposals moving forward. They are cdma2000 and UMTS Terrestrial Radio Access or **UTRA** (Universal Wireless Communication—136; UWC-136 has recently been dropped as a viable alternative). A brief overview of these technologies will be given next. More in-depth coverage of these systems is given in Chapters 5 and 6.

UMTS

Universal Mobile Telecommunications System Terrestrial Radio Access Network or **UMTS** Terrestrial Radio Access or UTRAN is the evolutionary pathway to 3G for GSM mobile systems. This system was proposed by ETSI and is supported by the UMTS Forum (see www.umts-forum.org) and several major manufacturers. This 3G system is slated to use present spectrum allocations and new frequency allocations in the 2-GHz band and to also employ combinations of wideband CDMA (**W-CDMA**) technology and either time division duplex (TDD) or frequency division duplex (FDD) CDMA technologies depending

upon spectrum availability. The use of TDD or FDD CDMA technology in conjunction with W-CDMA is to support the different UMTS service needs for symmetrical and asymmetrical services.

Another recent option to this evolution is the use of TD-SCDMA (time division – synchronous CDMA), a relatively new technology proposed by the China Wireless Telecommunications Standards (CWTS) group. The NTT DoCoMo system uses a prestandard form of W-CDMA technology for its popular FOMA (Freedom of Multimedia Access) system. Several other 3G systems are already operational in the European Union countries.

Cdma2000

This is the enhanced wideband version of CDMA. It is supported by the United States Telecommunications Industry Association (TIA) and the CDMA Development Group (CDG) (see www.cdg.org) and several major manufacturers. The major features of **cdma2000** are its backwards compatibility with CDMA IS-95B (a 2.5G technology), support for data services (data rates of up to 2 mbps), support for multimedia services (i.e., **Quality of Service** or QoS), and support for advanced radio technologies. A unique feature of cdma2000 is that it will support several different radio link bandwidths depending upon the required data rate. The first phase of the evolutionary pathway for cdma2000 technology is to implement what is known as 1xRTT technologies over a standard 1.25-MHz CDMA channel. 1xRTT can double the voice capacity of a cdmaOne network and will support packet data service at rates up to 144 kbps in a mobile environment. Cdma2000 1xEV is the next phase and it consists of two versions, cdma2000 1xEV-DO (data only) and 1xEV-DV (data and voice). 1xEV-DO can support peak data rates of 2.4 mbps on the downlink but only 153 kbps on the uplink and thus applications such as MP3 transfers and video conferencing are possible. 1xEV-DV supports integrated voice and simultaneous high-speed data packet multimedia services at speeds up to three mbps over an all-IP architecture for radio access and core network. Both systems are backward compatible with cdma2000 1xRTT and cdmaOne. The change-over from cdmaOne to 1xRTT has been ongoing in the United States.

UWC-136/EDGE

UWC-136 is the 3G proposal for the evolution of NA-TDMA cellular systems. This proposal was developed by the United Wireless Communications Consortium (UWCC) that consists of NA-TDMA manufacturers and service providers. As of this writing, the UWCC has been disbanded with its mission effectively taken over by the GSM Association. The TIA has already published TDMA 3G standards as TIA/EIA-136, Rev C. It appears at this time that most NA-TDMA operators have opted to follow the GSM/EDGE route to 3G cellular.

3G Mobile Network Evolution

For early second-generation mobile systems, services like voice and circuit-switched data were supported by the traditional cellular system components. With the advent of packet-switched data services, new functional elements had to be introduced into the network. For GSM/GPRS systems two new nodes are used to process all the data traffic. As GSM evolves to UMTS, the network will have to evolve again. Eventually, evolution toward an all-IP architecture for the core network will also occur in various phases. Cdma2000 also has an evolutionary roadmap for its all-IP network. This topic will be covered in more detail in Chapter 7.

3G Harmonization

In the hopes of achieving a quasi world standard for 3G wireless, harmonization activities between different regions and standards bodies are currently ongoing. In an effort to bring this task to fruition, two international bodies have been formed: the Third Generation Partnership Project (3GPP), to harmonize and also standardize the similar 3G proposals from ETSI and other W-CDMA proponents, and 3GPP2, for the

harmonization of cdma2000-based proposals from the Telecommunications Industries Association and others. Additionally, an Operators Harmonization Group has been formed to try to bring together the 3GPP and 3GPP2 initiatives to support a single end user terminal concept and global roaming. The last concept is referred to as G3G or Global 3G.

2.6 4G CELLULAR SYSTEMS AND BEYOND

Even before 3G cellular technologies have been fully rolled out, fourth-generation mobile communications (4G mobile) initiatives and technologies are being studied by academia and the wireless industry. 4G actually involves a mix of new concepts and technologies. Some of these new ideas are derived from 3G and are therefore evolutionary while some ideas involve new approaches and new technologies and are therefore revolutionary. The goal of 4G is the convergence of wireless mobile with wireless access communications technologies. A converged broadband wireless system appears to be the future trend in the wireless industry. This converged system will evolve in response to the issues of bandwidth efficiency, dynamic bandwidth allocation, quality of service, security, next-generation digital transceiver technologies, self-organizing networks, and future concerns that have yet to be recognized.

4G mobile networking will require an all-IP architecture and connectivity for anyone, anywhere, at anytime. Early 4G mobile network data rates are expected to reach over 20 mbps and eventually provide ATM speed wireless connectivity. Many in the wireless industry feel that eventually wireless ATM will provide the framework for the next generation of wireless communications networks.

Wireless ATM

The concept of wireless ATM was first introduced in the early 1990s as a way for a variety of mobile terminals to connect to an ATM network. In the late 1990s, a wireless ATM (WATM) working group was formed under the banner of the ATM forum (see www.atmforum.com). The group developed a vision of an end-to-end ATM network that had the ability to support a variety of wireless technologies for interconnectivity between various portions of the backbone network. However, a number of fundamental physical layer challenges to the technology derailed this effort from the fast track desired by the WATM group and put it on a slower track. The most severe problem faced by this technology was that ATM was designed for use over extremely reliable fiber-optic transmission channels. However, the wireless channel is inherently very unreliable and therefore imposes serious limitations on wideband transmissions. The slower track pursued by the WATM group involved active participation in the standards activities of wireless LAN industry groups, in particular HiperLAN/2, a European-based effort. Research on wireless ATM and wireless mobile ATM (wmATM) networks continues on a worldwide basis.

The All-IP Wireless Network

The extremely rapid acceptance of the Internet and wireless mobile technologies is paving the way toward new and innovative digital services for the mobile user over emerging high-bandwidth, high-speed mobile networks. The rapid transformation of wireless systems from voice-only networks to multimedia-capable digital networks will usher in the mobile information society much faster than anyone could have predicted only a few short years ago. With already more than a billion mobile phone users, many, including this author, predict that the use of some type of mobile appliance or end terminal will be the most common method of connecting to the Internet in the very near future. With the introduction of 3G technologies, the mobile industry has started moving toward that goal. Major efforts are underway by the service providers to supply services and applications to the mobile subscriber over a packet-switched IP (Internet Protocol) network. The ultimate goal is to eliminate circuit switching (the PSTN) and thus totally reconfigure the structure of the existing wireless cellular network.

Work has already begun on the theoretical design and initial standards work needed to implement an all-IP end-to-end system. Commonly referred to as fourth-generation (4G) wireless systems, 4G will allow the transport of high bit rate, rich multimedia content over an all-IP network. 4G networks will most likely evolve from the several different air interface standards currently being used for 3G. There are many technologic challenges that will have to be overcome to get to the point of having an all-IP network. Diverse mobile appliances or terminals connected by an assortment of access technologies will provide daunting engineering problems for the network system designer in terms of operations, management, interoperator billing, QoS issues, protocols, and so on. Not to mention the literal possibility of hundreds of billions of devices connected to such a network. Will it happen? That is not the correct question. When will 4G happen, is the question to be asked!

IEEE 802.20x

In late 2002, an IEEE 802 study group was formed at the request of the IEEE Computer Society. Under the category of Local and Metropolitan Area Networks, this new project has the title: "Local and Metropolitan Area Networks—Standard Air Interface for Mobile Broadband Wireless Access Systems Supporting Vehicular Mobility—Physical and Media Access Control Layer Specification." It is designated as IEEE 802.20.

According to the IEEE Wireless Standards Web site (see http://standards. ieee.org/wireless/) the project scope and purpose is as follows:

> **Project scope:** Specification of physical and medium access control layers of an air interface for interoperable mobile broadband wireless access systems, operating in licensed bands below 3.5 GHz, optimized for IP-data transport, with peak data rates per user in excess of 1 Mbps. It supports various vehicular mobility classes up to 250 Km/h in a MAN environment and targets spectral efficiencies, sustained user data rates and numbers of active users that are all significantly higher than achieved by existing mobile systems.
>
> **Project purpose:** The purpose of this project is to enable worldwide deployment of cost effective, spectrum efficient, ubiquitous, always-on and interoperable multi-vendor mobile broadband wireless access networks. It will provide an efficient packet based air interface optimized for IP. The standard will address end user markets that include access to Internet, intranet, and enterprise applications by mobile users as well as access to infotainment services.

This initiative, by the IEEE 802.20 study group for the standardizing of what are known as a new class of radio LANs (RLANs), is just one more piece of the puzzle as we move toward 4G mobile networks.

2.7 WIRELESS STANDARDS ORGANIZATIONS

The modern era of technology (the last three decades by this author's definition), has provided the telecommunications world with increased complexity, speed, and capacity of the available wireline, wireless, and fiber-optic facilities. This fact by itself has produced increased activity of standards bodies. Standardization is usually considered necessary for low-cost implementation and speed in bringing services to the market. Furthermore, with the present global nature of the telecommunications industry, standards are necessary to ensure interoperability of equipment from different vendors on a worldwide basis. Standards organizations usually consist of manufacturers, service providers, and users working together to promote physical characteristics for the anticipated telecommunications requirements of the future. With standards in place, users can develop applications that build upon the standards. There are several levels of standards organizations, with their sphere of influence depending upon their makeup. Standards bodies are sponsored at the implementation, national, regional, and international or global level.

Introduction

In the wireless telecommunications arena, we have seen that many regional standards for wireless mobile systems have evolved. At present, there is no one global wireless standard for cellular, or wireless LANs

for that matter. This is the nature of the beast, so to speak. In many cases, standards bodies have started to meet and plan the next generation of wireless technology before it is even technically feasible to implement it. In many cases, these efforts have been on a regional level and have continued to evolve in that fashion. Recently, in the form of IMT-2000, a global forum on the future of wireless cellular mobile systems was held. That forum mapped out a pathway for the evolution of 3G cellular that moves toward a single global standard. That process is still continuing but will most likely need many more years before it becomes a reality.

Implementation Groups

The process of standardization begins in an implementation group or a standards development organization. These groups generally consist of interested members from a particular manufacturing industry, the academic world and government entities, trade associations, industry service providers, and users. Some of the groups presently active in the wireless arena are IEEE 802, CDMA Development Group, UMTS Forum, Committee TR-45 of the TIA, GSM Association, and so on. See the IEEE Wireless Standards Web site for information about the activities of a typical implementation working group.

Regional Organizations

Regional standards organizations receive developed standards from implementation groups. The regional organizations are tasked with approving the standard. Usually, members of the pertinent subcommittee of the regional organization will vote on the standard. Some of the more well-known regional organizations are the European Telecommunications Standards Institute (ETSI), the Telecommunications Technology Committee (TTC) and the Association of Radio Industries and Businesses (ARIB) in Japan, the Telecommunications Technology Association (TTA) in Korea, the China Communications Standards Association (CCSA) in China, Committee T1 – Telecommunications (ANSI-T1) in the United States, and the EIA/TIA (Electronics Industries Alliance/Telecommunications Industry Association).

National Organizations

The most well-known national standards organization that exists in the United States is the American National Standards Institute or ANSI. The TIA and EIA develop North American wireless standards and forward them to ANSI for final approval as a national standard. Other national organizations have been already mentioned.

Global Organizations

Global standards organizations receive recommendations from regional organizations. These worldwide organizations give the final approval for an international standard. There are three global standards organizations: the International Telecommunications Union (ITU), the International Standards Organization (ISO), and the International Electrotechnical Commission (IEC).

QUESTIONS AND PROBLEMS

1. Explain the concept of time division duplex.
2. Assume that the transmitting antenna for the first mobile radio-telephone system in St. Louis, MO, was located on a tower at a height of 250 feet. Determine the range of this system assuming line of sight transmission and a receiving antenna height of 6 feet. Hint: Reference a typical communications systems text to find an equation for transmitting range that is given in terms of antenna heights.
3. Go to the FCC's Web site at www. FCC.gov and locate Technical Bulletin No. 53. Download it and bring a copy to class for discussion.

4. Search the Internet for Web sites about the early days of cellular telephone operation. Give the URLs of at least two Web sites devoted to this topic.
5. Explain how frequency division duplex operation was achieved by first-generation cellular systems.
6. Determine the downlink and uplink frequencies for AMPS channel 445 on the B-side channels. What type of channel is it?
7. Determine the downlink and uplink frequencies for AMPS channel 326 on the A-side channels. What type of channel is it?
8. What two AMPS system components provide the air interface?
9. Explain the purpose of the AMPS supervisory audio tones.
10. Describe the sequence of events that occurs when an AMPS cellular telephone is first turned on.
11. Of what use is the AMPS cellular service provider's system identification (SID) number?
12. What is the basic difference between a mobile-originated call and a mobile-terminated call?
13. What event triggers an AMPS handoff operation?
14. Why are supervisory audio tones and a signaling tone needed for the AMPS system?
15. How many D-AMPS subscribers can an AMPS channel support?
16. What is the fundamental difference between first-generation cellular systems and second-generation cellular systems?
17. List at least two advantages of the use of digital encoding for cellular telephone systems.
18. How do second-generation cellular systems support more than one user per channel?
19. What is a 2.5G cellular system?
20. What packet data transfer rate can the first implementation of CDMA cellular support?
21. What features do 3G cellular telephone systems provide?
22. Compare the UMTS 3G cellular system and the cdma2000 3G cellular system.
23. What is meant by harmonization in the context of 3G cellular telephone systems?
24. What are the basic characteristics of purposed 4G cellular telephone system?
25. What is the purpose of the IEEE 802.20 standards project?
26. What are regional standards organizations? What is their function?
27. What are national standards organizations? What is their function?
28. What is the function of the International Telecommunications Union?
29. Visit the Web site of the TIA. What is the function of the TR-34 Committee?
30. What organization puts the final stamp of approval on the IEEE 802.11 wireless LAN standards?

CHAPTER 3

Common Cellular System Components

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

- List the components of a wireless cellular network.
- Discuss the functions of the following cellular network hardware components: MS, RBS, BSC, and MSC.
- Discuss the functions of the following cellular network databases: HLR, VLR, AUC, EIR, and so forth.
- Discuss changes in the network components used to implement 3G wireless networks.
- Discuss the use of identification numbers with cellular network components.
- Explain the basics of SS7 signaling used in wireless cellular telecommunications networks.
- Explain the basic operations needed for call setup and call release.

As wireless cellular network technology has matured the system has become more sophisticated and complex in an effort to implement increased system functionality and cope efficiently with an ever increasing number of subscribers. This chapter takes a look at the various hardware network elements that are used to create a wireless cellular network. These network elements may be divided into three basic groups: the mobile or subscriber device that provides the user's link to the wireless network, the base station system that provides the wireless system's link to the subscriber over the air interface, and the wireless switching system that provides the interfaces to the PSTN and PDN and the correct information and connections to locate the subscriber and the databases needed to support system operations.

In this chapter, the structure and operation of 2G and 2.5G network hardware elements (i.e., SD, RBS, BSC, MSC, GMSC) is described and their relationship with other network elements is explored from both a hardware and software viewpoint. Also, an example of how the wireless network coverage area is expressed in logical terms is presented. The functions of the various network nodes that supply database information to the wireless network (HLR, VLR, AUC, EIR, etc.) are also presented to the reader. Additionally, this chapter gives some further insight into the signaling operations performed over SS7 that take place between wireless network elements that are used to set up the transfer of messages.

As these network devices and their functions are still fresh in the reader's mind, the future of cellular wireless is foreshadowed by a brief description of the generic system architectural model for 3G. The transformation of the wireless telecommunications network continues as it evolves toward an all-IP core network and radio access network (RAN). The presentation of more details about this topic is delayed until Chapter 7.

This chapter concludes with a section on the numbering and recommended identification systems used by wireless networks and several detailed examples of call setup and release operations. These examples are designed to tie together many of the hardware and network concepts presented within the chapter.

3.1 COMMON CELLULAR NETWORK COMPONENTS

The typical post–first generation (1G) wireless cellular telecommunications system as shown in Figure 3–1 consists of several subsystems or network elements designed to perform certain operations in support of the entire system. For 2G and 2.5G cellular networks, the air interface functions are typically performed by a fixed radio base station (RBS) and a mobile station (MS) or subscriber device (SD) that provide user mobility. The radio base station is usually controlled by a base station controller (BSC) and this portion of the cellular system is usually referred to as the base station system (BSS).

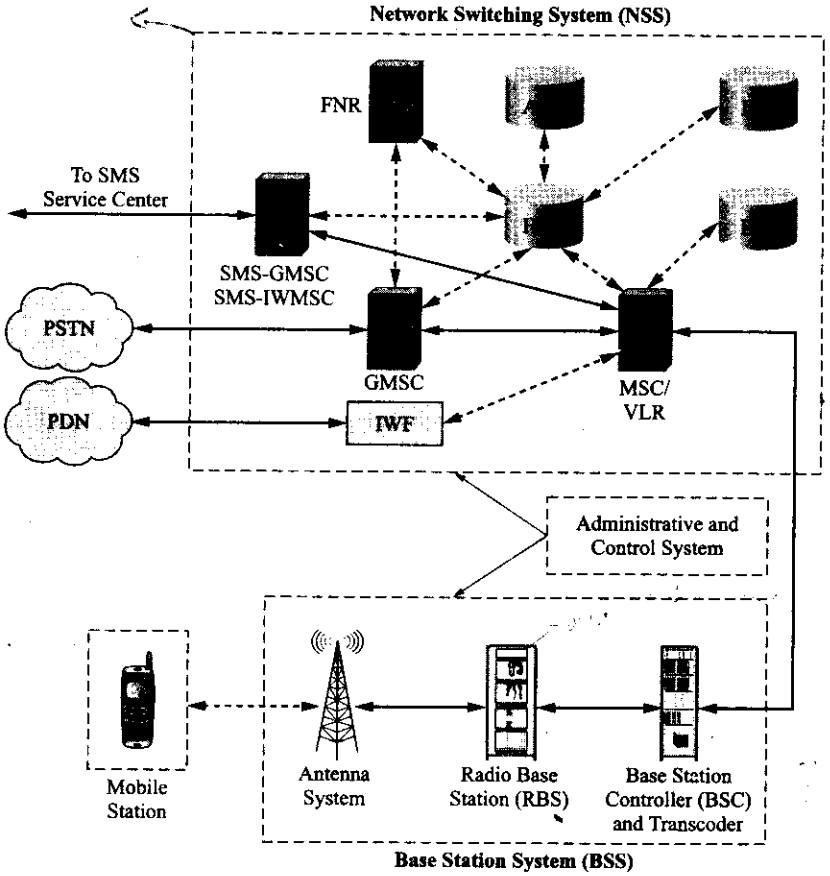

Figure 3–1 Typical wireless cellular system components.

The base station system is connected to a fixed switching system (SS) that handles the routing of both voice calls and data services to and from the mobile station or subscriber device. This switching system usually consists of a mobile switching center (MSC) and various databases and functional nodes used to support the mobility management and security operations of the system. The switching system is usually connected to the PSTN, the PDN, other public land mobile networks (PLMNs), and various data messaging networks through gateway switches (GMSCs). Other typical connections to the switching system are to network management systems and other accounting or administrative data entry systems.

The various network elements that make up the wireless system are interconnected by communications links that transport system messages between network elements to facilitate network operations and deliver the actual voice call or data services information. The rest of this section is devoted to descriptions of these

network elements and brief overviews of their basic functions. It should be pointed out again that all cellular wireless systems are standards based and therefore both the names of the system subunits and the communication interfaces between them are defined by the standard for the particular type of technology used by the system (GSM, NA-TDMA, CDMA, etc.).

In this chapter, the subject matter will be dealt with in as generic a way as possible using common terms and definitions. Later chapters devoted to particular systems will present the names of the system components and interfaces using the correct nomenclature for them as specified by the appropriate system standard.

Subscriber Devices

The first generation of wireless cellular systems provided connectivity to the PSTN for voice service. The initial term used in several standards for the mobile transceiver supplied to the cellular system users was mobile station. As cellular systems have matured and added ever faster data service delivery to the traditional teleservices available to the user, the term *subscriber device* (SD) has come to be used to describe the mobile transceiver for these newer systems. As the wireless network evolves toward an all-IP network, the expression used for the mobile transceiver is expected to morph one more time with the eventual adoption of the term **end terminal** (ET). This name change will be in keeping with the mobile station's ability to connect to an all-IP network and thus provide the functionality of an end terminal device.

The **subscriber device** is the link between the customer and the wireless network. The SD must be able to provide a means for the subscriber to control and input information to the phone and display its operational status. Additionally, the SD must be able to sample, digitize, and process audio and other multimedia (e.g., video) signals; transmit and receive RF signals, process system control messages; and provide the power needed to operate the complex electronics subsystems that provide the functionalities mentioned earlier. Therefore, as shown in Figure 3–2, the basic sections of the SD are as follows: some form of a man-machine interface, an RF transceiver section, a signal processing section, a system control processor, and a power supply/management section.

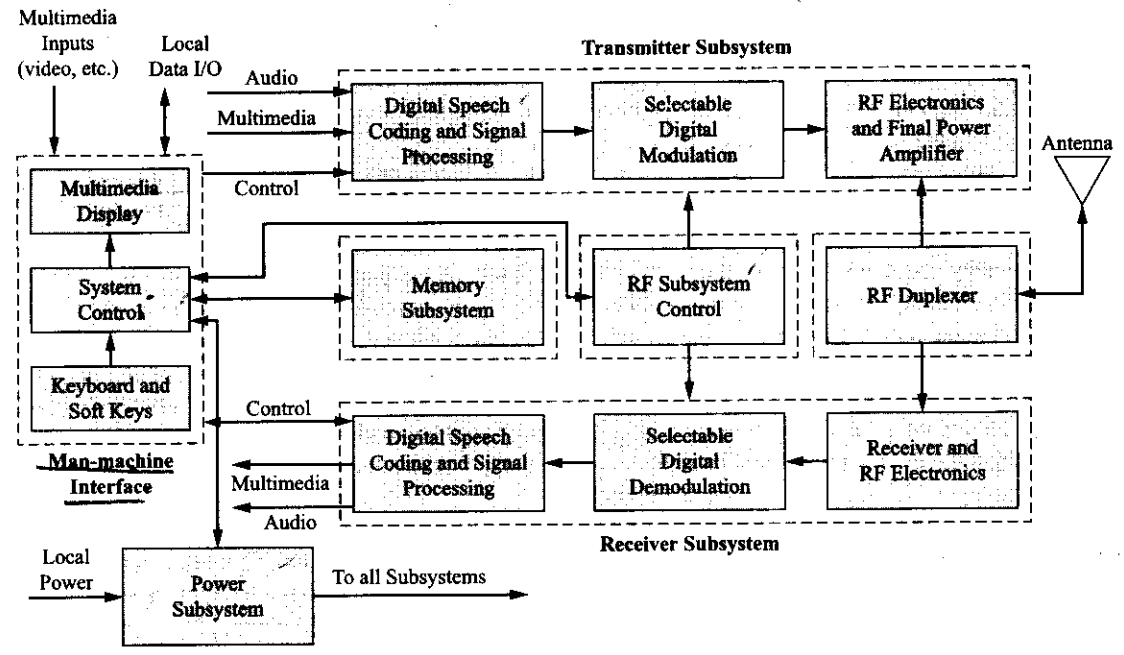

Figure 3–2 Typical subscriber device block diagram.

The man-machine interface can be as simple as a standard telephone keypad, an alphanumeric text display, and a microphone/speaker combination. Or, it may be more sophisticated with soft-key keypad functions and multimedia capability with a high-resolution color display and video camera or cameras for the transmission and display of video messages. Additional accessory interfaces usually also exist to provide the option of hands-free operation, battery charging, and a service port or a data port for connection to a PC.

The RF transceiver section contains the high-frequency RF electronics needed to provide the proper digital modulation and demodulation of the air interface RF signals and the ability to transmit and receive these RF signals. This section must also permit both variable power output and frequency agility under system software control.

The signal processing section of a subscriber device is usually based on digital signal processor (DSP) technology. Some of the functions performed by this section are speech sampling and coding, channel coding, and audio and video processing.

The system control processor provides overall subscriber device management. It implements the required interface with the other wireless network elements to provide radio resource, connection management, and mobility management functions through software control of the various functions and operations it must perform to set up and maintain the air interface radio link.

Finally, the power supply section provides the power to energize the entire system. Usually, the SD is battery operated with sophisticated algorithms built into the system to save and minimize power usage as much as possible in an effort to extend the battery life. When the battery becomes discharged, it may be recharged through a home accessory battery charger or through the accessory connector of one's car.

Base Station System Components

The **base station system** handles all radio interface-related functions for the wireless network. The BSS typically consists of several to many radio base stations (RBSs), a base station controller (BSC), and a transcoder controller (TRC). It should be noted that these last two network elements did not exist in the first analog cellular systems. In 1G systems the RBSs were connected directly to the MSC. The radio equipment required to serve one cell is typically called a base transceiver system (BTS). A single radio base station might contain three base transceiver systems that are used to serve a cell site that consists of three 120-degree sectors or cells. The radio base station equipment includes antennas, transmission lines, power couplers, radio frequency power amplifiers, tower-mounted preamplifiers, and any other associated hardware needed to make the system functional.

The base station controller's function is to supervise the operation of a number of radio base stations that provide coverage for a contiguous area (see Figure 3–3). It provides the communication links to the fixed part of the wireless network (PSTN) and the public data network (PDN) and supervises a number of air interface mobility functions. Some of these tasks include location and handoff operations and the gathering of radio measurement data from both the mobile device and the radio base station. The base station controller is used to initially set up the radio base station parameters (channels of operation, logical cell names, handoff threshold values, etc.) or change them as needed. The BSC is also used to supervise alarms issued by the radio base stations to indicate faults or the existence of abnormal conditions in system operation (including those of its own). For some faults the BSC can bring the reporting subsystem back into operation automatically (i.e., clearing the fault or alarm) whereas other faults require operator intervention in the form of an on-site visit by a field service technician.

The transcoder controller performs what is known as rate adaptation. Voice information that has been converted to a standard digital pulse code modulation (PCM) format is transmitted within the PSTN over standard T1/E1/J1 telephone circuits at 64 kbps. Both TDMA and CDMA systems use data rates of 16 kbps or less for the transmission of voice and control information over the air interface. The transcoder controller's function is to convert the PCM data stream to a format suitable for the air interface. **Vocoding** is

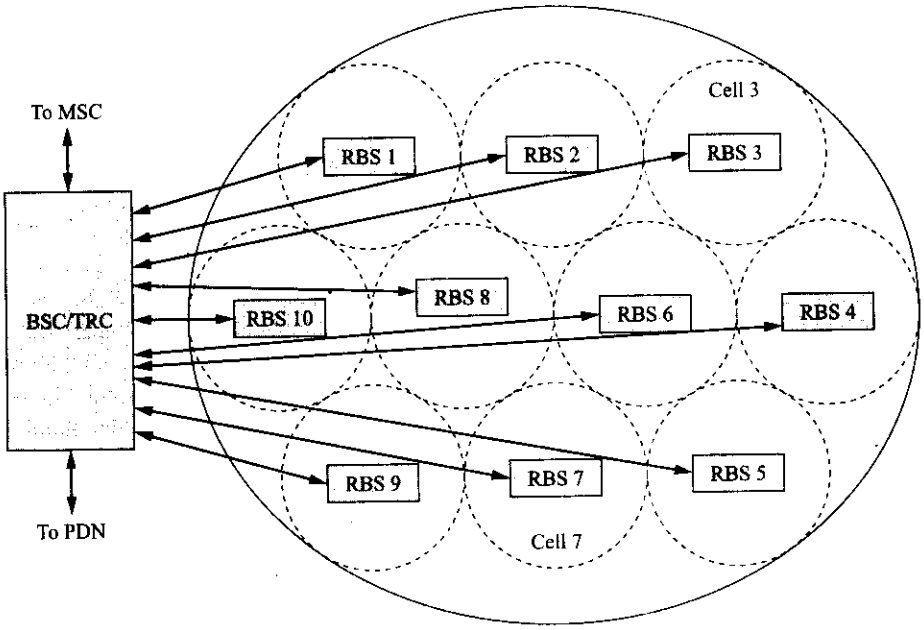

Figure 3–3 The base station controller's function.

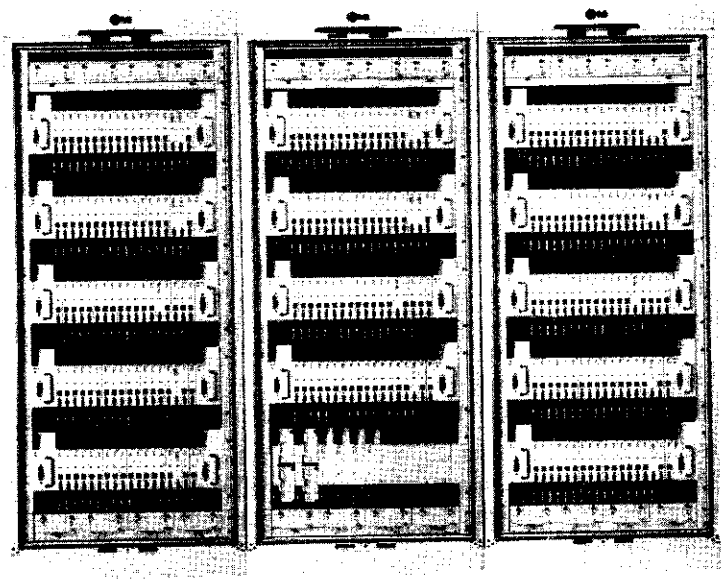

Figure 3–4 Typical cellular wireless equipment (BSC, TRC, and RBS) (Courtesy of LG Electronics Corp.).

another common term used for the process of converting audio to a digital format suitable for cellular transmission.

Physically, these units (BSC, TRC, and RBSs) are contained in standard radio relay rack enclosures. Figure 3–4 shows what a typical system looks like. Within the rack enclosure are subsystems devoted to functions such as power supply and control, environmental conditioning, switching, communications, processing, and so on. Additional hardware details about cellular base station systems will be presented in Chapters 5 through 8.

Radio Base Station

The **radio base station** consists of all radio and transmission interface equipment needed to establish a radio link with the MS. The typical RBS is composed of several subsystems that allow it to transmit to the MS on one frequency and to receive signals from the MS on another frequency. The two major wireless cellular systems used today for the air interface function are a form of either time division multiple access (TDMA) or code division multiple access (CDMA). The architecture and functionality of the air interface components of the RBS will depend upon the particular type of access system it is used in.

For TDMA systems, since frequency spectrum is a scarce resource, the primary function of the BSS is to optimize the use of available frequencies. The RBS supports this goal by having the ability to perform frequency hopping and support dynamic power regulation and the use of discontinuous transmission modes. All of these features tend to reduce interference levels within a TDMA system. For CDMA systems, all transmission is performed on the same frequency. However, precise timing, power control, and CDMA encoding and decoding are required to optimize system operation. The necessary subsystem components required for the proper functioning of a CDMA radio base station reflect this fact.

TDMA Radio Base Stations A typical TDMA radio base station consists of a distribution switch and an associated processor that is used to cross-connect individual timeslots of an incoming data stream to the correct transceiver units and provide overall system synchronization, multiple transceiver units (one per timeslot) with the ability to perform RF measurements on received signals, RF combining and distribution units to combine the output signals from the transceiver units and also distribute received signals to all the transceivers, an energy control unit to supervise and control the system power equipment and also to regulate the environmental conditions of the RBS, and power supply components (both rectified AC and battery-supplied DC) to provide power for system operation.

CDMA Radio Base Stations A typical CDMA radio base station consists of many of the same switching function, RF transceiver, power supply, and environmental conditioning components as the TDMA radio base station with the addition of a timing and frequency module that receives timing information from a **Global Positioning System** (GPS) receiver colocated with the RBS and channel cards that are responsible for the CDMA encoding and decoding functions on the forward and reverse links to and from the subscriber devices. For CDMA radio base stations, a typical design might consist of a main and a remote unit. The main unit provides all the functions except for RF power amplification. The two units are linked by fiber-optic communications cables and power supply cables. These cables supply all the signals needed by the high-power RF amplifier and the remote electronics that are typically mounted on a tower near the system antenna.

Base Station Controller

The **base station controller** functions as the interface between the mobile switching center and the **packet core network** (PCN) and all of the radio base stations controlled by the BSC. The PCN is a term used for the interface node (network element) between the BSC and the public data network. Figure 3–5 shows how the systems are interconnected.

Aside from the necessary power supply and environmental conditioning components, the BSC typically consist of several subsystems all colocated in a main cabinet or possibly several cabinets. The system organization tends to divide up these subsystems into those that are used to provide a connection or link between the MSC and the radio base stations and those subsystems that control the operation of these aforementioned units. The typical connection from BSC to the MSC or TRC (if it is not integrated into the BSC) is over standard T1/E1/J1 PCM links as is the connection from BSC to RBSs. A standard switching fabric is used within the BSC to direct incoming voice calls from the MSC to the correct RBS. Another switching fabric that can deal with subrate transmissions (less than 64 kbps) is usually also available within the BSC adding increased functionality to the system. If the TRC is colocated with the BSC, transcoding functions are also performed within the combined BSC/TRC unit.

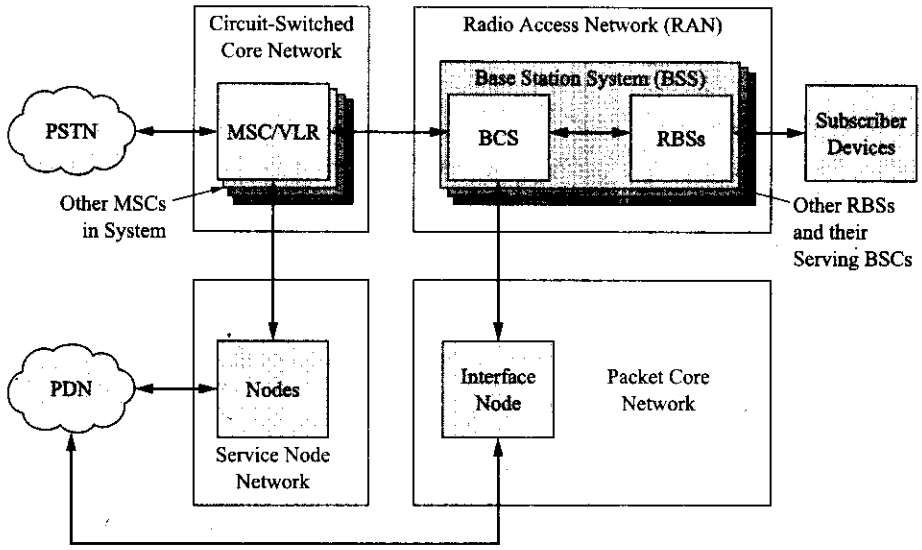

Figure 3–5 Typical CDMA wireless system (Courtesy of Ericsson).

The operation of each of these subsystems is controlled by processors under stored program control. Furthermore, the BSC system provides timing signals and connectivity to every subsystem within it and computer interfaces to the entire system for either network or element (subsystem) management functions.

Additionally, the BSC will supply signaling toward the MSC using message transfer part (MTP) protocol to transfer the messages over a PCM link connected to SS7 signaling terminals located within the MSC and the BSC. Signaling between the BSC and the RBSs is done over a PCM link using link access protocol on D-channel modified for mobile (LAPDm) signaling functions.

Connections to the PDN through an interface unit (PCN) connected to the BSC will be discussed in greater length and detail in Chapter 7.

Transcoder Controller

The **Transcoder Controller** (TRC) consists of subsystems that perform transcoding and rate adaptation. The TRC can be either a stand-alone unit or, more commonly, combined with the BSC to yield an integrated BSC/TRC. The TRC also can support the power saving option of discontinuous transmission. If pauses in speech are detected, the mobile station will discontinue transmission and the TRC will generate "comfort noise" back toward the MSC/VLR. An integrated BSC/TRC can typically handle many 100s of RBS transceivers.

Both TDMA and CDMA systems transmit speech over the air interface using digital encoding techniques that yield data rates of less than 16 kbps. The PSTN uses a PCM encoding scheme that yields a data rate for voice of 64 kbps. Therefore, voice messages coming from the PSTN must be transcoded to a rate suitable for the cellular system and, similarly, voice messages originating from a mobile station must be transcoded into a format suitable for the PSTN. This operation takes place in the TRC. The incoming PCM signal from the PSTN is converted back to an analog signal. At this point, 20-ms segments of the analog signal are converted to a digital code by a device known as a **vocoder.** The vocoder compares the 20-ms speech segment against a table of values. The entry in the table that is closest to the actual value is used to produce a code word that is much shorter than the corresponding PCM codes for the same 20-ms period. This compressed code word is what gets transmitted by the system. At the MS, the process is reversed to obtain an analog voice signal. For voice signals going in the opposite direction the steps are duplicated but in the reverse order. The obvious advantage to the use of vocoding is the reduced data rate needed for

speech transmission. Additional enhancements to this process have led to half-rate speech coders that can encode speech signals in only 8 kbps, and other variations on this theme.

Swiching System Components

As stated earlier, the switching system performs several necessary cellular network functions. It provides the interface (MSC) both to the radio network portion of the system (BSS) and to the PSTN and other PLMNs. It also provides an interface to the PDN and other network support nodes and gateways. Included in the switching system are functional databases (HLR, VLR, AUC/EIR, etc.) that contain information about the system's subscribers, their network privileges and supplementary services, present SDs locations, and other information necessary to locate, authenticate, and maintain radio link connections to the subscriber's devices. The following sections will provide brief overviews of the functions and operation of the various switching system subsystems and databases.

Visitor Location Register

The **visitor location register** (VLR) is a database that temporarily stores information about any mobile station that attaches to a RBS in the area serviced by a particular MSC. This temporary subscriber information is required by the MSC to provide service to a visiting subscriber. When an MS registers with a new MSC service area, the new VLR will request subscriber information from the MS's home location register (HLR). The HLR sends the subscriber information to the VLR and now if the MS either sends or receives a call the VLR already has the information needed for call setup. In a typical wireless network the VLR is integrated with the MSC to form an MSC/VLR thus reducing the amount of SS7 network signaling necessary to perform wireless network operations.

Mobile Switching Center

The **mobile switching center** (MSC) is at the center of the cellular switching system. It is responsible for the setting up, routing, and supervision of voice calls to and from the mobile station to the PSTN. These functions are equivalent to those performed by the traditional telephony circuit switch (e.g., 5ESS, DMS-100/200, and AXE 810) used in a central office by the wireline PSTN. The traditional equipment manufacturers of this type of switching system all sell a cellular version of their standard wireline switch. Most of these systems also combine VLR functionality, in addition to the telephony switching functions, yielding an integrated MSC/VLR system.

The basic functions performed by the MSC/VLR are as follows: the setting up and control of voice calls including subscriber supplementary services, providing voice path continuity through the use of the handoff process, call routing to a roaming subscriber, subscriber registration and location updating, subscriber data updating, authentication of MSs, delivery of short messages, signaling to other network elements (BSC, HLR, etc.) or networks (PSTN, PLMNs, etc.), and the performing of charging/accounting, statistical, and administrative input/output processing functions.

As shown in Figure 3–6, the typical MSC consists of the following components or subsystems devoted to network operations: a central processor and associate processors, group switch, traffic interfaces, timing and synchronization modules, and software to provide operations and maintenance (O&M) functions. The next several sections will provide some additional detail about the operation of a typical MSC.

MSC Interface and Switching Functions Today's "trunk" connections (i.e., high-capacity facilities) between local central office (CO) exchanges and gateways to long-distance provider facilities make available the transport of high bit-rate digital signals. These local and long-distance interoffice connections are most often supplied by fiber-optic cables that are carrying SONET-based optical signals at bit rates in the 100s of mbps range or higher (the STS-3 signal carried as OC-3 is 155.520 mbps). SONET is capable of transporting multiple T1/E1/J1 carriers and asynchronous transfer mode (ATM) traffic. The standard voice call is

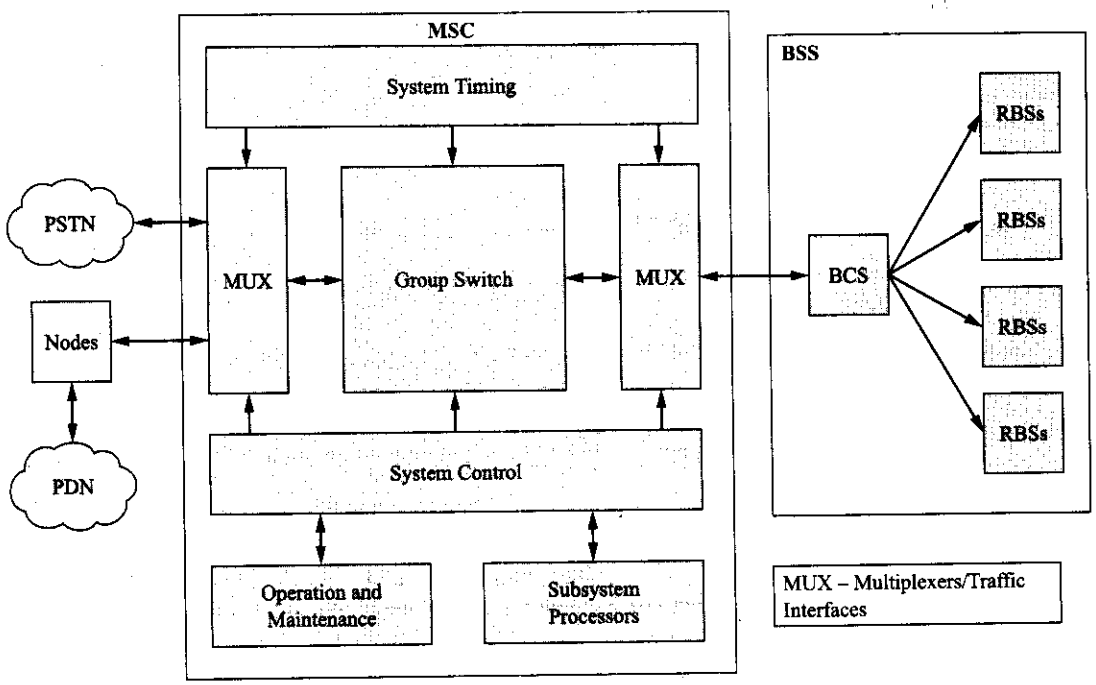

Figure 3–6 Typical MSC subsystems.

carried over these facilities as a DS0 signal that has a bit rate of 64 kbps. A T1/J1 carrier can transport twenty-four digitized voice calls and the E1 carrier has a capacity of thirty calls. The MSC can be thought of as just another central office exchange in that it has its own local exchange routing number(s) (i.e., N1/0N-NNX-XXXX where N1/0N is the three-digit area code and NNX is the exchange number). Therefore, the connection from the MSC to the PSTN or other PLMNs is usually provided in the same manner as other interoffice connections, over fiber trunk facilities or through traditional Tn/En/Jn carrier facilities depending upon the needed capacity.

Therefore, the MSC needs to provide the ability to multiplex and demultiplex signals to and from the PSTN. This functionality is built into the traffic interface subsystems (refer back to Figure 3–6). These interface units will bring the high bit-rate data streams down to the base T1/E1/J1 carrier signal after demultiplexing of the signals from the PSTN. Or conversely, they can be used to multiplex together many T1/E1/J1 signals to form a high bit-rate signal to be transmitted over a high-speed transmission facility back toward the PSTN (this operation is typically referred to as **backhaul**) or other networks as needed. The connection between the MSC and the base station controllers it services is also implemented with the same standard transmission T1/E1/J1 facilities or larger-capacity fiber facilities. Recently, cellular providers have been providing their own high-speed fixed point-to-point digital microwave backhaul networks with T1/E1/J1 or higher capacity from remote RBSs to BSCs and then from BSCs to the MSC location when traditional facilities are either not available or prove to be too costly to install and lease.

The **group switch** provides the same functionality in the MSC as it does in the PSTN local exchange. In both cases, the incoming voice calls on a particular T1/E1/J1 carrier arrive assigned to a particular timeslot. In order that the voice call can be directed to the correct BSC a combination space and timeslot interchange (TSI) switch must be used to redirect the voice call to both the correct output line and also to a free timeslot within the T1/E1/J1 carrier signal. The following example will describe the operation of a typical group switch in an MSC/VLR.

Example 3–1

A certain mobile subscriber is registered to a certain RBS in a cell that is located in an area that uses six BSCs to control the RBSs in that area. Show how the MSC directs an incoming call to the mobile subscriber if the MS's RBS is controlled by BSC #4.

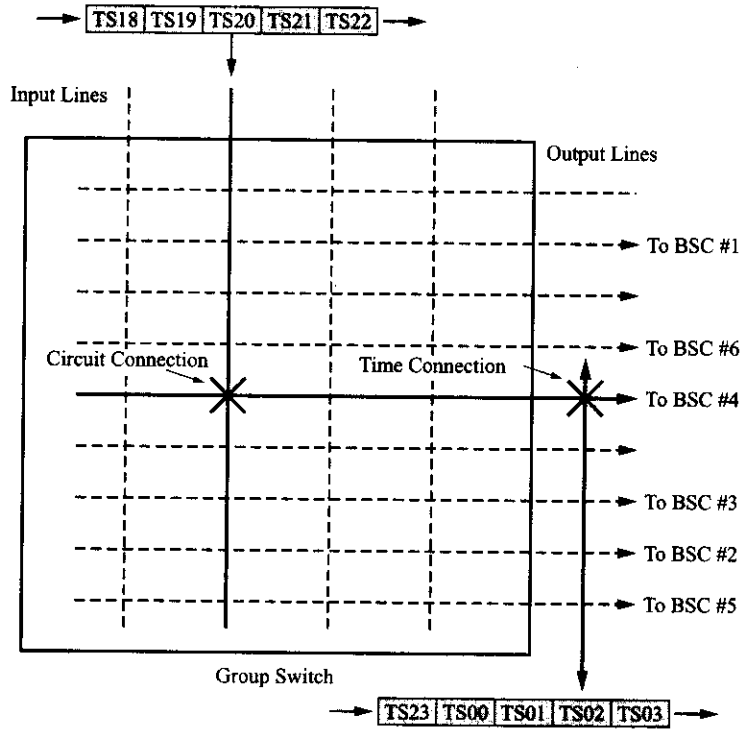

Figure 3–7 Operation of the group switch for Example 3–1.

Solution: Referring to Figure 3–7, assume that the incoming voice call occupies Timeslot #21 (any value from 0 to 23 could be used here) on a T1 carrier signal connected to a local exchange in the PSTN. After any necessary demultiplexing, the signal is applied to the group switch. The group switch processor implements a path that allows the signal to be redirected to available Timeslot #2 on the line connected to BCS #4 (this latter information is provided by the MSC). The switch performs this function as indicated by Figure 3–7 and the voice call is correctly routed toward BSC #4. The MSC and the BSC have been in contact by sending messages to one another over SS7 so that the BSC is aware of the new incoming voice call on Timeslot #2. Note that a functionally identical call path must also be established in the reverse direction to provide for duplex operation. There are duplicate subsystems available within the MSC to accomplish this task.

MSC *Signaling Functions* To coordinate the processing of calls both to and from the MS, the MSC and BSC must exchange messages using message transfer part (MTP) and the signaling connection control part (SCCP) of signaling system #7 (SS7). The MTP provides reliable transfers of signaling messages over standard (T1/E1/J1) digital transmission links running in parallel with digital traffic links (sometimes referred to as sidehaul connections). Refer back to Figure 3–5. One of the SCCP user functions is known as base station system application part or BSSAP. It is used for standard GSM signaling between a MSC and BSC.

The BSSAP protocol supports messages between the MSC and the BSS and also between the MSC and the MS. BSSAP is divided into two subparts: direct transfer application part (DTAP) that is used to send connection and mobility management messages between the MSC and the MS, and base station system management application part (BSSMAP) that is used to send messages between the MSC and the BSC related to the MS, a cell within the BSS, and the entire BSS.

MSC Database Functions As stated previously, the various functional databases contained within the cellular network switching system contain information about the system subscribers, their network privileges and supplementary services, present location and other information necessary to locate, authenticate, and maintain radio link connections to the subscriber's devices. Therefore, the MSC/VLR is continually sending and receiving data from the HLR, and AUC/EIR databases. The signaling and data transfer between the MSC/VLR and these databases is carried out using MTP and SCCP over SS7. More detail about these operations will be supplied with the descriptions of the databases themselves.

Home Location Register

The **home location register** (HLR) is a database that stores information about every user that has a cellular service contract with a specific wireless service provider. This database stores permanent data about the network's subscribers, information about the subscriber's contracted teleservices or supplementary services, and dynamic data about the subscriber's present location. The type of permanent data stored includes mobile station identification numbers that identify both the mobile equipment and the PSTN plan that it is associated with. This information would include a mobile station ID number that consists of a **country code,** either a national destination code or a number planning area code, and a subscriber number. Other ID numbers as defined and required by the particular wireless network are also stored by the HLR.

The HLR also plays a major role in the process of handling calls terminating at the MS. In this case, the HLR analyzes the information about the incoming call and controls the routing of the call. This function is usually supported by the transfer of information from the HLR to the VLR within the MSC where the subscriber's mobile is registered.

HLR Implementation and Operation An HLR can be implemented as a stand-alone network element or it can be integrated into an MSC/VLR to create an MSC/VLR/HLR system. The HLR itself consists of the following subsystems: storage, central processors, I/O system, and statistics and traffic measurement data collection. Additionally, SS7 signaling links are maintained between a network HLR and the MSC/VLRs and GMSC that compose a cellular network. Usually, a wireless service provider will have more than one HLR within a **public land mobile network** (PLMN) to provide the necessary redundancy to support disaster recovery. The information about subscriber subscriptions is usually entered into the HLR database through a service order gateway (SOG) or an operations support system (OSS) interface.

The HLR has two basic functions. It maintains databases of subscriber-related information. This information may consist of both permanent data such as subscriber-associated MS numbers and dynamic data such as location data. The HLR is able to support typical database operations like the printing and modification of subscriber data and the addition or deletion of subscribers. More complex operations like the handling of authentication and encryption information and the administration of MS roaming characteristics are also performed. The HLR also performs call handling functions such as the routing of mobile terminating calls, the handling of location updating, and procedures necessary for delivery of subscriber supplementary services.

HLR Subscription Profile A basic function of the HLR is to store a subscriber's profile. This profile defines a group of services that the subscriber has signed up for when first contracting for mobile service. The types of services available are typically referred to as **teleservices** (telephony, short message service, fax, etc.) and **bearer services** (i.e., data services). These services are typically grouped into basic service groups that are packaged for sales and promotion purposes. A user's profile stored by the HLR may be updated or

modified at any time with vendor-specific computer commands or more easily by clicking on graphical user interface (GUI) icons in a Windows-based application program.

Supplementary services are system functions like call waiting and call holding, multiparty service, calling line and connected line identification, call forwarding, call barring, and so on. Within each of these categories, there are many options that may be selected. These supplementary services may be programmed into a user's profile fairly easily as mentioned earlier. As well as the normal services that may be specified by a particular system standard, systems will typically offer vendor-specific supplementary services that are used in an attempt to provide some form of marketplace differentiation.

HLR/AUC Interconnection The authentication center (AUC) provides authentication and encryption information for the MSs being used in the cellular network. For GSM systems, so-called triplets are provided for the authentication of a mobile. Upon a request from a VLR, the HLR will be delivered a **triplet** (i.e., a ciphering key, a random number, and a signed response) for a particular mobile subscriber. The HLR receives the triplet information in response to a request to the AUC for verification of a subscriber. The HLR forwards the random number to the MSC/VLR where it is passed on to the mobile. The mobile performs a calculation using the random number and returns it to the MSC/VLR and from there to the HLR. If the results are the same as the signed response, the mobile is **authenticated** and it is now able to access the radio resources of the network. The AUC contains a processor, a database for the storage of key information for each subscriber, maintenance functions for subscriber information, and an interface for communications with the HLR. CDMA systems use a similar system for authentication.

Interworking Location Register

Interworking location registers (ILRs) are used to provide for intersystem roaming. The ILR allows a subscriber to roam in several different systems. For instance, in a wireless cellular system using an appropriate ILR, a subscriber could roam between an AMPS system and a PCS system. In this case, the ILR would consist of an AMPS HLR and parts of a PCS VLR.

Authentication Center and Equipment Identity Register

The authentication center (AUC) is a database that is connected to the HLR. The authentication center provides the HLR with authentication parameters and **ciphering keys** for GSM systems. Using the cipher keys, signaling, speech, and data are all encrypted before transmission over the air interface. The use of encryption provides over-the-air security for the system.

The equipment identity register (EIR) database is used to validate the status of mobile equipment. In GSM systems, the MSC/VLR can request the EIR to check the current status of an MS through the global database maintained by the GSM Association. This global database is updated daily to reflect the current status of an MS. The MS can be "black listed" indicating that it has been reported stolen or missing and thus not approved for network operation. Or, the MS might be "white listed" and therefore registered and approved for normal operation. The hardware necessary to perform AUC/EIR functions might be colocated within a wireless network.

Gateway MSC

The **gateway MSC** (GMSC) is an MSC that interfaces the wireless mobile network to other telecommunications networks. Although a cellular network might have numerous MSCs to facilitate coverage of a large geographical area, not all of these switching centers need to be connected to other wireline networks or other PLMNs. Usually this connection is made at one particular MSC and this MSC is now known as a gateway MSC or GMSC. To support its function as a gateway, the GMSC will contain an interrogation function for obtaining location information from the HLR of a subscriber. The GMSC will also have the ability to reroute a call to an MS using the information provided by the HLR. Charging and accounting functions are typically implemented in the GMSC.

Interworking Units

Interworking units (IWUs) are required to provide an interface to various data networks. These nodes are used to connect the base station controller and hence the radio base stations to various data services networks. This is necessitated by the fact that the MSC is a circuit-switched device and inappropriate for the transmission of data packets. Presently, for both TDMA and CDMA systems, these interworking units have evolved into specific functional nodes such as gateway GPRS support nodes (GGSNs) and **packet core network** (PNC) nodes, respectively. These IWUs will be discussed in greater detail in Chapter 7.

Data Transmission Interworking Unit An early interworking unit, the data transmission IWU (DTI), was used to allow the subscriber to alternate between speech and data during the same call. The main functions performed by the DTI were protocol conversion and the rate adaptation necessary for fax and data calls through a modem.

SMS Gateways and Interworking Units To provide **short message service** (SMS) (i.e., the sending of a text message consisting of up to 160 alphanumeric characters either to or from a mobile), two network elements are required in GSM networks: the short message service gateway MSC (SMS-GMSC) and the short message service interworking MSC (SMS-IWMSC). This first device is capable of receiving a short message from an SMS center (SC), interrogating an HLR to obtain routing information and message waiting data, and finally delivering the short message to the MSC of the receiving mobile. The second device is capable of receiving a mobile-originated short message from the MSC or an alert message from the HLR and delivering these messages to the subscriber's SMS center. Multimedia message service or MMS uses a different means of providing data transmission through the wireless network than SMS does. Again, more detail about SMS and MMS operations will be forthcoming in Chapter 7.

Network Management System

All modern telecommunication networks have some form of network management built into the system. This overarching management tool provides for overall network surveillance and support to the operation and maintenance of the entire network. A wireless service provider will usually have a **network operations center** (NOC) devoted to the use of this network management system (NMS) to provide 24/7 coverage of the system. Different equipment manufacturers have different names for these management systems; however, they all tend to have the same functionality. They provide fault management in the form of network surveillance, performance management, trouble management, configuration management, and security management.

Usually, the NMS has subnetwork management platforms that provide management of the circuit, packet, and radio networks. These subnetwork management platforms also provide configuration, fault, performance, and security management of their respective subsystems.

Other Nodes

Other nodes that may be connected to the switching system but are not really part of the telecommunications network itself are the SMS or service center (SC), the billing gateway (BGW), and the service order gateway (SOG). The service center acts like a store and forward center for short messages, the billing gateway collects billing information, and the service order gateway provides subscription management functions.

Service Center

The service center is used to facilitate the operation of short message service. It performs two functions: a mobile-originated short message is transferred from the cellular network to the SC for storage until the message can be transferred to its MS destination, or the SC stores a mobile-terminated short message from

some other short message entity (SME) that might or might not be a MS until it can be accepted by the intended MS.

Billing Gateway

The billing gateway (BGW) collects billing information from various wireless network elements (principally, the MSC and GMSC). The common term used for the information collected by the BGW is call data records (CDRs). As these call records are collected from the network elements they become files used by a customer administrative system to generate billing information for the system's subscribers. Information about monthly access fees, home usage and roaming usage charges, data and special services usage charges, and so on, are all used to generate a monthly bill for each subscriber.

Service Order Gateway

The **service order gateway** (SOG) is used to connect a customer administrative system to the switching system. This system is used to input new subscriber data to the HLR or to update current subscriber data already contained in the HLR. The SOG also allows access to the AUC and the EIR for equipment administration. When a customer initially signs a service contract with a cellular service provider, the information about the contract is entered into the customer administrative system. The administrative system sends customer service orders to the SOG. The SOG interprets the service orders and delivers the appropriate information to the correct network elements in the form of network service orders.

3.2 HARDWARE AND SOFTWARE VIEWS OF THE CELLULAR NETWORK

At this point in our discussion of the various hardware elements that are used to realize a cellular system, it will be instructive to examine a possible implementation of a typical 2G/2.5G/2.5G+/3G wireless system. How the components are actually physically laid out and connected to provide coverage to a particular area will be discussed. How the network elements are viewed by system software is slightly different however. This view will also be presented and contrasted with the hardware point of view. See Figure 3–8 for an illustration of a possible hardware layout used to cover a specific geographic area.

Figure 3–8 depicts a fairly large geographic area with a potential subscriber base of approximately 100,000 that is served by a cellular network consisting of two mobile switching centers and a total of six base station controllers. The reader is urged to try and relate the demographic and geographic features of his or her own hometown location to this example. For the sake of clarity, all the radio base stations (cells) for only one BSC are shown. All the details of the individual cells are not included at this time.

Hardware View of a Cellular Network

The area on the left side of the diagram is served by MSC-1 and thus will be known as the service area of MSC-1. The right side of the diagram is served by MSC-2 and is thus labeled as the service area of MSC-2. MSC-1 interfaces with three BSCs (BSC-1A, BSC-1B, and BSC-1C) that are used to cover the three areas that the MSC-1 service area has been subdivided into. Each of these BSCs has several to many RBSs serviced by it depending upon the population density and nature of the various areas (urban, suburban, business district, industrial, etc.). In some of the areas there may be both microcells and macrocells whereas other locations will just have macrocells. In the service area of MSC-2, there are three more BSCs (BSC-2A, BSC-2B, and BSC-2C).

The RBSs might be named to reflect their connection to a particular BSC (i.e., RBS-2A1, RBS-2A2, and so on for RBSs connected to BSC-2A). In this diagram, the GMSC provides the gateway connection to the PSTN for MSC-1 and MSC-2, and MSC-1 has the switching system databases colocated with it. PCM links

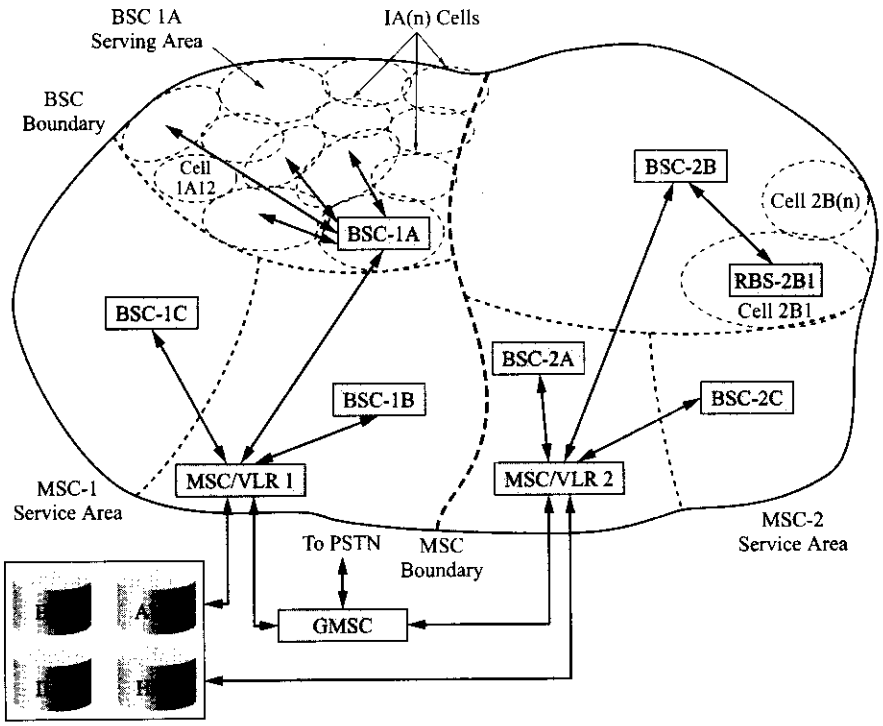

Figure 3–8 Hardware view of a cellular system.

exist between each RBS and its BSC, between each BSC and its MSC/VLR, and between the MSC/VLRs. These PCM links might be leased from the local telephone company or they may be implemented using microwave digital radio links installed by the service provider or a combination of both facilities. The gateway MSC is most likely linked to the PSTN by some form of high-capacity T-carrier or fiber span. Actual statistics about cell site locations and antenna statistics of cellular and PCS systems are available from the FCC's Web site. This information is contained in the universal licensing system database that may be found at http://wireless.fcc.gov/uls/.

Software View of a Cellular Network

The operations performed within the cellular network to complete calls, keep track of a mobile's location, and maintain radio links through handoff, but to name a few, are all directed by the network elements under program or software control. The cellular network therefore takes on a slightly different appearance to the system software. Physical objects and areas take on logical names to distinguish them from each other and to allow the software the ability to perform the required operations. Figure 3–9 shows the same geographic area as Figure 3–8; however, this time the cellular network is shown from a software viewpoint.

As shown in Figure 3–9, the network is defined by location area identity (LAI) numbers and cell global identity (CGI) numbers. The CGI numbers locate a particular cell whereas the LAI numbers define an area for paging. Because a mobile may have moved since its last location updating message (that would include the LAI number), an incoming call to the mobile will result in a page to every cell within the location area. If a mobile moves into another location area, it is required to automatically update its location with the VLR for the new location area.

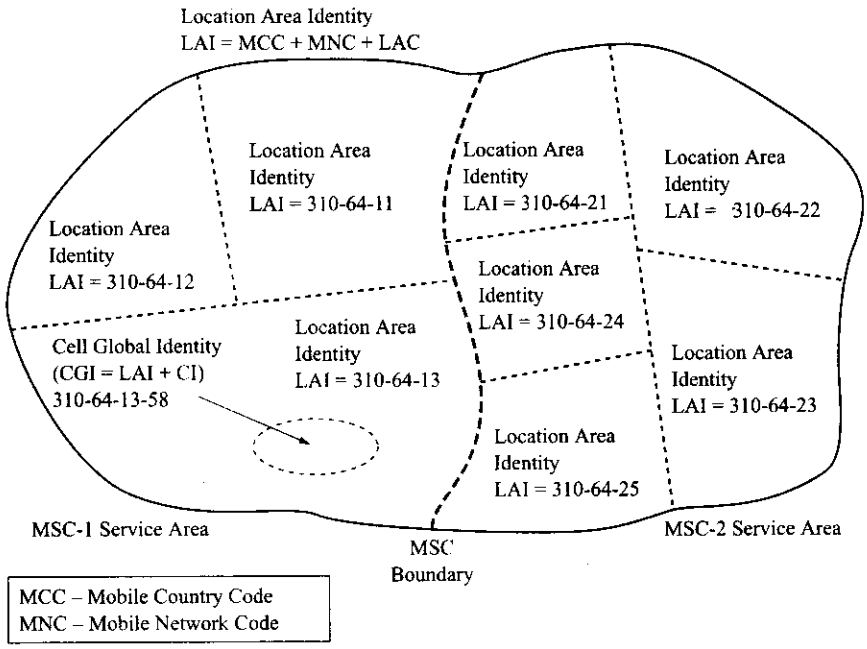

Figure 3–9 Software view of a cellular system.

3.3 3G CELLULAR SYSTEM COMPONENTS

For 3G cellular systems the network elements have transformed to reflect the transition of the system toward an all-IP or packet network. This system evolution has resulted in the transformation of the BSC function to that of a radio network controller (RNC). As shown in Figure 3–10, the function of the RNC node is to provide the interface between the wireless subscriber and the core networks. The core networks are the circuit core network for all circuit-switched voice and data calls and the packet core network (PCN) for all packet data calls. The RNC, although similar to the BSC, has additional functionality that distinguishes it from the BSC.

Each proposed 3G system uses the same designation of RNC instead of BSC. Much more detail will be provided about the components in 3G systems in Chapter 7.

3.4 CELLULAR COMPONENT IDENTIFICATION

To switch a voice call from the PSTN to a mobile subscriber the correct cellular network elements must be involved in the operation. It is therefore necessary to address these elements correctly or the operation will not be completed properly. The International Telecommunications Union (ITU), acting in its capacity as a global standards organization, has adopted several standards and recommendations to deal with these issues. Recommendation E.164 is known as the international public telecommunication numbering plan. This recommendation, adopted in 1997, details the numbers to be used for assigning PSTN telephone numbers on a global basis. This same recommendation is followed when assigning numbers to cellular telephones and provides a dialable number with which one can connect with the mobile through a wireless network. Furthermore, Recommendation E.212 deals with the numbering schemes for mobile terminals on a global basis. As stated before, the transmission of messages between cellular network elements used to facilitate cellular switching and control operations is accomplished through the use of SS7, in the same fashion as the PSTN. Therefore, network switching elements or processing nodes are associated with

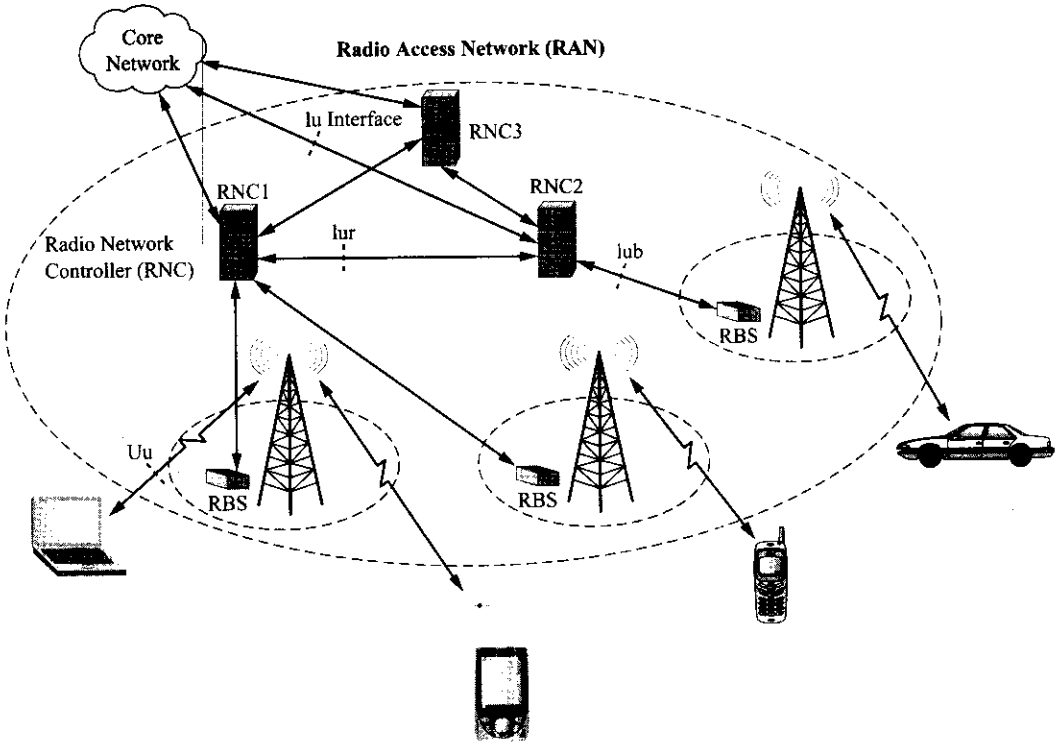

Figure 3–10 The 3G radio network controller.

addresses assigned to SS7 signaling points. These signaling point addresses are generated by the translation of E.164 and E.212 information into mobile global titles (Recommendation E.214) during the processing of operations by the cellular system elements.

This section will examine some of the basic numbering schemes used in wireless mobile networks for the different network elements that make up the system. Further details about specific systems will be offered in upcoming chapters.

Subscriber Device Identification

The mobile subscriber device (SD) can have several different system identification numbers associated with it. The identification information used depends upon the type of cellular technology (TDMA, GSM, or CDMA) employed by the network it is being used in and the scope of the network (e.g., national or international). The next few sections will expand upon this topic.

Mobile Station ISDN Identification Number

The mobile station ISDN (MSISDN) number is a dialable number that is used to reach a mobile telephone. There are slight variations in the MSISDN number depending upon whether one is in North America or in other parts of the world. Figure 3–11 provides a graphic of how these MSISDN numbers are formed.

As shown, in North America an MSISDN number consists of the following:

$$\text{MSISDN} = \text{CC} + \text{NPA} + \text{SN}$$

Where, CC = Country Code, NPA = Number Planning Area, and SN = Subscriber Number

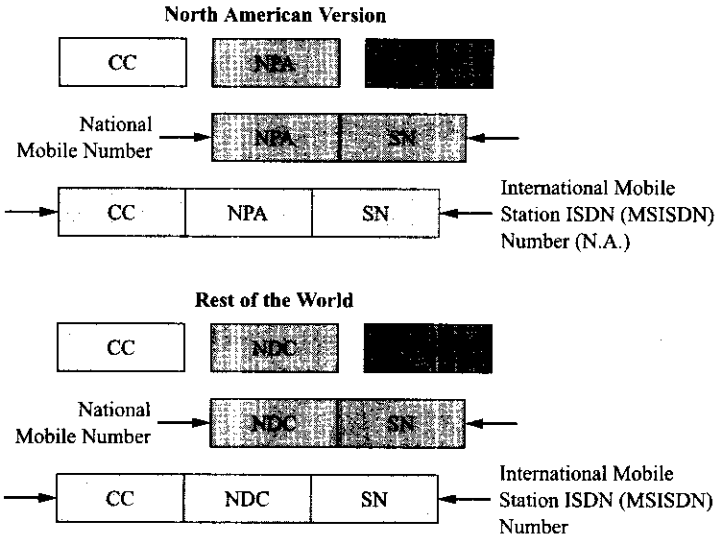

Figure 3–11 Formation of the MSISDN number.

Example 3–2

A cellular telephone subscriber signs up for service in Springfield, MA, USA. What is the subscriber's MSISDN?

Solution: Since the country code for the USA is +1 and the area code for Western Massachusetts is 413, the MSISDN will take the form

MSISDN = +1-413-732-XXXX

In the rest of the world an MSISDN number consists of the following:

MSISDN = CC + NDC + SN

Where, NDC = National Destination Code. The NDC is similar to the NPA but can also identify the type of network (fixed, wireless, etc.) being called.

International Mobile Subscriber Identity

For international public land mobile networks an international mobile subscriber identity (IMSI) is assigned to each subscriber. Figure 3–12 indicates how the IMSI is formed.

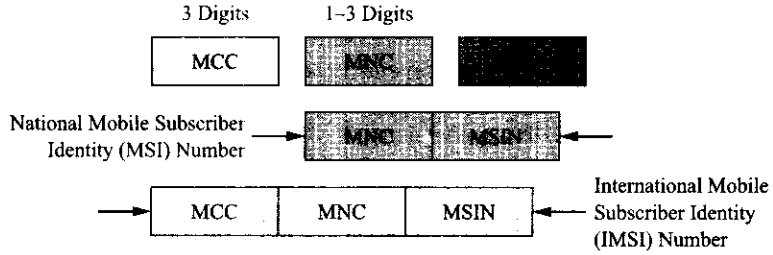

Figure 3–12 Formation of the IMSI number.

As shown, the IMSI number consists of the following:

IMSI = MCC + MNC + MSIN

Where, MCC = Mobile Country Code (see Recommendation E.212), MNC = Mobile Network Code, and MSIN = Mobile Subscriber Identification Number. For a GSM network the IMSI number is stored in the SIM (subscriber identity module) card that is inserted into the mobile telephone and provided to the subscriber by the service provider.

There is also a temporary mobile subscriber identity (TMSI) number that may be used instead of the IMSI. This TMSI number is used to provide security over the air interface and therefore only has local significance within an MSC/VLR area.

International Mobile Equipment Identity

For international mobile networks, an international mobile equipment identity (IMEI) number is defined and is used to uniquely identify a MS as a piece of equipment to be used within the network. Figure 3–13 indicates the structure of the IMEI number.

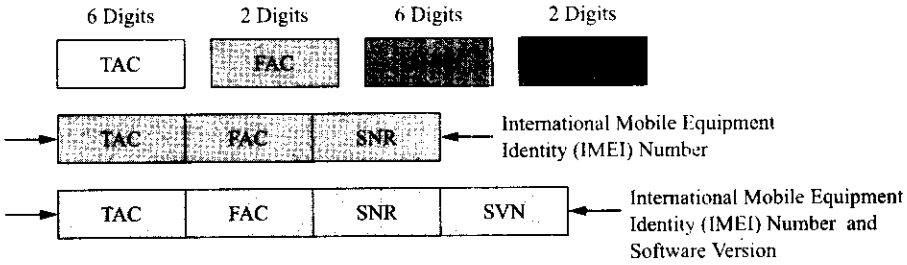

Figure 3–13 Formation of the IMEI number.

The IMEI can be modified to include information about the software version of the subscriber device operating system or application software within the identity number.

Cellular System Component Addressing

The rest of the cellular network hardware components that make up the switching system or the base station system have either signaling point (SP) addresses or some type of logical name assigned to them to distinguish them from similar components within the network. Some of the addresses are predetermined by the ITU Recommendations E.164 and E.212 and some are translated into new addresses that conform to Recommendation E.214. The logical names of devices are assigned by the system operator.

Additionally, physical areas of network coverage are also defined and given logical identification names and numbers to provide for the mobility management functions of the system or to define billing areas for regional or national service plans.

Location Area Identity

The location area identity (LAI) is used for paging an MS during an incoming (mobile terminating) call and for location updating of mobile subscribers. Figure 3–14 shows the structure of an LAI number.

As shown, the LAI consists of the following:

LAI = MCC + MNC + LAC

Where, again, MCC = Mobile Country Code, MNC = Mobile Network Code, and LAC = Location Area Code, which is 16 bits in length and therefore allows the network operator 65,536 different possible areas or codes within a network. The code is assigned by the mobile operator.

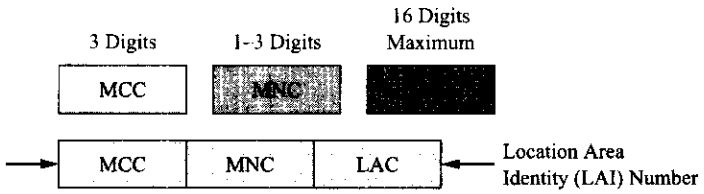

Figure 3–14 Formation of the location area identity number.

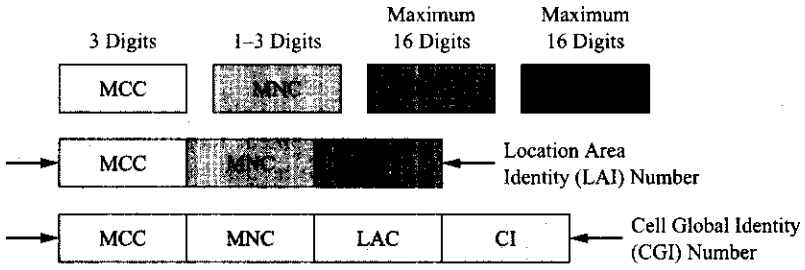

Figure 3–15 Formation of the cell global identity number.

Cell Global Identity

The cell global identity (CGI) is used for the unique identification of a cell within a location area. It is formed by adding 16 bits to the end of a LAI. This also allows for the possibility of 65,536 cell sites within a location area. Again, the code is assigned by the mobile operator. Figure 3–15 shows the structure of a CGI number.

Radio Base Station Identity Code

A radio base station identity code (BSIC) is used by the mobile operator to identify RBSs within the wireless network. This code allows an MS to distinguish between different neighboring base stations. The BSIC usually consists of a 3-bit network color code and a 3-bit base station color code.

Location Numbering

ID numbers may be assigned by the service provider to various regional or national areas to provide subscriber features such as regional or national calling plans.

Addressing Cellular Network Switching Nodes

Messages between both PSTN and PLMN network elements are sent over the SS7 network. To facilitate this operation, signaling point addresses or point codes are assigned to the network switching elements and processing nodes. Recall that one of SS7's functional elements, signaling connection control part (SCCP), provides the ability to communicate with wireless mobile network switching system databases without any speech connection. Additionally, message transfer part (MTP) provides the common platform to perform the required routing of the signal message to the correct destination over the various link sets that are available within the SS7 network.

ANSI-based SS7 network numbering for the network SPs consists of constructing 3-byte (24-bit) addresses that specify the telecommunications service provider (N = 0–255), the network cluster number (C = 0–255), and the member number (M = 0–255) in the following format: signaling point code (SPC) = N-C-M. Cluster numbers usually denote network geographic areas or sections and member numbers identify a specific node within the network. The service provider codes are set by Bellcore whereas the cluster and member numbers may be set by the service provider or Bellcore in special instances. Since a signaling

point might actually be a combined node that performs several functions (e.g., MSC/VLR or MSC/VLR/HLR), SS7 messages use one address for the node and another subsystem address (SSN) to indicate the individual element within the node (SSN = 6 for the HLR, SSN = 7 for the VLR, SSN = 10 for the AUC, etc.).

Global Title and Global Title Translation

A global title (GT) is an address of a fixed network element. However, the GT does not contain the information needed to perform routing in a SS7 system. The derivation of the correct routing information is performed by the SCCP translation function. The global title is used for the addressing of network nodes such as MSCs, HLRs, VLRs, AUCs, and EIRs in accordance with E.164.

Consider an incoming call to a GMSC. The first operation necessary by the wireless system is to locate the MS. The GMSC uses the MSISDN (MSISDN = CC + NPA/NDC + SN) number to point out the appropriate HLR. Usually the CC, NPA or NDC, and one or more digits of the SN are used from the MSISDN to create a GT to identify the HLR.

A mobile global title (MGT) is created during various location updating or mobility management functions initiated by the mobile. In the case of a GSM system, during a location updating function, the MSC/VLR only knows the mobile's IMSI number. This number is used to create an MGT number that points to the mobile's HLR. Figure 3–16 shows how the MGT number is formed. In this case, the E.164 part is used along with the E.212 part to form the E.214 MGT.

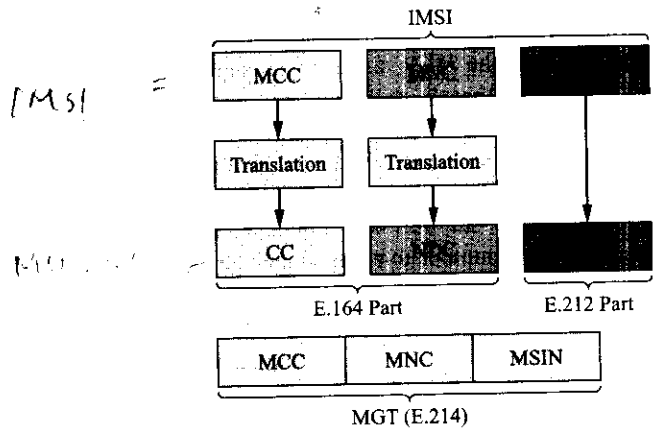

Figure 3–16 Formation of the mobile global identity number.

The necessary global title translation (GTT) is performed by a SCCP translation function to provide the correct signaling point address information for the subsequent routing of the message to the correct network node. This SCCP functionality permits MTP routing tables within individual signaling point locations to remain a manageable size as wireless telecommunications networks become larger and more complex.

The next section will provide some examples to tie all these concepts together.

3.5 CALL ESTABLISHMENT

The topic of **call establishment** was first introduced in Chapter 2 during an overview of the first-generation analog AMPS system. At that time, the reader was introduced to the many handshaking functions that were performed between the MS and the BS and between the BS and the MSC to complete call setup and handoff functions. Now that more detail about the network elements and databases of digital wireless cellular systems has been introduced, it would again be instructive to take a look at some of the basic wireless

network system functions that are involved with voice call establishment. The discussion of the delivery of digital services over these networks will be covered in Chapter 7.

Within the wireless mobile industry, the various possible voice call and data service circumstances that can occur are called traffic cases. Depending upon the type of traffic situation, radio resource management, connection management, and mobility management operations are needed to maintain the radio link between the mobile station and the base station system and the MSC. These management functions all reside at the Layer 3 or network layer level of the OSI model. The messages sent between the MS, RBS, BSC, and MSC using either LAPDm (modified for mobile LAPD) or MTP protocols implement Layer 2 or data link layer operations. The type of radio signals, modulation, and timing specifications used by the MS and the RBS are Layer 1 or physical layer characteristics. In Chapter 4, the suite of radio resource, connection, and mobility management operations will be presented in a slightly different view after the details of cellular architecture are presented.

Mobile-Terminated Call

The mobile-terminated call consists of the steps shown in Figure 3–17. Step #1: Any incoming call to a mobile system from the PSTN is first routed to the network's gateway mobile switching center (GMSC). Step #2: When the wireless mobile system detects an incoming call at the GMSC, the mobile system must first determine where the mobile is located at that particular moment in time. To determine the mobile's location, the GMSC will examine the mobile station's MDISDN to find out which home location register (HLR) the mobile subscriber is registered in. Using SS7 (SCCP), the MSISDN is forwarded to the HLR with a request for routing information to facilitate the setup of the call. Step #3: The HLR looks up which MSC/VLR is presently serving the MS and the HLR sends a message to the appropriate MSC/VLR requesting an MS roaming number (MSRN), so that the call may be routed. This operation is required since this information is not stored by the HLR; therefore, a temporary MSRN must be obtained from the appropriate MSC/VLR. Step #4: An idle MSRN is allocated by the MSC/VLR and the MSISDN number is linked to it. The MSRN is sent back to the HLR. Step #5: The MSRN is sent to the GMSC by the HLR. Step #6: Using the MSRN, the GMSC routes the call to the MSC/VLR. Step #7: When the serving MSC/VLR receives the call, it uses the MSRN number to retrieve the mobile's MSISDN. At this point the

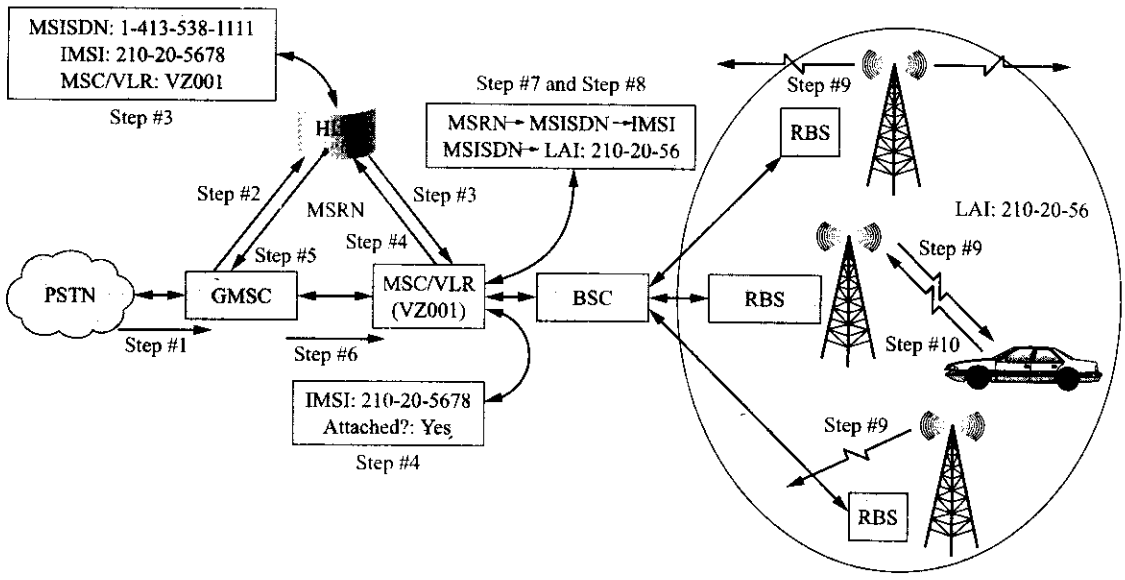

Figure 3–17 Mobile-terminated call operations.

temporary MSRN number is released. Step #8: Using the mobile's MSISDN, the MSC/VLR determines the location area where the mobile is located. Step #9: The MS is paged in all the cells that make up this location area. Step #10: When the MS responds to the paging message, authentication is performed and encryption enabled. If the authentication and encryption functions are confirmed, the call is connected from the MSC to the BSC to the RBS where a traffic channel has been selected for the air interface.

Mobile-Originated Call

A mobile-originated call consists of the steps shown in Figure 3–18. Step #1: The originating mobile subscriber call starts with a request by the mobile for a signaling channel using a common control channel. If possible, the system assigns a signaling channel to the mobile. Step #2: Using its assigned signaling channel, the MS indicates that it wants service from the system. The VLR sets the status of the mobile to "busy." Step #3: Authentication and encryption are performed. Step #4: The mobile specifies what type of service it wants (assume a voice call) and the number of the party to be called. The MSC/VLR acknowledges the request with a response. Step #5: A link is set up between the MSC and the BSC and a traffic channel is seized. The acquisition of the traffic channel requires several steps: the MSC requests the BSC to assign a traffic channel, the BSC checks to see if there is an idle channel available, if a channel is idle the BSC sends a message to the RBS to activate the channel, the RBS sends a message back to the BSC indicating that the channel has been activated, the MS responds on the assigned traffic channel, the BSC sends a message back to the MSC to indicate that the channel is ready, and finally the MSC/VLR sets up the connection to the PSTN. Step #6: An alerting message is sent to the mobile to indicate that the called party is being sent a ringing tone. The ringing tone generated in the PSTN exchange that is serving the called party is transmitted through the MSC back to the mobile. When the called party answers, the network sends a Connect message to the mobile to indicate that the call has been accepted. The mobile returns a Connect Accepted message that completes the call setup process.

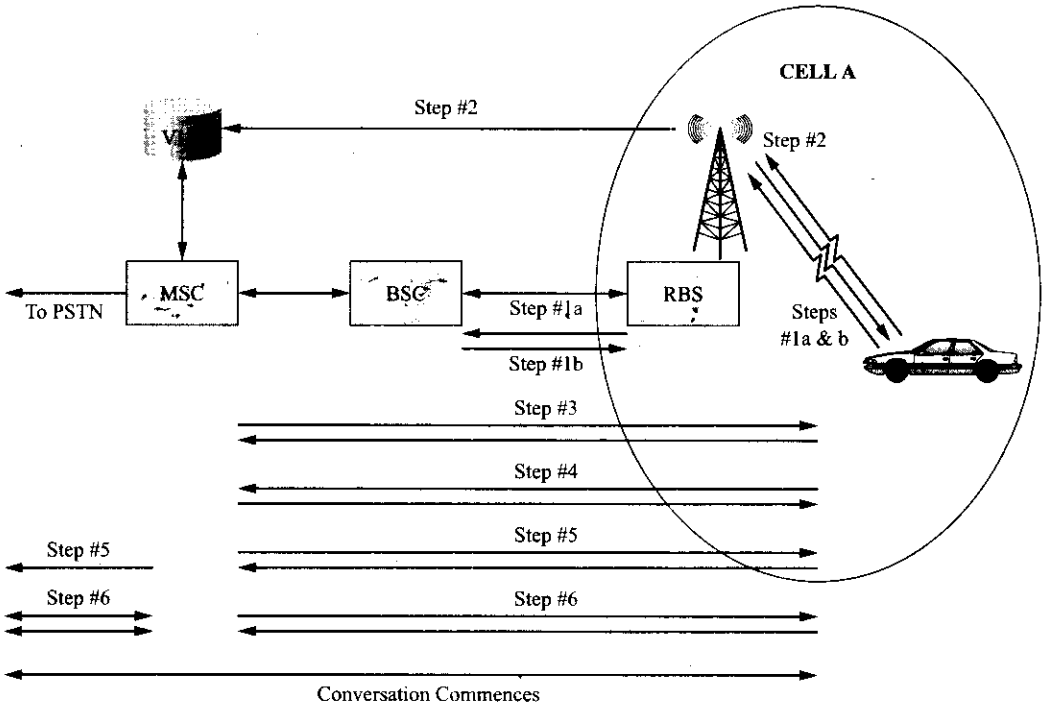

Figure 3–18 Mobile-originated call operations.

Call Release

Call release initiated by the mobile consists of the steps shown in Figure 3–19. Step #1: The mobile sends a Disconnect message to the RBS, the message is passed on to the BSC where it is sent through a signaling link to the MSC. Step #2: The MSC sends a Release message to the MS. Step #3: The MS sends a Release Complete message back to the MSC as an acknowledgement that the operation is complete. Step #4: The network initiates a channel release by sending a Clear Command message from the MSC to the BSC. The BSC sends the Channel Release message to the mobile through the RBS. Step #5: At this point, the BSC sends a Deactivate message to the RBS telling it to stop sending periodic messages to the mobile on a control channel. Step #6: When the mobile gets the Channel Release message, it disconnects the traffic channel and sends the LAPDm disconnect frame. The RBS sends an LAPDm acknowledgement frame back to the mobile. Step #7: A Release Indication message is sent from the RBS to the BSC. Step #8: The BSC sends an *RF* Channel Release message to the RBS that is acknowledged as shown, as soon as the RBS stops transmitting on the traffic channel. After a short period (regulated by a BSC system timer) a Clear Complete message is sent to the MSC from the BSC.

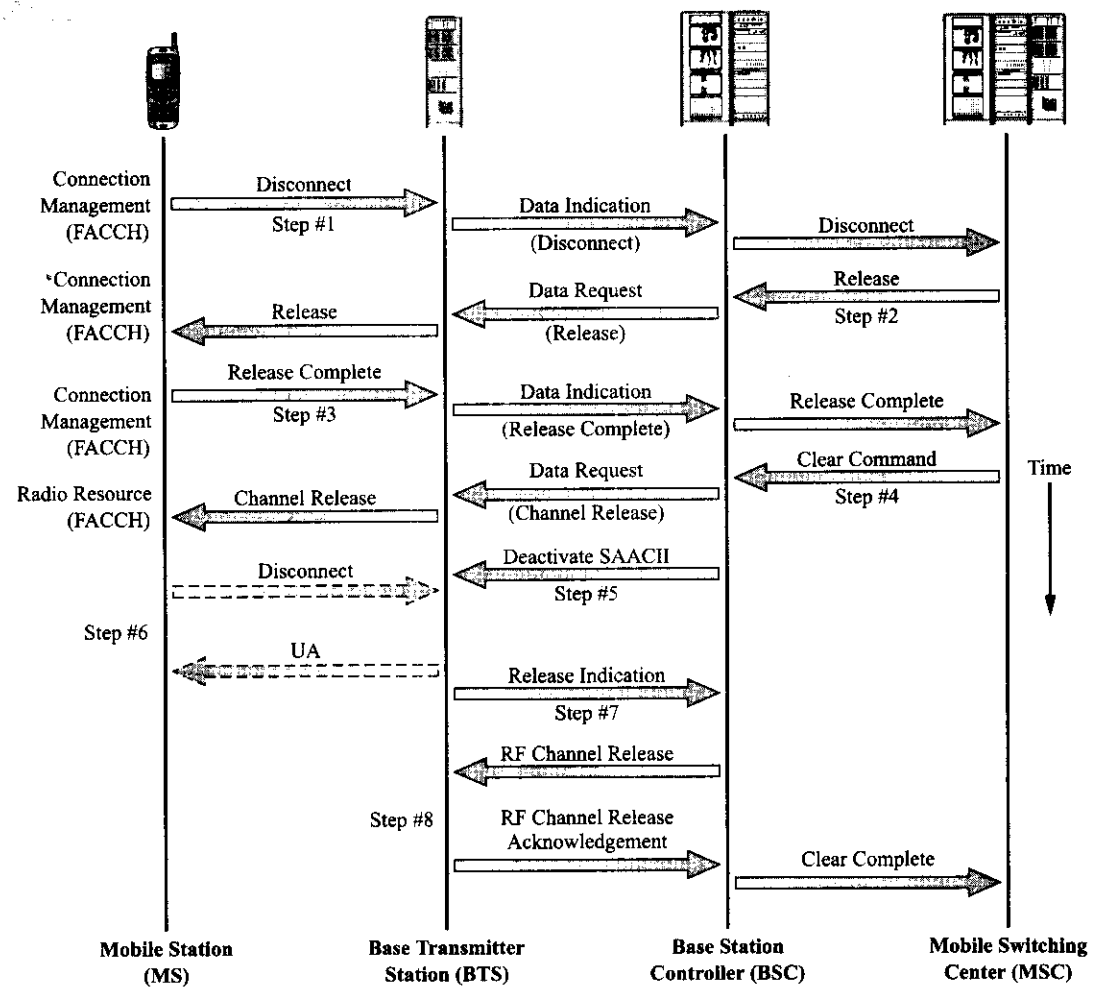

Figure 3–19 Call release operations.

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.