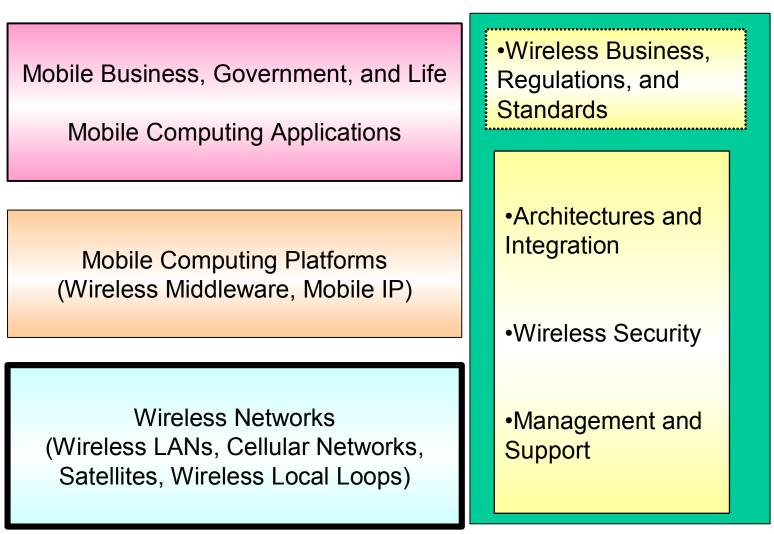
PART II

Wireless Networks

This part explains the wireless networks that use the air as the main transmission medium to interconnect mobile and fixed devices. These networks have several unique features: The bandwidths, and consequently data rates, of communication channels are restricted by government regulations; the communication channels between senders/receivers is often impaired by noise, interference and weather fluctuations; and the location of senders and receivers of information is unknown prior to start of communication and can change during the conversation. This part of the book starts with a discussion of the wireless communications principles and then shows how these principles are applied to wireless local, wide and metropolitan area networks (box with dark borders in the framework shown below).



5 Wireless Network Principles

Example: Wireless Network Design at a University

A medium-sized US university is aggressively pursuing a wireless agenda. The main campus of the university has about 20 buildings, including residential halls, in a 10-square-mile area in a suburb of a large city. There is a professional development center in the downtown area of the city, and the university also has an overseas office in Hong Kong. The university wants to provide wireless access in each building and also wants to develop an overall strategy for using wireless communications throughout the university community.

Here is a high-level summary of the overall approach (the terms used will be explained in this chapter):

- Wireless LANs, based on the IEEE 802.11 standard, were established in all buildings. Each building also had existing wired Ethernet LANs.
- Wireless Local Loop (WLL) was used on the main campus to interconnect many buildings on the main campus. A fiber optic metropolitan area network using FDDI already exists and currently interconnects older buildings.
- Paging networks based on ReFLEX were used throughout the university by construction workers and campus security personnel.
- The main campus was connected to the professional development center through a T3 Line.
- Satellites access for offshore instruction was rented from a satellite company.

The project involved a great deal of physical wireless network design issues such as antenna design for WLL, frequency allocations for the satellites and the WLL, and minimization of errors with maximum attempts to correct the errors. We will revisit this example at the end of this chapter (Section 5.11.1).

5.1 Introduction

This chapter discusses the basic principles of wireless communication networks. The following questions guide the discussion:

- How do wireless networks work and what is the basic vocabulary and classification of wireless networks (Section 5.2)?
- How are wireless frequencies allocated and regulated and what are the various frequency bands in which different wireless applications work (Section 5.3)?
- What are the basic approaches used in location management to keep track of wireless users who are always on the move (Section 5.4)?
- What roles does different wireless equipment (antennas, transmitters, receivers) play in wireless communications and what type of distortions are encountered by wireless wave propagation (Section 5.5)?
- What are different error detection schemes and why is error correction more important in wireless networks (Section 5.6)?
- Why are wireless communications becoming increasingly digital (Section 5.7)?
- What are the specialized signal encoding techniques and multiple access mechanisms (FDMA, TDMA, CDMA) and how do they assure that the wireless users do not interfere with each other (Section 5.9)?
- What is spread spectrum and what role does it play in wireless networks (Section 5.10)?

The discussion in this chapter is at an introductory level; however, some familiarity with computer networks is assumed. The short tutorial in Appendix A should be consulted before proceeding, if the reader is not sure. The topics discussed in this chapter span areas on which complete books have been written. The purpose of this chapter is to present a practical overview of the underlying principles on which the current and future wireless networks are based. Appendix B augments this discussion by taking a closer look at some of the physical communication topics such as error detection and correction, wireless propagation issues, and signal encoding. For yet more details on these topics, the following books should be consulted:[9]

- W. Stallings, *Wireless Communications and Networks* (Prentice Hall, 2002). I have used this book several times to teach wireless network courses to engineering students and technical professionals. It devotes four chapters to physical wireless communications (Antennas and Propagation, Signal Encoding Techniques, Spread Spectrum, and Coding and Error Control).
- T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. (Prentice Hall, 2001). This book is primarily intended for EE students and telecom engineers. It has an extensive discussion of physical communications (almost the entire book is devoted to layer 1 and 2 issues).
- J. Mark and W. Zhuang, *Wireless Communications and Networking* (Prentice Hall, 2003). This book is also intended for EE students and telecom engineers and is almost entirely devoted to physical networking issues.
- R. Bekkers and J. Smits, *Mobile Telecommunications* (Artech, 2000). This book covers, in addition to the physical networking issues, the standards and regulatory issues.

Chapter Highlights

[9] There are, of course, many other books on these topics but these are my favorites.

- Wireless networks are typically classified according to the distance. This classification broadly covers wireless local area networks (e.g., 802.11 networks), metropolitan area networks (e.g., wireless local loops), and wide area networks (e.g., cellular networks and satellites).
- Frequency allocations are important because designers need to know what frequency range to operate in. Wireless network operators cannot offer a service without knowing the operating frequency and any licensing restrictions for using the operating frequency.
- Lower-frequency bands are more congested because lower-frequency waves travel farther and need less powerful equipment to transmit (thus are cheaper to build). Because of the congestion, lower frequencies are auctioned frequently. Some companies have paid billions (yes, billions) of dollars to use frequencies for 3G cellular networks in these auctions.
- Location services are essential in wireless networks because the users of mobile services move around constantly, thus a network provider has to track the mobile users as they change their locations. These services are also important for 911 calls and mobile commerce.
- Multiple access mechanisms such as TDMA and CDMA allow multiple users to share the same medium without interfering with ach other. There are some some tradeoffs between CDMA and TDMA (no surprise!).
- Antennas and propagation help in understanding the type of errors and distortions the wireless signals have to go through in rural, urban, and rainy areas.
- Signal encoding and error correction are important because better encoding can improve data rate, and correction can aid recovery from propagation errors.

The Agenda
- How Do Wireless Networks Work
- Frequency Allocation and Location Services
- Physical Communications Considerations

5.2 How Do Wireless Networks Work – A Quick Overview

5.2.1 The Basic Operation of Wireless Networks

Wireless networks, as the name implies, interconnect devices without using wires – instead they use the air as the main transmission medium. Wireless networks are enjoying widespread public approval with a rapidly increasing demand. The increase in the number of cellular phones, palm pilots, PDAs, laptops, notebooks, and other handheld devices is phenomenal. To meet this demand, mobile communications technologies are emerging with digital speech transmission and the ability to integrate cordless systems into other networks. In the meantime, researchers are developing the next generation of technologies for several years to come.

Wireless communication systems consist of three basic elements: transmitters that generate and send signals, antennas that radiate the electromagnetic energy generated by the signals into the air, and receivers that receive and process the signals (see Figure 5-7). In some cases, transmitters and receivers are on same device, called *transceivers* (e.g., cellular phones). Antennas come in different forms and shapes (see Section 5.5). The unique features of wireless networks are:

- The bandwidths, and consequently data rates, of communication channels are restricted by government regulations. The government policies allow only a few frequency ranges for wireless communications.
- The communication channel between senders/receivers is often impaired by noise, interference and weather fluctuations.
- The senders and receivers of information are not physically connected to a network. Thus the location of a sender/receiver is unknown prior to start of communication and can change during active communication.

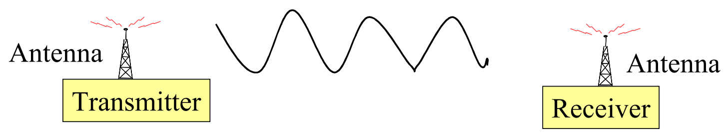


Figure 5-1: Wireless Communications

5.2.2 A Quick Scan of the Wireless Network Landscape

Figure 5-2 displays an overall classification of wireless networks in terms of distance covered, from very short range (10 meters) to very long range (thousands of miles). The classification broadly covers wireless local area networks, metropolitan area networks, and wide area networks. The figure also highlights a few examples with data rates and distance covered for each. Table 5-1 complements this figure by providing additional information (e.g., frequency band used). A brief overview follows; details are presented in the following chapters.

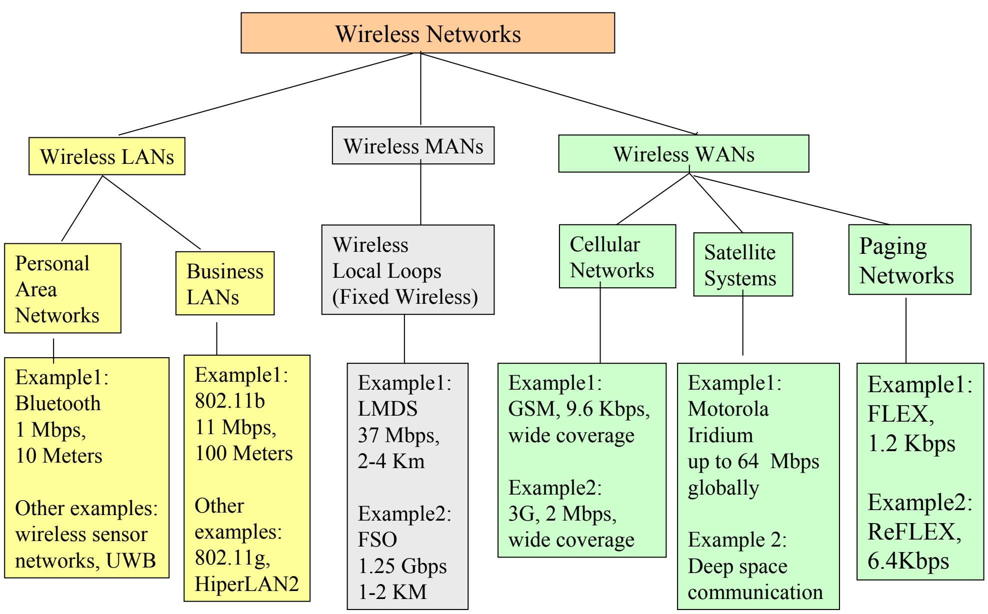


Figure 5-2: A View of Wireless Network Landscape

5.2.2.1 Wireless LANs (WLANs) and Wireless Personal Area Networks (WPANs)

WLANs allow workstations in a small area (typically less than 100 meters) to communicate with each other without using physical cables. These LANs are usually subdivided, for the purpose of convenience, into "business LANs" that can serve corporations, and personal area networks that are very short range networks for home and personal applications (this sub-classification does not always work, but is a good way to discuss the wireless LAN landscape).

The most common and widely used WLANs are the ones that belong to the IEEE 802.11 family. Although the original data rate was 1-2 Mbps, the data rates have been now extended to as much as 54 Mbps. There are three main extensions of the IEEE 802.11. The most widely used standard is the 802.11b (also known as Wi-Fi) WLAN that defines an 11 Mbps data rate for distances of up to 100 meters in the 2.4 GHz radio band. Another standard, 802.11b, uses the 5 GHz radio band and provides data rates up to 54 Mbps over shorter range (<50 meters). The latest development is the 802.11g standard that retains the same frequency range of 2.4 GHz of the popular Wi-Fi networks while increasing their data rates up to 54 Mbps. A practical implication of keeping the same frequency range is that the same access points can be used in the two networks. Thus a company using 802.11b can easily upgrade to 802.11g without additional costs. We should also mention the HiperLAN family of standards that were established by the European Telecommunications Standards Institute (ETSI) to provide both data and multimedia traffic. This family includes HiperLAN/1 and HiperLAN/2 that provide roughly the same data rates as the 802.11b and 802.11a/g, respectively. The differences between HiperLAN and 802.11 families are somewhat technical and will be discussed in Chapter 6.

Wireless Personal Area Networks (WPANs) are short-range (10 meters or less) radio networks for personal, home, and other special uses. Within the WPAN family, several specifications such as Bluetooth, wireless sensor networks, and UWB (Ultra Wideband) have emerged. Bluetooth is a wireless cable replacement standard that provides 1 Mbps data rate at 10 meters and less. It typically consists of a group of linked devices, such as a computer wirelessly connecting to a set of peripherals, known as a **"piconet."** Multiple piconets can be formed to provide wider coverage. Due to its relatively low data rates and very short distances, Bluetooth is being used in home appliances, "Bluetooth-enabled" cars and other such applications. UWB (Ultra Wideband) is a relatively new[10] technology and is stronger than the other short-range wireless systems (such as Bluetooth) because of its simpler device designs, lower power consumption and higher data rates. Another player in the short range radios is the wireless sensor networks (WSNs) that are formed between small, low-powered sensor devices mainly for monitoring and data collection purposes. Yet another player in short-range wireless, HomeRF, was primarily aimed at the needs of the small office and home office (SOHO) networks. This effort has been currently sidelined due to the popularity of other alternatives such as Bluetooth and UWB. See Chapter 7 for detailed discussion of these networks.

5.2.2.2 Wireless Metropolitan Area Networks (WMANs) – The Wireless Local Loop

WMANs span several miles and have been used in traditional packet radio systems for law-enforcement or utility applications. An interesting area of growth for wireless MANs is the

[10] As we will discuss in later chapters, UWB is not a new technology, it has been around for several years for military use, but has been "declassified in 2002 for commercial use.

wireless local loop (WLL) that is quite popular with the long distance telephone companies. WLLs are ***fixed wireless networks*** (i.e., the devices being connected are stationary) that deliver between 10 to 50 Mbps over several kilometers to address the last mile problem. The WLL providers use transmitters instead of wired local loops to deliver content (wireless TV, for example) to the subscribers. Several technologies exist for WLLs. The best-known examples of WLLs are LMDS (Local Multipoint Distribution Service) – a newer standard – and MMDS (Multichannel Multipoint Distribution Service) – an older standard. A relatively new entrant in the WLL market is Free Space Optics (FSO) that uses laser beams to deliver extremely high data rates (around 1 Gbps) over a few kilometers. FSO is gaining popularity because of its high security – it is difficult to intercept laser beams! In the last mile, wireless local loop technologies (LMDS, MMDS, and FSO) are providing strong competition to the wired local loops based on copper or fiber optic networks. Wireless solutions have the advantage that they can be installed quickly and less expensively. See Chapter 9 for discussion of WLLs.

5.2.2.3 Wireless WANs (WWANs) – Cellular Networks and Satellites

These networks provide wireless support over long distances. Traditional examples of wireless WANs are paging networks and satellite systems However, a great deal of business activity at present revolves around the cellular networks that provide support for cellular phones and other handheld devices such as PDAs and laptops.

Cellular networks, as we know, divide a geographic region into cells to serve a large number of users. Over the years, the cellular networks have evolved through several stages. The 1G (first generation) cellular systems, introduced in the early 1980s, use analog transmission and are primarily intended for voice applications. These networks are very slow – less than 1 killobit per second (Kbps). The 2G (second generation) wireless cellular systems, introduced in the late 1980s, use digital transmission and are also intended primarily for speech. However, they do support fax and email services at low bit-rate (8 to 9 Kbps) transmissions. GSM (Global System for Mobile Communications), the most prevalent example of 2G at present, operates at 9.6 Kbps. The 3G Systems represent the completely digital cellular facilities that can operate at 2 million bits per second. In addition to conventional voice, fax and data services, 3G promises to offer high-resolution video and multimedia services on the move. Although 3G systems are not widely deployed at the time of this writing due to their complexity and deployment cost, research in next generation (4G and 5G) cellular systems is underway at present to deliver 20 Mbps and higher. In addition to improvements in data rates, the cell sizes of newer cellular systems are getting smaller to maximize the re-use of a limited number of transmission frequencies. Smaller cells allow cellular companies to serve more customers by reusing the same frequencies, because two cells can use the same frequency for different conversations as long as the cells are not adjacent to each other. See Chapter 8 for more discussion of cellular networks.

A satellite is essentially a microwave repeater in the sky which receives signals from transmitting stations on earth and relays these signals back to the receiving stations on the earth. Satellites were first launched in 1962 and have ushered in the era of "celestial" and "global" communications. Satellites are used in areas such as telecommunications, weather reports, and scientific exploration. Although initially the domain of government, the private sector has become actively involved in satellites, with companies such as Hughes Space and Communications building a variety of satellites, and large corporations such as IBM, GE, and Walmart owning private satellites for corporate use. At present, communications through satellites is the foundation of global wireless networks in which users communicate with each other over long distances without a physical wire between them. From an end-user point of view, satellites can deliver around 50 Mbps over very wide (global) area networks. An interesting area of satellite communications is “Deep Space Satellites” that communicate over hundreds of thousands of miles. In fact, NASA is working on an Interplanetary Internet that would form an Internet between the satellites in the sky. See Chapter 9 for additional information about satellites.

A class of wireless wide area networks, not mentioned frequently in the popular press, is the paging networks that have been around for a long time. These networks are designed for short messages and are relatively slow. See Chapter 10 for additional information about paging networks.

5.2.3 Quick Comparison of Wireless Networks

Several wireless network technologies, ranging from short-range networks such as Bluetooth and UWB to deep space satellites that communicate over hundreds of thousands of miles, are currently available. With massive R&D in this area, new technologies are also being introduced and the existing technologies are being refined and improved on an ongoing basis. How can one evaluate the tradeoffs between the networks discussed so far and the others that will continue to appear over time? We will look at the underlying technologies in some detail in later chapters, but from an end-user point of view the two most important factors are data rates and the distance covered – you do not want to build a satellite system to communicate within a room (I hope not, but then you never know!). Another factor that is important for the wireless network providers and may be of some interest to the end-user is the radio frequency in which a particular wireless network operates. The main reason why radio frequency may be of interest to the end-user is that two wireless systems operating at the same frequency ranges cause interference, and also because changes in frequency bands require buying new hardware.

To facilitate quick comparison and analysis, Table 5-1 summarizes the characteristics of the available wireless network solutions in terms of the three factors: data rates, distance covered, and frequency used. It can be seen from this table that several technologies compete with each other directly or indirectly. We will do a more detailed comparative analysis as we go along by introducing other factors such as type of applications (e.g., voice, data), the level of mobility needed (e.g., no mobility, low mobility such as walking speed, or high mobility such as driving in a car).

Table 5-1: Wireless Technologies Alternatives

	Data Rate (Mbps)	Approximate Range (meters)	Radio Frequency (GHz)
Bluetooth	1 Mbps	10 meters	2.4 GHz
UWB	50 Mbps	<10 meters	7.5 GHz
IEEE 802.11a	Up to 54 Mbps	<50 meters	5 GHz (802.11a)
IEEE 802.11b	11 Mbps	100 meters	2.4 GHz
IEEE 802.11g	Up to 54 Mbps	100 meters	2.4 GHz
HiperLAN/2	Up to, 54	30 meters	5 GHZ
GSM	9.6 Kbps	Cell sizes 10 to 20 KM	Around 900 MHz
3G Cellular	Up to 2 Mbps	Cell sizes 5 to 10 KM	Between 1 GHz to 2 GHz
WLL (LMDS)	up to 37 Mbps	2 to 4 KM	Between 10 GHz to 100 GHz

FSO	100 Mbps to 2.5 Gbps	1 to 2 kilometers	teraHertz spectrum
Satellites	64 Kbps	thousands of miles	3 to 30 GHz

5.2.4 Wireless Versus Wired Networks

The wireless networks in the aforementioned categories are offering higher data rates than before. However, the wired networks are also offering higher data services. Thus, wireless networks not only have to compete with each other but they also have to compete with the fixed wired networks. Table 5-2 summarizes the typical data rates in the wireless versus wired world. As you can see, the wireless technology is much slower than the wired but is still used heavily because it offers greater flexibility to the users – speed isn't everything!

The main benefit of wireless, as discussed in Chapter 1, is that it is convenient and less expensive. However, several limitations and political and technical difficulties inhibit wireless technologies. Security is a major concern for corporate use of wireless networks. In addition, the lack of an industry-wide standard in wireless is a major problem. Consider, for example, frequency standards. The US and European agencies rarely agree on what frequency should be used for a new technology. The battle over frequencies for 4G networks is an example. Due to this, providers in Europe and in the US have to build different technologies and "gateways," and multi-frequency cards are needed to use the same technology in the US and Europe. In addition, there are several device limitations. For example, small LCD on a mobile telephone can only display a few lines of text, and browsers of most mobile wireless devices use wireless markup language (WML) instead of HTML.

The main problem is that wireless networks tend to be slower than the wired networks, as shown in Table 5-2. The data rate limitation of wireless is being addressed by the boadband wireless networks such as Free Space Optics and LMDS. Broadband wireless is suitable for graphics, video, and audio over wireless. Although broadband wireless networks provide higher data rates, they do not support mobility – they are fixed wireless. We will explain this issue also as we go along.

Table 5-2: Wireless Versus Wired Networks

	Local Area Networks (LANs)	**Metropolitan Area Networks (MANs)**	**Wide Area Networks (WANs)**
Wired	Wired LANs Ethernet (10-100 Mbps, 150 to 500 meters) Token Ring (4 -16 Mbps, 200 to 500 meters)	Wired MANs FDDI (100 Mbps, 50 Kilometers)	Wired WANs Dial up (56 Kbps) DSL/cable modems (200 Kbps-1 Mbps) ATM (44 Mbps to 140 Mbps) Frame Relay (44 Mbps) Higher data rates (over 100 Mbps) available
Wireless	Wireless LANs Bluetooth (1 Mbps, 10 meters) IEEE 802.11 LANs (11-54 Mbps, 100 meters)	Wireless MANs Wireless local loops (10-50 Mbps, 10 Kilometers)	Wireless WANs Current GSM systems at 9.6Kbps, future 3G systems at 2 Mbps Satellites at 64 to 100 Kbps

5.2.5 Technical Foundations of Wireless Networks

The following special issues need to be considered while examining wireless networks:

- **Applications Supported**. Different types of applications place different types of demands on the underlying wireless networks. For example, voice applications can tolerate more errors than financial transactions.
- **Location Management**. It is important to keep track of wireless users, since the location of a sender/receiver is unknown prior to the start of communication and can change during the conversation.
- **Frequency Allocations**. Wireless networks must compete with other media such as radio and television for frequency allocations because the bandwidths, and consequently data rates, of communication channels are restricted by government regulations.
- **Antennas and Propagation**. Specialized antennas, transmitters and receivers need to be designed to send and receive the wireless signals. In addition, once transmitted, the wireless signals encounter several propagation challenges due to scattering, fading and numerous other noises generated by the atmosphere (rain, thunderstorms, and different types of obstructions).
- **Specialized Signal Encoding Techniques**. Specialized techniques are needed to encode data into signals before it is transmitted. Sophisticated encoding techniques use a single signal to carry multiple data bits, thus significantly improving the data rate.
- **Error Detection and Correction**. Due to propagation challenges, wireless networks are prone to numerous errors. Thus errors must be detected accurately. But the bigger challenge is that the errors should be corrected as much as possible because a retransmission of an incorrect message may encounter the same type of errors.
- **Multiple Access Mechanisms**. Multiple users in the same frequency range need special mechanisms to avoid interference and collisions. Examples of these mechanisms are the popularly known CDMA (Code Division Multiple Access) and TDMA (Time Division Multiple Access) techniques used in cellular phones.

Table 5-3 summarizes the main ideas. We will discuss these factors in the balance of this chapter. Our old and faithful protocol stack, shown in Figure 5-3, illustrates at what layer of a network these issues need to be considered[11]. As can be seen, most wireless networking issues we will discuss are concerned with the lowest two layers. The Network and Transport layers are mostly handled by TCP/IP in most modern wireless networks (exceptions will be noted). The higher layers provide the voice and data applications discussed in previous chapters.

Table 5-3: Factors in Designing Wireless Networks

Factor	Why Important
Applications Supported	Most data applications require higher QoS than voice.
Frequency Allocations	Need to know what frequency range to operate. May need licenses to operate in several ranges.
Location Services	Essential because the users of mobile services change their location (e.g., a cellular phone). Also important for 911 calls.
Multiple Access Mechanisms	How users can share the same medium without interfering with ach other
Antennas and Propagation	Help in understanding the type of errors and how to deal with them

[11] If you have somehow fallen into this section and have never seen the network protocol stack based on the ISO/OSI model, then please stop, and take a look at Appendix A before proceeding.

Signal Encoding and Error Correction	Better encoding can improve data rate and correction can save valueable time.

HIGHER LAYERS (APPPLICATION, PRESENTATION, SESSION),
- Data Applications (e-Mail, Web Browsing, M-Business Apps
- Voice and Video Applications (Phone Calls, VOIP, Multimedia)

TRANSPORT LAYER
- End-to-End Message Delivery
- End-to-End Error Control

NETWORK LAYER
- Addressing and Routing
- Routing for Mobile Units (Location Management)

DATA LINK LAYER
- Error Detection and Correction
- Contention Management (Multiple Access Methods)
- Power Management and Control

PHYSICAL LAYER
- Frequency Allocations
- Transmission and Propagation (Fading, Scattering, etc.)
- Signal Encoding (Modulation/Demodulation)

Figure 5-3: Wireless Network Protocol Stack at a Glance

Why are Data Rates of Wireless Networks Increasing So Rapidly?

It is obvious to even the most casual observers that the data rates of wireless networks are improving rapidly. As we have seen in previous discussions, wireless LANs and wireless local loops are reaching the 50 to 100 Mbps range. Even the cellular networks have jumped from 1 Kbps to 2 Mbps since the mid-1980s. What is the reason? It turns out that there is not one breakthrough. Instead, several improvements in the following areas are being combined to deliver significant improvements in the data rates:

- **Use of digital transmissions**. Digital transmissions instead of analog transmissions are much more reliable and less error-prone.
- **Improvements in error correction.** Significant improvements have been made in correcting the transmission errors encountered in wireless. Simply stated, if a wireless transmission keeps encountering errors and re-transmitting, the data rates take a nosedive. Improvements in error correction, called Turbo Codes (discussed later in this chapter) have done wonders in this area.
- **Improvements in signal encoding and multiplexing**. Numerous improvements over the years have resulted in sending multiple bits per signal, thus increasing the data rate. For example, if you can send 4 bits on one signal, then the same transmission system will have 4 times the data rate.
- **Using packet switching.** Packet-switching systems break up a message into multiple packets and then send the packets on multiple parallel paths. For example, if a packet is broken into 5 packets and those packets are sent on 5 different paths, then the data rate can be improved, theoretically, 5 times.

Time to Take a Break
- How Do Wireless Networks Work
- Frequency Allocation and Location Services
- Physical Communications Considerations

.

5.3 Wireless Frequency Allocations and Regulations

5.3.1 Overview

Wireless communications use the "***radio frequency (RF)***" spectrum for transmitting and receiving information. The radio spectrum, shown in Figure 5-5, in fact represents the various frequencies supported by air and is a natural physical phenomenon. This spectrum includes radio waves, radar waves, microwaves, infrared waves, and X-rays. As we will see later, there is a correlation between wavelength and distance – higher frequencies do not travel far, due to media resistance (see Section 5.3.8). Only part of the radio frequency spectrum is available for telecommunications because some frequencies, especially in the lower frequency ranges, are consumed by other contenders such as ham radios. To accommodate telecommunications, there is a push to broaden the air frequency spectrum to higher frequency ranges. At present, telecommunications is considered to be technically feasible at frequencies as high as 300 GHz but many wireless systems, as we will see, operate in the 0.5 to 2 GHz zone.

Several factors are usually considered while allocating frequencies for wireless. First, the cost of components increases as you go to higher frequencies. Second, signal losses also increase as frequencies increase. These two factors are unfavorable to higher frequency ranges. But the third factor favors higher frequencies: communication systems at lower frequencies are disrupted regularly by man-made noise such as electrical motors, car ignition, and domestic appliances. Yet another factor is the bandwidth needed by each user. For example, each TV channel uses 6 MHz. As you go to higher data rates for end-users, higher frequency ranges are needed. Wireless carriers take these factors into account to determine an appropriate frequency range. The common frequency range at present, as stated previously, is in the 0.5 to 2 GHz range. Despite the problems with interference and noise, the lower frequency bands are used more often, for two main reasons: there is less loss on lower frequencies and it is cheaper to produce lower frequency waves (see Section 5.3.9).

Basic Terms

Signals, wireless or not, are represented as cyclic waves which may be discrete (digital) or continuous (analog). The following properties characterize cyclic waves (see the following figure):

- ***Frequency*** which shows the cycles per unit time of the wave
- ***Amplitude*** which shows the height of the wave
- ***Phase*** which shows how far, in degrees, the wave is from its beginning (phase 0).

The following communication terms are used widely:

- ***Hertz (Hz)*** = number of cycles per second. Frequency is measured in Hertz.
- ***Baud*** = number of signal changes per second
- ***Data rate*** = number of bits sent per second (bps). Data rate is equal to the baud rate if one bit is carried per signal. In general, data rate = baud rate x data bits per signal.
- ***Channel*** = a logical communication path. One physical wire can support multiple channels; each channel supports one user.
- ***Bandwidth*** = frequency range used by a signal, measured in Hz.
- ***Channel capacity*** = number of bits that can be transmitted per second. This is the same as data rate.

See Appendix A for additional discussion.

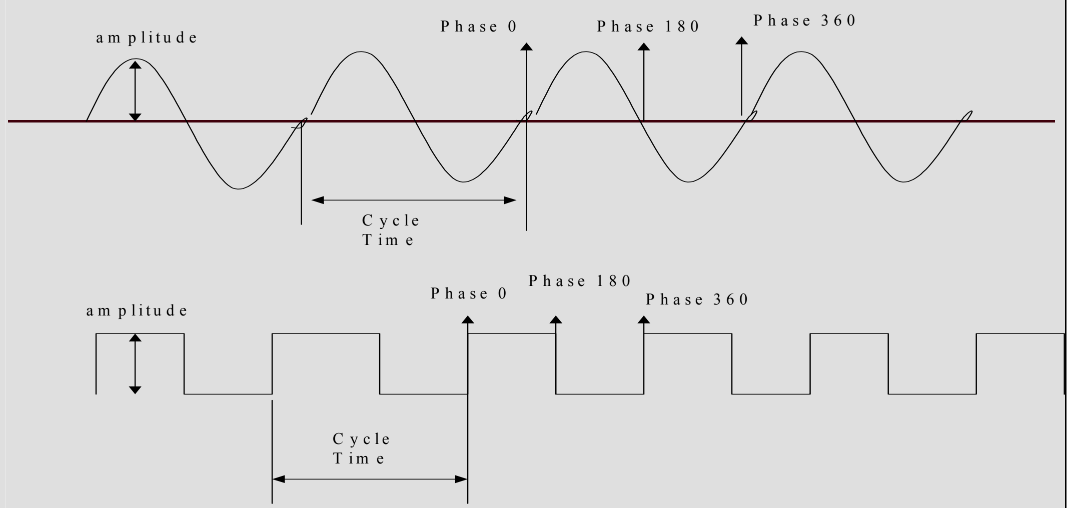


Figure 5-4: Cycle Wave Characteristics

5.3.2 Classifications of Transmission Media

A transmission medium is the physical path between the transmitter and receiver. The transmission medium is called ***guided media*** if the waves are guided along a solid medium, e.g., copper twisted pair, copper coaxial cable, or optical fiber. For example, the waves have to bend whenever and wherever the coaxial cable is bent around a path. Unguided media does not guide electromagnetic signals and is used commonly in wireless transmissions such as atmosphere and outer space. In the unguided media, transmission and reception are achieved by using antennas. We will look at antennas later. Wireless transmissions can be configured for:

- Directional transmission, i.e., the wireless communication is point-to-point. For example, a microwave transmitter can direct all of its signals to a receiver across a street.
- Omnidirectional transmission, i.e., the waves are transmitted equally in all directions. A radio station transmitter is an example.

5.3.3 General Frequency Ranges

Although radio frequencies can range from 0 to 500 GHZ or higher, the ITU (International Telecommunication Union) concentrates on radio frequencies that range from 9 KHz to 400 GHZ (see Figure 5-5). Within this range, the following are broad categories:

- Broadcast radio frequency range (30 MHz to 1 GHz). This is suitable for omnidirectional applications such as general radio.
- Microwave frequency range (1 GHz to 40 GHz). This frequency range is suitable for directional beams and is used for point-to-point transmission in several applications such as satellite communications.
- Infrared frequency range (roughly $3x10^{11}$ to $2x10^{14}$ Hz), useful in local point-to-point multipoint applications within confined areas.

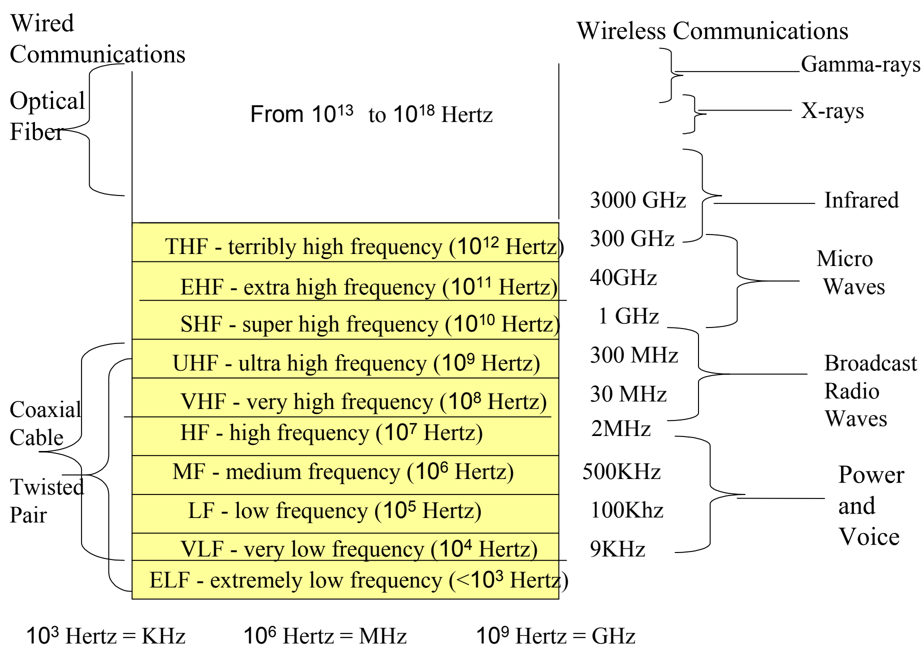


Figure 5-5: Radio Frequency Spectrum

We should also point out that radio waves are often referred to as radio carriers because they simply perform the function of delivering energy to a remote receiver. The data being transmitted is superimposed on the radio carrier so that it can be accurately extracted at the receiving end. This is generally referred to as ***modulation*** of the carrier by the information being transmitted. To extract data, a radio receiver tunes in (or selects) one radio frequency while rejecting all other radio signals on different frequencies. Modulation/demodulation techniques are discussed in Appendix A.

There are three major frequency bands of interest:

- Broadcast radio
- Microwave
- Infrared

5.3.4 Broadcast Radio (30 MHz to 1 GHz) Frequency Band

This frequency range is used in broadcast radio and consists of VHF and part of the UHF band that goes from 30 MHZ to 1GHz. Typical applications in this frequency range include FM radio and UHF and VHF television.

This is a congested frequency band and many new applications are emerging in this band. Due to new applications, the frequency range is expanded frequently.

Due to the low frequency range, the antennas in this band are omnidirectional, i.e., they broadcast in every direction. As we will see, this is the main differentiator from microwave communications in which the communication is directional. The antennas for broadcast radio are not required to be dish-shaped and need not be rigidly mounted to a precise alignment.

5.3.5 Microwave Frequency Band

This frequency band is frequently used in satellite and terrestrial communications.

5.3.5.1 Satellite Microwave (1 GHz to 20 GHz, typically)

A satellite is a microwave relay station that is used to link two or more ground-based microwave transmitters/receivers. The satellite receives transmissions on one frequency band (uplink), amplifies or repeats the signal, and transmits it on another frequency (downlink).

Satellites are used in a variety of ways. For example, television distribution from TV content providers consists of satellites that transmit from the content provider to the local cable TV station, which then transmits to your house over cable. PBS, for example, uses satellites exclusively to distribute content. Satellites are also used for long-distance telephone transmission between telephone exchange offices. In addition, several large private businesses such as IBM set up networks through satellites and then lease channels (these are expensive.)

5.3.5.2 Terrestrial Microwave (20 GHz to 40 GHz)

These wireless communications most commonly use a parabolic "dish," typically 3 meters in diameter. The dish is fixed rigidly and focuses on a narrow beam of transmission. The main characteristic is that this type of communication achieves line-of-sight transmission to the receiving antenna (relays are used in between). To facilitate line-of-sight transmission over longer distances, the dishes are located at substantial heights above ground level.

Terrestrial microwaves are commonly used in long-haul telecommunications service, i.e., instead of fiber or coax cables. The main advantage of microwave is that it requires fewer repeaters; but its limitation is that it only works on line of sight. Thus it is used commonly for short point-to-point links between buildings (e.g., closed circuit TV, wireless LANs, and as bypasses to local telephone companies – a business can directly send its information to a local control office without the local telephone company).

The most common frequency band used in terrestrial microwave is 4 GHZ. This can deliver up to 200 Mbps to end users.

5.3.6 Infrared Frequency Band

These waves operate in the terribly high frequency (THF) range that covers 300 GHz to 3000 GHz. The most significant property of infrared waves is that they do not penetrate solid objects such as walls. Due to this limitation they are used in point-of-sight (POS) applications such as TV remote control, garage door openers and other remote control devices.

Commonly Used Frequencies in Wireless Systems

- Cellular networks: Mostly around 900 MHz
- IEE 802.11 LANs: 2.4 GHz (802.11b, 802.11g) and 5 GHz (802.11a)
- Satellite systems: 3 to 30 GHz
- Wireless local loops: 10 to 100 GHZ
- Infrared wireless LANs; 300 GHz to 400 THz

5.3.7 Licensed versus Unregulated Frequency Bands

There are basically two approaches in using wireless frequencies: use an unlicensed band or use a frequency that is regulated. The following two unlicensed bands in the US have been defined (counterparts exist in other countries):

- **Industrial, Science, and Medicine (ISM):** This is in the 2.4 GHz range. The ISM band is very heavily crowded at present. In fact, most wireless LANs such as Bluetooth and 802.11 use these unlicensed bands.
- **Unlicensed National Information Infrastructure (UNII):** This is in the 5.2 GHz range. Initially set aside in 1997 for wireless Internet access, the UNII band is relatively free at the time of this writing. One of the reasons is that signals in this band do not travel as far as the lower-frequency ISM band (see the next section for explanation).

Most other frequency bands are regulated by different agencies (see 5.3.10). There are some tradeoffs in using unregulated versus regulated frequency bands. Naturally, the unregulated bands give the suppliers freedom, but then everyone can use these bands. Thus the issues of privacy and interference from other players are more serious. Regulated bands require permission but tend to reduce the interference from other suppliers.

5.3.8 Relationship Between Wireless Frequency and Distance Covered

A very important relationship exists between wireless frequency and distance covered. The relationship is:

$$d = k/f$$

where d = distance covered, f = frequency used, and k = constant that depends on environmental factors. Thus, the distance covered is inversely proportional to the frequency being used. This implies that higher the frequency, the shorter is the distance covered. Consider, for example, two systems; one operates at 1 GHz while the other at 30 GHz. Then the one at 1 GHZ can travel farther than the one at 30 GHz. The underlying reason is found in physics; the higher frequencies have lower wavelengths and hence encounter more resistance from the surrounding medium. This has several important implications in wireless systems:

- Xrays and infrared waves (very high frequency) do not travel far, while lower frequencies are typically used for long range transmissions.
- Lower frequencies are more congested and highly competed for because the distance can be longer and the power requirements are lower. Almost all providers want to transmit at lower frequency ranges. This has led to frequency auctions (see the sidebar “Wireless Frequency Auctions”).
- Many new developments such as 4G wireless are hampered by the lack of availability in lower frequencies. Some experiments of 4G have been done at 60 GHz due to a lack of lower frequencies. But at such high frequencies, the signals travel only around 100 meters. In this scenario, millions of cells would be needed to provide cellular service in a large city – not a very entertaining thought.

Frequency Allocations for Space Exploration

The first manned spaceflight, was accomplished by the USSR in 1957. This was accomplished by Russia's Sputnik that transmitted at frequencies of 20.07 MHz and 40.002 MHz radiated by four whip antennas on the craft's exterior. These frequencies were selected in spite of a ruling by the International Geophysical Year (http://www.nas.edu/history/igy/) that any satellite placed in orbit should transmit at a frequency of 108 MHz. It is not clear why such low frequencies were chosen by Russia. Some believe that 20 MHz was chosen so the amateur ham radio operators around the world could see its transmission, providing the USSR with free worldwide updates as to Sputnik's location.

As part of the US space program that landed man on the moon, the crew of Apollo 11 and the other Apollo missions, communicated with earth via a 2.3 GHz radio band known as the S-band. The S-band was part of NASA's Deep Space Network, which was established to handle the communication needs of all deep space missions (see Chapter 9 for details of Deep Space Network).
As an interesting side note, some of the unmanned spacecraft launched in that era are still functioning and communicating with Earth even today. For example, the U.S. Pioneer 6 and Pioneer 8 spacecraft that were launched around 1965, still transmit home from their solar orbits. This is a tribute to the early space age engineers,

5.3.9 Why Low Frequencies are More Congested

Lower frequencies are usually much more congested and contested than higher ones. In fact, within a frequency band, the lower spectrum of frequencies is used more often than higher ones. For example, within the microwave spectrum (1 GHz to 40 GHz), the lower end of frequencies (5-6 GHz) are used more frequently. Similarly, between the two unlicensed bands, the lower (ISM) is much more congested than the higher (UNII) band. There are two main reasons for this:

- Lower frequencies suffer less loss and travel farther. The loss is proportional to (d/λ), where d is distance covered and λ is the wavelength (see the formula for loss below). Thus given the same distance, loss is proportional to frequency (recall that frequency = 1/λ).
- It is cheaper to produce lower-frequency waves. Very simple antennas of your cellular phone or laptop can generate signals in the lower frequency. But you need powerful antennas for higher-frequency waves.

The main disadvantage of lower frequencies is that they suffer interference from nature (e.g., rain, thunderstorms) as well as other neighbors who also insist on using the same frequency ranges.

Formula for Loss:

$$\frac{P_t}{P_r} = \frac{(4\pi d)^2}{\lambda^2} = \frac{(4\pi f d)^2}{c^2}$$

- P_t = signal power at transmitting antenna
- P_r = signal power at receiving antenna
- λ = carrier wavelength in meters
- f = carrier frequency = 1/λ

- d = propagation distance between antennas (in meters)
- c = speed of light

For this equation to work, d and λ have to be in the same units (e.g., meters). This equation is very useful in calculating signal loss for given wavelength and distance, determining the distance between two antennas for a given wavelength and power loss, etc.

5.3.10 Frequency Regulations

Low frequencies from 9KHz to 300 MHZ are in high demand (especially VHF: 30-300 MHZ) because they can be supported without expensive antennas and can travel far, as discussed previously. The lower frequencies are used frequently for broadcast radio and are thus more in demand. In addition, cellular phones also use lower frequencies (around 900 MHz).

Due to this popularity, regional, national, and international issues in regulating frequencies occur more commonly concerning lower frequencies. In fact, lower frequencies are auctioned in several places, especially in Europe. In particular, auctions for 3G cellular frequency licenses have been very lucrative. For example, in some countries, interested parties have bid in excess of US $30 billion for 3G bands. In the summer of 2000 alone, European countries spent in excess $190 billion to buy the spectrum needed to run 3G phone services over the European continent (see the sidebar "Wireless Frequency Auctions – The 3G Windfall").

There are different types of procedures for military, emergency, air traffic control, etc. Some frequency bands are reserved for these uses. Different agencies, as discussed in Chapter 1, license and regulate the frequency allocations. Examples are:

- Federal Communications Commission (www.fcc.gov) that regulates frequency allocations in the US
- The British www.open.gov.uk/radiocom that is used in the UK
- European Telecommunications Standards Institute (ETSI) for European countries
- Radio Communications Bureaus to handle frequency interference across national borders.
- ITU (International Telecom Union). Headquartered in Geneva, consisting of several sectors such as ITU-R (radio communications), ITU-T (standards), and ITU-D. (Development).
- The National Telecommunications and Information Administration (NTIA – www.ntia.gov), is part of the United States Commerce Department.

A good discussion of different regulatory bodies and their role in frequency allocations can be found in Bekkers [2000].

Wireless Frequency Auctions – The 3G Windfall

Many auctions are held around the globe, especially in Europe, to sell available wireless frequency bands. See, for example, the FCC site (http://wireless.fcc.gov/auctions/data/bandplans.html).

Some "entrepreneurs" buy a bunch of desirable wireless frequencies in the low frequency range and then sell them to the highest bidder. This practice is similar to the Internet domain name selling where some people buy a bunch of appealing domain names and then sell them for profit. Very clever.

Auctions for 3G cellular frequency licenses have been very active around the globe because in order to provide 3G phone services, the providers must own the license to operate in those frequency ranges. As mentioned previously, some companies have bid for 3G bands in billions of US dollars. Consequently, one of the biggest costs for the wireless carriers in moving toward 3G networks has been the bidding on spectrum auctions. With this much money invested in 3G networks, the wireless carriers have to come up with new business models, products and services to attempt to recoup these huge expenditures.

While wireless carriers have spent a lot of money, many countries have made a great deal of money by auctioning their 3G frequency bands. For example, Britain raised $32.2 billion (22 billion pounds) from five operators after seven weeks of bidding n 2000 when the 3G frequencies were heavily sought after. The German auction raised $37 billion from six operators after two weeks of bidding in September 2000. According to 3G Newsroom (http://www.3gnewsroom. com), five countries (Germany, Britain, USA, Italy, and South Korea) raised a total of over $100 billion.

While many countries made a great deal of money, some took a different approach. For example, Singapore offered competitive rates for its 3G licenses starting at only $87 million, which was later lowered to $58 million. The primary driver for this surprisingly low price was that Singapore understood the softening of the demand for 3G. It also felt that many companies were hesitant to invest large sums for 3G and might decide not to bid at all if the price were too high. This lower price offering was met enthusiastically by the Singapore telecom industries and resulted in an accelerated adoption of 3G in Singapore. Similarly, other countries such as Norway and Finland have adopted a beauty content approach and have given away spectrum licenses for free or for reduced prices to accelerate the adoption of new cellular networks.

Sources:

http://www.3gnewsroom.Com

http://www.med.govt.nz/pbt/auctions/conclusionpt1.html

http://www.fool.com/news/2000/german3g000816.htm?source=EDSTRM

http://www.3gnewsroom.com/3g_news/jan_01/news_0122.shtml

Evans, N. *Business Agility: Strategies For Gaining Competitive Advantage Through Mobile Business Solutions.* Prentice Hall, 2001.

5.4 Location Management in Wireless Networks

5.4.1 Main Drivers and Approaches

Wireless systems have to keep track of the users as they move around. The basic reason for this is rather obvious: the cellular network needs to know where you are if it has to direct a call to you. In addition to cellular networks, WLANs also need to know if you move from one LAN cell to another. But besides this obvious reason, there are other drivers for location management:

Commercial drivers: Location-sensitive information can be bundled into mobile commerce and other application as a value-added service. Mobile operators are expecting significant revenues by offering a number of location-based services such as positional commerce. Most applications in this area require between 100- to 500-meter accuracy.

Regulatory drivers: Regulatory agencies are requiring mobile operators to provide accurate locations for emergency purposes and public safety. For example, the US FCC has mandated support for E911 services – this requires that when a mobile user calls 911, the mobile operator should be able to automatically locate the caller with high accuracy (100 meters for 67 percent of calls, 300 meters for 95 percent of calls). See the sidebar "E911 Services for Cellular Providers." There is a similar European Union requirement for E-112 emergency support.

Technological drivers: Deployment of 2.5G and 3G networks and better handsets such as Communicators and Media Phones enable new services that could include location as a value-added service. Availability of high-speed networks with high-resolution terminals are driving developments of new location-based services.

To meet these demands, a wide range of techniques for location management have been introduced. Examples of the techniques include:

- **Cell ID-based location**. This is the oldest and by far the most commonly used technique. In this case, a mobile user is assigned an ID of the cell that the user is in. This Cell ID is stored in a database. As the user moves from one cell to another, the location database is updated accordingly.
- **Other location services**. Although Cell ID-based services work quite well, they are not very accurate because they can only tell which cell a user is in. This is not acceptable for the E911 mandate because a cell could be 10 miles wide and it is not enough to know that you are somewhere in this 10 mile area. Other techniques include Assisted Global Positioning System (AGPS), Angle of Arrival (AOA), and variants of Time of Arrival (TOA).

These techniques yield different location accuracies (typical ranges are 50 meters to several kilometers). The choice depends on the accuracy needed and the ease of implementation and use, among other factors. Another consideration is that the location services have to perform well because there is a tradeoff between mobility and data rate (see the sidebar "Mobility Versus Data Rate"). In the meantime, cell-ID-based location services continue to thrive.

Mobility Versus Data Rate

There is a general tradeoff between data rate and mobility of the users. In general, high mobility means lower data rates. For example, 3G networks provide data rates that range from 300 Kbps to 2 Mbps. The lower data rate (around 500Kbps) is for high-mobility cellular services (the ones who want to use their phones while traveling in cars and trains). The higher data rate (around 1 Mbps) is for low-mobility services (walking within a building, for example).

The technical reason for this tradeoff is very simple. High mobility services need frequent handoffs and update of location services – a slow process. For example, if you are using your cellular phone while traveling in a train, then you are going through many cells quickly. Take, for example, travel from Philadelphia to New York (a distance of around 100 miles). With average cell size of about 3 miles, you will go through more than 30 cells. Thus you cannot expect high data rates with so many handoffs and updates of location services.

Another issue that impacts the data rate of highly mobile systems is the fading encountered due to scattering, reflection, and diffraction (Section 5.5.7). Fading and other impairments require extensive error detection and correction schemes that tend to slow down the data rates.

5.4.2 Cell-based Location Services

The Cell ID-based location services were developed in Europe to support the GSM cellular networks but have been generalized since then**.** Figure 5-6 shows the basic idea. The cellular network is comprised of many ***"cells,"*** which typically cover 5 to 20 miles in area. The users communicate within a cell through wireless communications. A ***Base Transceiver Station (BTS***), also known as a ***Base Station (BS)***, is used by the mobile units in each cell by using wireless communications. One BTS is assigned to each cell. Regular cable communication channels can be used to connect the BTSs to the ***Mobile Switching Center (MSC),*** also known as the ***Mobile Telecommunications Switching Office (MTSO)***. The MSC is the heart of cellular networks – it determines the destination of the call received from a BTS and routes it to a proper destination, either by sending it to another BTS or to a regular telephone network. Keep in mind that the communications are wireless within a cell only. The bulk of cell-to-cell communication is carried through regular telephone lines. Although different techniques can be used, for example polling the neighboring cells to see where you are, the most common technique relies on the MSC to use two databases called ***Home Location Register (HLR)*** and ***Visitor location Register (VLR)*** to locate the mobile users.

The cellular system uses Cell ID to keep track of the cellular user by coordinating the home location register (HLR) and visiting location register databases (HLR). Here is a simplified view. Suppose you live in Morristown (New Jersey); your HLR will indicate that Morristown is your home town and the Morristown VLR will have an entry for you because you make most of your calls, one assumes, from Morristown. Every time you turn on your cellular phone in Morristown, the MSC verifies that you are still in Morristown, so no location updates take place. Let us assume that you are now driving from New York to Philadelphia. Upon arrival in Philadelphia, you turn on your cellular phone to call home. Basically, now the HLR and VLR will be updated to indicate that you are in Philadelphia so that the incoming calls can be routed to you. We will take a closer look at the actual mechanics of this system in Chapter 8.

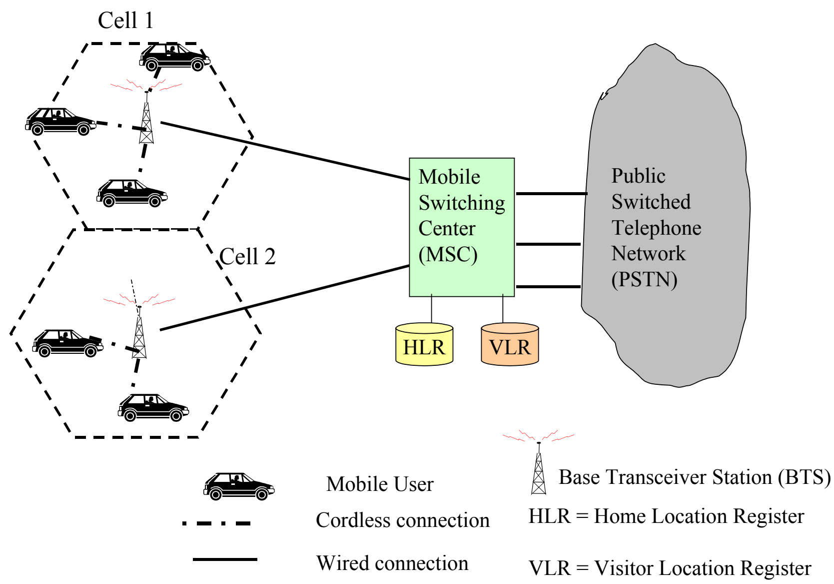


Figure 5-6: A Cellular Communication Network

5.4.3 Other Approaches

Many other positioning technologies have been developed and continue to be developed at the time of this writing. Here is a quick synopsis of the best known. Many of these methods utilize the same mathematical triangulation principles to calculate locations by measuring the difference in time it takes for a signal to reach a set of known objects.

5.4.3.1 Angle of Arrival (AOA).

In this case, the angle at which radio waves from your device "attack" an antenna is used to calculate the location of the device. This is somewhat similar to police work where they can tell the location of a suspect by studying how the bullet entered a body (a horrible but illustrative example). In reality the angle of arrival is observed at multiple (about 3) base stations. Knowing the arrival angle at the three base stations, the location can be calculated.

5.4.3.2 Network Assisted Global Positioning System (AGPS).

In this case, a GPS chip is installed inside a phone to track the location of the user. The mobile phone is also equipped with an antenna that receives signals being sent from the US's GPS satellites. The signals originate from the satellites at precisely the same time and propagate toward earth at the speed of light. When the mobile user device receives the signal, it calculates its location based on the difference in time it takes for each of the satellites' signals to reach the mobile user's special phone.

5.4.3.3 Time of Arrival (TOA)

In this case, the time taken between the device and the antenna is used to calculate the location of the device. Instead of time taken, some approaches use power dissipated. For example, these techniques compare the transmitted power from a BTS versus received power on the handset to estimate the distance and location. We will mainly discuss time of arrival.

The following two approaches are best known for TOA:

Enhanced Observed Time Difference (EOTD). This approach uses the same principle as AGPSs, but instead of a signal originating from a satellite, signals are sent from cellular base stations and are then picked up by specialized mobile phones that receive the signals and calculate a location. EOTD is dependent on the visibility of base stations. In both AGPS and EOTD, the location is calculated in the handset.

Uplink Time Difference of Arrival (UTDOA). UTDOA is the opposite of EOTD and AGPS in that the mobile device is the source of the signal, and signal detection and calculation is done by equipment located in a wireless operator's base station. With UTDOA the mobile device does not need to be modified for the system to work, meaning it operates with all existing mobile phones and devices. Also it has a distinct advantage in that the signal processing done at the base station is much powerful and better, owing to the availability of processing resources at the base station.

E911 Services for Cellular Providers

The FCC has mandated that all cellular providers offer accurate location information for the users of E911 locations. The basic requirement is that the cellular providers should be able to locate within, "for network-based solutions: 100 meters for 67 percent of calls, 300 meters for 95 percent of calls." Although the cellular providers are reluctant to spend the extra money to adhere to this new FCC regulation, there is a growing public demand for adherence to this regulation. Thus they have no choice. Although many other commercial drivers are pushing the development of location services, support for E911 is by far the biggest driver for highly accurate location services.

AGPS is one of the most promising technologies for supporting E911. However, AGPS requires additional capabilities on the handset. These features increase the weight of the phone as well as draining the limited battery power. In addition, older models of phones cannot take advantage of the new GPS systems. Other methods, such as UTDOA and EOTD, use triangulation methods to determine the location of a cell phone user by the time it takes their signal to reach multiple towers located at different places. However, difficulties arise when the cell phone user is traveling in a vehicle between cell towers. Special algorithms are needed to counteract the problem of moving targets.

References:

http://contact.bellsouth.com/email/bbs/phase2/how911works.html

http://www.wirelessweek.com/article/CA4845

5.4.4 Tradeoffs between Different Location Services

Table 5-4 shows a comparison of the main location-based services in terms of two parameters: accuracy and mobile device capabilities needed. Other parameters can be added, if needed. Accuracy refers to the radius in which a location technology can pinpoint the location of a wireless phone. This is naturally the most important criterion. As stated previously, most commercial applications require between 100- to 500-meter accuracy, while emergency services such as E911 require less than 100-meter accuracy. As seen from Table 5-4, AGPS is the most accurate while Cell ID is least accurate. The variance in accuracy depends on the type of area being covered – unobstructed areas give higher accuracy than densely populated ones. With AGPS, for example, typical accuracies are in the 50-meter range but in unobstructed environments can be in the 10 – 20 meter range. The reason for increase in indoor and urban situations is that the signal processing power of the mobile unit is not powerful enough to detect highly attenuated signals from distant or blocked cell sites. Cell ID is the least accurate of all of the location technologies because accuracy is a function of cell sizes, i.e., how closely spaced cell towers are. Thus this approach is essentially useless in non-urban environments where cell sites may be many kilometers apart. In urban centers where cell sizes are small, accuracies can be between 250 and 500 meters.

Table 5-4: Comparison of Location Services

Location Service Approach	Accuracy	Mobile Device Capabilities Needed
Cell-ID Based	Depends on cell size (between 5 to 20 Km)	None – Network-Based
AOA (Angle of Arrival)	100 to 200 meters	None – Network Based
AGPS (Assisted Global Positioning System)	10 to 50 meters	Yes –Mobile Device Assisted
E OTD (Enhanced observed time difference)	100 to 150 meters	Yes – Mobile Device Assisted
UTDOA (Uplink time difference of arrival)	50 to 100 meters	None –Network Based

The capabilities needed by mobile devices to support a given service is also an important criterion. This depends on if a service is network- or mobile device-based. In a network-based service, the mobile device does not need any capabilities, but in a mobile device-based service, the device needs a special card and associated software. Both AGPS and EOTD require special handset capabilities to calculate a location, limiting the service to only those users who are willing to upgrade or replace their handsets. Network-based technologies such as UTDOA and Cell ID locate all existing and future phones on the operators' network. Because these services can calculate locations without any enhancement to the mobile device, the devices can be a simple wireless phone. Both AGPS and EOTD require additional components in the phone or device and thus increase the overall cost of the device.

To summarize, UTDOA is suited for applications that require high accuracy and availability in indoor, urban and suburban areas; EOTD is suited for suburban applications requiring moderate accuracy; AGPS is suited for applications requiring high accuracy in suburban and rural environments; and Cell ID is best for low-accuracy services in dense urban areas. Beyond performance, the ability of UTDOA to combine good performance with the ability to locate all handsets makes it an appealing choice. If highly accurate location (10 to 20 meters) is the main requirement, then AGPS is definitely the winner. However, the following needs to be considered for AGPS:

- GPS has a long Time To First Fix (initial time taken to locate a user), which ranges between 40 and 60 seconds. This could be too long in an urgent situation. This problem is overcome with AGPS as the assistance data, which are sent to the handset, significantly reduces this time to less than 5 seconds.
- AGPS begins to suffer in densely populated areas because the GPS satellite signal becomes extremely weak after traveling thousands of miles before it reaches the earth. Further attenuations due to buildings and signs may kill the GPS signal. Although improvements have been made to overcome this issue, in many highly dense areas, AGPS systems simply cannot calculate the distance of a location.

Finally, in rural environments, all of the location technologies mentioned show nearly perfect accuracy. However, as population density increases along with physical obstructions such as buildings and signs, accuracy decreases – especially with AGPS, due to its weak signals. This is not an issue with the other location technologies as their close proximity to the signal source allows locations to be processed in all environments. Developments in smart antennas, discussed later, can help improve the location services because the signals of such antennas follow the users as they move around – thus providing powerful signals to support location-based services.

Sources of Information for Location Based Services

Jagoe, A. *Mobile Location Services: The Definitive Guide*. Prentice Hall, 2002.

Turner, A. *Mobile Location Services*. McGraw Hill, 2002.

http://www.palowireless.com/lbs/

http://www.trimble.com/gps/why.html

http://www.gps-practice-and-fun.com/index.html

http://contact.bellsouth.com/email/bbs/phase2/how911works.html

http://www.nokia.com/pc_files_wb2/mposition_mobile_location_services.pdf

Time to Take a Break

- ✓ How Do Wireless Networks Work
- ✓ Frequency Allocation and Location Services
- Physical Communications Considerations

5.5 Wireless Antennas and Propagation

5.5.1 Introduction

We already know that the wireless communication systems consist of three basic elements: transmitters that generate and send signals, antennas that radiate the electromagnetic energy generated by the signals into the air, and receivers that receive and process the signals (see Figure 5-7). This section, an abbreviation of [Stallings 2002, Chapter 5], introduces the basic ingredients of transmitters/receivers, antennas, and the propagation of wireless waves.

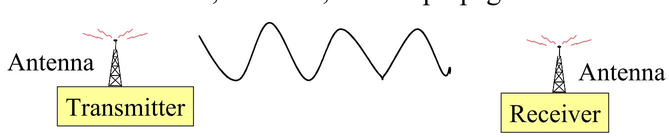


Figure 5-7: Wireless Communications

5.5.2 Transmitters/Receivers

Figure 5-8 shows a conceptual view of transmitters. A receiver consists of similar components and performs similar operations but in the reverse order. Let us briefly review the transmitters. Suppose you want to generate a signal that is sent at 900 MHz (a typical frequency for mobile units), and the original source, let us say a guitar player, generates a signal at 300 MHz. The amplifier strengthens the initial signal so that it is strong and clear. The oscillator creates a carrier wave of 600 MHz, and the mixer combines the original signal with the oscillator and produces a 900 MHz frequency signal. The purpose of the oscillator-mixer pair is to take the input signal and transmit it at the desired frequency. Thus the oscillator produces a frequency that fills the gap – it generates a 600 MHz signal that when mixed with 300 MHz produces the desired 900 MHz signal. Thus the mixer also provides modulation, etc.

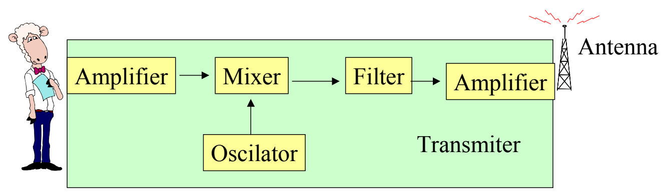


Figure 5-8: Transmitters

The filter, as the name implies, makes sure that only the licensed frequencies are passed along. Different frequency signals can be generated while mixing (known as "higher harmonics"). Filters – required by the FCC – assure that the radio stations, for example, only transmit at the licensed frequencies and do not interfere with others. Another amplifier is used to further strengthen the signal before sending it out to the "cruel and uncertain world" of wireless media where rain, thunderstorms, and other hostile things can happen to it. Receivers, as indicated above, perform similar operations but in the reverse direction.

5.5.3 Antennas

An antenna is an electrical conductor or system of conductors to send/receive RF (radio frequency) signals. Transmission radiates electromagnetic energy into space and reception collects electromagnetic energy from space. In two-way communication, the same antenna can be used for transmission and reception. The manner in which antennas distribute and collect energy from the surrounding space has a profound influence on the efficient use of frequency spectrum, the cost of establishing new networks, and the service quality provided by the wireless networks.

Power is radiated from an antenna in all directions but it may be stronger in one direction than others. Radiation patterns are the graphical representation of radiation properties of an antenna and are commonly used to characterize the performance of antennas. Reception patterns are the receiving antenna's equivalent to radiation patterns. Different types of antennas produce different radiation patterns. Furthermore, antennas can be coupled into groups or arrays. In such a system, different types of interference among the individual antenna elements can be combined to produce a strong radiation pattern. Intelligent antenna systems range from simple diversity techniques to fully adaptive antenna array systems. Radiation patterns broadly fall into two categories (Figure 5-9):

- **Isotropic** (omnidirectional) transmissions that radiate power equally in all directions. This happens in lower frequency ranges and is used in broadcast radio communications. For example, wireless access points in wireless LANs use omnidirectional antennas to communicate with mobile devices such as laptops. The unfocused radiation pattern of omnidirectional antennas reaches users with only a small fraction of the total radiated energy. While viable for simple RF environments where no specific knowledge of the users' location is available, this approach has many limitations. Omnidirectional antennas waste frequencies because they keep transmitting in all directions even if all users lie only in one direction. Omnidirectional strategies attempt to overcome environmental challenges by simply increasing transmit power. This is not a good solution because it generates greater interference for other users. These limitations have required the development of other types of antennas.
- **Directional** transmissions that radiate power in one direction. Directional transmissions are common in higher frequency ranges where preferential transmission and reception directions can be predefined between senders and receivers. This results in more power being focused in the preferred direction, and less power in other possibly interfering directions. These antennas have the natural limitation that if an object moves out of the directional beam, it cannot receive the signal. Common examples of directional antennas can be found in satellite communications.

Naturally, it is desirable to develop antennas that can provide directional services to follow the customers as they move around. This is done through smart antennas, discussed later.

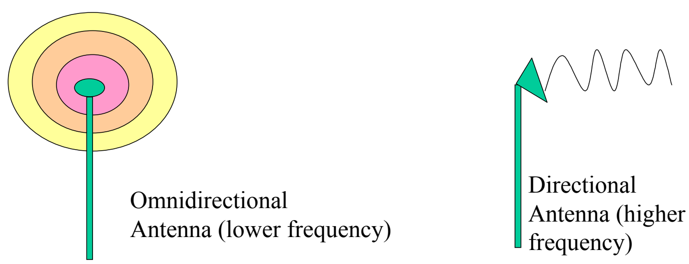


Figure 5-9: Antenna Types

Antenna gain is a measure of directionality of antennas. It simply represents the power output, in a particular direction, compared to that produced in any direction by a perfect omnidirectional antenna with isotropic radiation patterns. For example, a gain of 1 indicates that the antenna is ominidirectional, while a gain of 3 in one direction indicates that the directionality of an antenna is 3 times greater than that of an omnidirectional antenna. A gain of 10 or higher requires very heavy directional and focused beams. The gain depends on wavelength and the physical size and shape of the antenna.

Very Small Antennas

Antennas can be extremely small, almost invisible in some cases. For example, the RFID (radio frequency identification) tags have a chip and an antenna. The RFID tag is I very thin plastic label somewhat similar to the commonly used bar code label. The antenna on an RFID tag is so small that it can be printed on the tag with carbon-based inks. RFID tags (chip plus antenna) are also called "transponders". Due to the small antennas, they are not very powerful – their range is usually a few centimeters. Thus, the RFID reader (the receiver) has to be very close to the tag to read it.

5.5.4 Smart Antennas

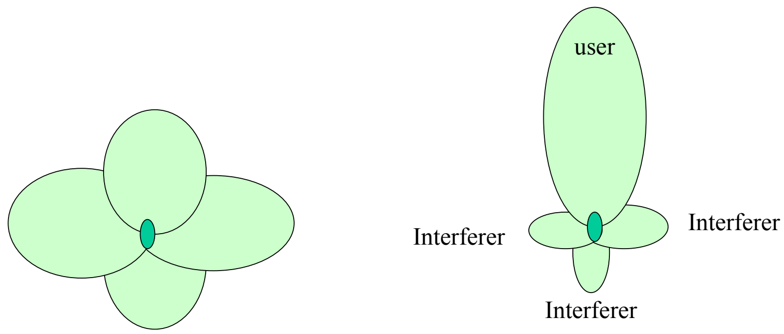
The basic idea of a smart antenna is to propagate signals to follow objects as they move around or to choose the transmission directions that minimize noise. A smart antenna system combines multiple antenna elements with processing capabilities to automatically optimize its radiation/reception pattern to adjust to the environment. Generally co-located with a base station, a smart antenna system intelligently uses an antenna array (many antenna elements) to transmit and receive in an adaptive manner. In other words, such a system can automatically change the directionality of its radiation pattern to follow user populations as they move around. This can dramatically increase the performance characteristics (such as capacity) of a wireless system. Another main benefit of smart antennas is improvement in transmission quality because they can increase desired signal power and reduce interference as needed. Smart antennas can also be used to improve the location services discussed previously, because they follow the users by determining the angle of arrival. Thus, smart antennas can be used for fulfillment of the FCC requirements of E911 and to offer location-based services (i.e., positional commerce).

Smart antennas attempt to increase the antenna gain according to the location of the user and are based on a mixture of:

- Switched-beam systems with a number of fixed beams at an antenna site – the beam with least interference and best signal strength is chosen.
- Adaptive antennas with an array of antennas that can adjust patterns based on noise, interference, and location of objects

A switched-beam antenna system has a fixed set of beam patterns with a highly directional main beam. These antenna systems detect signal strength, choose from one of several predetermined fixed beams and switch from one beam to another as the mobile device moves around a cell. These systems are very easy to build because the beams already exist – you only need to switch from one to another. Switched-beam systems offer many of the advantages of more sophisticated smart antenna systems at a fraction of the complexity and expense. An example of the coverage pattern is given in Figure 5-10a.

The main limitation of these systems is that the beams of the antenna system can only be chosen from already available positions – new and arbitrary directions cannot be created on the fly. Another disadvantage is that the subscriber's signal strength varies as the user moves from the center of the beam to the edge of the coverage region of that beam. Despite these limitations, switched beam systems are popular because of their low cost and the few interactions needed with the base station receiver. Thus they do not add extra burden to the base stations.



a) Coverage Pattern (Top View) for a Switched Beam Antenna with 4 Elements

b) Coverage Pattern (Top View) for an Adaptive Antenna giving Preferential Treatment for the User and Minimizing the Interferers

Figure 5-10: Two Types of Smart Antennas (Switched Beam and Adaptive Array)

Adaptive antennas are based on a more sophisticated approach. They use a variety of new signal-processing algorithms to locate and track various types of signals of the moving devices to minimize interference and maximize signal strength. As compared to the switched beam system, the adaptive system provides optimal gain while tracking and minimizing interfering signals, as shown in Figure 5-10b. The ability to track users smoothly with main lobes and "intruders" (interferers) with nulls ensures that the transmission link capacity is not wasted. Discussion of the actual algorithms used in adaptive systems can be found in [Liberti 1999].

There is a great deal of research and development activity in this area. See, for example, the book by J. Liberti and T. Rappaport, *Smart Antennas for Wireless Communications* (Prentice Hall, 1999). The following two publications also contain very useful information:

- W. Schuttengruber, A. Molisch, E. Bonek, "Smart Antennas for Mobile Communications," The Institute of Communications and Radio-Frequency Engineering (2001), http://www.nt.tuwien.ac.at/mobile/research/smart_antennas_tutorial/
- "Smart Antenna Systems," ArrayComm, International Engineering Consortium (2003), http://www.iec.org/online/tutorials/smart_ant/

5.5.5 Propagation of Wireless Waves

Radiations after leaving an antenna are propagated by using ground-wave, sky-wave, or line-of-sight propagation (see Figure 5-11). For most wireless systems, line-of-sight propagation is of main interest because ground-wave and sky-wave propagations operate at very low frequencies (between 2 to 30 MHz) while most wireless communication systems operate at 900 MHz and above. A quick overview of the three is provided for completeness. We will take a closer look at line-of-sight propagations later.

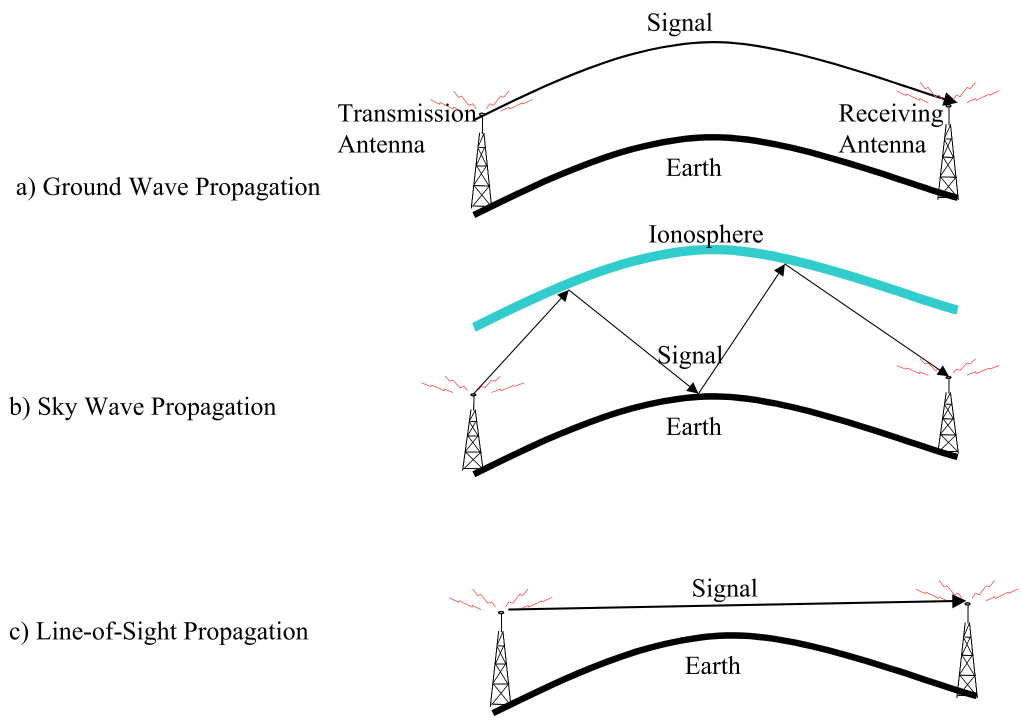


Figure 5-11: Propagation Modes

5.5.5.1 Ground-Wave Propagation

Ground-wave propagation, shown in Figure 5-11a, follows the contour of the earth. The waves can propagate over considerable distances by using this propagation mode. The frequencies in this mode are very low (below 2 MHz). At these low frequencies, the electromagnetic waves tilt towards the earth due to currents induced in the earth's surface, thus following the contour of the earth. A common example of ground-wave propagation is AM radio.

5.5.5.2 Sky-Wave Propagation

In this mode, the signal is reflected from the ionized layer of the atmosphere back down to earth, as shown in Figure 5-11b. The signal can travel a number of hops, back and forth between the ionosphere and earth's surface, and can travel far by just bouncing back and forth. The reflection effect is caused by ***refraction*** – a change of direction of electromagnetic waves due to change in media density. Refraction is caused by a bending of microwaves by the atmosphere. Basically, the velocity of electromagnetic waves is a function of the density of the medium – when a wave changes media, the speed also changes and the wave bends at the boundary between media. A common example of refraction is the bending of light waves as they go through different media. For example, if you insert a pencil in a glass of water, it looks bent. Sky-wave propagation for wireless systems works in higher frequencies (2 to 30 MHz) and is used commonly in amateur radio and CB radio. Other well-known applications are international radio services such as BBC and VOA (Voice of America).

5.5.5.3 Line-of-Sight (LOS) Propagation

For this propagation mode, the transmitting and receiving antennas must be within the line of sight (see Figure 5-11c). This mode works for microwave transmission and satellite communication at high frequency ranges (above 20 MHz). LOS propagations work because signals above 30 MHz are not reflected by the ionosphere and thus penetrate the ionosphere to reach satellites. However, *refraction* does cause bending of microwaves by the atmosphere. How? Velocity of electromagnetic waves is a function of the density of the medium; thus when a wave changes medium, its speed changes. Due to this, the wave bends at the boundary between mediums, resulting in refraction.

Line-of-Sight (LOS) is used very frequently in wireless communications. For example, cellular networks, wireless local loops, wireless LANs, and satellite communications all use LOS propagations (see Table 5-5). Several equations help communications engineers determine the distances between antennas (see Section 5.5.6). LOS propagations suffer from different types of losses due to distortions, noise and fading (see Section 5.5.7). Due to these losses, error correction is very important in LOS wireless communications (see Section 5.6).

Table 5-5: Summary of Wireless Frequency Ranges, Applications, and Propagations

Frequency Range	Type of Waves	Typical Applications	Propagation
< 2 MHz	Extremely Low to Medium Frequencies (Power and Voice Waves)	AM radio	Ground-wave propagation
2 MHz to 30 MHz	High Frequency (Broadcast Radio Waves)	Amateur radio, CB radio, and international radio services such as BBC and VOA (Voice of America).	Sky-wave propagation
30 MHz and 300 MHz	Very High Frequency (Broadcast Radio Waves)	VHF television, FM broadcast and two-way radio	Line-of-sight propagation
300 MHz to 3000 MHz	Ultra High Frequency	UHF television, cellular phones, wireless LANs (see Notes)	Line-of-sight propagation
3 to 30 GHZ	Super High Frequency (Microwaves)	Satellites, wireless local loops, terrestrial microwave links	Line-of-sight propagation
30 to 300 GHZ	Extremely High Frequency (Microwaves)	Wireless local loops, experimental links	Line-of-sight propagation
300 GHz to 400 THZ	Infrared	Infrared LANs	Line-of-sight propagation

Notes: Cellular phones typically operate around 900 MHz and Wireless LANs typically use the ISM band (2.4 GHz).

5.5.6 Line of Sight – a Closer Look

As shown in Table 5-5, line-of-sight propagation is used very heavily in common wireless applications, ranging from wireless LANs to satellites. How far can the LOS signal travel? The following equation shows the distance for line of sight with no intervening obstacles:

$$d = 3.57\sqrt{h}$$

Where

d = distance between antenna and horizon (km)

h = antenna height (m)

Thus, if there is an antenna 100 meters high, it can transmit up to approximately 35 km. However, other factors such as refraction (the bending of waves during propagation) modify the distance covered. The effective, or radio, line of sight is given by:

$$d = 3.57\sqrt{\mathrm{K}h}$$

Where K = adjustment factor to account for refraction. As a rule of thumb, K = 4/3. The following equation is used to calculate the maximum distance between two antennas for LOS propagation:

$$3.57\left(\sqrt{Kh_1} + \sqrt{Kh_2}\right)$$

Where

h_1 = height of antenna one

h_2 = height of antenna two

These equations can be used to determine the LOS distances for different situations.

5.5.7 Line-of-Sight Transmission Impairments

Line-of-sight propagations suffer a variety of impairments while traveling between the transmitters and receivers. Examples of these impairments are:

- Attenuation and attenuation distortion
- Free space loss
- Noise
- Other impairments such as atmospheric absorption

Attenuation. Strength of signal falls off with distance over transmission medium. Attenuation factors for unguided media are: a) received signal must have sufficient strength so that circuitry in the receiver can interpret the signal, b) signal must maintain a level sufficiently higher than noise to be received without error, and c) attenuation is greater at higher frequencies, causing distortion. For the first two factors, signals of high strength are used and amplified/repeated to regain strength. The general approach for the third factor is to use amplifiers that strengthen higher frequencies.

Free Space Loss. As a signal traverses the free space (air), it disperses in different directions and consequently loses strength. Free space loss is the primary reason for attenuation for satellite signals because the signal attenuates as it disperses over long distances. Even if no other loss occurs, free space loss weakens the signal because it is being spread over a larger and larger area.

Noise. Signals traveling through the air are distorted due to different types of noises such as crosstalk (unwanted coupling between signal paths such as signals from one user interfering with others) and impulse noise (irregular pulses or noise spikes typically caused by external electromagnetic disturbances such as lightning or faults and flaws in the communications system). Another type of noise, called thermal noise, causes distortions due to agitation of electrons (called white noise). Thermal noise is present in all electronic devices, and transmission media, and cannot be eliminated because it is a function of temperature. Thermal noise is particularly significant for satellite communication and causes weakness of satellite signals.

Other Impairments. There are other impairments such as atmospheric absorption (water vapor and oxygen contribute to attenuation) and multipath reflections due to obstacles that reflect and scatter signals so that multiple copies with varying delays are received. These impairments cause fading and attenuation, discussed next.

5.5.8 Fading in Mobile Environments

Fading, the variations of received signal due to the changes in transmission paths or media, presents many serious technical problems in wireless communications. As the mobile units travel through different environments surrounded by buildings, cars, and other objects, fading occurs frequently. The faded signals vary significantly from their original form due to the impairments. In particular, multi-path propagation, as illustrated in Figure 5-12, leads to the following types of impairments:

- Reflection – occurs when signal encounters large surfaces. The surface is large relative to the wavelength of the signals.
- Diffraction – occurs at the edge of an impenetrable body that is large compared to the wavelength of the radio wave.
- Scattering – occurs when incoming signal hits an object whose size is in the order of the wavelength of the signal or less.

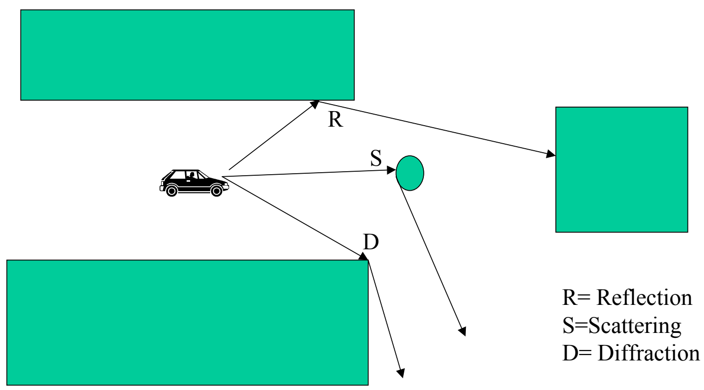


Figure 5-12: Multi-path Propagation

Different types of fading can occur due to a combination of these and other impairments. For example, *fast fading* happens when you are talking on a mobile phone while traveling in a busy downtown area. In this case, the radio signals heavily scatter, diffract, and reflect from different objects around you. However, if you are in an open field such as an empty parking lot, then the fading is slow due to decreased scattering, reflection, and diffraction.

5.5.9 Effect of Fading and Approaches to Deal With It

An unwanted effect of scattering, reflection, and diffraction is that the received signal can become distorted and/or diffused. In particular, the following two are the major side effects:

- Multiple copies of a signal may arrive at different phases. If phases add destructively, the signal level relative to noise declines, making detection more difficult.
- Intersymbol interference (ISI). One or more delayed copies of a pulse may arrive at the same time as the primary pulse for a subsequent bit.

To complicate matters further, the same type of fading does not happen at all frequencies. In particular, signal fading due to multi-path propagation affects only a few of the frequencies in the total channel bandwidth, thus leading to "partial" fading.

A number of approaches, discussed in later sections, are used to address these issues. These include adaptive equalization (for ISI, mainly), FEC (forward error correction), spread spectrum, and other techniques used in current wireless networks. One of the simplest and most common ways to address fading and errors due to multi-path signal propagation is through the use of ***antenna diversity*** at the base transceiver station. The basic idea of diversity is that different channels experience independent fading events. Thus we can compensate for these errors by sending the same signal over different channels. Access points used in wireless LANs commonly use diversity techniques. In addition, spread-spectrum coding techniques, as we will see, use different frequency channels to reduce the effects of signal fading due to multi-path and Doppler effects (see the sidebar "The Doppler Effect"). In spread-spectrum systems the data is spread over several frequencies, thus greatly reducing the effects of fading on any one frequency for a given time interval.

To summarize, one of the greatest challenges that all wireless networks face is the challenging transmission characteristics of typical wireless channels. Extreme variability in received signal strength (fading) and arrival time is typically experienced due to multi-path propagation effects, delays and the Doppler effect. As stated above, multi-path propagation effects occur as a result of transmitted signals being reflected off objects (buildings, bridges, signs, cars, etc.) before they reach the receiver. Because of this, the path length of signals varies and multiple copies of the same signal arrive at the receiver with different phase offsets. Multiple copies and ISI degrade the BER (bit error rate) of the network. Due to these distortions, wireless systems need to detect and recover from errors by using the techniques discussed below.

The Doppler Effect

Based on the work by the Austrian mathematician and physicist, Christian Doppler (1803-1853), the Doppler effect represents the **shifts** in the frequency of electromagnetic radiation emitted by a moving object. Consider, for example, a police car approaching you. The pitch of the siren changes as the vehicle approaches you and then goes away from you. First the pitch becomes higher, then lower. This change in pitch results from a shift in the frequency of the sound waves, as illustrated in the following picture. As the vehicle approaches you, the sound waves from its siren are compressed towards the observer, resulting in an increase in frequency or pitch. As the vehicle recedes, the sound waves are stretched, causing the siren's frequency to decrease. Similarly, the electromagnetic radiation emitted by a moving object also exhibits the Doppler effect. The radiation emitted by an object moving toward an observer is squeezed; its frequency appears to increase. In contrast, the radiation emitted by an object moving away is stretched. Doppler shifts occur in satellites, stars, galaxies, clouds and other objects as they move towards or away from an observer. Shifts in frequency result not only from relative motion – they are also associated with very strong gravitational fields.

The Doppler effect is used in wireless communications, especially in satellite systems, to estimate the speed of the satellites and to take into account the gravitational forces. Though receivers and transmitters can be mobile here on Earth, their relative velocities are small enough to minimize the Doppler effect. However, as both the earth and spacecraft move rapidly relative to each other in space communications, the Doppler effect becomes non-negligible. If only one spacecraft were present, recovering the signal would be straightforward; one would simply stretch or compress the signal based on the spacecraft's velocity. However, with multiple spacecrafts, the Doppler effect makes it more challenging to design multiple access schemes given that each spacecraft exhibits its own Doppler effect and Doppler effects will vary even for a single spacecraft throughout its mission.

Sources:

http://archive.ncsa.uiuc.edu/Cyberia/Bima/doppler.html http://www.gmi.edu/~drussell/Demos/doppler/doppler.html

5.6 Error Detection and Correction

Wireless systems must be designed to cope with data transmission errors that may be introduced due to attenuation, noise, and fading. Two approaches are used: error detection and error correction. Error detection codes are used to detect the presence of an error. Once an error is detected, techniques based on ***automatic repeat request (ARQ) protocols*** are used so that the block of data with errors is discarded. Later, the transmitter retransmits that block of data. This approach is quite useful in wired networks but is not suitable for wireless systems. The reason is quite simple: if there is an error in transmission, then it is possible that re-transmission will also be in error. For example, if the error is due to bad weather (rain and thunderstorm), then retransmissions will also have errors. In general, error detection and retransmission is inadequate for wireless applications – it is much better to try to recover from errors as much as possible, for the following reasons:

- The error rate on wireless links can be high. This could result in a large number of retransmissions.
- The same errors could be encountered in retransmissions. Errors due to bad weather, noise, and fading in highly mobile environments means that retransmissions will also have errors.
- Wireless systems have long propagation delays compared to transmission time. Thus, wireless systems are slower, hence retransmissions can be slow and expensive. For example, retransmissions in a satellite network can cause significant delays.

To automatically recover from errors, error correction codes, or forward correction codes (FECs), are designed to detect and correct errors. A common approach in error recovery is Hamming code (see the sidebar, "Error Correction Example"). A more recent approach, called Turbo Codes, is gaining popularity because it promises a great deal of improvement in error-free communications [Guizzo 2004].

Detailed discussion of various error detection and correction schemes is beyond the scope of this chapter due to space limitations. See Appendix B for additional information on this topic.

Error Correction Example

Different error correction techniques are available. Block Code Error Correction is among the most widely used technique. It uses Hamming distance (the number of different bits) between two bit patterns to correct errors. The following is a quick example to illustrate the key points.

Hamming distance is the number of different bits between 2 n-bit binary sequences. For example, given two sequences: v_1=011011 and v_2=110001, then the Hamming distance d(v1, v_2)=3.

The basic idea of block code error correction is the following:

- For each data block, create a codeword.
- Send the codeword across the communication line.

- If the code is invalid, then look for data with the shortest Hamming distance because it may possibly be the correct code.

Consider the following example:

Datablock (k=2)	Codeword (n=5)
00	00000
01	00111
10	11001
11	11110

Then instead of the datablock, the codeword is sent. Suppose you receive codeword 00100. This is an error because no such code exists. The closest is 00000 (only one bit different). Thus the receiver will recover from this error by assuming that the data block of 00 was sent.

5.7 Analog Versus Digital Communications in Wireless Systems

There is a significant growth in popularity of digital techniques for sending analog data. This is especially true in wireless systems where most current and future carrier systems are digital. For example, the 2G and higher cellular systems are digital and almost all wireless LANs use digital transmissions. Conceptually, a communication network is analog or digital depending on

- Data received/generated by the transmission facility
- Techniques used to handle attenuation over long distances

Analog communication networks receive/generate analog data and use amplifiers to handle attenuation (see Figure 5-13a). The analog data is either generated directly by a source (e.g., a human) or is converted to analog by modems. The main problem with analog communication networks, as shown in Figure 5-13a, is that the amplifiers do not know the content of the inputs – they amplify whatever is received, including the noise.

In a digital communication network, the data received/generated by the transmission facility is digital and repeaters are used in the transmission facility over long distances to recover the patterns of 1’s and 0’s (see Figure 5-13b). The digital data may originate from a digital device such as a computer or it may be digitized before being fed into the transmission facility. Repeaters cannot be used for analog data. They assume that the input is digital, thus when the signal is attenuated, the repeaters regenerate the original bit patterns (see Figure 5-13b). Repeaters are used to filter the noise. The communications industry is evolving toward the use of digital communication networks. These networks carry digital images of computer bits, voice, video, facsimile, graphics, and many other types of data.

Digital communication networks are more attractive for several reasons:

- Digits are more rugged and free of noise because it is easier to detect 1’s and 0’s even in distorted messages. For example, if amplitudes of 9 volts and 5 volts are used to indicate 1’s and 0’s, respectively, then a receiver can detect a 1 if the amplitude is greater than 5 and a 0 for 5 volts or less. The analog data, once distorted, cannot be recovered.

- Repeaters are used instead of amplifiers and reduce additive noise. Repeaters along a transmission path can detect a digital signal and retransmit a clean (noise-free) signal. These repeaters prevent accumulation of noise along the transmission path. In contrast, if a distorted analog signal is amplified then the distortion is also amplified.
- Digital communication is especially suitable for computer networks because data bits can be directly fed into a communication medium without any modulation/demodulation.
- A single medium (e.g., a cable) can multiplex voice, data and video because they all appear as bits. This allows an organization to develop one backbone network to support all telephone lines, televisions and computers.
- Digital communications are becoming more economical largely due to the availability of chips that can digitize the analog signals efficiently. The theories of digital communications have been around for a number of years but were not economically feasible. It is expected that the costs of digital communications will continue to decrease due to the advances in very large system integration (VLSI).
- Digital communications are more secure than analog communications because digital data streams can be scrambled (encrypted) by using sophisticated computer techniques. The encrypted bits can be deciphered (decrypted) only by equally sophisticated decryption devices/algorithms. The encryption/decryption on analog data (such as human voice) is not sophisticated. This is why "secure" telephone conversations first convert voice to digits before encryption.
- Newer technologies such as optic fibers benefit from digital transmission, and advances in voice digitization are reducing the bandwidth requirements for voice-signals. The combined effect is that digital communications are the favored area of investigation and advancement.
- Better multiplexing techniques can be used with digital networks. For example, instead of the inefficient FDM (Frequency Division Multiplexing), the more efficient TDM (time division multiplexing) can be used in digital networks (see Section 5.9 for a discussion of FDM and TDM). In general, conversion to digital signaling allows use of more efficient digital switching techniques.
- More sophisticated schemes, such as spread spectrum, discussed in Section 5.10, can be used in digital networks.

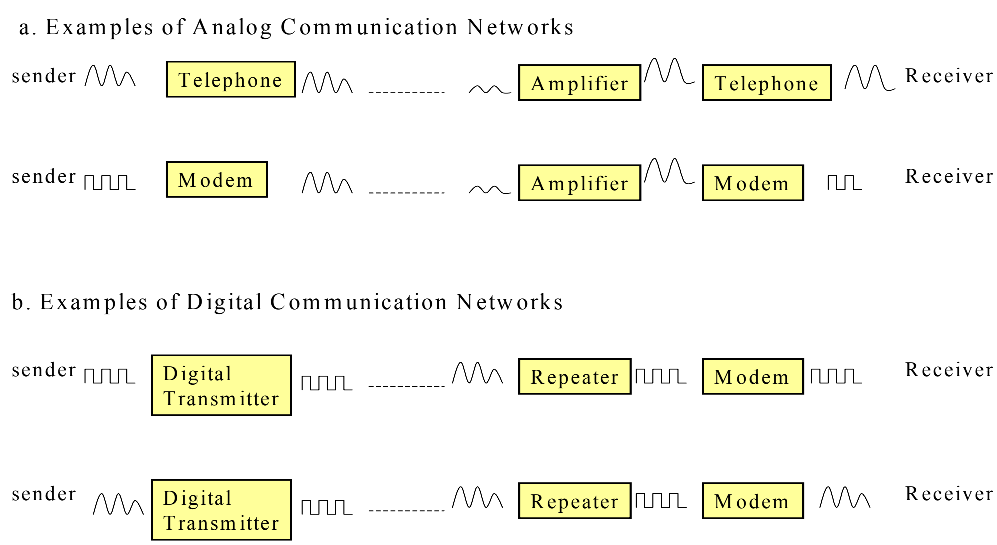


Figure 5-13: Digital versus Analog Communications

5.8 Signal Encoding Techniques

5.8.1 Overview

Wireless systems, just like wired systems, need to encode data into signals that carry the data over the wireless medium. Specifically, the data generated by wireless transmitters needs to be encoded before transmission over wireless networks for propagation, attenuation, and all other distortions discussed previously. Simply stated, data conveys meaning to a user and a signal is an encoding of data in some electromagnetic format. For example, a human voice is data that is encoded as digital signals by modems before transmission.

Data can be digital or analog; signals can also be digital or analog. Figure 5-14 shows the techniques used to convert (encode) data into signals. The technique employed depends on the format of data (analog or digital) and the encoded signal (analog or digital). Figure 5-14 shows that telephones are used to encode analog data to analog signals and digital transmitters are used to encode digital data to digital signals. These conversions can be simple (i.e., a 0 data bit appears as no voltage and a 1 appears as some voltage) or more sophisticated. But how are the digital data bits transmitted by using the analog waves? As shown in Figure 5-14, modems are used to modulate (convert) digital data to analog signals and then demodulate them back to digital data. Similarly, a codec (coder-decoder) is used to convert analog data to digital signals.

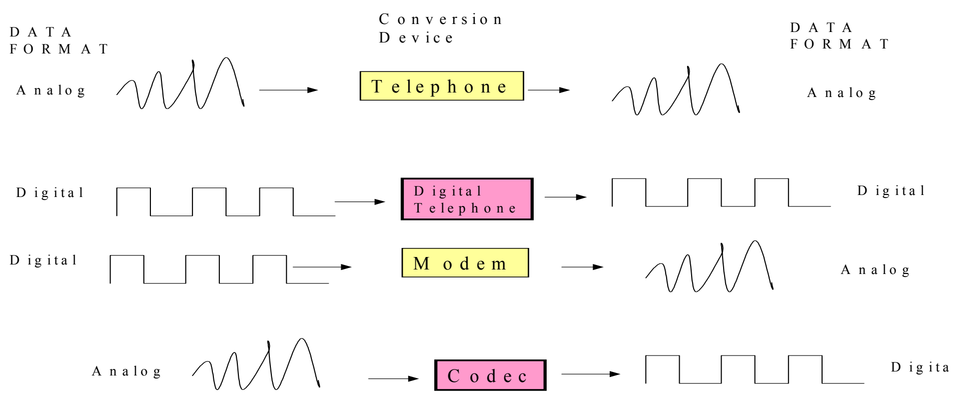


Figure 5-14: Examples of Data to Signal Conversions

Different types of schemes are used for encoding as indicated by the following options (see Table 5-6):

- **Analog data, analog signal**: One of the oldest techniques, in use since the early days of telephone systems, employed for voice transmission over voice-grade lines. The circuitry is simple because analog data in electrical form can be transmitted easily and cheaply.
- **Digital data, analog signal**: In use since the dawn of computer communications in the 1960s, this technique is still used by modems to transmit computer data over dial-up lines. Used widely since some transmission media will only propagate analog signals, e.g., optical fiber and unguided media.
- **Digital data, digital signal**: Used commonly at present to transmit computer data over digital facilities. The equipment is less complex and expensive than digital-to-analog modulation equipment.

- **Analog data, digital signal**: Used to transmit voice and other analog data over digital circuits. This permits use of modern digital transmission and switching equipment.

A wide range of encoding schemes (AM, FM, ASK, FSK, PSK, PCM), discussed later in this section, are used at present. As to how well a receiver can interpret an encoded signal, and how the encoding schemes can be compared and contrasted, see the sidebar, "Signal Encoding Evaluation."

Table 5-6: Signal Encoding Schemes

	Analog Signal	Digital Signal
Analog Data	Amplitude modulation (AM) Frequency modulation (FM) Phase modulation (PM)	Pulse code modulation (PCM) Delta Modulation
Digital Data	Amplitude-shift Keying (ASK) Frequency-shift Keying (FSK) Phase-shift Keying (PSK)	Digital data can be directly fed into a communication medium without any modulation/demodulation.

Signal Encoding Evaluation

What determines how successful a receiver will be in interpreting an incoming signal? The following are the key criteria:

- Signal-to-noise ratio (SNR): An increase in SNR decreases bit error rate, thus a higher SNR is good for better reception.
- Data rate: An increase in data rate increases bit error rate, thus higher data rates are not good for reception.
- Bandwidth: An increase in bandwidth allows an increase in data rate, thus it also leads to higher data rates and lower quality of reception.

The main factors used to compare different encoding schemes are:

- How well it performs in the presence of noise and attenuation. For example, amplitude modulation is more susceptible to noise than others.
- How much bandwidth it consumes. For example, frequency modulation consumes more bandwidth than others.
- How much does it cost? The higher the signal rate and the complexity of the encoding scheme, the greater the cost.

5.8.2 Analog Data to Analog Signal Encoding

This type of signal encoding is quite old and is only used in older wireless systems to transmit voice over analog wireless circuits such as 1G cellular networks. ***Amplitude Modulation (AM),*** shown in Figure 5-15**,** is commonly used. Basically, the amplitude of the data signal (the inner signal) is modified by the carrier signal as shown in Figure 5-15**. *Frequency Modulation (FM)*** is used to represent different data signals with different frequencies, and ***Phase Shift Modulation (PSM)*** is used to represent different data signals with different phases. Since these techniques are very similar to the techniques used in digital data to analog signal encoding, we will examine them more closely in the next section.

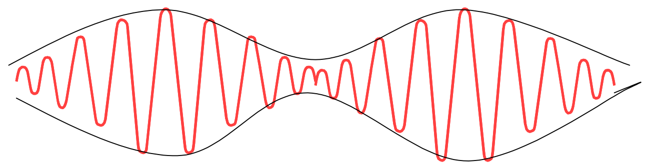


Figure 5-15: Amplitude Modulation

5.8.3 Digital Data to Analog Signal Encoding: Modems and Interfacing Devices

A ***modem*** (modulator/demodulator) is a hardware device which converts digital data to analog signals and vice versa. A modem is also called ***data circuit equipment (DCE)*** or "dataset" in communication systems because it interfaces between voice and data communication systems. Common modulation and demodulation techniques used in modems are (see Figure 5-16):

- ***Amplitude-shift Keying (ASK)***, where the 0 and 1 bits are represented by the height of the amplitude. For example, 5 volts may be used to represent bit 0 and 9 volts may be used to represent bit 1. This technique is quite simple but is susceptible to sudden gain changes due to noise and attenuation. For this reason, it is not heavily used in wireless systems.
- ***Frequency-shift Keying (FSK),*** where the 0 and 1 bits are represented by two different frequencies, say 1000 cycles per second and 2000 cycles per second, respectively. FSK is less susceptible to errors than ASK and is used for high-frequency (3 to 30 MHz) radio transmissions.
- ***Phase-shift Keying (PSK),*** not shown in Figure 1-5, where a certain phase (e.g., 90) is used to represent bit 0, and a change to bit 1 is indicated whenever the phase of the signal changes. PSK is also less error-prone than ASK. In addition, PSK uses bandwidth more efficiently. Thus many wireless systems use PSK and its variants.

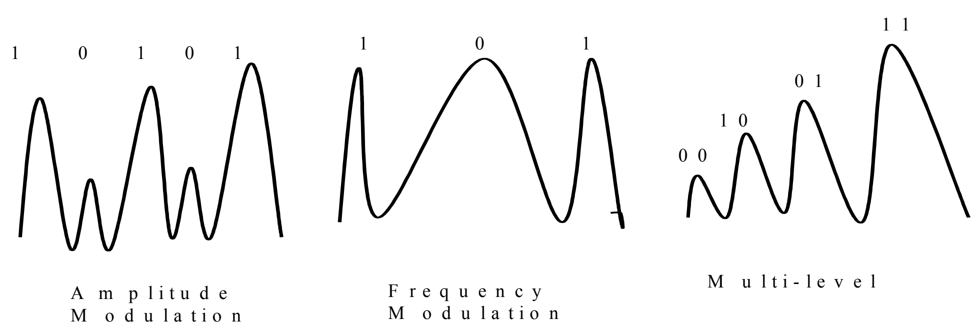


Figure 5-16: Signal Modulation/Demodulation

If one signal level carries more than one data bit, as shown in Figure 5-16, then the modulation is called ***multilevel modulation***. This technique is used in modems to increase data rates. For example, some modems have a switch which can be used to increase the data rate from 28000 bps (bits per second) to 56000 bps. This switch essentially starts sending two bits per signal instead of one, thus doubling the bits per second. Multilevel modulation can be used in FSK and PSK systems also. For example, Binary FSK (BFSK) uses only 2 frequencies (one for zero and the other for 1) but a Multilevel FSK (MFSK) may use 4 or 8 frequencies to represent 2 or 3 bits, respectively. Similarly, Binary PSK (BPSK) uses only 2 frequencies (one for zero and the other for 1) but a Multilevel PSK (MPSK) may use 4 or 8 frequencies to represent 2 or 3 bits, respectively.

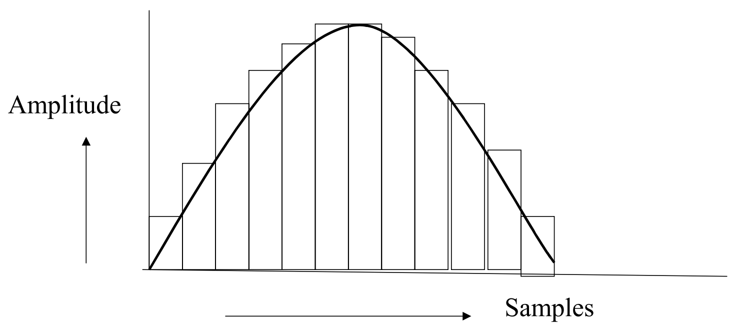
From a performance point of view, ASK is more susceptible to noise than FSK or MFSK systems. As compared to PSK, FSK systems use more bandwidth because more frequencies are needed to carry more bits. Thus, PSK techniques are used more often.

5.8.4 Analog Data to Digital Signal (Pulse Code Modulation – PCM)

Digital data can be directly fed into a digital communication facility without any modulation/demodulation. However, the voice and video signals are analog and need to be digitized through A/D (analog/digital) converters known as codecs (coders-encoders). The A/D technique commonly used is ***pulse code modulation (PCM)***. However, some newer and more efficient schemes are being used. Delta Modulation (DM) is such a scheme.

5.8.4.1 Pulse Code Modulation (PCM)

Figure 5-17 illustrates the basic principle of PCM. The analog signal is sampled several times before transmission (Figure 5-17 shows 12 samples) where each sample represents the amplitude of the analog signal. These samples as sent as digital data and then reconstructed into the original signal on the receiving side by using approximations. Naturally, more samples lead to better signal reconstruction.



This shows 12 samples, each sample represents the amplitude of the wave. These samples as sent as digital data and then reconstructed into the original signal on the receiving side.

Figure 5-17: Pulse Code Modulation (PCM)

The theoretical principle of digitizing an analog signal is that the signal is sampled at twice the signal bandwidth to faithfully reconstruct the signal. This is based on the following sampling theorem formula:

No.Number of digital samples per second = 2 x bandwidth of analog signal

To digitize voice, for example, 8000 samples are taken in a second because the maximum bandwidth of the human voice is 4000 Hz. The pulse code modulation (PCM) digitizing technique consists of the following steps:

- **Sampling**: The analog signal is digitized at 8000 per second. This is based on the sampling theorem (sample rate should be higher than twice the highest frequency). Each analog sample is assigned a binary code.

- **Quantizing**: Each sampled signal amplitude is converted to a level. PCM allows 128 levels of signal amplitude. Analog samples are referred to as pulse amplitude modulation (PAM) samples. The digital signal consists of a block of *n* bits, where each *n*-bit number is the amplitude of a PCM pulse.
- **Encoding**: the amplitude is represented by bits. One byte (8 bits) is used to encode 128 levels (7 bits to represent 128 levels and 1 bit for supervisory and control use). By quantizing the PAM pulse, the original signal is only approximated at the receiving side. This leads to quantizing noise.

Thus one voice channel is converted to 64,000 bits per second (8000x8 data bits = 64,000 bps of data). This is the main reason why we see most digital transmission facilities provide 64 Kbps channels (each 64 Kbps channel can be used to carry one human telephone conversation). A 64 Kbps channel is termed a DS-0 channel.

Let us illustrate PCM by using another example in which one minute of voice signal needs to be stored after digitizing. First, there will be a need for 60x8000 = 480,000 samples. Since each sample occupies 8 bits (1 byte), the total storage required for one minute of digitized voice is 480,000 bytes. This illustrates that digitizing voice has important performance considerations – PCM would generate 8000x8 bits, i.e., 8000 bytes of data for only one second of human voice. Thus transmission of long conversations over communication lines poses large storage requirements, especially if the voice needs to be stored at intermediate nodes. Many variants of PCM have been developed to address these problems [Stallings 2002].

5.8.4.2 Delta Modulation (DM)

Analog input is approximated by a staircase function that moves up or down by one quantization level (δ) at each sampling interval. The bit stream approximates a derivative of the analog signal (rather than amplitude); i.e., 1 is generated if the function goes up, otherwise 0. This significantly reduces the data carried; i.e., instead of 8 bits per signal, 1 bit is carried (this can increase data rate 8 times). Another advantage of DM over PCM is the simplicity of its implementation.

Two parameters are important for delta modulations: a) size of step assigned to each binary digit (δ), and b) sampling rate. The accuracy is improved by increasing the sampling rate. However, this increases the data rate, which in turn can lead to bit errors (you cannot win!).

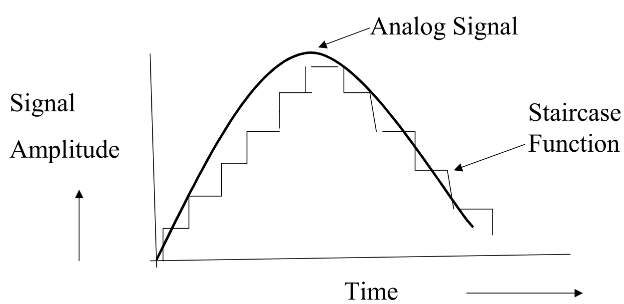


Figure 5-18: Delta Modulation

5.9 Multiple Access Mechanisms (FDMA, TDMA, CDMA)

5.9.1 Overview

In wireless systems, just like wired systems, the users of channels need special mechanisms to avoid interference and collisions. The three mechanisms used most commonly are frequency division multiple access (FDMA), time division multiple access (TDMA) and code division multiple access (CDMA). These three schemes are illustrated in Figure 5-19.

Before going into the details, an example may help. Suppose you are in a party where multiple people want to talk to each other. In an FDMA scheme, the different groups will break up into small subgroups and each subgroup will communicate in parallel. In a TDMA scheme, the people who want to talk to each other will take turns (i.e., use time slices). In CDMA, everyone will start talking at once, but only some people will understand each other because of a code. For example, people speaking Spanish will be tuned to Spanish only, French to French, etc.

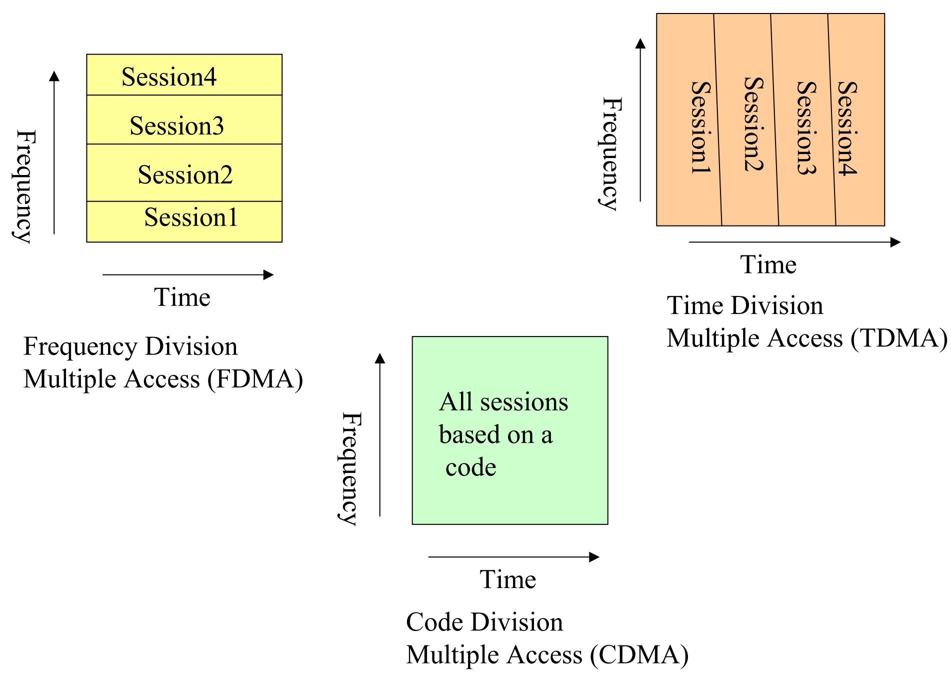


Figure 5-19: Multiple Access Techniques

5.9.2 Frequency Division Multiple Access (FDMA)

In this case, the frequency spectrum is subdivided so that each user gets a dedicated channel. This is one of the simplest and oldest schemes. However, it is not very efficient because a channel could be wasted if no one in that channel is talking. For example, FM radio divides the spectrum into 30 Khz channels, and FDMA divides each 30 Khz channel into 3 users (10 Khz each).

Base station cost for FDMA is high with very limited capacity. Due to these limitations, very few wireless systems use FDMA at present. Most applications of FDMA are in satellites.

5.9.3 Time Division Multiple Access (TDMA)

In this case, frequency bands available to the network are divided into time slots, with each user having access to one time slot at regular intervals. TDMA thereby makes more efficient use of available bandwidth than FDMA. TDMA has been available since 1992, with numerous modifications and improvements made since then.

The basic idea of TDMA is quite simple. Each subscriber transmits at different times; thus the data bits or small clumps of bits from different sources are interleaved on the same communication medium. Framing is used to identify which bits belong to which source so that the signals can be separated at the receiving side. Many variants of TDMA have been developed over the years. Digital switches such as the PBXs (private branch exchanges) perform TDMA so that one line can be shared by many digital users. In some cases, computing devices perform time division multiplexing so that one cable can be shared by many devices.

In cellular phone systems, the communication is divided into 6 millisecond (ms) frames, each divided into 1 ms time slots for 6 users. Each time slot has a header and data. Errors may corrupt headers and cause loss of time slots. In some cases the whole frame is lost. Call quality of TDMA is similar to FDMA but it can handle more calls. Several extensions of TDMA have been proposed and implemented (it can support 15 users per voice channel).

TDMA has been used by major US carriers such as AT&T Wireless Services, BellSouth and Southwestern Bell. A common technology is IS-136 TDMA, which exists in North America at both the 800 MHz and 1900 MHz bands. TIA standard IS-54 defines the TDMA interface between a mobile station and cell-site radio. It uses PCM (pulse code modulation) for speech encoding and a variant of PSK (phase-shift key) for modulation. To learn more about industrial implications of IS-136 TDMA technologies, visit the web sites of AT&T Wireless Services, Bellsouth Cellular Corp., and Ericsson.

5.9.4 Code Division Multiple Access (CDMA) Technology

CDMA differs from the other two technologies by its use of coding techniques for transmitting voice or data over the air. Rather than dividing the RF spectrum into separate user channels by frequency slices or time slots, CDMA technology separates users by assigning them digital codes within the same broad spectrum. The receivers only tune to those codes and ignore the others. The assigning of codes to different users is also known as spread spectrum, a technique very heavily used in wireless LANs as well as cellular networks. Spread spectrum, explained in Section 5.10, transmits different data bits on different signals, based on a secret scheme. The receiver must know the code to understand the signal. Specific characteristics of CDMA are:

- Groups of bits from digitized speech are tagged with a unique code that is associated with a cellular call.
- Several cellular calls are combined and transmitted over 1.25MHz and then reassembled on the receiver side.
- Receiver detects a signal by tuning to correct phase position between incoming and locally generated signals from code.
- Speech coder operates at a variable rate (fully when user is talking).
- It adjusts for near-far power adjustments (nearer stations generate less powerful signals).
- When powered on, the mobile system knows the CDMA frequency, so it tunes to that frequency and searches for a pilot signal (pilot signals represent base stations).
- Mobile station will pick the strongest pilot and register.

- When moving from cell to cell, new pilot is picked up.

Advantages of CDMA technology include high user capacity and immunity from interference by other signals. Like TDMA IS-136, CDMA operates in the 1900 MHz band as well as the 800 MHz band. The CDMA technology used in North America is based on the IS-95 protocol standard first developed by QUALCOMM . The major US carriers using CDMA are AirTouch, Bell Atlantic/Nynex, GTE, Primeco (PCS consortium of AirTouch, Bell Atlantic/Nynex and USWest), and Sprint PCS (consortium of Sprint, Comcast, Cox and TCI). To learn more about industrial aspects of CDMA technology, visit the websites of qualcomm, AirTouch, CDMA, Development group (www.cdg.org), sprint PCs, and PrimeCo.

CDMA and its variants are at present very popular and are targeted as the core technologies for the 3G wireless systems.

5.9.5 TDMA Versus CDMA Controversy

TDMA and CDMA are accepted TIA (Telecom Industry Association) standards (IS-54 and IS-95). Hardware vendors are lobbying hard on both sides. There are many, many variants of TDMA and CDMA in the industry. In addition, there are many performance reports with conflicting and confusing results. In general, the following seems to hold:

- Call clarity: CDMA appears to be better but is questioned.
- Network capacity: CDMA may be more efficient than TDMA.
- Privacy: CDMA codes provide more privacy.
- Economy: TDMA allows same equipment for multiple users.
- Maturity: TDMA is very mature (in use since 1992) but CDMA is catching up.
- More features; TDMA offers more but CDMA can do it also.

Although the TDMA-versus-CDMA debate continues, CDMA is taking a clear lead in future cellular networks, as the 3G and 3G+ systems are almost exclusively based on CDMA.

5.10 Spread Spectrum

5.10.1 Overview

Spread-spectrum technology was developed for military and intelligence operations during the Second World War. The basic idea is that the message is “spread” over a range of frequencies in a manner to make it jam-resistant. Although used in many wireless systems, spread spectrum is most widely used in wireless LANs. These LANs transmit in the unlicensed industrial, scientific, and medical (ISM) bands designated by the FCC. Due to the unlicensed nature of ISM, it is important to minimize interference in this band.

Spread-spectrum technology is a wideband radio frequency technique that basically transmits different data bits on different signals, based on a secret scheme. The receiver must know the parameters of the spread-spectrum signal (the code) being broadcast to understand the signal. If a receiver is not tuned to the right frequency based on the code, a spread-spectrum signal looks like background noise. Thus spread sequence automatically includes security, which is why this scheme was used in the Second World War for secure communications. In spread-spectrum systems, more bandwidth is consumed than in the case of narrowband transmission, but the tradeoff produces a signal that is secure and louder. There are two types of spread-spectrum radio: frequency-hopping and direct-sequence.

5.10.2 Frequency-Hopping Spread Spectrum

Frequency-hopping spread-spectrum (FHSS) uses a narrowband carrier that changes frequency in a pattern known to both transmitter and receiver. For example, the transmitter can transmit 8 bits of data at frequency f1, send the next 12 bits at frequency f2, and then go back to 8 bits at f1. Properly synchronized, the net effect is to maintain a single logical channel. To an unintended receiver, FHSS appears to be short-duration impulse noise. Wireless LANs such as Bluetooth use FHSS. The technical foundation of FHSS is based on interesting history – see the sidebar, "How an Actress and Self-Playing Pianos Led to FHSS."

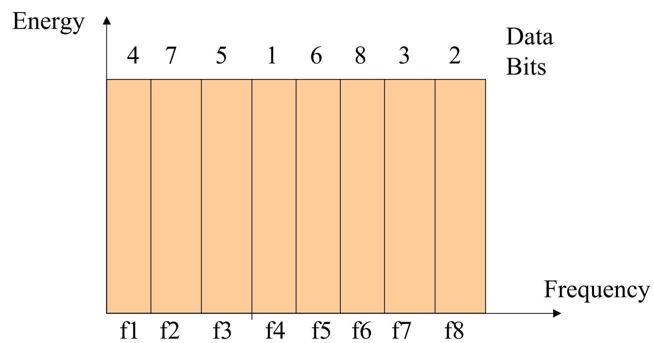


Figure 5-20: Frequency-Hopping Spread Spectrum

Figure 5-20 illustrates FHSS. In this case, bit 4 is transmitted on frequency f1, bit 7 on f2, bit 5 on f3, bit 1 on f4, etc. It is called frequency-hopping because the frequency of a radio transmission changes many times a second and leaps around in an apparently random fashion. The signal hops around and spreads itself across a range of the radio spectrum, rather than remaining on a single frequency. Thus a radio signal is made almost impossible to intercept. Only a receiver programmed with the same random sequence would be able to follow the signal as it hopped from one frequency to another. The main characteristics of FHSS are:

- Signal is broadcast over seemingly random series of radio frequencies.
- Signal hops from frequency to frequency at fixed intervals.
- Frequency sequence is dictated by a "spreading code" that shows what bits will be transmitted on what frequency.
- Receiver knows the spreading code and uses this code to hop between frequencies in synchronization with transmitter to pick up the message .

A large number of frequencies are used in FHSS to carry data. This results in a system that is quite resistant to jamming. The jammer must jam all frequencies. With fixed power, this reduces the jamming power in any one frequency band.

The main advantage of FHSS is that eavesdroppers hear only unintelligible blips. In addition, attempts to jam the signal on one frequency succeed only at knocking out a few bits.

How an Actress and Self-Playing Pianos Led to FHSS

Hedy Lamarr, a Hollywood actress who died in 2000, played an unlikely off-screen role as a technological pioneer, co-inventing FHSS in the 1940s – an early incarnation of spread-spectrum wireless technology.

Hedy Lamarr and her husband, George Antheil, were jointly awarded a patent in 1942 for a "secret communication system," currently known as frequency-hopping spread spectrum. Hedy got involved in this project in the 1930s when she accompanied her previous husband, Fritz Mandl, an Austrian arms dealer, to numerous meetings and dinners, and became familiar with the problem of sending control signals to a torpedo after it was launched from a ship. Hedy divorced her husband, and went to America in 1937 where she became a successful actress. She met and married George Antheil, an experimental composer. George was known for composing music for "player pianos," mechanical instruments that play back music encoded as holes punched in a roll of paper. Lamarr and Antheil worked together to develop a variant of spread-spectrum technology that made radio transmissions extremely difficult to jam or intercept.

The idea was to do "frequency-hopping" by changing the frequency of a radio transmission many times a second. This caused the signal to leap around in a random fashion so that it was almost impossible to intercept. Only a receiver programmed with the same random sequence would be able to understand the signal. To encode and control the frequency-hopping pattern, Lamarr and Antheil proposed using a punched-paper roll – like that of a player piano. Their system hopped between 88 different frequencies, the number of keys on a modern piano. The player-piano rolls in the transmitter (aboard the ship) and receiver (in the torpedo) would be started at exactly the same moment, and would then stay synchronised after launch, providing a secure radio link from the ship to the torpedo.

Lamarr and Antheil offered their idea to America's armed forces, but it was not taken seriously. It was nearly 20 years before a radio based on FHSS was constructed by using electronic components instead of the mechanical components and paper rolls. It was used to secure communications during the Cuban missile crisis in 1962. FHSS is currently the basis of Bluetooth, a short-range wireless protocol for wireless personal area networks, and is also used in other wireless LANs.

Source: "Player-Piano Pioneer," *The Economist*, print edition, Jun 19th 2003, available from: http://www.economist.com/science/tq/displayStory.cfm?story_id=1841130

5.10.3 Direct Sequence Spread Spectrum

Direct sequence spread spectrum (DSSS), a variant of spread spectrum, generates a redundant bit pattern for each bit to be transmitted. This bit pattern is called a chip (or chipping code). The longer the chip, the greater the probability that the original data can be recovered (and, of course, the more bandwidth required). Even if one or more bits in the chip are damaged during transmission, statistical techniques embedded in the radio can recover the original data without the need for retransmission. To an unintended receiver, DSSS appears as low-power wideband noise and is rejected (ignored) by most narrowband receivers. The IEEE 802.11 Wireless Ethernet LAN uses DSSS as well as FHSS. In addition, DSSS is used in Code-Division Multiple Access (CDMA), a popular standard for cellular phones. Let us look at the following example to understand CDMA plus DSSS (two for the price of one!):

Let us assume that a chipset is of size 6 and code is a sequence of 1s and -1s. Then for a "1" bit, a chip pattern could be <c1, c2, c3, c4, c5, c6>. For a "0" bit, a complementary code is used: <-c1, -c2, -c3, -c4, -c5, -c6>. The receiver knows the sender's code and performs an electronic decode function. Consider, for example, two different users, Joe and Pat, with the following patterns:

Joe's code = <1, –1, –1, 1, –1, 1>

```
To send a 1 bit from Joe's handset = <1, –1, –1, 1, –1, 1>

To send a 0 bit from Joe's handset = <–1, 1, 1, –1, 1, –1>
```

Pat's code = <1, 1, –1, – 1, 1, 1>

```
To send a 1 bit from Pat's handset = <1, 1, –1, –1, 1, 1>

To send a 0 bit from Pat's handset = <–1, –1, 1,1, – 1, –1>
```

The receiver receiving with Joe's code uses the scheme: (Joe's code) x (received chip pattern). This produces a 1 if Joe sent a 1 and a 0 if Joe sent a 0.

Similarly, the receiver receiving with Pat's code uses the scheme: (Pat's code) x (received chip pattern). This produces a 1 if Pat sent a 1 and a 0 if Pat sent a 0.

The main characteristics of Direct Sequence Spread Spectrum (DSSS) are:

- Each bit in the original signal is represented by multiple bits in the transmitted signal.
- Spreading code spreads the signal across a wider frequency band.
- The spread is in direct proportion to the number of bits used.
- One technique combines a digital information stream with the spreading code bit stream using exclusive-OR.

5.11 Short Case Studies and Examples

5.11.1 University Wireless Network Revisited

Let us revisit the university campus wireless network example introduced in the beginning of this chapter. Figure 5-21 shows a conceptual view of the network. The network shows how the main campus, the professional development center, and the overseas office in Hong Kong are interconnected. Access to the public Internet is also shown.

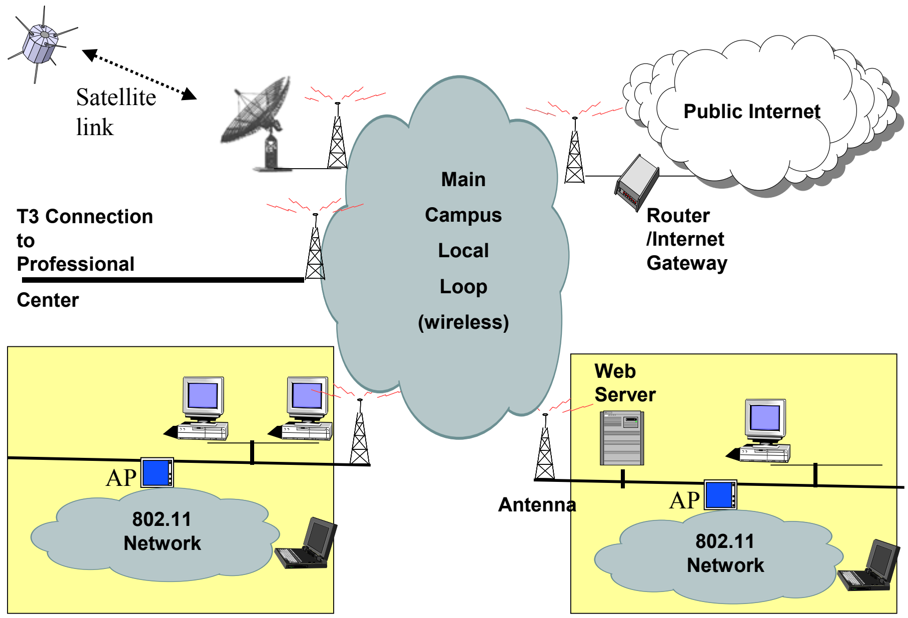


Figure 5-21: Example of a University Network

Many design choices have been made in this overall architecture. We will justify these choices in later chapters. The main decisions are:

- Wireless LANs, based on the IEEE 802.11 standard, are established in all buildings (only two are shown for the purpose of illustration). Each building also has its wired Ethernet LANs and some buildings have their own Web/LAN servers. A wired backbone network exists in each building that interconnects the individual LANs and the wireless access points (APs).
- The heart of the campus network is a Wireless Local Loop (WLL), based on LMDS technology, that is used on the main campus to interconnect many buildings on the main campus. Each building has an antenna for participating in the WLL. The wired FDDI network is not shown in this diagram (we are into wireless!).
- The main campus is connected to the professional development center through a T3 Line. Note that the T3 line is connected to the WLL antenna.
- Access to the public Internet is provided through a router/Internet gateway that is connected to a WLL antenna. Due to workload, more than one such connection may exist.
- Satellite access for offshore instruction is also provided through a WLL antenna.

This network does not show all the decisions. For example, addition of a paging network is left to the reader as an exercise.

For physical wireless network design, the following decisions were made:

- A mixture of omnidirectional and directional antennas were used in the WLL. Smart antennas were not considered at present.
- The frequency allocations for the satellites were decided by the satellite provider and the frequency allocations for WLL were based on the LMDS range (30 GHz).
- To minimize errors and provide secure communications, a combination of diversity and FHSS techniques were used in the WLL.

5.11.2 A Village Area Network – Using Motorcycles as Communicating Devices[12]

In the hills of northeastern Cambodia, several rural villages have been interconnected with one another to form a "village area network (VAN)." This VAN also connects the villages with the Cambodian government, medical specialists and the Internet by using wireless networks, albeit by using motorcycles (see later). The goal is to allow far-off villagers in Cambodia to exchange emails, access medical services, and review the government regulations that influence the villagers. The Cambodian experiment is one out of many initiatives aimed at providing wireless Internet connectivity to distant locations in developing countries. The project is being pioneered by a Boston company called First Mile Solution (FMS). FMS has in fact developed a Village Area Networking Kit to replicate this experience to other villages. The VAN toolkit is expected to be priced at $500 to $600, a fraction of the cost of the electricity and communications infrastructure that would otherwise be necessary to just deliver email to the villages.

The main technical challenge in providing Internet access for far-flung rural locations is that it is virtually impossible to deploy wired networks due to the topography. This "last kilometer" problem is solved by using wireless communications. The main solution consists of a wireless LAN (WLAN) that was implemented using a hub with an omnidirectional antenna to communicate with kiosks equipped with directional antennas. To operate properly, the kiosks as well as the hub had to be in line of sight. If the hub and a kiosk were not in a line of sight, then another kiosk that was in line of sight was used to establish a hop creating a multi-hop network. The main application initially is email.

The most interesting part of this design is how the email data collected at the WLAN hubs (schools) is transferred to the regional centers. There is no satellite or any other type of link between the hubs and the regional centers. Thus five motorcycle riders travel their routes five days a week, visiting each hub and downloading and uploading email wirelessly. The system, developed by FMS, uses a receiver box powered by the motorcycle's battery. The driver rolls slowly past the school, without even stopping, to download the village's outgoing email and to deliver incoming email. The school's computer system and antenna are powered by solar panels. Newly collected data is stored for the day in a computer that is strapped to the back of the motorcycle. At the end of the day, the motorcycles end up at the provincial capital where an advanced school is equipped with a satellite dish, allowing a bulk email exchange with the outside world. Thus the motorcycles are used as a mobile link between the village school and the provincial capital school.

The results of this testing are very promising. By using this VAN, the village teachers were able to communicate with regional offices to send and receive reports and directives. The village health care center provided better quality as reports and photos of symptoms were emailed regularly to the regional hospital. Some of the reports were sent to the Harvard Medical School for examination and guidance. Villagers were also able to review news and prices of the crops that they have been planting for the season. Complaints to the municipality were also sent via email and the villagers also marketed their own products and engaged in trade with suppliers in close-by villages or cities.

The deployment of this project helped revolutionize a remote village by using wireless networks to provide the services that were never available to the inhabitants before.

Source:

[12] Suggested by Bashar Saleh.

- Popkin, J., *Closing Rural Cambodia's 'Digital Divide' on Motorcycles*, Gartner Research Organization, Document Code: 118507, November 18, 2003.
- Also featured as "Rural Cambodia, Though Far Off the Grid, Is Finding Its Way Online" in *NY Times* print edition, Jan 26, 2004.

5.11.3 Wireless Nurse Call System

Hospitals traditionally have relied on nurse call systems that use overhead paging systems to notify caregivers of patient events, and dispatch caregivers to a given patient's room. These traditional nurse call systems have some problems. The first problem is that the caregivers have to locate a telephone when they hear the overhead paging system informing them that they are required in a certain location. This can consume precious time, especially in the case of a medical emergency. Worse yet, if there is no response to a page by a caregiver, the staff member issuing the page cannot be sure whether the intended caregiver is unaware of the page, or is aware but is unavailable to respond. In addition, the constant overhead paging announcements can be disruptive to other caregivers and depressing to the patients who are forced to listen to the pages. Finally, overhead paging systems increase operational costs because they require additional staff to man the nurse call system and issue pages.

To address these issues, wireless pagers have been introduced that notify caregivers that they are needed without employing overhead paging. A significant improvement is the use of "enterprise wireless voice/data systems," which allow mobile staff within a business campus to communicate using wireless handsets. Healthcare organizations, with their many mobile caregivers, have been using such systems and are now also using the Computer-Telephony Integration (CTI) to make other applications possible. By using these technologies, mobile caregivers can be automatically notified of patient events directly on their wireless handsets. Notification of a particular caregiver can be specified to correspond to those patient rooms or beds that they are responsible for during a particular shift. In addition, call forwarding/redirection can be used to allow an alternate caregiver to respond to a patient should the assigned caregiver be unavailable. This is possible because the enterprise wireless system is also connected to a PBX (Private Branch Exchange).

Technically, these enterprise wireless systems use wireless "cells" deployed within buildings. Some of these wireless systems, like Nortel Networks Companion products, operate in the 1.9 GHz radio spectrum, and at very low power output (in the order of 30 milliwatts). Developed specifically for patient care, these "PCI" systems are designed so that they do not interfere with medical equipment. This is accomplished through lower power output and the operational frequency range beyond the 900-megahertz spectrum, where many medical devices are designed to operate. As a result, PCI–based wireless system are becoming commonly accepted in healthcare settings, where medical monitoring equipment susceptible to interference is often present.

Source: www.nortelnetworks.com/solutions/ health/collateral/wp3685-a.pdf

5.12 Concluding Comments

This chapter has given a broad overview of the underlying principles that are common to all wireless networks, from wireless local area networks to satellite communication systems. Topics have included frequency allocations and location services that are important issues for wireless providers but also impact the end users. A discussion of multiple access mechanisms such as TDMA and CDMA has attempted to highlight how multiple users can share the same wireless medium without interfering with ach other. A quick overview of physical wireless communication issues has been included to emphasize the role of antennas, signal encoding and error correction in the generation, transmission, and reception of wireless signals.

5.13 Review Questions and Exercises

1) What are different types of wireless networks? Capture their key characteristics through a table that expands Table 5-1.

2) What are the main differences between wireless and wired networks?

3) What are the technical foundations of wireless networks? Summarize the key points.

4) What are the frequency ranges that are most frequently used in wireless networks and why are low frequencies more congested?

5) What are the main unregulated frequency bands and what roles do the regulatory bodies play in frequency allocations?

6) Find a case study of wireless frequency allocation from the recent literature.

7) What are the various location management systems and what are the tradeoffs? Which technique is most useful for E911 and why?

8) What are the different types of antennas and what is an antenna gain?

9) What is a smart antenna? Why is it considered smart? What is its advantage over the traditional antennas?

10) What are the propagation issues in wireless networks and why they are important?

11) Supposing you have two antennas, each 100 meters high. How far apart they should be for an LOS communication?

12) Discuss the various line-of-sight impairments and the approaches used to address them.

13) What is fading? What contributes to it? What are the after-effcts and what are the main approaches used to recover from fading errors?

14) What are the various error detection and correction schemes? Why is error correction more important in wireless?

15) What are different types of signal encoding techniques and what are the tradeoffs?

16) Compare and contrast CDMA with TDMA.

17) What is spread spectrum? Explain through an example.

5.14 References

Bates, R. *Wireless Broadband Handbook*. McGraw Hill, 2001.

Bekkers, R. and Smits, J. *Mobile Telecommunications*. Artech, 2000.

Gralla, P. *How Wireless Works*. Que, 2001.

Guizzo, E. "Closing in on the Perfect Code." *IEEE Spectrum*, March 2004, pp. 36-42.

Jagoe, A. *Mobile Location Services: The Definitive Guide*. Prentice Hall, 2002.

Liberti, J. and Rappaport, T. *Smart Antennas for Wireless Communications*. Prentice Hall, 1991.

Mark, J. and Zhuang, W. *Wireless Communications and Networking*. Prentice Hall, 2003.

Rappaport, T. *Wireless Communications: Principles and Practice*. 2nd ed. Prentice Hall, 2001.

Sampei, S. *Applications of Digital Wireless Technologies to Global Wireless Communications*. Prentice Hall, 1998.

Stallings, W. *Wireless Communications and Networks*. Prentice Hall, 2002.

Stojmenovi, I., et al. *Handbook of Wireless Networks and Mobile Computing.* John Wiley, 2002.

Walker, J., ed. *Advances in Mobile Information Systems*. Artech House Mobile Communications Library, December 1998.