

# Communication Networks

## INTRODUCTION

We first expand on multiplexing, duplexing, and multiple access schemes along with their merits, drawbacks and applications. Various random-access and controlled-access schemes are briefly discussed. The focus then turns mainly towards some aspects of wired communication networks, such as circuit-switched and packet-switched networks, and telephone and cable networks. We then briefly introduce cryptography and network security, mainly the fundamentals of private-key and public-key cryptography and digital signature. After studying this chapter and understanding all relevant concepts, students should be able to achieve the following learning objectives:

1. Define multiplexing and characterize its applications.
2. Identify differentiating aspects of FDM and TDM techniques.
3. Distinguish the difference between synchronous TDM and asynchronous TDM.
4. Understand communication modes.
5. Know the difference between the duplexing methods FDD and TDD methods.
6. Elaborate on the need for multiple access schemes.
7. Assess FDMA, TDMA, and CDMA schemes and compare them.
8. Analyze unslotted and slotted ALOHA random-access methods.
9. Detail the differences among various CSMA schemes.
10. Portray how controlled-access methods work.
11. Provide basic descriptions of circuit-switched and packet-switched networks.
12. Expand on various forms of network topologies.
13. Discuss routing and flow-control functions.
14. Appreciate the evolution and current status of wired LANs.
15. Grasp how digital data transmission over telephone networks operates.
16. Explain how digital data transmission over cable networks works.

## CONTENTS

|                                                   |     |
|---------------------------------------------------|-----|
| Introduction.....                                 | 457 |
| 11.1 Multiplexing .....                           | 458 |
| 11.2 Duplexing .....                              | 461 |
| 11.3 Multiple Access.....                         | 464 |
| 11.4 Random Access .....                          | 468 |
| 11.5 Controlled<br>Access .....                   | 473 |
| 11.6 Wired<br>Communication<br>Networks.....      | 474 |
| 11.7 Network<br>Security and<br>Cryptography..... | 485 |
| Summary and<br>Sources .....                      | 490 |
| Problems.....                                     | 491 |

17. List the network security requirements.
18. Differentiate between private-key cryptography and public-key cryptography.
19. Understand the need for digital signatures and how they work.
20. Connect various aspects of multiuser communications.

## 11.1 MULTIPLEXING

*Multiplexing* is a technique in the physical layer, whereby the simultaneous transmission of several independent signals from different sources is carried out. With multiplexing, there generally exists a priori resource allocation, and the resource sharing is usually done within the confines of a local site, but over a single common communication channel. Multiplexing is done when the bandwidth of the channel is higher than those of individual sources. At the transmitter, signals are combined (multiplexed) together, while keeping them apart enough to avoid interference, so at the receiver, they can be separated (demultiplexed). Multiplexing is thus done by a multiplexer in the transmitter and a demultiplexer in the receiver. There are three common types of multiplexing, namely, frequency-division multiplexing, time-division multiplexing, and wavelength-division multiplexing.

### 11.1.1 Frequency-Division Multiplexing

*Frequency-division multiplexing* (FDM) is a multiplexing technique that combines many signals into a single, high-bandwidth signal. In FDM, the bandwidth of a link is greater than the combined bandwidths of the signals to be transmitted, and the available transmission channel bandwidth is thus divided into a number of nonoverlapping frequency bands. In FDM, signals from all sources are transmitted simultaneously, but each occupying a different frequency band. FDM has generally been viewed as an analog multiplexing technique. FM stereo broadcasting represents a typical application of FDM. However, in principle, a digital system can employ FDM as well, such as cable telephony, data upstream and downstream, and TV channels over co-axial cable, and DSL telephony and data upstream and downstream over twisted-wire pair. A block diagram of a typical FDM system is shown in [Figure 11.1a](#).

In FDM, the incoming lowpass signals are all filtered to make them frequency limited. The removed high-frequency components do not significantly contribute to signal representation, but they are capable of interfering with other signals. The band-limited signals individually modulate carriers to shift their frequency ranges to occupy mutually exclusive frequency bands and guard bands separate

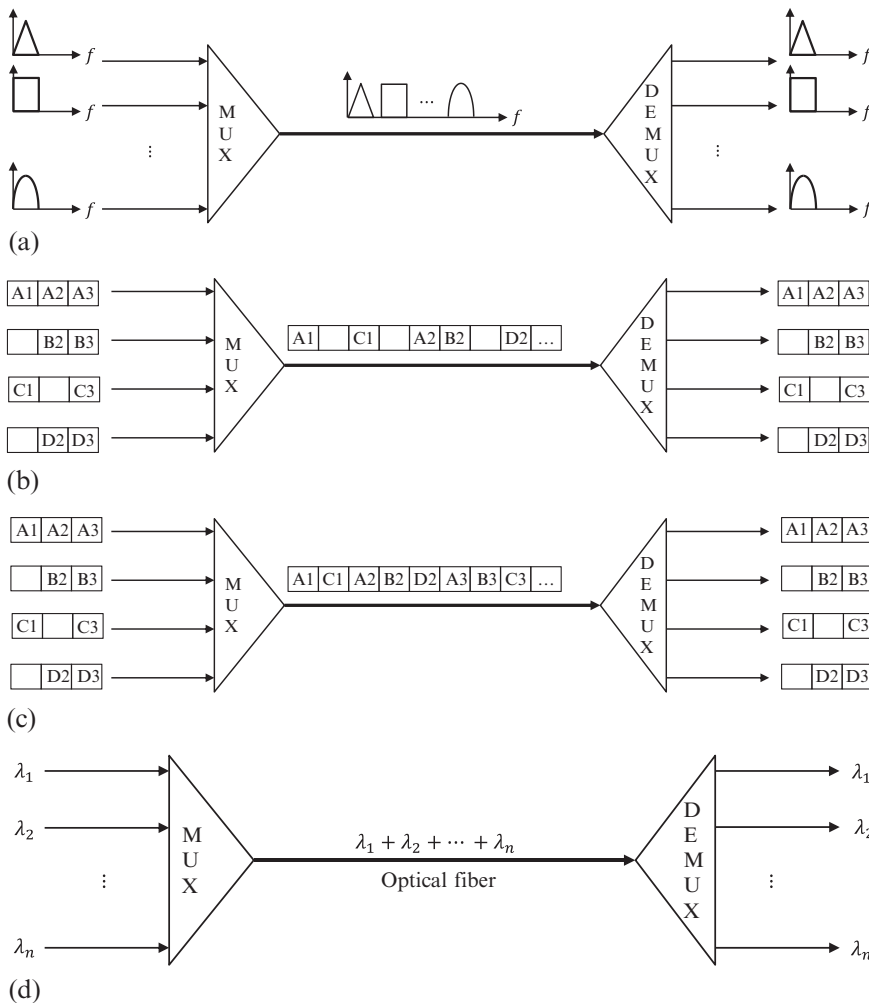


**FIGURE 11.1** Multiplexing: (a) frequency-division multiplexing, (b) synchronous time-division multiplexing, (c) asynchronous time-division multiplexing, and d) wavelength-division multiplexing.

adjacent channels. The larger the guard band, the easier it is to design the filters to separate the signals, but the price being paid is the reduction in the number of channels. The SSB modulation is a good choice for FDM. The modulated signals are bandpass filtered to restrict the modulated signals to prescribed range. The resulting signals are combined to produce a composite multiplexed signal to be transmitted over the channel. At the receiver, a bank of bandpass filters separates the individual modulated signals. Individual demodulators recover the original baseband signals.

### 11.1.2 Time-Division Multiplexing

*Time-division multiplexing* (TDM) is a digital multiplexing technique for combining several low-data-rate signals into one high-data-rate signal. In TDM, the available transmission channel is time-shared by a number of different sources. In TDM, the digital data or the digitized data from different sources are interleaved in time into a single digital data signal to be transmitted over the channel. A prime application of TDM is T1 carrier transmitting 24 digitized voice signals at 1.544 Mbps.

There are two types of TDM: synchronous TDM and statistical TDM. Block diagrams of a synchronous TDM and a statistical TDM are shown in [Figures 11.1b](#) and [11.1c](#), respectively. In *synchronous TDM*, each source has an allotment in the output even if it does not have data to transmit. In *statistical TDM*, a source is given access to the multiplexer only during periods of activity, this in turn can greatly enhance channel efficiency, and statistical TDM is suited for applications, such as speech communications.

In synchronous TDM, each source is timed by the same clock that provides timing within the TDM. Each input signal is first band-limited to avoid aliasing. The LPF outputs are then sequentially sampled at the transmitter by a rotary switch, known as a *commutator*, to form a signal consisting of samples of the input signals periodically interlaced in time. The multiplexed signal is then transmitted. At the receiver, the samples from individual sources are separated and distributed by a rotary switch called a *decommutator*, which operates in full synchronization with the commutator at the transmitter. In synchronous TDM, the data flow of each source is divided into units of several bits, where each unit occupies one input time slot. A round of data units from each source is collected into a frame (i.e., a frame consists of many slots), where each slot corresponds to a source. Frame synchronization in TDM is a major requirement. Either a synchronizing signal is transmitted as one of the multiplexed signals or it can be an additional bit, known as a *framing bit*, through which synchronization is done. In the case of T1 carrier, we have a frame consisting of a framing bit and 192 information bits ( $= 24 \text{ slots} \times 8 \text{ bits per slot}$ ) and a transmission rate of 1.544 Mbps ( $= 193 \text{ bits per frame} \times 8000 \text{ frames per second}$ ).

### 11.1.3 Wavelength-Division Multiplexing

*Wavelength-division multiplexing* (WDM) is a multiplexing technique to combine optical signals. In WDM, the available fiber-optic transmission channel is shared by a number of different light sources. WDM is conceptually quite similar to FDM. The advantage of WDM is to exploit the full capacity of the fiber-optic cable by allowing multiple beams of light at different frequencies

to be transmitted on the same fiber-optic cable. A prime application of WDM is the SONET (Synchronous Optical Network) standard developed in North America. A block diagram of a synchronous WDM system is shown in Figure 11.1d.

Optical signals at different optical wavelengths (colors) are combined by the multiplexer at the transmitter to form a single light to be transmitted through the high-speed fiber-optic cable, and the splitting of the light sources is done by demultiplexer at the receiver. The combining and splitting of light sources are done by various optical devices, such as prism and diffraction gratings. The channel spacing in WDM systems must be large enough to prevent interference between adjacent channels. Most WDM systems use 50-GHz channel spacing, such as ITU-G.692 that accommodates 80 channels over 4 THz.

## 11.2 DUPLEXING

Before we introduce the duplexing methods, it is important to highlight various communication modes. Communication modes can be categorized into three distinct types: simplex, half-duplex, and full-duplex (also known as duplex), as shown in Figure 11.2.

In *simplex mode*, transmission is unidirectional (i.e., transmission always occurs only in one direction). The simplex mode uses the entire capacity of the channel to send data in one direction. Examples of simplex communication are baby monitoring, public address, and commercial broadcasting.

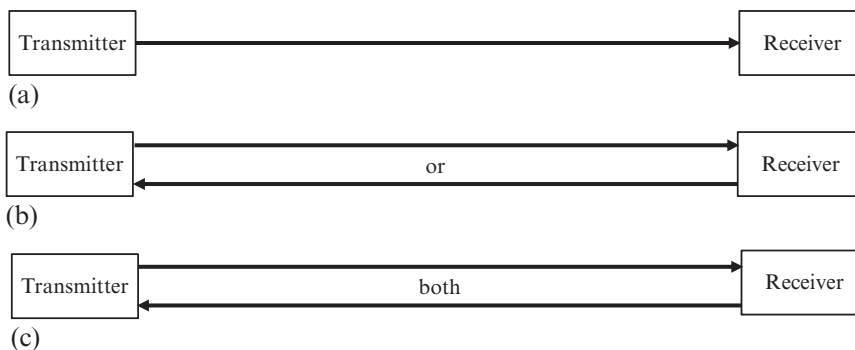


**FIGURE 11.2** Communication modes: (a) simplex communication, (b) half-duplex communication, and (c) full-duplex communication.

In *half-duplex mode*, transmission is bi-directional, but only in one direction at a time. The end devices can both transmit and receive, but not at the same time. In half-duplex mode, the entire capacity of the channel is taken over by whichever of the two devices is transmitting at the time. Examples of half-duplex communication are walkie-talkies and CB radios as well as dial-up fax machines.

In *full-duplex mode*, transmission is simultaneously bi-directional. The end devices can transmit and receive simultaneously. A prime example of full-duplex communications is the PSTN that provides two-way communications. ITU-T voice-band dial-up data modems can also provide full-duplex communications in one of the following three ways:

- Using four telephone wires, with a pair for each direction, such as V.26 modem.
- Dividing up the bandpass spectrum of a telephone channel into two disjoint frequency bands, with an exclusive band for each direction, such as V.22 bis modem.
- Employing echo cancellation (i.e., subtraction of the echo of its own transmission from the incoming signal to recover the signal sent by the other side), while allocating the entire spectrum to both ends, such as V.32 modem.

The means to provide two independent channels for two-way communications is called **duplexing**. A duplexing scheme is a crucial element of a wireless communication system, and thus needs to be carefully defined from outset in a system design. There are two distinct methods to achieve duplexing: frequency-division duplexing (FDD) and time-division duplexing (TDD), where each has its own merits. FDD is a legacy method, and has been used in both analog and digital communication systems. FDD examples include all communication satellite systems, many cellular mobile systems, such as AMPS and GSM, as well as public safety standards, such as TETRA. TDD is exclusively used in digital communication systems. TDD examples include digital cordless phones, such as DECT, wireless Bluetooth devices, and IEEE802.11 WLAN (Wi-Fi). Some 4G mobile systems, such as IEEE802.16 WiMAX and LTE, can employ both FDD and TDD.

### 11.2.1 Frequency-Division Duplexing

In *frequency-division duplexing* (FDD), two disjoint bands of frequency are provided to users. With FDD, two different carrier frequencies, one from each band, are assigned to a user; one carrier frequency for transmission from the user (also known as upstream, return link, or uplink) and one carrier frequency for reception by the user (also known as downstream, forward link, or downlink), as shown in [Figure 11.3a](#). A guard band is required between the two

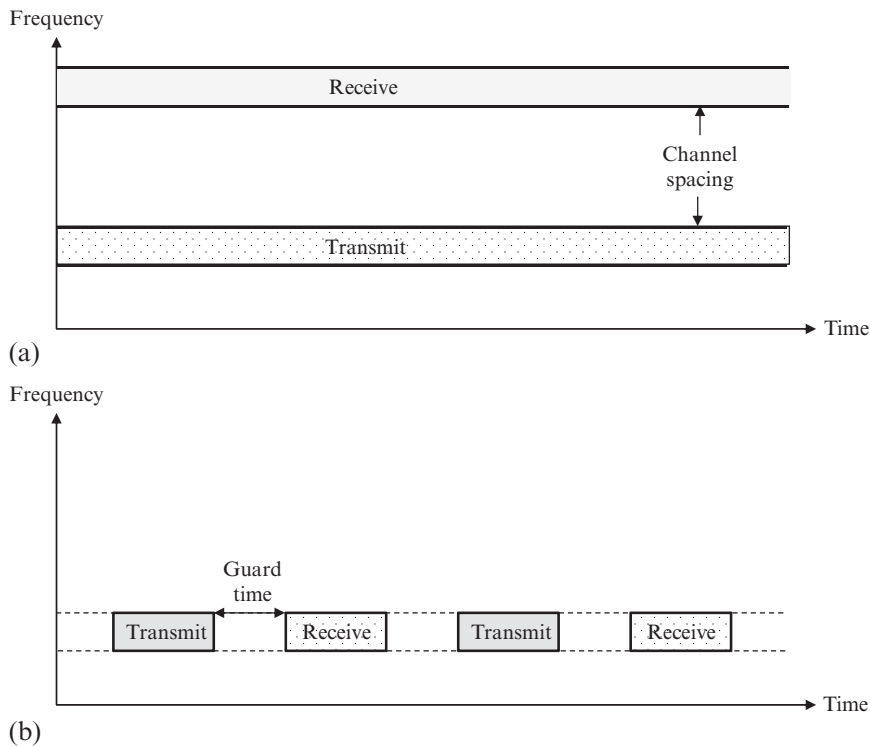


**FIGURE 11.3** Duplexing: (a) frequency-division duplexing and (b) time-division duplexing.

bands. The difference between the two assigned frequencies, also known as the *frequency split*, remains constant, and must be large enough to allow the use of low-cost techniques to separate the two signals; otherwise, the transmit and receive signals may be adversely affected by one another.

Since the same antenna is generally employed for both transmission and reception, a *duplexer* is needed to separate the transmitted and received signals. This additional hardware complexity in the handheld transceivers can be a cost-driving factor. The frequency plan for FDD, as set by a regulatory body or limited by the functionality of the available equipment, warrants fixed equal bandwidths, simply because it is not normally possible to reallocate spectrum from one direction to another. This thus makes it ideal for symmetric communications, such as voice applications, with almost identical information flows in both directions. Although FDD provides true simultaneous transmission and reception, it is not always spectrally efficient, since the traffic may be asymmetric.

### 11.2.2 Time-Division Duplexing

In *time-division duplexing* (TDD), time rather than frequency is used to separate the transmission and reception of the signals, and thus a single frequency is assigned to a user for both directions. TDD provides quasi-simultaneous bidirectional flow of information. Duplexers are therefore not required, and the cost of a TDD system is not very high, the reason lies in the fact that the transmitter and receiver use the same components, such as filters and mixers. With TDD, two time slots, one for upstream (transmission) and one for downstream (reception), are assigned to a user, with a short data burst in each direction, as shown in [Figure 11.3b](#). A guard time between transmit and receive streams is required.

Time split between the forward and reverse channels is so small that the transmission and reception appear to be simultaneous and continuous to users. The guard time is intended for a time allowance for the round-trip propagation delay (i.e., the time that it takes the signal to travel the geographical distance between the transmitter and receiver and back). This time interval must be sufficiently long so the transmit and receive signals do not clash. TDD is thus employed where generally the distance between the transmitter and receiver is small; otherwise, the channel efficiency drops as the time guard then needs to be rather long. TDD can increase the number of time slots in favor of one direction over the other. This capability to increase capacity in one direction (i.e., to dynamically handle asymmetric traffic) is an advantage, since it can efficiently accommodate the Internet applications.

## 11.3 MULTIPLE ACCESS

*Multiple access*, also known as *channelization*, allows remote sharing of a common transmission medium by many users, in which the medium is partitioned into separate communication channels and then dedicated to particular users upon their requests. Multiple access techniques can be classified into three broad categories: frequency-division multiple access (FDMA), time-division multiple access (TDMA), and code-division multiple access (CDMA). These major access techniques are used to share the available bandwidth in a wireless communication system. In general, channelization techniques are not suitable for the transmission of bursty data traffic, and instead lend themselves to continuous, steady traffic data flow. Since a fixed assignment strategy would be wasteful of resources and degradation to capacity, access channels are not assigned on a permanent basis. Initial channel assignment is generally done by a random-access channel. FDMA, TDMA, and CDMA are access methods in the data-link layer.

### 11.3.1 Frequency-Division Multiple Access

Frequency-division multiple access is the oldest of all multiple access schemes, and was used in the first demand assignment system for satellites (i.e., single channel per carrier (SCPC)) and the first-generation of cellular mobile systems (i.e., the advanced mobile phone system (AMPS)). In *frequency-division multiple access* (FDMA), the available channel bandwidth is divided into many nonoverlapping frequency bands, where each band is dynamically assigned to a specific user to transmit data. In an FDMA system, signals, while occupying their assigned frequency bands, can be transmitted simultaneously and continuously without interfering with each other. In FDMA, there is a central controller that allocates the frequency band to users, solely based on their needs. This is usually done during the call set up. Once a band is allocated to a user, it then belongs to the user exclusively for the continuous flow of information during the call. To prevent interference, the allocated bands are separated from one another by small guard bands. In other words, FDMA allows the users to transmit simultaneously, but over disjoint frequency bands, a user exploits a fixed portion of the band all the time, as shown in [Figure 11.4a](#). FDMA is best suited for connection-oriented applications; it is, however, inefficient in terms of utilization of power and bandwidth. If an FDMA channel is not in use by the user, then it sits idle and cannot be used by other users. It has poor spectral efficiency, since guard bands must be employed to avoid overlapping of the adjacent channels, and this in turn reduces the channel capacity.

Bandpass filters are used to confine the transmitted energy within the assigned band and tight RF filtering is required to minimize adjacent channel interference. Duplexers must be employed since both the transmitter and receiver operate at the same time. FDMA is a low-cost technology to implement, and has low transmission overhead. Synchronization in FDMA is simple, once it is established during the call set up, it can be easily maintained, as transmission occurs continuously. At the hub or base station, due to the significant susceptibility to nonlinear effects of power amplifiers, such as signal spreading and intermodulation, either each FDMA channel must employ its own amplifier or a highly-linear amplifier with a significant back-off power is required for the transmission of the composite signal. FDMA is generally used in combination with other multiple access schemes, where the spectrum is divided into large sub-bands. Each sub-band serves a large group of users, where within a group, another multiple access method can be employed. Examples include employing FDMA/TDMA in the GSM cellular mobile systems or in satellite systems offering VSAT applications.

### 11.3.2 Time-Division Multiple Access

Time-division multiple access is a multiple access scheme, which is widely used in VSAT and broadband satellite systems and the GSM cellular mobile systems.

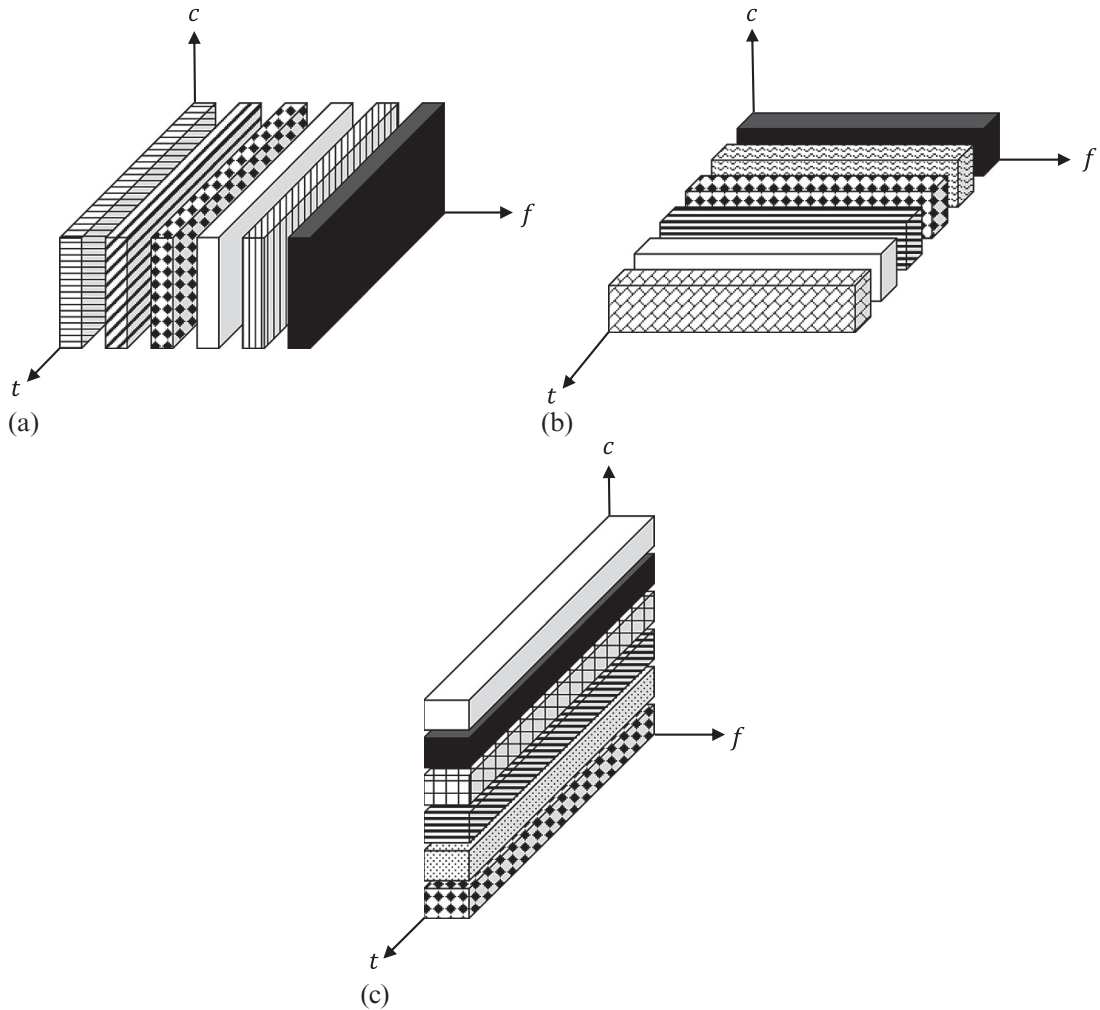


**FIGURE 11.4** Multiple access: (a) frequency-division multiple access, (b) time-division multiple access, and (c) code-division multiple access.

In *time-division multiple access* (TDMA), the available channel bandwidth in its entirety is used by every user, but the users take turns in making use of the channel in a timely manner. In other words, the channel is sequentially time-shared among many users through nonoverlapping time slots in a circular manner (i.e., one after the other). A user is allowed to transmit in a buffer-and-burst mode only in its allocated time slot. The transmission for any user is thus noncontinuous, and that in turn requires transmission to be in digital form. The noncontinuous transmission also results in low battery consumption, since the transmitter can be turned off when not in use, which is most of the time. All users

employ the same frequency band to transmit their signals, as shown in [Figure 11.4b](#). In TDMA/FDD systems, the frame structure for forward transmission is the same as that for reverse transmission. However, in TDMA/FDD systems, such as GSM, for a given user, there are several time slots of delay between the forward and reverse time slots, a duplexer is thus not needed in a transceiver.

Synchronization in TDMA is a critical requirement. To accomplish synchronization for geographically-dispersed users with different propagation delays, guard times are required to ensure transmissions do not overlap. Another source of overhead is a preamble consisting of pre-determined bits to allow the receiver to synchronize to the transmitted bit stream. In TDMA, a central controller sends a certain bit stream to all users with which all transmitters synchronize their clocks. Because of the higher data rate and subsequently the shorter symbol duration, adaptive equalization is usually required. In TDMA, it is possible to allocate different numbers of time slots per frame to different users for their different traffic patterns (i.e., the bandwidth-on-demand feature can be accommodated). In a mobile environment, TDMA yields a lower level of co-channel interference, as only a small percentage of users are transmitting, and the hand-off process is more efficient, since the transmission is not continuous. Since the burst transmission rate is quite high, a frequency band with a large bandwidth is required. In general, the spectrum is divided into many disjoint bands (i.e., FDMA) and then in a given band, TDMA is employed. A prime example of FDMA/TDMA is the GSM cellular mobile system, in which a frequency band of 200 kHz is allocated to eight TDMA users, where the burst rate is about 271 kbps.

### 11.3.3 Code-Division Multiple Access

*Spread spectrum* is a multiple access through which a transmission bandwidth several orders of magnitude greater than the minimum required RF bandwidth is used. This is achieved by using a pseudorandom noise-like spreading code in the transmitter to spread the bandwidth (i.e., to make it into a wideband signal), and employing the same code in the receiver to despread the received signal so as to recover the original message signal. Spread spectrum is not very efficient when used by a single user, but when there are many users sharing the same bandwidth without interfering with one another, spread spectrum systems become bandwidth efficient. Spread spectrum provides a high degree of immunity to multipath interference. Two forms of spread spectrum are used: *frequency-hopping spread spectrum* (FHSS) and *direct-sequence spread spectrum* (DSSS), also known as *code-division multiple access* (CDMA). In FHSS, the spectrum is widened by changing the carrier frequency in a pseudorandom manner. FHSS is used in Bluetooth wireless devices. Frequency hopping does not cover the entire spread spectrum at any time, as such we can select the rates at which

hopping can occur. In the *slow-frequency hopping*, the symbol rate is an integer multiple of the hop rate, whereas in the *fast-frequency hopping*, the hop rate is an integer multiple of the symbol rate. An FHSS system provides a level of security.

In DDSS, the baseband modulated signal is created by multiplying the modulating signal, which consists of information bits, by a pseudorandom sequence, whose bandwidth is much larger than that of the signal itself, thereby spreading the bandwidth. In CDMA, all users employ the same frequency band to simultaneously transmit (i.e., there is no clear separation in time or frequency). In CDMA, each user is assigned a unique randomized code sequence, which is different from and orthogonal to (i.e., uncorrelated with) all other codes. A code is used to generate the randomized noise-like high-bit-rate signal that is mixed with the information signal to spread the spectrum and the receiver uses the code to recover the desired signal. As shown in Figure 11.4c, users transmit over the entire frequency band at the same time, while using their own distinct codes. Because of bandwidth re-use, CDMA is bandwidth efficient, and delivers gradual degradation in performance as the number of users is increased. In a CDMA system, there is a *near-far problem* that occurs when the received power levels from different users do not appear to be equal. It is thus a critical requirement to provide a precise automatic power control mechanism. Power control is implemented by rapidly sampling the radio signal strength indicator levels and then sending a power change command over the forward link. As channel data rates are very high, the symbol (chip) duration is very short and much less than the channel delay spread. The reception can be improved by collecting the time-delayed versions of the signal. CDMA is employed in cellular mobile systems, such as CDMA-2000.

## 11.4 RANDOM ACCESS

*Random access* involves dynamic sharing of the transmission medium among many users so as to better match bursty data traffic. Transmission is random among users and since there is no scheduled time, access to the medium is based on contention. Users access the channel when they have one packet or more to transmit. Simultaneous transmissions from multiple users result in collision. The resulting conflict must thus be resolved by some channel protocol for retransmission of the packets. The random-access schemes are simple to implement and under light traffic can provide low-delay packet transfer. Under heavy traffic, the throughput and packet delays both can suffer.

### 11.4.1 ALOHA

The unslotted ALOHA protocol, also called pure ALOHA or original ALOHA, is a simple random-access scheme, which was designed for a radio (wireless) LAN

in the early seventies. In *unslotted ALOHA* scheme, whenever a user has a packet, it is transmitted. However, because there is only one communication channel, there is the possibility of collision among packets from different users. When the data from two or more users collide, they become garbled and are lost to the system. The unslotted ALOHA scheme relies on acknowledgments, if the acknowledgement does not arrive after a time-out period, the user then assumes that the packet or the acknowledgement has been destroyed and retransmits the packet. When the time-out period passes, each transmitter waits a random amount of time before resending its packet. In other words, the unslotted ALOHA scheme requires the transmitter to use a *back-off algorithm*, which chooses a random number in a certain retransmission time interval. The randomness can help spread out the transmissions and thus reduce the likelihood of additional collision among transmitters. In principle, when traffic is very light, the probability of collision is very small and retransmissions are thus hardly required.

In an unslotted ALOHA, there are two distinct modes. In one mode, collision occurs occasionally, this occurs when the traffic load is light. In the other mode, when the traffic is heavy, there is a surge of collisions and this in turn results in a snowball effect (i.e., the increased number of backlogged transmitters increases the likelihood of additional collisions). As shown in Figure 11.5a, the vulnerable

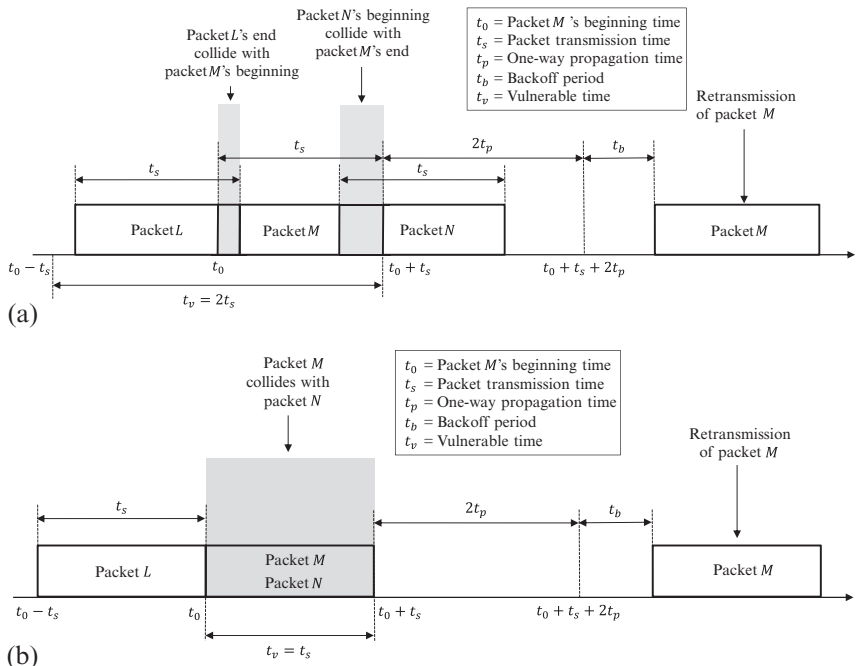


**FIGURE 11.5** Vulnerable time: (a) unslotted ALOHA and (b) slotted ALOHA.

time that is the length of time in which there is a possibility of collision is twice as long as the time required for the transmission of a packet. In other words, for a transmitter to send its packet collision free, no other transmitter should transmit less than half of the vulnerable time before this transmitter starts transmission and no transmitter should start sending during half of the vulnerable time after this transmitter has sent its packet.

Noting that the probability of a successful transmission (i.e., the probability of no collision) is the same as the probability of no transmissions during the vulnerable time, the *throughput*  $S$  (i.e., the average number of successfully transmitted packets) is as follows:

$$S = Ge^{-2G} \quad (11.1)$$

where  $G$  is the average number of packets generated during one packet transmission time. The maximum value of the throughput  $S$  is  $\frac{1}{2e} \cong 18.4\%$  that occurs at  $G = 0.5$  and corresponds to a total arrival rate of exactly one packet per vulnerable period. This is a reflection of the fact that two or more arrivals in a vulnerable period result in a collision.

For a given value of the throughput  $S$ , there are two different values of the load  $G$  corresponding to the two possible modes. One mode is when the arrival rate  $G$  is modest (i.e.,  $G \approx S$ ); in other words, when the traffic is considered to be on the light side, and the other is when the traffic is regarded to be on the heavy side (i.e.,  $G \gg S$ ); in other words, when many transmitters are backlogged. It is, however, important to note that there is no clear demarcation between the two modes. In unslotted ALOHA, there is no rule that defines when the transmitter can transmit, as such it has a vulnerable time twice the packet transmission time and its performance is thus rather poor.

The slotted ALOHA scheme can improve the ALOHA performance by reducing the probability of collisions. This is achieved by constraining the transmitters to send their packets in synchronized fashion. In *slotted ALOHA*, the time is divided into equal time slots, where each time slot is equal to the packet transmission time. In slotted ALOHA, the transmitters keep track of transmission time slots and are forced to send their packets only at the beginning of a time slot. This means that if a transmitter misses the beginning of a time slot, it must wait until the beginning of the next time slot. Of course, there is still the possibility of collision if two transmitters try to send at the beginning of the same time slot. However, as shown in [Figure 11.5b](#), the vulnerable time that is the length of time in which there is a possibility of collision is then equal to the time required to send a packet.

Noting that the probability of a successful transmission, that is the probability of no collision, is the same as the probability of no transmissions during the

vulnerable time, the throughput  $S$  (i.e., the average number of successfully transmitted packets) is as follows:

$$S = Ge^{-G} \quad (11.2)$$

where  $G$  is the average number of packets generated during one packet transmission time. The slotted ALOHA scheme, similar to the unslotted ALOHA scheme, may display a bimodal behavior. The maximum value of the throughput  $S$  is  $\frac{1}{e} \cong 36.8\%$  that occurs at  $G = 1$  and corresponds to a total arrival rate of exactly one packet per vulnerable period. This is a reflection of the fact that two or more arrivals in a vulnerable period result in a collision and if one packet is generated during one packet transmission time, then 36.8% of these packets reach their destination successfully.

## 11.4.2 CSMA

An ALOHA scheme with an essentially uncoordinated access to transmission medium has a low maximum throughput, which is primarily due to packet collisions. In *carrier sense multiple access* (CSMA), the chance of collision is minimized by sensing the medium for the presence of a carrier signal from other transmitters. CSMA is based on the principle of sense before transmit, that is, each transmitter in the network must first check the state of the medium (i.e., determine whether there is an ongoing transmission) before sending a packet. However, because of the propagation delay (i.e., the time needed for a signal to go from one end of the medium to the other), it is still possible for a collision to occur. When a transmitter sends a packet, it takes time to reach every other transmitter and for every transmitter to sense it. A transmitter may sense the medium and find it idle, only because the packet sent by another transmitter has not yet been received. Thus, the vulnerable period consists of one propagation delay, and if no other transmitter sends a packet during this period, the transmitter will in effect have the channel for its transmission simply because no other transmitter will transmit thereafter.

As shown in [Figure 11.6](#), when a transmitter finds the channel busy in CSMA, there are three different methods that it can employ to portray its behavior: 1-persistent method, nonpersistent method, and  $p$ -persistent method. All three approaches are sensitive to the end-to-end propagation delay of the medium.

In *1-persistent CSMA*, if the channel is busy, the transmitter with a packet to send continuously senses the channel until the channel becomes idle. As soon as the channel is sensed idle, the transmitter sends its packet (with probability 1). This method has the highest chance of collision, for two or more transmitters with ready packets may find the channel idle and send their packets immediately. Transmitters that are involved in a collision perform the back-off algorithm to

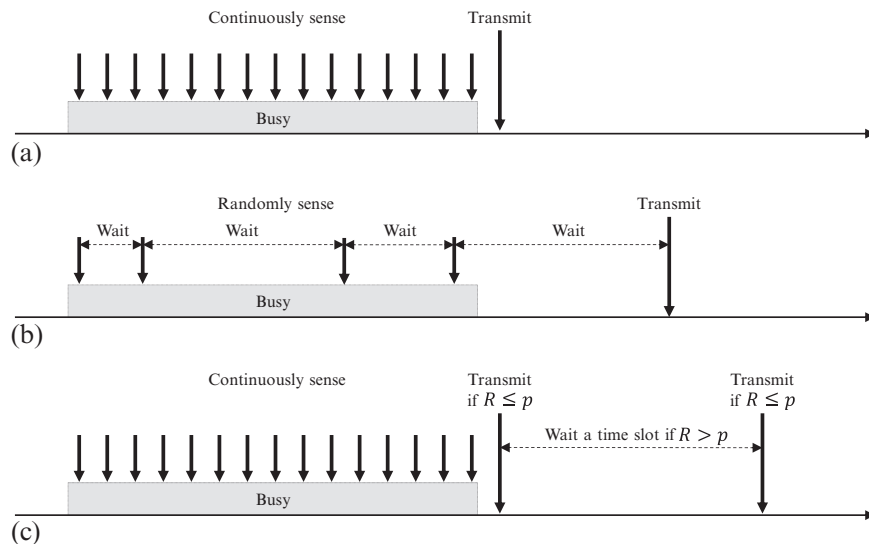


**FIGURE 11.6** Behavior of persistence methods: (a) 1-persistent, (b) nonpersistent, and (c) p-persistent.

schedule a future time to resense the channel. Since the transmitters behave in a greedy fashion, 1-persistent CSMA has a relatively high collision rate.

In *nonpersistent CSMA*, if the channel is busy, the transmitter with a packet to send immediately runs the back-off algorithm, and reschedules a future resensing time. By immediately rescheduling a resensing time and not persisting, the chance of collision is reduced relative to 1-persistent CSMA. This is due to the fact that it is unlikely that two or more transmitters will wait the same amount of time and retry to send immediately. This method, vis-à-vis 1-persistent CSMA, yields a longer delay and lower network efficiency. The reason lies in the fact that the channel may remain idle when there may be transmitters with packets to send.

Suppose the channel has time slots with a slot duration equal to the round-trip propagation time. In *p-persistent CSMA*, a transmitter with a packet to send senses the channel, if the channel is busy, it persists with sensing until the channel becomes idle, if the channel is idle, with probability  $p$ , the transmitter sends its packet, and with probability of  $1 - p$ , it decides to wait for the beginning of the next time slot and senses the channel again. If the channel is then idle, with probability  $p$ , it transmits, and if the channel is busy, the transmitter waits for a random period of time (i.e., a back-off period) as if a collision has occurred. This approach combines the advantages of the other two methods. This method spreads out the transmission attempts by the transmitters that have been

waiting for transmission to be completed and hence to increase the likelihood that a transmitter with a packet successfully seizes the channel.

In CSMA, if two transmitters with packets to send begin sending at the same time, each will transmit its complete packet, even though they collide. In the *carrier sense multiple access with collision detection* (CSMA/CD) scheme, the amount of wasted bandwidth can be reduced by terminating its transmission as soon as a collision is detected and the transmitter then attempts to transmit later after waiting for a random amount of time. CSMA/CD, which is commonly used in wired LANs, such as Ethernet, yields a throughput that is greater than those of the ALOHA schemes.

In a wireless environment, either transmission or sensing is done, that is both simultaneous transmission and sensing is not possible and the collision detection time is twice the end-to-end propagation delay. CSMA/CD cannot be used in a wireless environment of a shared channel. In wireless networks, the collision detection of CSMA/CD is unreliable due to the *hidden-node problem*, i.e., when a node is visible from a wireless access point, but not from other nodes communicating with that access point. In *carrier-sense multiple access collision avoidance* (CSMA/CA), before a transmitter sends its packet, it has to broadcast a signal onto the network in order to listen for collision scenarios and to tell other devices not to broadcast. Collisions are avoided by forcing transmitters to send reservation messages when they find the channel is idle. When a packet is transmitted, it is then done in its entirety.

## 11.5 CONTROLLED ACCESS

To achieve high throughputs in medium access control, some form of coordination is required. In *controlled access*, sharing a transmission medium is based on scheduling, i.e., a transmitter cannot send unless it has been authorized to do so. The systems employing controlled-access schemes are sensitive to the reaction time in the form of propagation delay and network latency.

### 11.5.1 Reservation

In a *reservation* system, a transmitter needs to make a reservation before sending data. Time is divided into cycles that can be variable in length. Each cycle consists first of a reservation interval and then an interval during which the transmitters with reservations send data. The number of mini-slots in a reservation interval corresponds to the number of transmitters in the system. Transmitters use their corresponding mini-slots in the reservation interval to broadcast that they have a packet to transmit in a corresponding cycle. By listening to the reservation interval, the transmitters can determine the order of transmissions in the corresponding cycle.

If the number of transmitters in a network is very large, then the overhead associated with the reservation interval becomes significant. To alleviate this problem, a mini-slot is not allocated to each transmitter. Instead, transmitters contend for a reservation mini-slot by using a random-access technique, such as slotted ALOHA. In fact, this is the way packet service is provided over GSM cellular telephone networks.

### 11.5.2 Polling

In *polling*, transmitters take turns accessing the medium. At any given time, only one transmitter can send data, after it is done, the right to transmit is passed to another transmitter. There are various methods for passing the right to transmit from one transmitter to another. In one method, a central controller periodically polls the transmitters to determine their service request. The central controller polls the transmitters sequentially in a circular fashion or another polling order. Polling can also be implemented in a distributed fashion on networks with a ring topology, which is known as token passing method. A token goes around the ring; once a transmitter has it, then the transmitter has the right to access channel and send its packet data. In a way, receiving a token corresponds to receiving a polling message. When the transmitter has no more data to send, it releases the token, thus passing it to the next transmitter in the ring.

If the numbers of transmitters is very large and the traffic is bursty, the time required to poll all transmitters in the network can be an excessive overhead burden. One efficient way for rapidly polling a large number of transmitters with light traffic is called *probing* or *tree search*. In that, rather than polling individual transmitters, groups of transmitters are polled. A group polling message is broadcast to all transmitters in a group, all transmitters in the group that have packets to send then respond, if multiple transmitters respond, a collision will occur. Subsequently, that group is sub-divided into two sub-groups, the process of probing will then proceed, until the transmitter with a packet is isolated, and then it transmits its packet.

## 11.6 WIRED COMMUNICATION NETWORKS

A *network* is a set of devices connected by a number of communication links to provide information transfer among users located at different places. In assessing data-communication networks, there are a number of measures to consider. Performance is an important measure that can be evaluated by several networking metrics, including:

- **Throughput:** The actual rate at which information, in the form of bits and packets, is successfully transmitted over the network.

- **Delay:** It consists of processing, buffer, transmission, and propagation delays associated with transfer of data packets.
- **Jitter:** It represents the delay variation experienced by received data packets.
- **Accuracy:** It is primarily measured by bit error rates and packet-loss rates.

There are of course other measures of high importance, such as network reliability, availability, and security, to name just few.

### 11.6.1 Circuit-Switched and Packet-Switched Networks

A *switched network* consists of a series of interlinked nodes by which temporary connections among nodes, including the intermediate switches and the end devices, can be made. There are two distinct methods of switching. One is circuit switching that occurs at the physical layer. The other one is packet switching, which in turn is divided into two sub-categories. One is the virtual-circuit packet switching that is normally done at the data link layer, and the other one is the datagram packet switching that normally occurs at the network layer. It is not necessary for the layers to operate in the same switching method.

A *circuit-switched network* is made of a set of switches connected by physical links, in which each link is divided into multiple channels. Communication via circuit switching involves three phases: setup phase, data transfer phase, and teardown phase. In circuit switching, the resources need to be reserved during the setup phase, and the resources remain dedicated, without interruption, for the entire duration of data transfer until the teardown phase.

Since the link is exclusively dedicated to the communicating pair, even during intervals when the communication path happens to be idle, circuit switching is not efficient. Despite its low efficiency, a circuit-switched network has minimal delay, as during data transfer the data are not delayed at each switch. Also, in a circuit-switched network, data rate is guaranteed. The PSTN is a prominent application of circuit switching, as it can provide interconnection for voice signal exchange between telephones, data transfer between dial-up modems, and image transfer between fax machines.

In a *packet-switched network*, there is no resource reservation and resources are allocated on demand. The allocation is done on a first-come, first-served basis. Packet switching is a method of transmitting data in which long messages are sub-divided into packets. The size of the packet is determined by the network and the governing protocol. Each packet is passed from source to destination through intermediate nodes. When a switch receives a packet, the packet must wait (i.e., be stored) if there are other packets being processed. There are two categories of packet-switching; one is *datagram*, also known as *connectionless packet switching*, and the other is the *virtual-circuit packet switching*, also known as *connection-oriented packet switching*.

In a *datagram network*, each packet, referred to as a datagram, is routed independently of all others. It is normally done at the network layer. In a datagram network, the destination address in the header of a packet remains the same during the entire transmission of the packet, and a switch uses a routing table that is based on the destination address to route each packet. Although call setup time is avoided, which is an advantage for short messages, each packet may experience a wait at a switch before it is forwarded. Since not all packets in a message necessarily travel through the same switches (i.e., the same path) the delay is not uniform for the packets of a message. The packets may arrive out of order and resequencing may be required at the destination. By sharing the transmission line among multiple packets, a high utilization, at the expense of queuing delays, can be achieved. The Internet Protocol (IP) and the User Datagram Protocol (UDP), which are the core members of the Internet protocol suite, both employ connectionless packet switching.

A *virtual-circuit packet-switched network* is a cross between a circuit-switched network and a datagram network. On the one hand, as in a circuit-switched network, there are call setup and teardown phases in addition to the data transfer phase, and on the other hand, as in a datagram network, data are packetized. However, the address in a header defines what the next switch should be and not the final destination. In a virtual-circuit packet switched network, all packets follow the same path established during the connection, and arrive in sequence. Due to not requiring routing decisions for each packet, packet transmission and delivery may take less time. For sustained packet flows, such as long file transfers, the Internet downloads, audio and video streaming, virtual-circuit packet switching is appropriate. Transmission Control Protocol (TCP), used in the Internet protocol suite, provides reliable connection-oriented packet switching.

## 11.6.2 Topology

In the context of communication networks, *topology* refers to the way in which a network is physically laid out, that is, how the end-point devices, also known as network nodes, attached to the network are inter-linked. Topologies can be broadly categorized as follows: star, mesh, bus, ring, and tree. As each of them possesses its own merits and drawbacks, a hybrid topology may be implemented in a communication network, such as star clusters used with both ring and bus topologies to create the star-ring and the star-bus hybrid topologies. The focus of topologies discussed here is on wired networks, as it is easy to implement any type of topology in a wireless network.

In a *star topology*, network nodes are directly connected by point-to-point links to a central node, usually called a hub, as shown in [Figure 11.7a](#). Typically, each node requires two links to connect to a central hub, one for transmission and

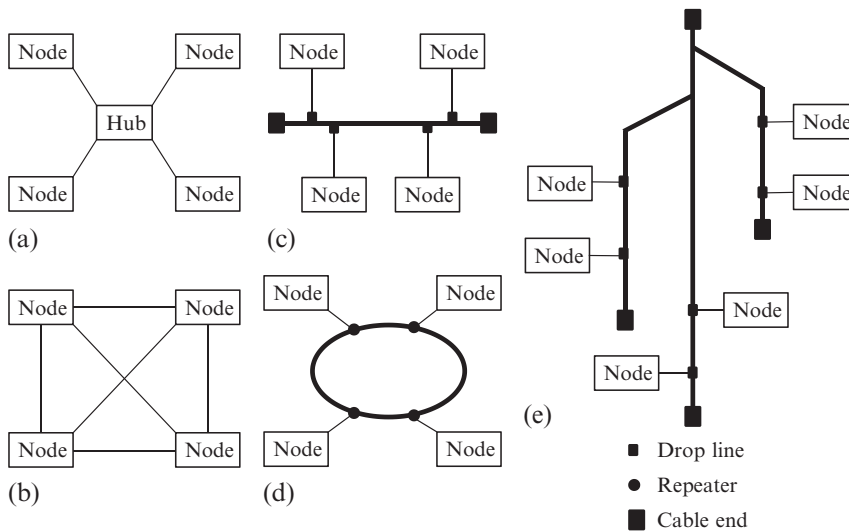


**FIGURE 11.7** Topology: (a) star, (b) mesh, (c) bus, (d) ring, and (e) tree.

one for reception. In a star topology, there is no direct link between the nodes. In case of node-to-node communications requirement, a node must send the data to the hub, which then relays to another node, hence double-hub communications. Applications of the star topology may include the local loops in the PSTN, satellite VSAT networks, and 1-Gbps Ethernet LANs.

A star topology has the benefit of easy installation and reconfiguration of access nodes. This reduces the overall network costs and makes network management much easier. It also has the advantage of being robust, when one link fails, all other links remain active, which in turn brings about easy fault identification and isolation. In a star topology, a variety of network media can be used, making it easier to merge legacy and new networks. In a star topology, transmissions are virtually collision free. The huge drawback of star is its vulnerability to the hub failure that can cause a complete suspension of network operation. To enhance reliability, hub redundancy is an important consideration for this topology. Also, in a wired network environment, cable costs can be higher than they would be in some other topologies.

In a **mesh topology**, network nodes are all connected by dedicated point-to-point links to one another. In a mesh network of  $n$  nodes,  $n(n-1)/2$  two-way connections are required, as shown in Figure 11.7b. Each device must have  $n-1$  input/output ports. Applications of the mesh topology may include the connection of regional offices in the PSTN, satellite telephone networks using SCPC to provide telephony in remote and isolated areas, and connecting main computers of a hybrid network consisting of other topologies.

Dedicated links in a mesh topology avoids the sharing problems. Mesh topology is very resilient to failure, for there are always many alternative routes in case a link becomes unavailable. Dedicated lines between any two nodes can bring about privacy and security. Point-to-point links make fault identification and isolation easy, as traffic can be re-routed to avoid links with problems. Despite its many benefits, mesh topology is implemented in a limited scope. This is due to the fact that the amount of cabling and the number of input/output ports is quite significant, and the installation and reconfiguration are quite difficult. Implementation of a mesh wired network is thus quite costly, whereas a wireless mesh network is virtually cost free.

In a *bus topology*, network nodes are all connected to a long cable called bus. A transmission from a node propagates along the length of the bus in both directions and can be received by all other nodes, as shown in [Figure 11.7c](#). A bus topology can provide point-to-multipoint connections. To avoid transmission collision and monopolization of transmission by a node, a mechanism to regulate transmission is thus required. An application is early low-speed Ethernet LANs.

As the signal travels along the backbone bus, it becomes weaker, and that can limit the number of nodes and appropriate hardware interfacing. In a bus topology, the problem of ringing (i.e., signal reflection from the cable ends) needs to be resolved, for they can cause degradation in signal quality. In a bus topology, installation is easy with modest cabling. Additional nodes, which can be easily added by more taps to the network backbone, can slow the network for all nodes. Moreover, a fault in the bus cable stops all transmission, and reconfiguration is difficult, simply because new nodes may require modification of the backbone bus. Privacy and security is poor, for the signal from a node is broadcast to all.

In a *ring topology*, each node is allowed to have dedicated point-to-point connections with only two neighboring nodes in a closed loop, as shown in [Figure 11.7d](#). The links are unidirectional (i.e., the data is passed along the ring in one direction, from one node to another, until it reaches its destination). When a node receives a signal from a neighboring node that is intended for another node, it regenerates the bits and passes them along the ring. The IBM Token Ring local area network employs ring topology.

In a ring topology, a single node cannot monopolize the network and its resources. Since in a ring topology, adding or removing a node requires changing only two connections, installation reconfiguration, and fault isolation are all easy. In fault isolation, an alarm is issued once a signal is not received by a node within a specified period. A break in the ring, such as a disabled node, can bring the network down. This is viewed as a significant drawback, and a dual ring may thus be used. Also, adding nodes to the ring can disrupt the entire ring.

**Tree topology** is a general form of bus topology. The tree layout begins at a point known as the head-end. Each cable starting at the head-end may have branches, and branches in turn may have branches, as shown in [Figure 11.7e](#). The structure is hierarchical, and the nodes at higher levels are thus more important. Applications may include cable television network and some lightly-loaded LANs.

A tree topology is not robust to failure, because there is only one path between any pair of nodes. A transmission from a node is broadcast to all, i.e., it is received by all other nodes; it is therefore not very secure. To resolve this issue and also prevent a node to transmit continuously, nodes must transmit data in blocks, known as frames, where each frame has its own destination address and control information.

### 11.6.3 Routing and Flow Control

**Routing** is concerned with the paths travelled by packets through the network. The objective is to determine the best paths in a network for packets to follow from a source to a destination. **Flow control** is concerned with the timing and rate of packet transmissions, and its objective is to ensure that the source does not overwhelm the destination with packets. We briefly outline some aspects of routing and flow control from a high-level perspective, and the details of various algorithms implementing these two major network functions are beyond the scope of this chapter.

The criteria to find the best path connecting a source and a destination may include the minimum number of links (hops), the minimum overall end-to-end delay, the highest user or network throughput, the lowest overall cost, the highest overall reliability, the lowest packet loss rate, the minimum buffer requirements, the maximum number of users that can be supported, and network fairness (i.e., users are receiving a fair share of system resources). However, minimum-delay routing is often used and routers thus exchange information to get a picture of the delays throughout the network. This in turn helps to balance loads, thus decreasing local congestion.

In order to identify the set of links that lead to the best path, a routing algorithm with global knowledge about the state of the network is required. Overall, an effective routing algorithm is to accommodate as many goals as possible. They may include rapid and accurate delivery of packets, adaptability to node or link failure as well as to varying traffic loads, ability to avoid congested links, and low overhead.

Flow control in datagram and virtual-circuit packet networks is done differently. In connectionless networks, sources do not interact with the network prior to the transmission of data. It is thus possible that sources overload the

network, causing buffer overload and packet loss. The sources with packet loss retransmit the lost packets, and that in turn will cause more buffers to overflow and require further retransmissions. This leads to a decrease in the network throughput to an unacceptable level. There are various techniques for flow control, and we briefly describe the one implemented in the Internet. The network nodes that experience congestion, that is, buffer overflow, send the network nodes causing the congestion special control packets requesting the transmit nodes to decrease the rate at which packets are injected into the network. However, for this scheme, there is no standard to specify under what conditions control packets should be sent and how the receiving nodes must react to control packets.

Flow control in connection-oriented networks may be done at two levels. The first level is when a connection request is made, and the network responds whether it has the required resources to provide the connection. The second level is when a connection is made, in that the packet rate and the burstiness of sources are controlled by the network.

#### 11.6.4 Local Area Networks

A *local area network* (LAN) links a number of telecommunications equipment, while spanning a relatively small area. LANs are capable of transmitting at very high rates with limited distances. Wired LANs of various types have been in operation for the past several decades. Today's wired LANs are several orders of magnitude faster than the earlier ones. Today, a LAN not only connects computers and smart phones and allows sharing of resources in a small geographical area, but more importantly, a wired LAN is now connected to the Internet.

Of all wired LANs, which existed over many years, the only survival is *Ethernet*. As such, our discussion of wired LANs is limited to Ethernet. The IEEE 802.3 standard, known as Ethernet, comprises four generations of Ethernet: i) standard (traditional) Ethernet at 10 Mbps, ii) fast Ethernet at 100 Mbps, iii) 1-Gigabit Ethernet at 1 Gbps, and iv) 10-Gigabit Ethernet at 10 Gbps.

All IEEE LANs sub-divide the data link layer, i.e., the second layer of the communication network protocol, into two sub-layers: logical link control (LLC) and medium access control (MAC). The *MAC sub-layer*, the lower portion of the data-link layer which interfaces with the physical layer and performs part of framing function, defines the specific access method for each type of LAN. For instance, it defines CSMA/CD for Ethernet LANs and the token-passing for Token Ring LANs. The *LLC sub-layer*, the upper part of the data-link layer which interfaces with the network layer, handles flow control and error control

and part of framing tasks. This allows interconnectivity between different LANs, as it makes the MAC sub-layer transparent.

Table 11.1 presents some of the main characteristics associated with various Ethernet LANs. All Ethernet LANs employ baseband signaling, in that, the bits employ a line code to send the digital signal on the line. In 10-Mbps and 100-Mbps Ethernet LANs, the control access method to the sharing medium is CSMA/CD with 1-persistent method. 1-Gbps and 10-Gbps Ethernet LANs, however, operate in a switched full-duplex mode, in that there is no need for CSMA/CD method, as each node is connected to the switch via two separate links and there is no need for collision detection.

| Table 11.1 Ethernet LANs     |                                           |                |                |              |
|------------------------------|-------------------------------------------|----------------|----------------|--------------|
| LAN types                    | Transmission medium                       | Maximum length | LAN topology   | Line coding  |
| <b>10-Mbps Ethernet LAN</b>  |                                           |                |                |              |
| 10Base5                      | Thick coaxial                             | 500 m          | Bus            | Manchester   |
| 10Base2                      | Thin coaxial                              | 185 m          | Bus            | Manchester   |
| 10Base-T                     | Twisted pair                              | 100 m          | Star           | Manchester   |
| 10Base-F                     | Optical fiber                             | 2 km           | Point-to-point | Manchester   |
| <b>100-Mbps Ethernet LAN</b> |                                           |                |                |              |
| 100Base-TX                   | Two unshielded twisted pairs              | 100 m          | Star           | 4B5B + MLT-3 |
| 100Base-FX                   | Two optical fiber multimode               | 2 km           | Star           | 4B5B + NRZ-I |
| 100Base-T4                   | Four unshielded twisted pairs             | 100 m          | Star           | Two 8B-6T    |
| <b>1-Gbps Ethernet LAN</b>   |                                           |                |                |              |
| 1000BaseSX                   | Two optical fiber multimode               | 550 m          | Star           | 8B-10B + NRZ |
| 1000BaseLX                   | Two optical fiber single-mode             | 5 km           | Star           | 8B-10B + NRZ |
| 1000BaseCX                   | Two shielded twisted pairs                | 25 m           | Star           | 8B-10B + NRZ |
| 1000BaseT                    | Four unshielded twisted pairs             | 100 m          | Star           | 4D-PAM5      |
| <b>10-Gbps Ethernet LAN</b>  |                                           |                |                |              |
| 10GBase-SR                   | Two optical fibers multimode at 850 nm    | 300 m          | Point-to-point | 64B-66B      |
| 10GBase-LR                   | Two optical fibers single-mode at 1310 nm | 10 km          | Point-to-point | 64B-66B      |
| 10GBase-EW                   | Two optical fibers single-mode at 1550 nm | 40 km          | Point-to-point | SONET        |
| 10Gbase-X4                   | Two optical fibers multimode at 1310 nm   | 0.3 - 10 km    | Point-to-point | 8B-10B       |

### 11.6.5 Telephone and Cable Networks

*Plain Old Telephone Service* (POTS), developed over a century ago, is the voice-grade telephone service that is based on analog signal transmission over a pair of twisted-wire connecting the user's telephone set and the central office. In the PSTN, the communication link between the customer's premises and the central office is known as the *last mile* or the *local loop* providing access to the end user. During the past half century, the PSTN has also been used, at a growing rate, for the transmission and reception of digital data, such as e-mail, fax, and data file transfer. The device that performs the conversion between the digital signal and a standard form suitable for transmission over a telephone channel in the PSTN is known as a *voice-band dial-up modem*. The fundamental role of such a modem is to transmit digital data over the telephone network in a transparent fashion, as the PSTN treats the transmitted digital data like the analog voice.

A large array of voice-band dial-up modems, including the ITU-T V-series modems over the telephone network, was developed between the early sixties and the late nineties. There are still millions of them being used in rural and remote areas to provide connectivity to the Internet. The ITU-T V-series recommendations on data communication over the telephone network specify the protocols that govern approved modem communication standards and interfaces. These standard modems were designed to have different capabilities, such as half-duplex and full-duplex communications, symmetric and asymmetric configurations, as well as adaptive bit rates. Their evolution started with the introduction of V.21 modem standard employing binary FSK operating at 300 bps and ended up with the development of V.34 modem standard employing 240-QAM with trellis-coding operating at 33.6 kbps, over a hundred times faster than the first generation of dial-up modems.

As discussed in Chapter 2, Shannon stated that the channel capacity for a band-limited, power-limited channel with additive white Gaussian noise (AWGN) is given by the following formula:

$$C = W \log_2 \left( 1 + \frac{S}{N} \right) \quad (11.3)$$

Note that the channel bandwidth  $W$  is measured in Hz and the signal-to-noise ratio at the receiver input in linear scale (not in dB). In a typical telephone line, the bandwidth  $W$  is roughly 3.4 kHz and the line signal-to-noise ratio is about 30 dB. Using (11.3), the capacity of a typical telephone network then turns out to be around 34 kbps. It is thus fair to say that V.34 modem approached the capacity that Shannon promises for a typical telephone channel.

In response to providing faster modems over the PSTN, the V.90 *modem standard* was developed in the late nineties. V.90 modem, also known as

**56K modem**, may be used only if the other end is using digital signaling, such as the digital modem at the Internet service provider (ISP). This higher data rate is due to the higher SNR, which in turn is due to no quantization noise, for there is no sampling involved in downlink. To arrive at the 56-kbps figure, there are 8000 samples per second and 8 bits per sample; however, one of the bits in each sample is used for control purposes, hence 56 kbps. The analog modem at the user's end for uplink is a V.34 modem standard. V.90 was anticipated to be the last standard for modems operating near the channel capacity of POTS lines to be developed. V.90 is generally used in concert with the V.42bis compression standard.

**Digital subscriber line (DSL)** refers to a group of digital technologies that can provide high-speed digital signal transmission over the existing twisted-wire pair in local loops. The term DSL is primarily a set of standards that define the central office interface. Through the central office, the user is connected directly to the broadband backbone network. The twisted-wire pair used in the local loop, originally designed for voice analog signals, is intentionally loaded, at regular intervals across the local loop, with coils. The extra inductance makes the frequency response fairly flat across the effective voice band. This helps to improve the voice quality, but at the expense of significant attenuation beyond 3.4 kHz. DSL uses an adaptive technology, in that it uses a bit rate based on the condition of the local loop. Factors such as the distance between the user location and the switching office, the size of the wire, the signaling type, the level of crosstalk interference from other lines, and the level of signal-to-noise ratio can all affect the available bandwidth.

DSL, which includes the family of ADSL, VDSL, HDSL, and SDSL, is often referred to as *xDSL*. A widely-used DSL technology offered by a number of service providers is the ADSL. **ADSL** stands for *asymmetric DSL*, in that, the speed in the downstream direction is higher than the speed in the upstream direction. The twisted-wire pair in the local loop potentially has a bandwidth of about 1.1 MHz, but the filter used at the end office where each local loop terminates limits the bandwidth to 4 kHz. Once the filter is removed, the entire 1.1 MHz bandwidth can be utilized for data transmission and voice communications. ADSL allows the subscriber to use voice and high-speed data at the same time, as shown in [Figure 11.8a](#). This is achieved by using frequency division multiplexing to divide the bandwidth into a voice channel, an upstream channel (25–138 kHz), and a downstream channel (138–1104 kHz). The lowest part of the band up to 25 kHz is reserved for POTS, which nominally requires a bandwidth of 4 kHz. The ADSL bandwidth is separated by a LPF for voice and a HPF for data, where both have cutoff frequencies of 25 kHz.

The ADSL uses **discrete multitone (DMT) modulation**, similar to OFDM. In that, multiple carrier signals at different frequencies are employed. The available bandwidth is divided into 256 parallel DMT channels, each consisting of

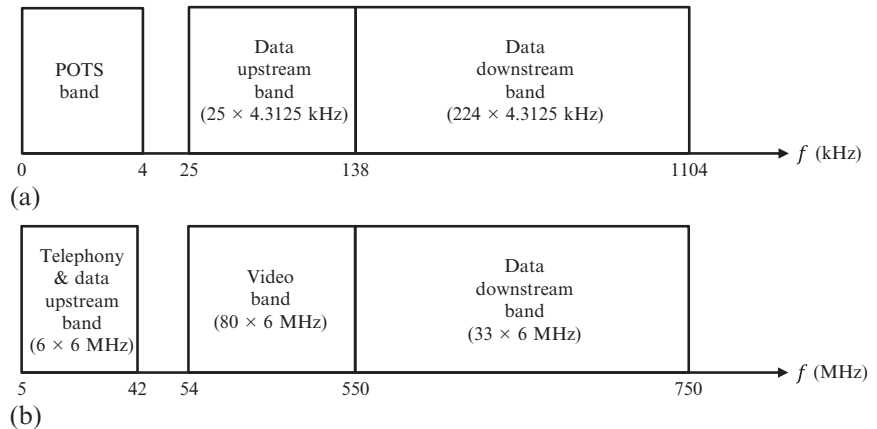


**FIGURE 11.8** Spectrum allocation: (a) ADSL and (b) CATV.

4.3125 kHz and capable of carrying about 60 kbps. It would be possible to have an upstream data rate of 1.44 Mbps and a downstream data rate of 13.4 Mbps. In practice, however, transmission impairments prevent attainment of these data rates. Current implementations of ADSL operate at from 1.5 to 9 Mbps downstream and at from 16 to 640 kbps downstream in a bandwidth of up to 1.1 MHz and over a distance of 2.7 to 5.5 km. VDSL does even better, as it can support data rates of 13 to 52 Mbps downstream and 1.5 to 2.3 Mbps upstream in an aggregate bandwidth up to 30 MHz and over a distance of 0.3 to 1.5 km.

Many decades ago, cable TV started to distribute broadcast TV signals to locations where reception was poor, and it was called Community Antenna Television (CATV). CATV used coaxial cable and due to signal attenuation and thus use of dozens of amplifiers, the cable network was unidirectional. During the past couple of decades, the second generation of cable network, called a hybrid fiber-coaxial (HFC) network, was developed. The use of fiber-optic cable reduces the need for amplifiers down to very few. This move from traditional to hybrid infrastructure made the cable network bi-directional. Yet, the last part of the HFC network, from the fiber node to the subscriber is still a coaxial cable.

As shown in Figure 11.8b, the coaxial cable part of HFC has a bandwidth ranging from 5 MHz to 750 MHz, consisting of 6-MHz channels. To provide bi-directional connectivity, the bandwidth is divided into three bands: video, upstream data and cable telephony, and downstream data. The video band occupies frequencies from 54 MHz to 550 MHz, thus accommodating more than 80 channels. The downstream-data band occupies the upper band from 550 MHz to 750 MHz, consisting of 33 channels. The upstream-data band occupies the lower band from 5 to 42 MHz, comprising 6 channels.

Both upstream and downstream bands are time shared by many subscribers. The channels are shared, statically or dynamically, by a group of subscribers. In short, a subscriber with data must contend for the channel with others. The same situation exists for downstream, the channel must be shared among a group of users using multicasting concept. The throughput per 6-MHz channel is about 10 to 30 Mbps for the upstream and about 42 Mbps for the downstream.

## 11.7 NETWORK SECURITY AND CRYPTOGRAPHY

Private, dedicated communication links among users are quite rare; instead, there are always public communication networks shared by many users. The security threats that can arise in public networks are numerous, the array of attacks is constantly widening, and network security is continually becoming more challenging. Cryptography is an effective means of combating network threats, especially in e-commerce and m-commerce applications. Although the area of cryptography is very mathematical, only a very brief and simple introduction to this topic is made here.

To make a network secure, there are network security requirements that must be met. They include:

- **Confidentiality**, i.e., ensuring the transmitted message is hidden from unauthorized parties.
- **Authentication**, i.e., verifying the communicating parties are those they claim to be.
- **Integrity**, i.e., confirming that the message content has not been tampered with.
- **Nonrepudiation**, i.e., not being able to deny the transmission between the two parties has taken place.

**Cryptography** is the science of making secret communication to make messages secure. By **encryption**, an original message, called **plaintext**, is transformed into a coded message, called **ciphertext**. The reverse process is called **decryption**. The algorithm used for encryption and decryption is often called a **cipher** and the process of encryption and decryption requires a **secret key**, without which the unauthorized parties must not be able to recover the original message.

A key is a number (value) that the cipher operates on. To encrypt a message, an encryption algorithm, an encryption key, and the plaintext are needed and to decrypt it, a decryption algorithm, a decryption key, and the ciphertext are required. The encryption and decryption algorithms are public, that is, anyone can access them, but the keys are secret and thus need to be protected.

There are two broad categories of cryptography: *private-key cryptography*, also known as *secret-key cryptography* and *symmetric-key cryptography*, and *public-key cryptography*, also referred to as *asymmetric-key cryptography*.

It is important to emphasize that there is a fundamental question as how secret keys in symmetric-key cryptography and public keys in asymmetric-key cryptography are efficiently and securely distributed and maintained. As the two widely used ways to distribute keys, we will briefly discuss the key distribution center (KDC) for symmetric-key cryptography and a certification authority (CA) for asymmetric-key cryptography. It is important to note that it is also possible to create a session key between two users without using a KDC or a CA. One such method is Diffie-Hellman key agreement, through which the two users create a secret shared key by using a series of exchanges over a public network.

11.7.1 Private-Key Cryptography

The private-key cryptography uses the same key for both the encryption at the transmitter and the decryption at the receiver, as shown in Figure 11.9a. In secret-key cryptography, decryption algorithm is the inverse of encryption algorithm, and the number of symmetric keys for  $N$  users to communicate is  $\frac{N(N-1)}{2}$ , as each pair must have a unique symmetric key. The number of keys grows quadratically, thus making these systems unfeasible for larger-scale use. The

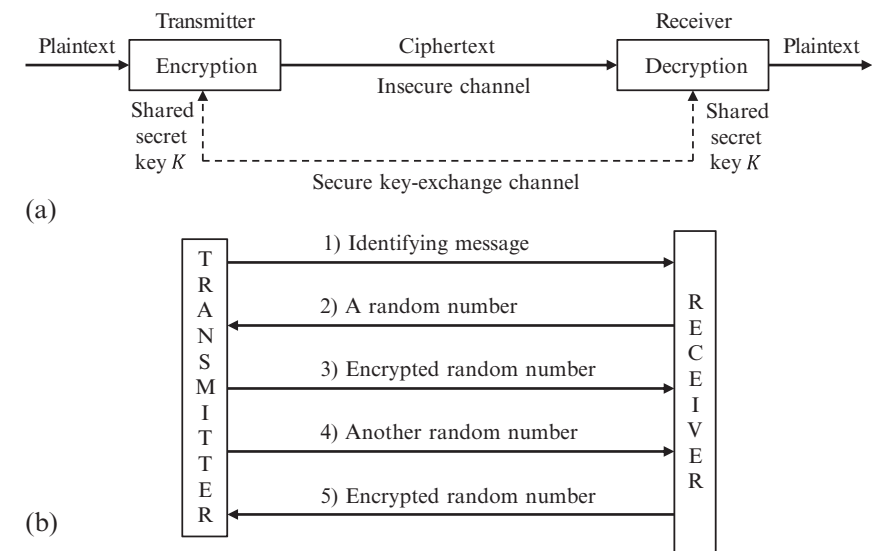


FIGURE 11.9 Private-key: (a) cryptography and (b) authentication.

symmetric-key cryptography is more efficient and widely practiced than asymmetric-key cryptography, and they are often used for long messages, for they also require less time to encrypt. An example of asymmetric-key cryptography is the Data Encryption Standard (DES). At the transmitter, a 64-bit plaintext is created into a 64-bit ciphertext, and at the receiver, a 64-bit ciphertext is converted back to a 64-bit plaintext, noting that in DES, the same 56-bit key is used for both encryption and decryption.

Figure 11.9b shows a secret-key authentication. The transmitter first sends a message identifying itself, the receiver, as a challenge, then sends back a message containing a random number. The transmitter then sends, as a response, an encrypted version of the random number. The receiver decrypts it by using the shared key. If the encrypted number is the same random number as the one sent earlier, then the receiver has authenticated the given transmitter. Similarly, the transmitter can then authenticate the receiver, of course, using another random number.

In symmetric-key distribution, a trusted third party, referred to as a *key distribution center* (KDC), is used. In symmetric-key cryptography, each user establishes a shared secret key with the KDC. The secret keys, created by the KDC, are used exclusively between the KDC and the users, and not among the users themselves. When a user wants to transmit secretly with another one, the transmitter then asks the KDC for a session (temporary) secret key to be used between the two users. A session symmetric key between two parties is used only once.

## 11.7.2 Public-Key Cryptography

The public-key cryptography is based on personal secrecy rather than sharing secrecy. In public-key cryptography, as shown in Figure 11.10a, two different keys, a public key and a private key are used. It is a salient requirement that it must not be possible to determine the private key from the public key. In general, the public key is small and the private key is large. A pair of keys can be used many times. The number of keys for  $N$  users to communicate is  $2N$ . The algorithm is complex and more efficient for short messages.

Figure 11.10b shows a public-key authentication. The transmitter first sends a message identifying itself. The receiver, as a challenge, then sends back a message containing a random number that has been encrypted by using the transmitter's public key. The transmitter then uses its own private key to determine the random number, and then responds by sending that random number. To develop a public-key cryptographic system is a complex task. A widely-used public cryptographic system is RSA (named after its inventors, Rivest, Shamir, and Adleman). The RSA algorithm is a block cipher based on modular arithmetic and factorization of a number that is the product of two large prime numbers.

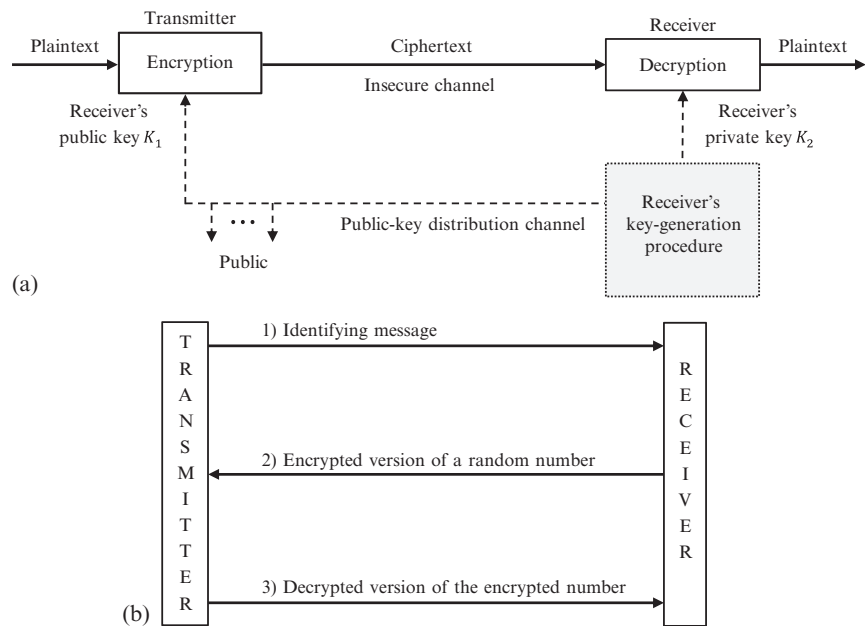


FIGURE 11.10 Public-key: (a) cryptography and (b) authentication.

In asymmetric-key distribution, a federal or state organization, known as a *certification authority* (CA), is used. In asymmetric-key cryptography, public keys, like secret keys, need to be distributed securely; the process can otherwise be subject to forgery. The CA first checks the identification of a user, asks the user's public key and writes it on the certificate. To prevent the certificate itself from being forged, the CA signs the certificate with its own private key, which is difficult to be forged. The user uploads the signed certificate. Anyone who wants a user's public key downloads the user's signed certificate and then uses the CA's public key to obtain the user's public key.

### 11.7.3 Digital Signatures

Message authentication protects two communicating parties from any third party; it does not, however, protect the two parties against each other. In situations where something more than authentication is needed, the digital signature is the most attractive solution, in addition, it can also prevent denial and forgery. Digital-signature signing and verification processes are shown in Figure 11.11. Public-key cryptography can be used to provide nonrepudiation by producing a digital signature. A *digital signature* uses a pair of private-public keys belonging to the sender to provide message integrity and message authentication.

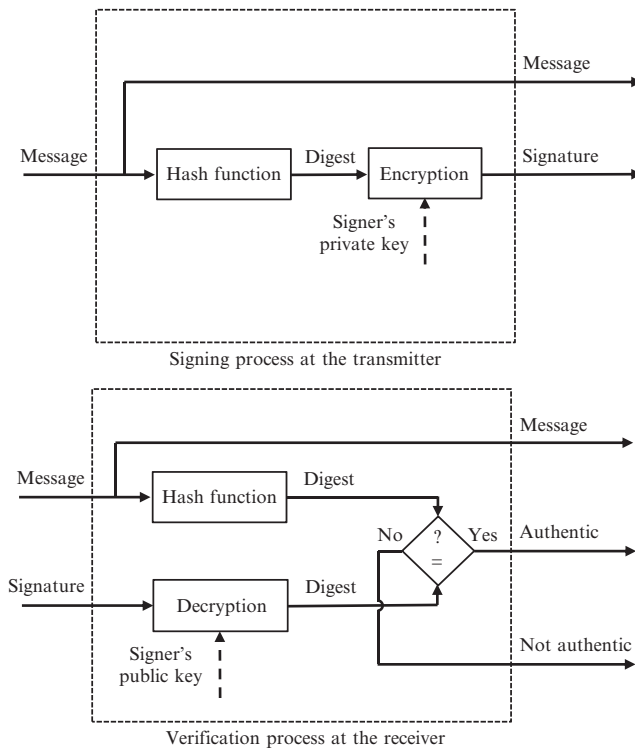


**FIGURE 11.11** Digital-signature signing and verification processes.

In a digital-signature system, on the one hand the messages are normally long, but on the other hand we must use asymmetric-key cryptography systems that are very inefficient when dealing with long messages. The solution is to create a *digest* of a message through a hashing function (e.g., a checksum is produced that is much smaller than the message). Therefore, the transmitter first produces a digest, then encrypts it using its own private key to produce the signature, and then sends the signature along with the message to the receiver. At the receiver, the signer's public key is applied to the signature to get the transmitted digest, and the digest is also directly determined from the message. If both digests are the same, then only the given transmitter could have issued the message and the message has maintained its integrity.

It is important to highlight the distinction between how private and public keys are used in digital signatures, vis-à-vis in public-key cryptography, for confidentiality. In a digital signature, the signer (transmitter) signs the message digest with the signer's private key and the verifier (receiver) verifies with the signer's public key. In contrast, in cryptography, the public and private keys of the receiver are used in the process.

## Summary and Sources

The area of communication networks is a broad area, and only few topics of importance in this chapter were covered briefly. We first provided brief descriptions of various forms of common multiplexing techniques, duplexing methods, and multiple access schemes. We then presented random and controlled-access schemes, along with their merits and drawbacks. The focus then turned towards telephone and cable networks. The chapter ends with network security and cryptography, in which security requirements are noted and private and public cryptography are highlighted.

There is no single book to fully discuss all areas covered in this chapter. However, there are excellent books on communications networks, such as Leon-Garcia and Widjaja [1], Forouzan [2], and Stallings [3]. There are also many references on various multiplexing, duplexing, and access methods [4–14].

- [1] Leon-Garcia and I. Widjaja, *Communication Networks - Fundamental Concepts and Key Architectures*, second ed., McGraw-Hill, ISBN: 0-07-119848-2, 2004.
- [2] A. Forouzan, *Data Communications and Networking*, fifth ed., McGraw-Hill, ISBN: 978-0-07-337622-6, 2013.
- [3] W. Stallings, *Data and Computer Communications*, ninth ed., Prentice-Hall, ISBN: 978-0-13-139205-2, 2011.
- [4] G.J. Mullett, *Basic Telecommunications: The Physical Layer*, Delmar, Cengage Learning, ISBN: 978-1-4018-4339-7, 2003.
- [5] Gitlin, J.F. Hayes and S.B. Weinstein, *Data Communications Principles*, Plenum Press, ISBN: 0-306-43777-5, 1992.
- [6] Sklar, *Digital Communications*, second ed., Prentice-Hall, ISBN: 0-13-084788-7, 2001.
- [7] J.D. Gibson, *The Communications Handbook*, second ed., CRC Press, ISBN: 0-8493-0967-0, 2002.
- [8] P.M. Shankar, *Introduction to Wireless Systems*, John Wiley and Sons, 2002. ISBN: 0-471-32167-2.
- [9] J.F. Hayes, *Modeling and Analysis of Computer Communications Networks*, Perseus Publishing, 1984. ISBN: 0306417820.
- [10] D.P. Agrawal and Q. Zeng, *Introduction to Wireless and Mobile Systems*, third ed., Delmar, Cengage Learning, 2011. ISBN: 978-1-4390-6205-0.
- [11] R.L. Freeman, *Fundamentals of Telecommunications*, second ed., IEEE-Wiley, 2005. ISBN: 0-471-71045-8.
- [12] International Telecommunications Union, *Handbook on Satellite Communications*, third ed., John Wiley & Sons, 2002. ISBN: 978-0-471-22189-0.
- [13] T.S. Rappaport, *Wireless Communications*, second ed., Prentice-Hall, 2002. ISBN: 0-13-042232-0.
- [14] J.J. Spilker Jr., *Digital Communications by Satellites*, Prentice-Hall, 1977. ISBN: 0-13-214155-8.

## Problems

- 11.1 Suppose 1000-bit packets are transmitted on a shared channel of 10 Mbps. Determine the throughput for the following two cases:
  - (a) The system is an unslotted ALOHA system with a packet rate of 1000 packets per second.
  - (b) The system is a slotted ALOHA system with a packet rate of 2000 packets per second.
- 11.2 Propose a strategy to improve the performance of probing algorithm.
- 11.3 Describe all major aspects of the TDMA used in GSM.
- 11.4 Describe all major aspects of the CDMA used in IS-95.
- 11.5 Describe all major aspects of the FDMA in AMPS.
- 11.6 Show bit rate equivalence of FDMA and TDMA.
- 11.7 Compare the message delay in FDMA and TDMA.