CHAPTER 1

OVERVIEW OF BROADBAND WIRELESS NETWORKS

COPYRIGHTED MATERIAL

Mobility and flexibility make wireless networks effective extensions and attractive alternatives to wired networks. Wireless networks provide all the functionality of wired networks, but without the physical constraints of the wire itself. However, the wireless link possesses some unique obstacles that need to be solved. For example, the medium is a scarce resource that must be shared among network users. It can be noisy and unreliable where transmissions from mobile users interfere with each other to varying degrees. The transmitted signal power dissipates in space rapidly and becomes attenuated. Physical obstructions may block or generate multiple copies of the transmitted signal. The received signal strength normally changes slowly with time because of path loss, more quickly with shadow fading and very quickly because of multipath fading. The most distinguishing issues in wireless network design are the constraints placed on bandwidth and power efficiency.

The broadcast nature of wireless transmission offers ubiquity and immediate access for both fixed and mobile users, clearly a vital element of quad-play (voice, video, data, and mobile) services. Moving from one location to another does not lead to disruptive reconnections at the new site. Wireless technology overcomes the need to lay cable, which is difficult, expensive, and time consuming to install, maintain, and especially, modify. Providing wireline connectivity in rural or remote areas runs the risk of someone pulling the cable (and accessories such as amplifiers) out of the ground to sell! A wireless network avoids underutilizing the access infrastructure. Unlike wired access (copper, coax, and fiber), a large portion of wireless deployment costs is incurred only when a customer signs up for service. The Fiber-to-the-Home (FTTH) Council reported that in September 23, 2008, there were 13.8 million FTTH networks in North America but the adoption rate is only 3.76 million (about 27%) even though many of these homes are located in strategic neighborhoods. The take up rate improved marginally to 34% (7.1 million connected homes) with 20.9 million homes passed on March 30, 2011. The cable industry's capital expenditure over the last 15 years is estimated at $172 billion. Broadband usage for cable services fared better but still fall below 50%. According to the National Cable and Telecommunications Association (NCTA), there were 129.3 million homes passed by cable video service in June 2011 (which translates to over 96% of U.S. households passed), but the take up rate is 45.5%. These numbers are unlikely to

---

Broadband Wireless Multimedia Networks, First Edition. Benny Bing.
© 2013 John Wiley & Sons, Inc. Published 2013 by John Wiley & Sons, Inc.

increase significantly in future with high-speed wireless and free broadcast services becoming widely available.

Terrestrial wireless access may offer portable and mobile service without the need for a proprietary customer premise equipment (CPE), such as a set-top box. This facilitates voice, TV, and Internet connectivity inside and outside the residential home. For instance, such connectivity can be made available on virtually any open space (e.g., on a fishing boat!), on fast moving vehicles and trains, and even when the subscriber moves to a foreign location. The ability to connect disparate end-user devices quickly and inexpensively remains one of the key strengths of wireless. New smartphones and tablets all come with two or more wireless network interfaces but no wired interfaces, thus making wireless connectivity indispensable. These devices demand higher wireless rates to support multimedia applications, including high-quality video streaming, which is in contrast to low bit rate voice applications supported by legacy cellular systems.

Because cellular systems cover long distances, they involve costly infrastructures, such as base stations (BSs) and require users to pay for bandwidth on a time or usage basis. Each BS may potentially serve a large number of mobile handsets. Coordination between BSs as users move across wireless coverage boundaries is achieved using a mobile backhaul, which also carries a variety of user traffic. The BS may prioritize near and far handsets. For example, the BS can reduce interference by transmitting at a lower power to closer handsets. In contrast, on-premise and geographically limited wireless local area networks (wireless LANs) require no usage fees, employ lower transmit power, and provide higher data rates than cellular systems. Wireless LANs are built around cheaper access points (APs) that connect a smaller number of stationary user devices, such as laptops or tablets to a wired network. However, achieving reliable high-speed wireless transmission is a challenging task. Besides the need to overcome traditional issues, such as multipath fading and interference from known and unknown sources, broadband wireless transmission also demands new methods to support highly efficient use of limited radio spectrum and handset battery power. This chapter discusses several fundamental topics related to broadband wireless networks. These include environmental factors, frequency bands, multicarrier operation, multiple antenna systems, medium access control, duplexing, and deployment considerations.

1.1 INTRODUCTION

Mobile broadband represents a multibillion dollar market. Service providers, including incumbent cable/telephone wireline providers, can increase the number of subscribers significantly by leveraging on broadband wireless solutions (e.g., in areas not currently served or served by competitors). The performance of a broadband wireless network is heavily dependent on the characteristics of the wireless channel, such as signal fading, multipath distortion, limited bandwidth, high error rates, rapidly changing propagation conditions, mutual interference of signals, and the vulnerability to eavesdrop and unauthorized access. Moreover, the performance observed by each individual user in the network is different and is a function of its location as well as the location of other interacting users. In order to improve spectral efficiency and hence, the overall network capacity, wireless access techniques need to be closely integrated with various interference mitigation techniques including the use of smart antennas, multi-user detection, power control, channel state tracking, and coding. Broadband wireless networks must also adequately address the combined requirements of wireless and multimedia communications. On one hand, the network must allow users to share the limited bandwidth resource efficiently to achieve higher rates. This implies two criteria: maximizing the utilization of the radio frequency spectrum and minimizing the delay experienced by the users. On the other hand, because the network supports multimedia traffic, it is expected to handle a wide range of bit rates together with various types of real-time and non–real-time traffic attributes and quality of service (QoS) guarantees.

More than 60% of Americans are using a wireless device to talk, send email, take pictures, watch video, listen to music, and play online games. Compressed video is a key traffic type that needs to be accommodated due to the emergence of many personal smartphones and tablet computers. Despite the smaller displays, many of these devices can support high-definition (HD) video with 720p (1280 × 720 pixels) picture resolutions. Highly efficient video coding standards, such as H.264/MPEG-4 Advanced Video Coding (AVC), are normally used to compress these videos for wireless delivery. This enables efficient use of radio spectrum, but the higher compression efficiency may also result in higher bit rate variability. In addition, compressed video is very sensitive to packet loss, with very limited time for packet retransmission, and wireless channels tend to be more error-prone than wired networks. Although wireless rates are typically lower than its wired counterparts, serving bandwidth-intensive applications, such as HD videos, may not always be an issue since such videos can be downloaded in the background. Users tend to watch videos in their own time, rather than according to broadcast schedules. However, real-time video applications (e.g., Skype video chat) may pose a problem depending on bandwidth availability.

Figure 1.1 shows the evolution of wireless access standards. Since the late 1990s, there were numerous digital cellular standards supporting second-generation (2G), 2.5 generation (2.5G), and third-generation (3G) services. These standards can be broadly categorized under code division multiple access (CDMA) or time division

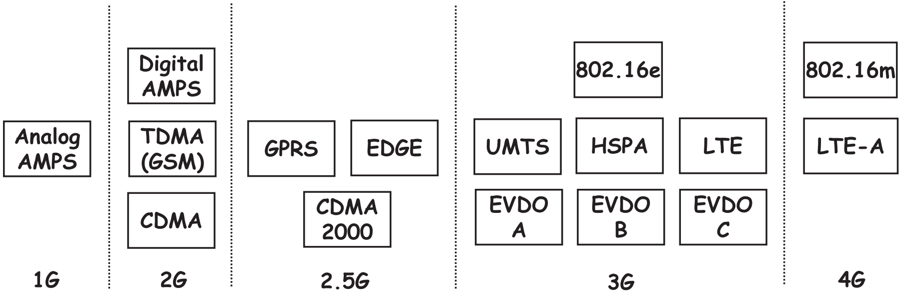


Figure 1.1 Evolution of wireless access standards.

multiple access (TDMA), and there was much debate on the individual merits and capacity of these systems. However, high-speed fourth-generation (4G) wireless standards are converging towards multicarrier transmission as the defacto method and currently there are only two 4G standards. Unlike legacy digital cellular services that primarily support voice and low rate data, the demand for 4G wireless is driven by millions of personal devices that require high-speed Internet connectivity. These sleek devices exude mass appeal due to their usability, and many devices employ an open software platform for users to program their own applications or download other applications. 4G wireless is the missing link that allows multimedia applications running on these devices to become portable, thus enabling on-the-go entertainment.

1.2 RADIO SPECTRUM

To encourage pervasive use of a wireless technology, the operating radio frequency (RF) band should be widely available. Locating a harmonized band is a difficult task because spectrum allocation is strictly controlled by multiple regulatory bodies in different countries. These include the Federal Communications Commission (FCC) in the United States, the European Committee of Post and Telecommunications Administrations (CEPT), Ofcom in the United Kingdom, the Radio Equipment Inspection and Certification Institute (MKK) in Japan, the Australian Communications and Media Authority (ACMA), and others.

1.2.1 Unlicensed Frequency Bands

Many wireless networks operate on unlicensed frequency bands, as illustrated in Figure 1.2. Many of these bands are available worldwide. Since the allocated spectrum is not licensed, large-scale frequency planning is avoided and ad hoc

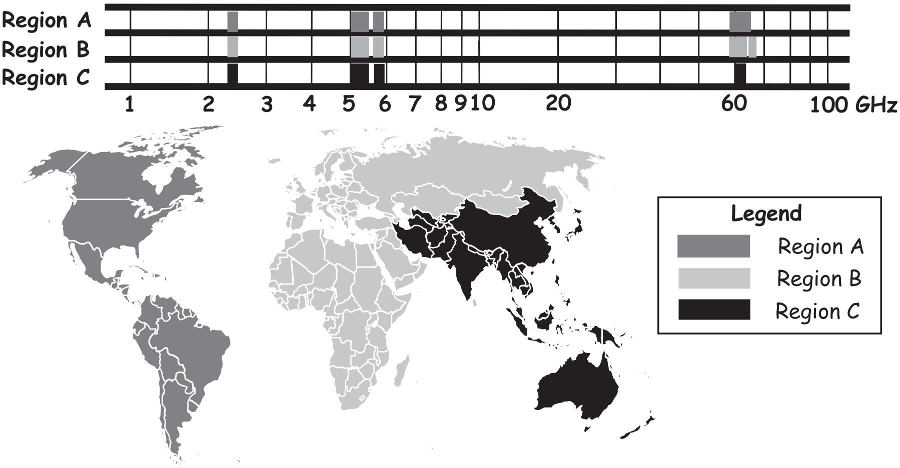


Figure 1.2 Popular unlicensed frequency bands at 2.4 GHz, 5 GHz, and 60 GHz.

deployments are possible. Perhaps the most popular band is the 2.4 GHz industrial, scientific and medical (ISM) band, which has been adopted by the IEEE 802.11 wireless LAN standard. The rules for operating in this band were first released by the FCC in 1985, which also includes the 900 MHz and the 5.8 GHz bands. Another band that has become popular is the 5 GHz Unlicensed National Information Infrastructure (U-NII) frequency band. The large amount of radio spectrum in this band enables the provision of high-speed Internet and multimedia services. The rules for operating in the 5 GHz U-NII band were released by the FCC in 1997. The band is subdivided into three blocks of 100 MHz each, corresponding to the lower, middle, and upper U-NII bands. The FCC subsequently expanded the middle U-NII band when 255 MHz of bandwidth was added in 2003. Thus, the 5 GHz U-NII band offers substantially more bandwidth than the 2.4 GHz band (580 MHz vs. 83.5 MHz). Recently, the 60 GHz unlicensed band has emerged, providing about 7–8 GHz of bandwidth, which is significantly higher than the 2.4 and 5 GHz bands.

1.2.2 The 2.4 GHz Unlicensed Band

The 2.4 GHz channel sets and center frequencies are shown in Table 1.1. The operating power for the 2.4 GHz band is limited to 100 mW (U.S.), 100 mW (Europe), and 10 mW/MHz (Japan). If transmit power control is employed, up to 1 W operation is permissible in the United States. A radiated power of 30 mW is normally used in 802.11 wireless LANs. In France, the output power for outdoor operation between 2454 and 2483.5 MHz is restricted to 10 mW. The total available bandwidth is 72 MHz in the United States, but generally higher in Europe and other parts of the world (83.5 MHz total bandwidth). The 2.4 GHz channels are defined on a

TABLE 1.1 2.4 GHz Channel Sets and Center Frequencies

Channel ID	North America (GHz)	Europe, Japan, and Australia (GHz)
1	2.412	2.412
2	2.417	2.417
3	2.422	2.422
4	2.427	2.427
5	2.432	2.432
6	2.437	2.437
7	2.442	2.442
8	2.447	2.447
9	2.452	2.452
10	2.457	2.457
11	2.462	2.462
12		2.467
13		2.472

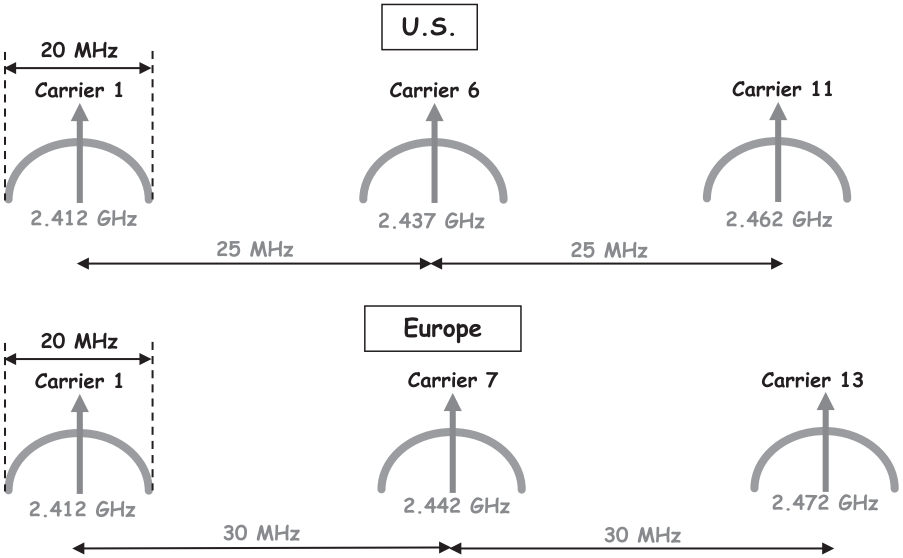


Figure 1.3 2.4 GHz channel spacing in the United States and Europe.

5 MHz channel grid (i.e., each channel has a bandwidth of 5 MHz). This results in 11-13 nonoverlapping channels. However, some wireless systems such as 802.11 wireless LANs require a higher channel bandwidth of 20 MHz. This gives rise to 11-13 overlapped 20 MHz channels or only 3 nonoverlapping 20 MHz channels in the 2.4 GHz band, as illustrated in Figure 1.3. These nonoverlapping channels need not be assigned sequentially and not all channels need to be assigned.

Interference must be carefully evaluated in the 2.4 GHz band especially when deploying large-scale 802.11 networks, such as in conference or exhibit halls and airport terminals. Besides 802.11 devices, many other devices operate in the same GHz band. They include 802.15.1 (Bluetooth) and 802.15.4 (ZigBee) devices, as well as digital cordless phones and microwave ovens that do not have in-built interference avoidance mechanisms. Microwave ovens remain one of the detrimental sources of interference in the 2.4 GHz band due to the high operating power (typically 750 to 1000 W). These ovens are present in almost every home and office. Due to the difficulty in controlling interference, devices must observe etiquette during transmission so that incompatible systems may co-exist. Basic elements of etiquette are to listen before transmit (to detect ongoing transmissions), and to limit transmit time and transmit power. The carrier (activity)-sensing mechanism used in 802.11 provides natural etiquette support.

1.2.3 The 5 GHz Unlicensed Band

Unlike the 2.4 GHz band, 5 GHz channels are defined on a 20 MHz grid. The channel sets, center frequencies, and operating power for the 5 GHz band are shown

TABLE 1.2 5 GHz U-NII Channel Sets, Center Frequencies, and Operating Power

Frequency band	Channel numbers	Center frequency (GHz)	Maximum output power in the United States	Maximum output power in Europe (EIRP)
5.15 to 5.25 GHz (100 MHz) (U-NII lower band or U-NII-1)	36	5.18	50 mW[a] (2.5 mW/MHz) (Indoor)	200 mW[c]
	40	5.20		
	44	5.22		
	48	5.24		
5.25–5.35 GHz (100 MHz) (U-NII middle band or U-NII-2)	52	5.26	250 mW[a] (12.5 mW/MHz) (Indoor/outdoor)	200 mW[c]
	56	5.28		
	60	5.30		
	64	5.32		
5.470–5.725 GHz (255 MHz) (U-NII-2 extended) (worldwide)	100	5.50	250 mW[a] (12.5 mW/MHz) (Indoor/outdoor)	1 W[c]
	104	5.52		
	108	5.54		
	112	5.56		
	116	5.58		
	120	5.60		
	124	5.62		
	128	5.64		
	132	5.66		
	136	5.68		
	140	5.70		
5.725–5.825 GHz (100 MHz) (U-NII upper band or U-NII-3)	149	5.745	1 W[b] (50 mW/MHz) (Outdoor point-to-point)	4 W[c]
	153	5.765		
	157	5.785		
	161	5.805		

[a] Maximum antenna gain is 6 dBi.

[b] Maximum antenna gain is 23 dBi.

[c] Requires DFS.

in Table 1.2. The 5 GHz channel assignment and bandwidth are shown in Figure 1.4. The U-NII lower and upper bands (four channels each) are normally employed in the United States, whereas the U-NII middle band is normally supported in Europe. Note that the center frequencies of the 20 MHz channels start (and end) at different points in each band. The highest number of contiguous channels of 11 is available in the 5.470–5.725 GHz middle band. A significant part of the 5.8 GHz ISM band (ranging from 5.725 to 5.850 GHz) has been absorbed in the upper U-NII band. However, ISM channel 165 with a center frequency of 5.825 GHz falls outside of the U-NII band. Thus, the total available bandwidth at 5 GHz (ISM + U-NII) is 580 MHz, giving 24 nonoverlapping channels. The upper U-NII band holds the most

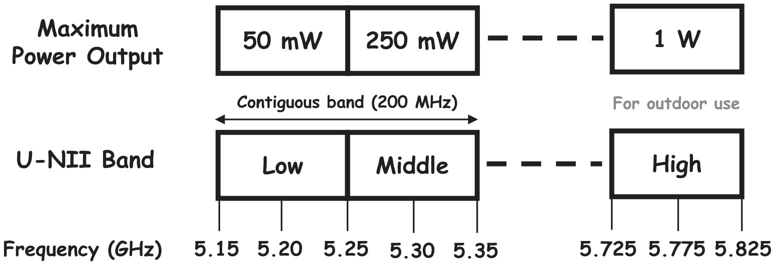


Figure 1.4 5 GHz channel assignment and bandwidth.

promise, as it allows the possibility of longer operational range without the need for a range extender.

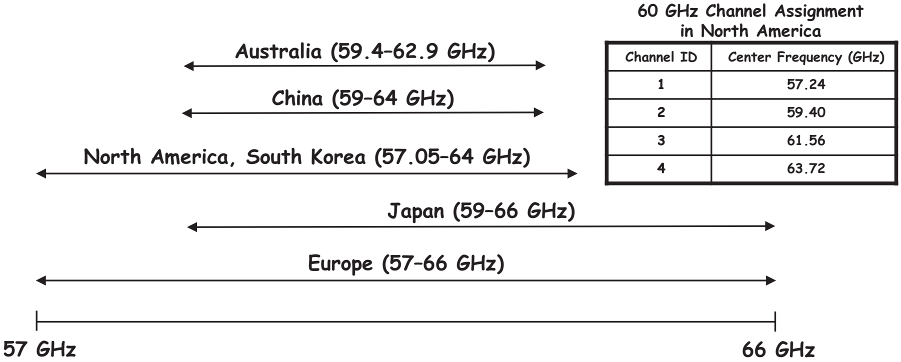
Some 5 GHz wireless channels overlap with radar frequencies. Unlike the 2.4 GHz band, however, most interference sources at 5 GHz are located outdoors and may be attenuated sufficiently if penetrated indoors. Nevertheless, dynamic frequency selection (DFS) and transmit power control (TPC) are interference mitigation techniques recommended for unlicensed 5 GHz operation. With DFS, a 5 GHz device can automatically detect radar transmissions and change to a different channel. With TPC, the transmit power can be reduced by several dB below the maximum permitted level.

1.2.4 The 60 GHz Unlicensed Band

A higher frequency band generally implies a higher amount of available bandwidth. At high frequencies, oxygen absorption in the atmosphere leads to rapid fall off in signal strength. Although the operating range becomes limited, this facilitates frequency reuse (i.e., bandwidth reclamation) in high-capacity, picocell (very small cell) wireless systems. There are two oxygen absorption bands ranging from 51.4 to 66 GHz (band A) and 105 to 134 GHz (band B). There are also peaks in water vapor absorption at 22 and 200 GHz. The oxygen absorption is lower in band B than band A, while the water vapor attenuation is higher. These observations suggest that band A is more suitable for communications than band B. The Millimeter Wave Communications Working Group first recommended the use of the 60 GHz band in a report released by the FCC in 1996 [1]. The 60 GHz frequency band is also available worldwide (Figure 1.5). It is uniquely suited for carrying extremely high data rates (multi-Gbit/s) over short distances. Transmit power is typically limited to 10 mW, and the band is subdivided into four nonoverlapping channels of 2.16 GHz each. Nearly 7 GHz of unlicensed spectrum is available in the United States and Japan, whereas 8 GHz of bandwidth is available in Europe. Currently, there are no coexistence issues.

1.2.5 Licensed Frequency Bands

The popular 3GPP Evolved Universal Terrestrial Radio Access (E-UTRA) licensed frequency bands are listed in Table 1.3. The downlink (DL) band is employed in the



60 GHz Channel Assignment in North America

Channel ID	Center Frequency (GHz)
1	57.24
2	59.40
3	61.56
4	63.72

Figure 1.5 60 GHz frequency allocation.

TABLE 1.3 Major E-UTRA Frequency Bands

E-UTRA band	Band name	UL band (MHz)		DL band (MHz)		Duplex mode
1	2.1 GHz	1920	1980	2110	2170	FDD
2	PCS 1900	1850	1910	1930	1990	FDD
3	1800 MHz	1710	1785	1805	1880	FDD
4	AWS	1710	1755	2110	2155	FDD
5	850 MHz	824	849	869	894	FDD
7	2.6 GHz	2500	2570	2620	2690	FDD
8	900 MHz	880	915	925	960	FDD
9	1700 MHz (Japan)	1749.9	1784.9	1844.9	1879.9	FDD
10	1.7/2.1 GHz	1710	1770	2110	2170	FDD
12	Lower 700 MHz	699	716	729	746	FDD
13	Upper C 700 MHz	777	787	746	756	FDD
14	Upper D 700 MHz	788	798	758	768	FDD
17	Lower B, C 700 MHz	704	716	734	746	FDD
18	850 MHz (Japan)	815	830	860	875	FDD
19	850 MHz (Japan)	830	845	875	890	FDD
20	CEPT800	832	862	791	821	FDD
21	1500 MHz (Japan)	1447.9	1462.9	1495.9	1510.9	FDD
24	U.S. L band	1626.5	1660.5	1525	1559	FDD
33	TDD 2000 lower	1900	1920	1900	1920	TDD
34	TDD 2000 upper	2010	2025	2010	2025	TDD
35	TDD 1900 lower	1850	1910	1850	1910	TDD
36	TDD 1900 upper	1930	1990	1930	1990	TDD
37	PCS center gap	1910	1930	1910	1930	TDD
38	IMT Ext	2570	2620	2570	2620	TDD
40	2300 MHz	2300	2400	2300	2400	TDD
41	U.S. 2600	2496	2690	2496	2690	TDD
42	3500 MHz	3400	3600	3400	3600	TDD
43	3700 MHz	3600	3800	3600	3800	TDD

Source: 3GPP TS.104 V10.2.0.

transmission from the BS to the handset. Conversely, the uplink (UL) band is employed in the transmission from the handset to the BS. High-speed packet access (HSPA) systems are deployed in the major E-UTRA cellular bands. For example, there are over 400 tri-band 850/1900/2100 MHz HSPA devices that support global roaming. Many HSPA devices also support legacy global system for mobile communications (GSM), general packet radio service (GPRS), and Enhanced Data rates for GSM Evolution (EDGE), giving rise to quad-band 850/900/1800/1900 MHz devices. A combination of higher spectrum (e.g., 1800/1900/2100 MHz) for improved capacity and sub-1 GHz spectrum (e.g., 700/850/900 MHz) for improved coverage in rural areas and urban in-buildings, is highly desirable. However, with 4G wireless standards employing larger bandwidths (such as 40 MHz), it is important to be able to use bands that offer wider bandwidths. Thus, the International Telecommunication Union (ITU) identified the 2.6 GHz band for supporting mobile broadband services. This extension band is large enough to allow operators to deploy wideband channels to achieve faster data speeds. In addition, some 700 MHz spectrum (also known as digital dividend spectrum) is released for 4G wireless services as analog TV broadcasters migrate to more efficient digital TV platforms. In the 2007 ITU World Radio Conference, the allocation of 700 MHz spectrum for mobile service has been harmonized in the following regions:

- 698–806 MHz for the Americas
- 790–862 MHz for Europe, Middle East, and Africa
- 698–862 MHz or 790–862 MHz for Asia.

The United States is currently the only country in the world that can build ubiquitous wireless Internet access and communications using the 700 MHz and 2.5 GHz (2.496 to 2.69 GHz) Educational Broadband Service (EBS) spectrum.

1.3 SIGNAL COVERAGE

Wireless networks employ either radio or infrared electromagnetic waves to transfer information from one point to another. The use of a wireless link introduces new restrictions not found in conventional wired networks. The quality of the wireless link varies over space and time. Objects in a building (e.g., structures, equipment, and people) can block, reflect, and scatter transmitted signals. In addition, problems of noise and interference from both intended and unintended users must also be solved. While wired networks are implicitly distinct, there is no easy way to physically separate different wireless networks. Well-defined network boundaries or coverage areas do not exist since users may move and transmissions can occur in various locations of the network. Wireless networks lack full connectivity and are significantly less reliable than the wired physical layer (PHY). Thus, one of the most important aspects of wireless system design is to ensure that sufficient signal levels are accessible from most of the intended service areas. To support mobility, separate wireless coverage areas or cells must be properly overlapped to ensure service continuity. Estimating signal coverage requires a good understanding of the communication channel, which comprises the antennas and the propagation medium. Usually, additional signal power is needed to maintain the desired channel quality and to offset the amount of received signal power variation about its average level. These power variations can be broadly classified under small-scale or large-scale fading effects. Small-scale fades are dominated by multipath propagation (caused by RF signal reflections), Doppler spread (caused by relative motion between transmitter and receiver), and movement of surrounding objects. Large-scale fades are characterized by attenuation in the propagation medium and shadowing caused by obstructing objects. These effects are explained in the following sections.

1.3.1 Propagation Mechanisms

Signal propagation patterns are unpredictable and changes rapidly with time. Consequently, signal coverage is not uniform, even at equal distances from the transmitter. A transmitted RF signal diffuses as it travels across the wireless medium. As a result, a portion of the transmitted signal power arrives directly at the receiver, while other portions arrive via reflection, diffraction, and scattering. Reflection occurs when the propagating signal impinges on an object that is large compared to the wavelength of the signal (e.g., buildings, walls, and surface of the earth). When the path between the transmitter and receiver is obstructed by sharp, irregular objects, the propagating wave diffracts and bends around the obstacle even when a direct line-of-sight (LOS) path does not exist. Finally, scattering takes place when obstructing objects are smaller than the wavelength of the propagating signal (e.g., people, foliage).

1.3.2 Multipath

Among the various forms of radio signal degradations, multipath fading assumes a high degree of importance. Multipath is a form of self-interference that occurs when the transmitted signal is reflected by objects in the environment such as walls, trees, buildings, people, and moving vehicles. When a signal takes multiple paths to reach the receiver, the received signal becomes a superposition of different components (Figure 1.6), each with a different delay, amplitude, and phase. These components form different clusters, and, depending on the phase of each component, interfere constructively and destructively at the receiving antenna, thereby producing a phenomenon called multipath fading. Such fading produces a variable bit error rate that can lead to intermittent network connectivity and significant delay variation (jitter).

Multipath fading represents the quick fluctuations in received power and is therefore commonly known as fast (or Rayleigh) fading. In addition, it is often classified as small-scale fading because the rapid changes in signal strength only occur over a small area or time interval. Multipath fading is affected by the location of the transmitter and receiver, as well as the movement around them. Such fading tends to be frequency selective or frequency dependent. Of considerable importance to

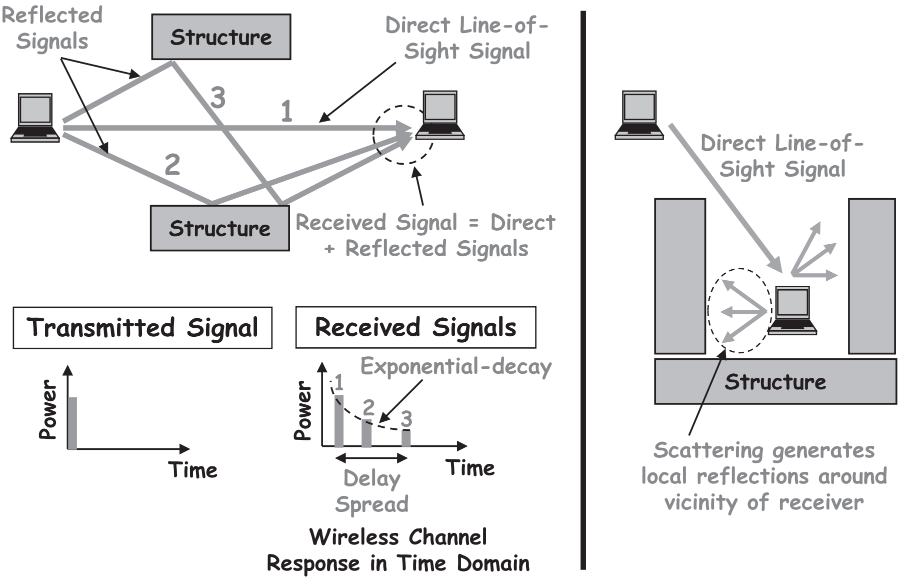


Figure 1.6 Multipath propagation and signal scattering.

wireless network designers is not only the depth but also the duration of the fades. Fortunately, it has been observed that the deeper the fade, the less frequently it occurs and the shorter the duration when it occurs. The severity of the fades tends to increase as the distance between the transmitter and receiver, and the number of reflective surfaces in the environment, increase. Multipath fading can be countered effectively using diversity techniques, in which two or more independent channels are somehow combined. The motivation here is that only one of the channels is likely to suffer a fade at any instant of time.

Since multipath propagation results in varying travel times, signal pulses are broadened as they travel through the wireless medium. This limits the speed at which adjacent data pulses can be sent without overlap, and hence, the maximum information rate a wireless system can operate. Thus, in addition to frequency-selective fading, a multipath channel also exhibits time dispersion. Time dispersion leads to intersymbol interference (ISI) while fading induces periods of low signal-to-noise ratio (SNR), both effects causing burst errors in wireless digital transmission. Figure 1.7 shows the impact of ISI. In this case, the same delay spread is assumed. Thus, while multipath propagation causes fast fading at low data rates, at high data rates (i.e., when the delay spread becomes comparable with the symbol interval), the received signals become indistinguishable, giving rise to ISI. Lowering the data symbol rate and/or introducing a guard time interval (also known as a cyclic prefix [CP]) between symbols can help mitigate the impact of time dispersion.

The performance metric for a wireless system operating over a multipath channel is either the average probability of error or the probability of outage. The average probability of error is the average error rate for all possible locations in the

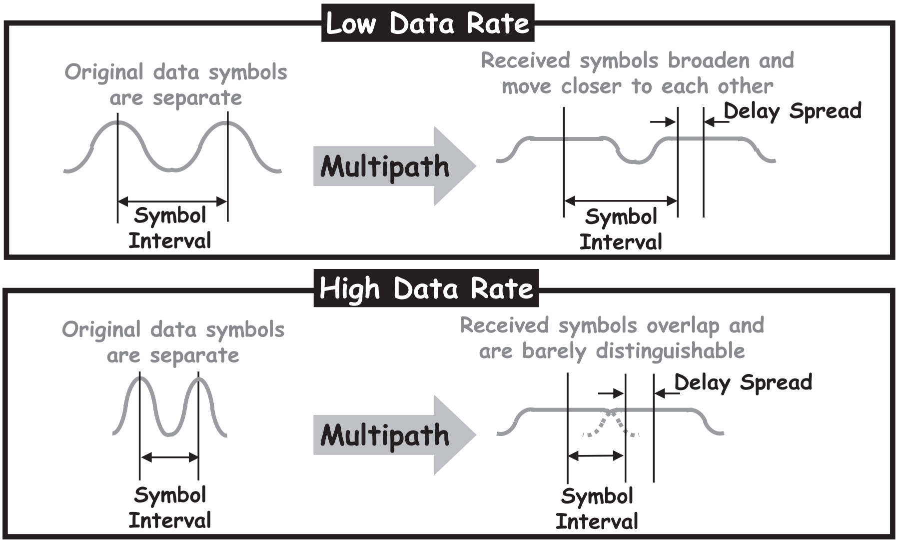


Figure 1.7 Intersymbol interference.

cell. The probability of outage represents the error probability below a predefined signal threshold for all possible locations in the cell.

1.3.3 Delay Spread and Time Dispersion

Delay spread is caused by differences in the arrival time of a signal from the various paths when it propagates through a time-dispersive multipath channel. The net effect of the arrival time difference is to spread the signal in time. The delay spread is proportional to the length of the path, which is in turn affected by the span of the propagating environment, as well as the location of the objects around the transmitter and receiver. The delay spread decreases at higher frequencies due to greater signal attenuation and absorption. A negative effect of delay spread is that it results in ISI. This causes data symbols (each representing one or more bits) to overlap in varying degrees at the receiver. Such overlap results in bit errors that increase as the symbol period approaches the delay spread. The effect becomes worse at higher data rates and cannot be solved simply by increasing the power of the transmitted signal. To avoid ISI, the duration of the delay spread should not exceed the duration of a data symbol, which carries a set of information bits.

The root mean square (rms) delay spread is often used as a convenient measure to estimate the amount of ISI caused by a multipath wireless channel. The maximum achievable data rate depends primarily on the rms delay spread and not the shape of the delay spread function. The rms delay spread in an indoor environment can vary significantly from 30 ns in a small room to 250 ns in a large hall. In outdoor environments, a delay spread of 10 μs or less is common (for a range of 1000 ft or less), although for non-LOS cases, a delay spread in the order of 100 μs is possible. If the product of the rms delay spread and the signal bandwidth is much less than

1, then the fading is called flat fading. If the product is greater than 1, then the fading is classified as frequency selective.

1.3.4 Coherence Bandwidth

A direct consequence of multipath propagation is that the received power of the composite signal varies according to the characteristics of the wireless channel in which the signal has traveled (Figure 1.8). More importantly, multipath propagation often leads to frequency-selective fading, which refers to nonuniform fading over the bandwidth occupied by the transmitted signal. The fades (notches) are usually correlated at adjacent frequencies and are decorrelated after a few megahertz. The severity of such fading depends on how rapidly the fading occurs relative to the round-trip propagation time on the wireless link. The bandwidth of the fade (i.e., the range of frequencies that fade together) is called the coherence bandwidth. This bandwidth is inversely proportional to the rms delay spread. Thus, ISI occurs when the coherence bandwidth of the channel is smaller than the modulation bandwidth. If the coherence bandwidth is small compared with the bandwidth of the transmitted signal, then the wireless link is frequency selective, and different frequency components are subject to different amplitude gains and phase shifts. Conversely, a wireless link is nonfrequency selective if all frequency components are subject to the same attenuation and phase shift. Frequency-selective fading is a more serious problem since matched filters that are structured to match the undistorted part of the spectrum will suffer a loss in detection performance when the attenuated portion of the spectrum is encountered. Either the data rate must be restricted so that the signal bandwidth falls within the coherence bandwidth of the link or other techniques such as spread spectrum must be used to suppress the distortion. The delay spread caused by multipath is typically greater outdoors than indoors due to the wider coverage area. This gives rise to a higher coherence bandwidth in indoor environments. For example, an indoor channel with a delay spread of 250 ns corresponds to a coherence

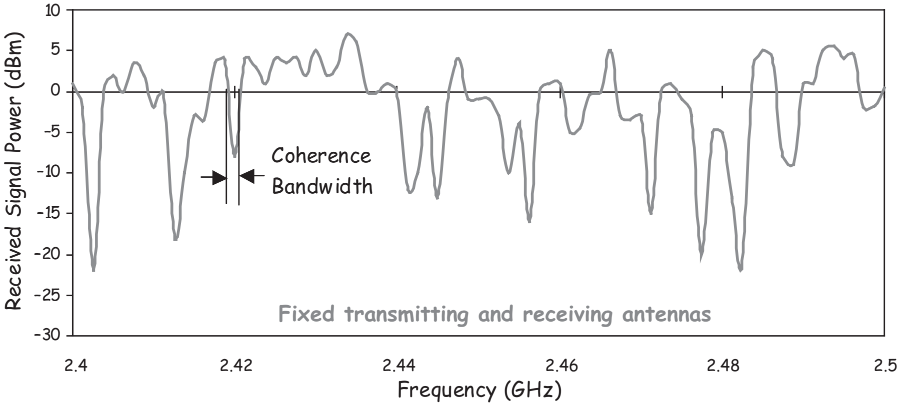


Figure 1.8 Signal fading characteristics (2.4 GHz band).

bandwidth of 4 MHz. An outdoor channel with a larger delay spread of 1 μs implies a smaller coherence bandwidth of 1 MHz.

Signals with bandwidth larger than the coherence bandwidth of the channel may make effective use of multipath by resolving (isolating) many independent signal propagation paths to provide better SNR at the receiver. This is exploited by some multiple antenna systems. On the other hand, multipath interference can be avoided by keeping the symbol rate low, thereby reducing the signal bandwidth below the coherence bandwidth. Although a wideband receiver can resolve more paths than another receiver with a narrower bandwidth, this may be done at the expense of receiving less energy and more noise per resolvable path.

1.3.5 Doppler Spread

Doppler spread is primarily caused by the relative motion between the transmitter and receiver. It introduces random frequency or phase shifts at the receiver that can result in loss of synchronization but affects LOS and reflected signals independently. Reflected signals affected by Doppler shifts are perceived as noise contributing to intercarrier interference (ICI) in multicarrier transmission. The Doppler effect may also be due to the movement of reflecting objects (e.g., vehicles, humans) that causes multipath fading. In an indoor environment for instance, the movement of people is the main cause of Doppler spread. A person moving at 10 km/h can induce a Doppler spread of ±22 Hz at 2.4 GHz. The Doppler spread for indoor channels is highly dependent on the local environment, providing different shapes for different physical layouts. On the other hand, outdoor Doppler spreads consistently exhibit peaks at the limits of the maximum Doppler frequency. Typical values for Doppler spread are 10–250 Hz (suburban areas), 10–20 Hz (urban areas), and 10–100 Hz (office areas).

Just as coherence bandwidth is inversely related to the delay spread, coherence time is defined as the inverse of the Doppler spread. The coherence time determines the rate at which fading occurs. Fast fading occurs when the fading rate is higher than the data symbol rate. The coherence time is a key parameter that affects channel feedback mechanisms in high-speed mobile systems. For example, the delay in sending channel feedback information from the handset to the BS may exceed the coherence time. This renders the feedback outdated by the time the BS processes the information. The relationship between coherence time and coherence bandwidth is shown in Figure 1.9. This relationship forms the basis for fading channel classification.

1.3.6 Shadow Fading

Besides multipath fading, large physical obstructions (e.g., walls in indoor environments, buildings in outdoor environments) can cause large-scale shadow fading. In this case, the transmitted signal power is blocked and hence severely attenuated by the obstruction. The severity of shadow fading is dependent on the relative positions of the transmitter and receiver with respect to the large obstacles in the propagation environment, as well as the number of obstructing objects and the dielectric

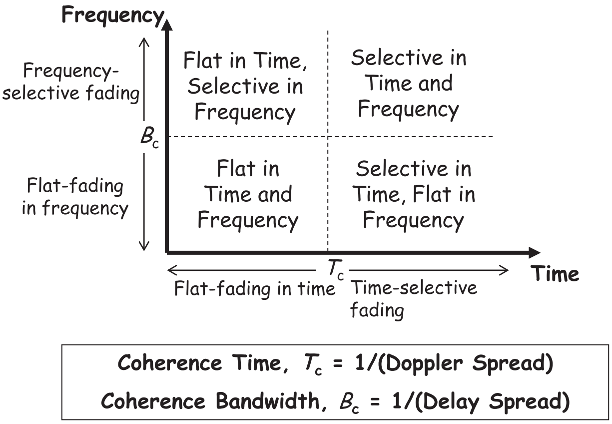


Figure 1.9 Coherence time and coherence bandwidth.

properties of the objects. Unlike multipath fading (which is usually represented by a Rician or Rayleigh distribution), shadow fading is generally characterized by the probability density function of a log-normal or Gaussian distribution. Increasing the transmit power can help to mitigate the effects of shadow fading although this places additional burden on the handset battery and can cause interference for other users.

1.3.7 Radio Propagation Modeling

The transmitted signal power normally radiate (spread out) in all directions and hence attenuates quickly with distance. Thus, very little signal energy reaches the receiver, giving rise an inverse relationship between distance and path loss. Depending on the severity, the decay in signal strength can make the signal become unintelligible at the receiver. Radio propagation analysis allows the appropriate power or link budget to be determined between the RF transmitter and receiver. It can be very complex when the shortest direct path between the transmitter and receiver is blocked by fixed or moving objects, and the received signal arrives by several reflected paths. The degree of attenuation depends largely on the frequency of transmission. For example, lower frequencies tend to penetrate objects better, while high frequency signals encounter greater attenuation.

For clear LOS paths in the vicinity of the receiving antenna, signal attenuation is close to free space. This is the simplest signal loss model where the received signal power decreases with the square of the distance between the transmitter and receiver. For instance, the signal strength at 2 m is a quarter of that at 1 m. At longer distances away from the receiving antenna, an increase in the attenuation exponent is common (Figure 1.10). In this case, the signal attenuation is dependent not only on distance and transmit power but also on reflecting objects, physical obstructions, and the amount of mutual interference from other transmitting users. Small changes in position or direction of the antenna, shadowing caused by blocked signals and moving obstacles (e.g., people and doors) in the environment may also lead to drastic fluctuations in signal strength. Similar effects occur regardless of whether a user is

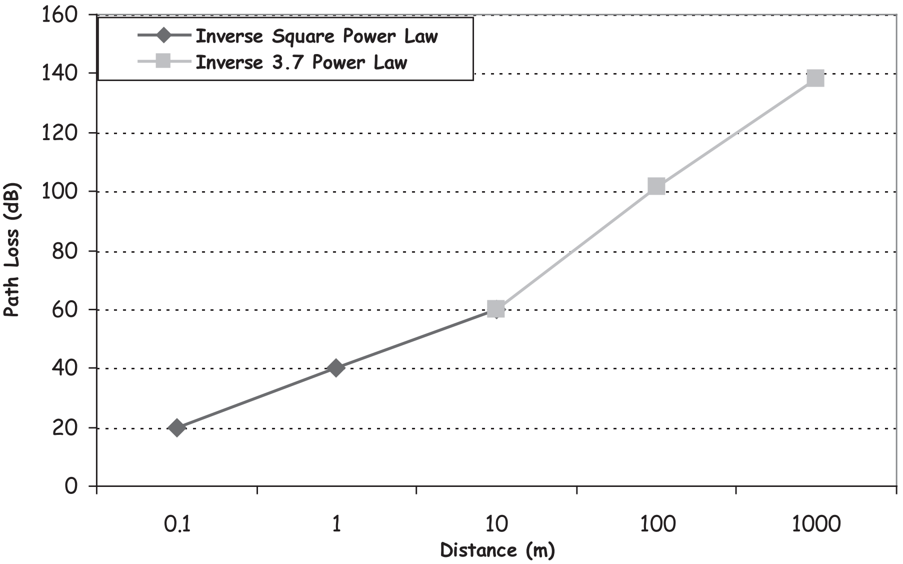


Figure 1.10 Signal attenuation for omnidirectional antenna with spherical (isotropic) radiation pattern.

stationary or mobile. Hence, while the free-space exponent may be relevant for short distance transmission (e.g., up to 10 m), the path loss is usually modeled with a higher-valued exponent of 3 to 5 for longer distances. For indoor environments, where objects move very slowly, fading is primarily due to the receiver. Thus, the path loss (a distance-related phenomenon) is independent of fast fading (a time-related phenomenon).

An accurate characterization of the propagation mechanism is difficult since this is greatly influenced by a number of factors, such as antenna height, terrain, and topology. Radiowave propagation modeling is usually based on the statistics of the measured channel profiles (time and frequency domain modeling) or on the direct solution of electromagnetic propagation equations based on Maxwell's equations. The most popular models for indoor radio propagation are the time domain statistical models. In this case, the statistics of the channel parameters are collected from measurements in the propagation environment of interest at various locations between the transmitter and receiver. Another popular method involves ray tracing, which assumes that all objects of interest within the propagation environment are large compared with the wavelength of propagation, thus removing the need to solve Maxwell's equations. Its usefulness is ultimately dependent on the accuracy of the site-specific representation of the propagation environment.

Modeling the channel characteristics of narrowband and wideband signals is different. For narrowband signals, the emphasis is on the received power whereas for wideband communications, both the received signal power and multipath characteristics are equally important. A further distinction exists between models that describe signal strength as a function of distance as opposed to a function of time.

The former is used to determine coverage areas and intercell interference while the latter is used to determine bit error rates and outage probabilities.

1.3.8 Channel Characteristics

Many wireless systems, including multiantenna systems, show large performance improvements when the channel characteristics are known. The extent to which these gains are achieved depends on the accuracy with which the receiver can estimate the channel parameters. In practical implementations, the channel characteristics are often affected by fading, which can be time arying. In some cases, the channel is assumed to be reciprocal. This implies that the channel model in the forward direction (directed at the receiver) is identical to that in the reverse direction (directed at the transmitter). To learn the behavior of the wireless channel and build a model, a channel sounding method can be used. The transmitter provides a periodic multitone test sequence to excite the channel. At the receiver, the arriving test signal is correlated with a local copy of the test sequence. Due to the impulse-like autocorrelation function of the test sequence, the correlator output at the receiver provides the measured channel impulse response, which must be acquired real time for correct estimation of the signal path statistics. In general, the channel impulse response is stationary at short distances, but can be highly variable over longer distances. It can also be time varying due to the presence of nonstationary objects (e.g., people and vehicles).

1.3.9 Gaussian Channel

The Gaussian channel provides an upper bound on the wireless system performance and accurately describes many physical channels, including time-varying channels. Often referred to as the additive white Gaussian noise (AWGN) channel, it is typically used to model the noise generated in the receiver when the transmission path is ideal (i.e., LOS). The noise is modeled with a constant power spectral density over the channel bandwidth and a Gaussian probability density function. If the user is stationary, the model is valid even when multipath propagation is present. In this case, the channel is approximated as Gaussian with the effects of multipath fading represented as a path loss.

The Gaussian channel model has received considerable attention because of its importance for single-transmitter and single-receiver channels. However, the Gaussian model may not be suitable when applied to multiple access channels where users transmit intermittently. In this case, the desired model is one where the number of active users on the network is a random variable, and the Gaussian signals are conditioned on the fact that some user is transmitting. In addition, Doppler spread, multipath fading, shadowing, and mutual interference from multiple transmitting users make the channel far from Gaussian.

1.3.10 Rayleigh Channel

There are two kinds of channel fading, namely, long-term (log-normal) and short-term (Rayleigh) fading. Long-term or slow fading is characterized by the envelope

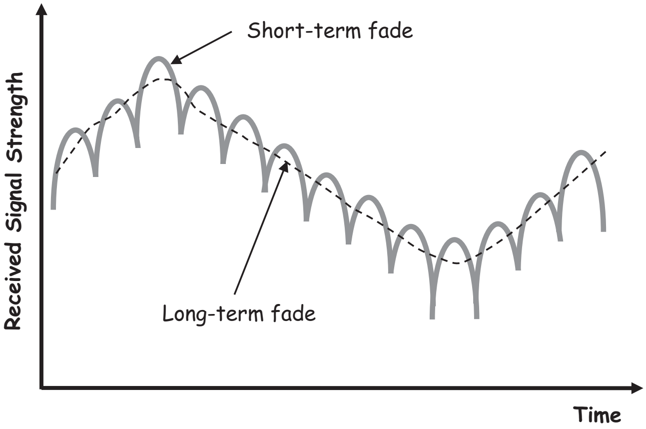


Figure 1.11 Short- and long-term fading characteristics.

of the fading signal, which is related to the distance and the received power. Short-term or fast fading is primarily caused by reflections of a transmitted signal and refers to the rapid fluctuations of the received signal amplitude. Short-term fading occurs when there is no direct LOS signal component. Conversely, if there is a LOS component, long-term fading occurs. In addition, second-order statistics on the fades exist and are usually functions of time (e.g., level crossing rates, average duration of fades, and distribution of fade duration). The received signal is a product of the long-term fading and short-term fading characteristics. The short-term fading signal is superimposed on an average value that varies slowly as the receiver moves (Figure 1.11).

1.3.11 Rician Channel

In some wireless channels, a dominant path (normally the LOS path) exists between the transmitter and receiver, in addition to many scattered (nondirect) paths. This dominant path reduces the delay spread and may significantly decrease the fading depth, thus requiring a much smaller fading margin in system design. The probability density function of the received signal envelope is said to be Rician. Rayleigh fading can be considered as a special case of Rician fading. This is because Rician fading is essentially a superposition of a Rayleigh distributed signal and a LOS signal. The strong LOS signal reaches the receiver at about the same time as the reflected signals generated by local scatterers near the receiver. Rician channels are highly desirable because they require a much smaller fade margin and supports significantly larger bandwidths than Rayleigh channels. The Rician fading function is very similar to the Rayleigh fading function with the exception of two additional variables: the arrival angle of the LOS signal and the *K* factor. The LOS angle specifies the relative location of the direct path with respect to the faded spectrum by changing the static frequency shift. The *K* factor represents the energy ratio of the LOS signal to the scattered signal. Thus, a higher value for *K* shows a greater proportion of signal energy in the dominant path.

1.4 MODULATION

The maximum data rate is not only bounded by the multipath characteristics of the channel but also the modulation technique used. With a bandwidth-efficient modulation technique, a greater number of bits can be transmitted in each symbol interval, resulting in higher data rates without increasing the symbol rate. As shown in Figure 1.12, in multilevel modulation, each symbol is mapped into multiple data bits. A 2 Mbit/s bit stream using 64-QAM modulation requires a symbol rate of 33 Ksymbol/s. Unfortunately, a higher-order modulation, such as 64-QAM, demands a high SNR for good performance and may not be appropriate for noisy or power limited wireless systems. The modulation is normally chosen together with a coding scheme that provides some level reliability in the transmission of the bit stream. As we shall see later, the modulation and coding scheme (MCS), hence the data rate, is often adjusted dynamically to cope with changes in the SNR. This process is called link or rate adaptation, which is not required in wired networks. Other considerations for selecting a MCS include bandwidth (spectral) and power efficiency, resistance against multipath, cost/complexity of implementation, envelope variation, and out-of-band radiation.

1.4.1 Linear versus Constant Envelope

Constant envelope (e.g., binary phase-shift keying [BPSK] and quadrature phase-shift keying [QPSK]) and linear modulation (e.g., quadrature amplitude modulation [QAM]) are common digital modulation methods employed in wireless transmission. Constant envelope modulation involves only the phase, whereas linear modulation requires both the phase and the amplitude to be modulated (thus resulting in a

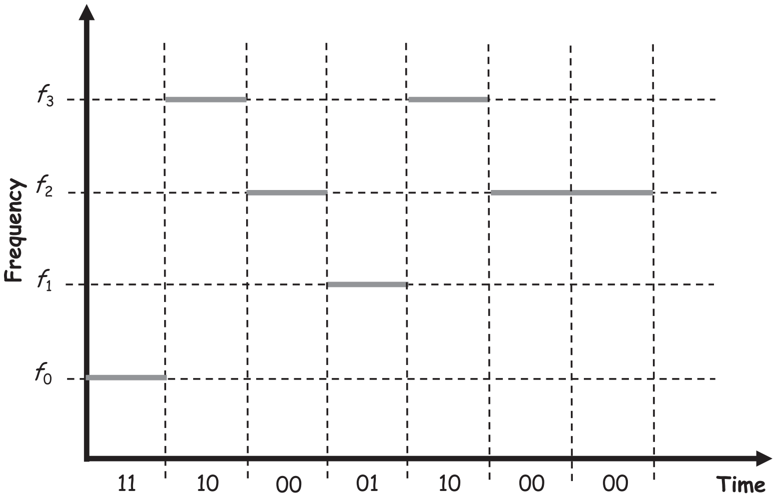


Figure 1.12 Multilevel modulation.

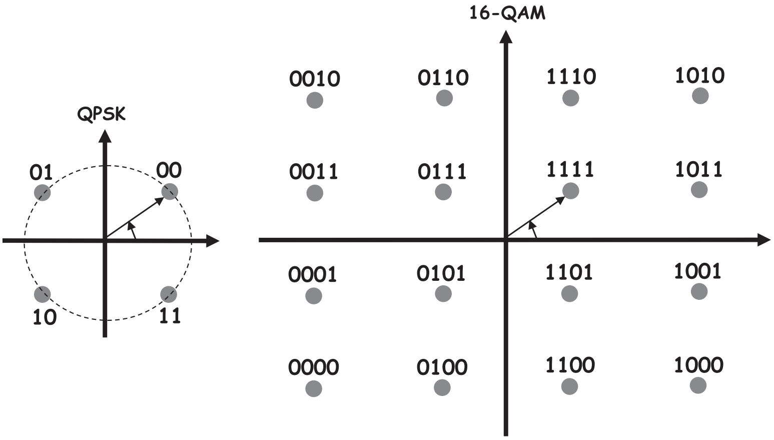


Figure 1.13 16-QAM and QPSK constellations.

non-constant envelope). For QAM, a constellation diagram can be plotted where each constellation point specifies an amplitude and a phase of the modulation, and is mapped to a modulated data symbol. For example, 16-QAM codes 4 bits per symbol, and there are 16 points on the constellation diagram, whereas QPSK comprises 4 points, each representing 2 bits per symbol (Figure 1.13). Thus, the symbol rate is equal or less than the actual bit rate. A lower symbol rate reduces instances of ISI.

Constant envelope modulation provides gains in power efficiency since it allows more efficient power amplifiers to be used. On the other hand, linear modulation is spectrally more efficient than constant envelope modulation but requires the use of linear power amplifiers. The concern for linearity is primarily due to the possibility of higher intermodulation and out-of-band power emissions. There is considerable debate as to whether multilevel modulation, such as QAM, provides any real gain in frequency reuse situations, since it requires an increase in SNR. This results in an increase in frequency reuse distance, which more than offsets the modulation improvement.

1.4.2 Coherent versus Noncoherent Detection

There are two basic techniques of detecting data symbols from the demodulated signals, namely, coherent (synchronous) and noncoherent (envelope) detection. In coherent detection, the receiver first processes the signal with a local carrier of the same frequency and phase, and then cross-correlate with other replicated signals received before performing a match (within a predefined threshold) to make a decision. In a multipath environment, phase synchronization is not easy, and hence, noncoherent detection is preferred. This can be achieved by comparing the phase of a received symbol with a previous symbol to obtain the relative phase difference. Since noncoherent detection does not depend on a phase reference, it is less complex but exhibits worse performance when compared with coherent methods. Although no absolute phase reference is needed for the demodulation of differential phase modulation, the receiver is sensitive to carrier frequency offset.

1.4.3 Bit Error Performance

The ratio of the energy per bit to the noise power spectral density (E_b/N_0) is an important parameter in wireless digital communications. E_b/N_0 is commonly used with modulation and coding designed for noise-limited rather than interference-limited communication, and for power-limited rather than bandwidth-limited communications (e.g., spread spectrum systems). It normalizes the SNR and is also known as the "SNR per bit." It is especially useful when comparing the bit error rate (BER) performance of different digital modulation schemes without taking bandwidth into account. E_b/N_0 is equivalent to the SNR divided by the link spectral efficiency in bit/s/Hz, where the transmitted data bits are inclusive of overheads due to error coding and network protocols, such as medium access control (MAC) protocols. Thus, E_b/N_0 is generally used to relate actual transmitted power to noise. The noise spectral density N_0, is usually expressed in W/Hz. Therefore, E_b/N_0 is a nondimensional ratio.

1.5 MULTIPATH MITIGATION METHODS

If symbol rates greater than the inverse of the delay spread are required, then several techniques can be used to compensate for ISI caused by a time-dispersive channel and regenerate the data symbols correctly. They include equalization, multicarrier transmission, and wideband transmission, error control, and antenna diversity.

1.5.1 Equalization

Equalizers essentially correct the phase differences between the direct and reflected signals (Figure 1.14) by subtracting the delayed and attenuated images of the direct signal from the received signals. To do this, equalizers employ a training sequence at the start of each packet transmission to derive the impulse response (transfer characteristic) of the channel. This response must be frequently remeasured as the wireless channel can change rapidly in both time and space. When applied to slowly varying indoor wireless links, equalization typically requires a minimum overhead of 14–30 symbols for every 150–200 data symbols. For rapid variations in the wireless link, equalization can be difficult because these problems have to be alleviated using a long equalizer training sequence. Nevertheless, by using joint data and link estimation algorithms, data symbols can still be reliably recovered even after the link undergoes fading.

The simplest equalization technique (also known as linear equalization) amplifies the attenuated part and attenuates the amplified part of the spectrum. This technique attempts to invert and neutralize the effects of the medium. This can be achieved by passing the received signal through a filter with a frequency response

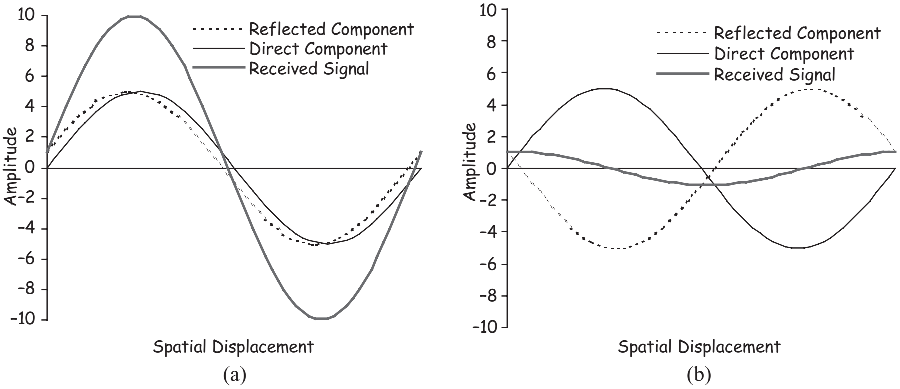


Figure 1.14 Phase correction needed for multipath signals. (a) In-phase received signal. (b) Out-of-phase received signal.

that is the inverse of the frequency response of the medium. For time-varying wireless channels, the linear equalizer can be made adaptive by employing an adaptive filter at the receiver whose frequency response adapts to the inverse of the channel response, further suppressing the multipath components. Although linear equalizers are the simplest to implement, their performance is limited because the inverse filter of the wireless channel may not exist in all cases (e.g., channels with nulls) or may cause the signal to be restored to its original form at the expense of boosting the received noise to a very high level. For example, when additional interference or frequency-selective fades are present at the input to the equalizer, the equalizer must minimize distortion without adversely enhancing noise. In this case, more sophisticated equalizers are needed.

The decision feedback equalizer (DFE) is a nonlinear equalizer that does not enhance the noise because it estimates the frequency response of the link rather than inverting it. It can also handle channels with nulls. Like spread spectrum transmission, a DFE can isolate the arriving paths and take advantage of them as a source of implicit diversity to improve performance. A DFE feeds past decisions through a feedback filter and subtracts the results from the input to the decision device. This nonlinear process results in less noise enhancement than linear equalization. A DFE also exploits decisions about the received data to cancel ISI from the received signal. It is easier to implement than a linear equalizer since the input to its feedback filter is a data symbol. Hence, convolution can be implemented based solely on additions, and no multiplications are required. However, DFEs are prone to error propagation and are useful only when the medium exhibits reasonably low bit error rate without coding.

The maximum likelihood sequence estimator (MLSE), also known as the Viterbi equalizer, operates by testing hypotheses of the transmitted data sequence, combined with knowledge about the channel response, against the signal that was received. The signal that gives the closest match is chosen. Sequence detection offers performance improvements over traditional equalization when the channel response is known and time-invariant. MLSE equalizers are well suited for spread spectrum systems that employ different code sequences to distinguish the transmission from multiple users. Although these equalizers achieve the lowest error probability, they are complex and may not be feasible at very high data rates.

1.5.2 Multicarrier Transmission

At high data rates, the computation complexity for an equalizer increases. In addition, the overhead for channel estimation increases when the channel is time-varying. Multicarrier or multitone systems rely on sending several data streams in parallel at low symbol (signaling) rates that combine to the desired high rate transmission. For example, a 100 Msymbol/s data stream (corresponding to 100 Mbit/s using BPSK modulation) can be converted to 100 streams of 1 Msymbol/s using 100 subcarriers. In doing so, multipath delay spread is compensated without the need for complex equalization. The principle behind multicarrier schemes is that since ISI caused by multipath is only a problem for very high symbol rates (typically above 1 Msymbol/s), the data rate can be reduced until there is no degradation in performance due to ISI. The overall frequency channel is divided into a number of subchannels (which can range from 64 in indoor networks to 2048 in outdoor networks), each modulated on a separate subcarrier at a lower symbol rate than a single carrier scheme using the entire channel. Hence, the delay spread now impacts a smaller proportion of the lengthened symbol time. Since each subchannel is narrow enough to cause only flat (nonfrequency-selective) fading, this makes a multicarrier system less susceptible to ISI. By adding a guard interval (GI) to each symbol, such interference can be completely eradicated.

Multicarrier transmission is spectrally efficient because the subcarriers can be packed close together. It also allows considerable flexibility in the choice of different modulation methods. On the negative side, multicarrier transmission is more sensitive to frequency offset and timing mismatch than single-carrier systems. Fast synchronization of multiple subcarriers at the receiver is crucial in reducing processing delays and overheads. This is normally achieved by allocating a specific set of pilot subcarriers for synchronization purposes. The location of the pilots may be changed from symbol to symbol and the power of these subcarriers is boosted to minimize errors from recovering the training signal. In general, either the overall data rate is increased or a larger delay spread can be tolerated if a larger number of subcarriers is used.

Like multilevel modulation, a major drawback of multicarrier systems is the high peak-to-average power ratio (PAPR). The power amplifier in a multicarrier system is more costly since it must back off more than a single-carrier system, especially for user devices operating near the edge of a cell due to the higher transmit power. The large amplifier backoff requires the use of a highly linear (and inefficient) amplifier with precise gain control, which leads to high power consumption. Such an amplifier is also needed to suppress ICI when subcarriers become misaligned due to multipath propagation. Since a lower number of subcarriers leads to reduced PAPR, the number of subcarriers and the subcarrier spacing have to be chosen appropriately. In addition to linear amplifiers, signal distortion techniques, such as peak clipping or filtering, spectrum shaping or windowing (which reduces output spectrum), may have to be implemented to counter the PAPR effect. Low-phase noise oscillators and high-resolution analog-to-digital converters are also needed.

1.5.3 Orthogonal Frequency Division Multiplexing

Orthogonal frequency division multiplexing (OFDM) is a special form of multicarrier transmission involving orthogonal subcarriers. Although its adoption has become widespread in many wireless network standards, OFDM also forms the basis of some wireline standards, including discrete multitone digital subscriber line (DSL) and HomePlug powerline communication systems. OFDM employs equally spaced subchannels, each containing a subcarrier that carries a portion of the user information. An OFDM system uses less bandwidth than an equivalent frequency division multiple access (FDMA) system because OFDM subchannels are overlapped. In this way, OFDM offers some degree of frequency diversity since if the subchannel becomes degraded by interference, noise, or fading, the data rate is reduced but the connection is maintained. Error control further enhances the reliability of each subcarrier. OFDM can be combined with frequency hopping, allowing nonsequential transmission of subcarriers, thereby achieving a greater degree of frequency diversity and interference averaging.

In OFDM, because data symbols are transmitted in parallel rather than serially, the duration of each OFDM symbol is longer than the corresponding symbol on a single carrier system of an equivalent data rate. This reduces the effect of ISI. In addition, the orthogonality between the subcarriers mitigates ICI. The baseband frequency of each subcarrier is then carrier-modulated using the same RF carrier. By switching subcarriers after each symbol time, frequency selective fading is minimized. If the symbol interval is T and the number of subcarriers is N, then the subcarriers are spaced at intervals of $1/T, 2/T, \ldots, N/T$. Thus, the baseband frequencies of the OFDM subcarriers become an integral multiple of the base frequency ($1/T$). This leads to orthogonal subcarriers where the peak of one modulated subcarrier can overlap on the null of every other subcarrier without causing unnecessary interference within the subchannel (Figure 1.15). When the receiver samples at the center frequency of each subcarrier, the desired signal is obtained. Only sinusoidal subcarriers can ensure orthogonality, allowing the receiver to distinguish between the waveforms.

The use of overlapped subchannels results in higher spectral efficiency. As can be seen in Figure 1.16, three modulated subcarriers can occupy the same bandwidth as two modulated subcarriers. Unlike FDMA, OFDM allows separation of subcarriers without a guard band. OFDM reduces ISI compared with high-speed single carrier transmission because the symbol rate can be reduced over multiple subcarriers, leading to flat fading, which is easier to deal than frequency-selective fading. A larger number of subcarriers allow the symbol rate for each subcarrier to be reduced further. This can, in turn, mitigate a larger multipath delay spread in outdoor environments. For example, the 802.16 wireless access standard employs 256 OFDM subcarriers compared with 64 subcarriers in the 802.11 standard.

At the transmitter, if there are N subcarriers (N samples), the time-domain subcarrier sequence $f(n)$ can be generated from the frequency-domain sequence $F(k)$ using the inverse discrete Fourier transform (DFT):

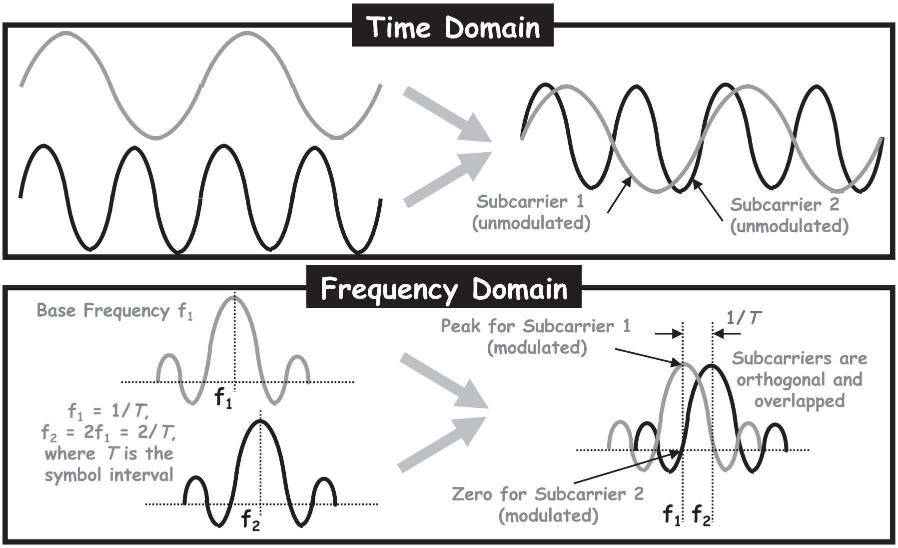


Figure 1.15 OFDM subcarriers in time and frequency domains.

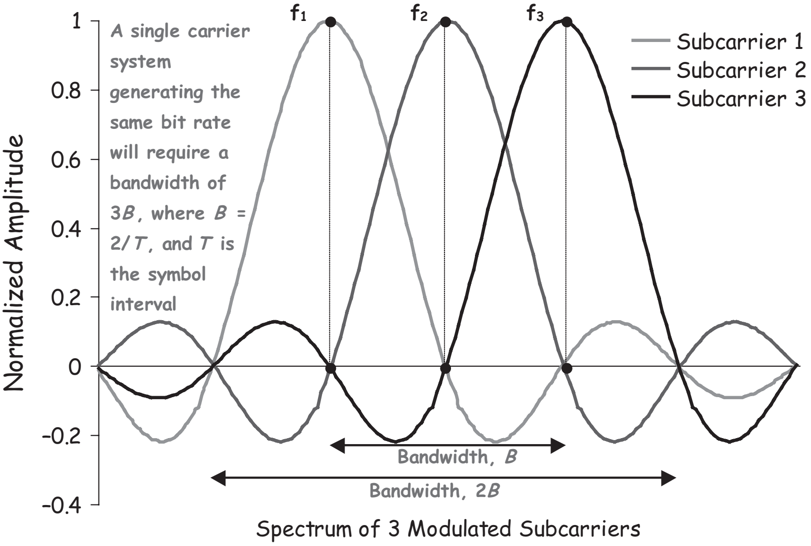


Figure 1.16 Spectral efficiency of overlapped OFDM subcarriers.

$$f(n)=\sum_{k=0}^{N-1} F(k)e^{+\frac{j2\pi nk}{N}}, \quad n=0,1,2,\ldots,N-1. \tag{1.1}$$

At the receiver, the DFT can be used to generate *F*(*k*).

$$F(k)=\sum_{n=0}^{N-1} f(n)e^{-\frac{j2\pi nk}{N}}, \quad k=0,1,2,\ldots,N-1. \tag{1.2}$$

The resulting Fourier spectrum has discrete frequencies at *k*/*NT*, where *T* is the sampling interval in the time domain, and *N* is the number of samples. The sampling frequency in the frequency domain is 1/*T*. The DFT operations in OFDM can be implemented using fast Fourier transform (FFT), which can be adapted to different data rates and link conditions. The FFT size corresponds to *N*. FFT allows efficient digital signal processing implementation and significantly reduces the amount of required hardware compared with FDMA systems. For example, OFDM achieves reduced processing complexity compared with single carrier systems, achieving log*N* per sample complexity using FFT. As long as the channel is linear, the subcarriers can be demodulated without interference and without the need for analog filtering to separate the received subcarriers.

Due to the simplicity of equalizer process, OFDM reduces the need to maintain a LOS path. A CP absorbs late-arriving symbols in the presence of delay spread caused by multipath. The CP is essentially a GI that is longer than the delay spread, thereby preventing one symbol from interfering with the next. A short GI reduces the operating range of the OFDM system. It is an important component of practical OFDM implementations because it leads to simple equalization. A copy of the last portion of the data symbol is usually appended to the front of the next symbol to form the CP. Figure 1.17 shows the typical OFDM transmitter and receiver implementation.

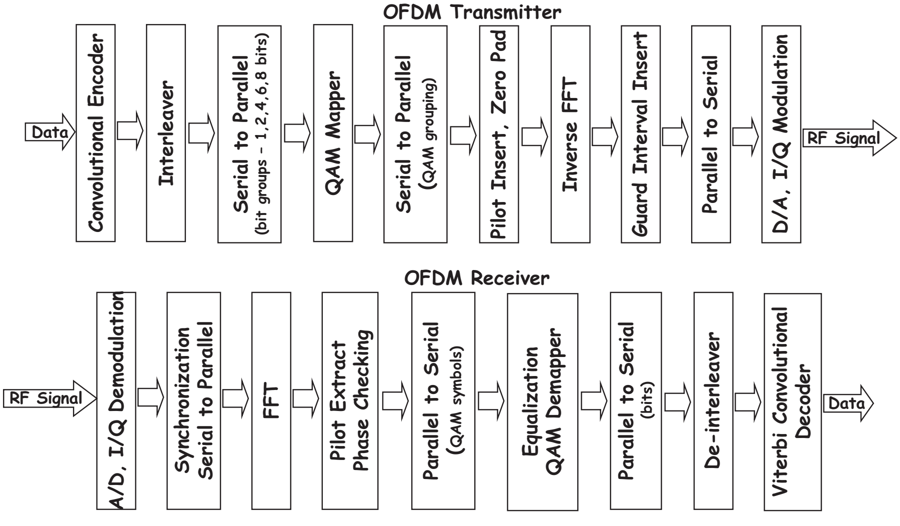


Figure 1.17 OFDM transmitter and receiver.

1.5.4 Wideband Systems

In wideband systems, such as spread spectrum or ultra-wideband (UWB) systems, the transmission bandwidth is much greater compared with the signal bandwidth. Spread spectrum is a powerful combination of bandwidth expansion (beyond the coherence bandwidth) and coding, whereby the former combats ISI and the latter allows individual symbols smeared by multipath to be received correctly. The general concept of spread spectrum has been in existence for more than 50 years and was originally developed by the military for reliable and secure communications. Spread spectrum refers to signaling schemes that are based on some form of coding (that is independent of the transmitted information) and which use a much wider bandwidth beyond what is required to transmit the information. The wider bandwidth means that interference and multipath fading typically affect only a small portion of a spread spectrum transmission. The received signal energy is therefore relatively constant over time, and frequency diversity is achieved. This in turn produces a signal that is easier to detect provided the receiver is synchronized to the parameters of the spread spectrum signal. Spread spectrum signals are able to resist intentional jamming since they are hard to detect and intercept, thus ensuring some level of PHY security. Instead of an equalizer, spread spectrum systems often use a simpler device, called a correlator, to counter the effects of multipath.

There are two spread spectrum techniques: direct-sequence spread spectrum (DSSS) and frequency-hopping spread spectrum (FHSS). In DSSS, the data stream is combined with a higher speed digital code. Each data bit is mapped into a common pattern of bits known only to the transmitter and intended receiver. The bit pattern is called a pseudonoise (PN) code, and each bit of the code is called a chip (the term chip emphasizes that one bit in the PN code forms part of the actual data bit). The sequence of chips within each bit period is random, but the same sequence is repeated in every bit period, thus making it pseudorandom or partially random. The chipping rate of an n-chip PN code is n times higher than the data rate. Thus, a high rate for the chipping sequence results in a very wide bandwidth. DSSS spreads the energy (power) of the signal over a large bandwidth. The energy per unit frequency is correspondingly reduced. Hence, the interference produced by DSSS systems is significantly lower compared with narrowband systems. This allows multiple DSSS signals to share the same frequency channel. To an unintended receiver, DSSS signals appear as low-power wideband noise and are rejected by most narrowband receivers. Conversely, this technique diminishes the effect of narrowband interference sources by spreading them at the receiver.

As an example, 802.11b devices employ DSSS using the 11-chip Barker PN code (Figure 1.18). Information bits are spread out 11 times by the PN code before being modulated. The chip rate is 11 times faster than the data rate. Chips are despread by the same PN code at the receiver and mapped back into the original data bit (Figure 1.19). If the PN code generated by receiver is perfectly synchronized to the transmitted signal, the despreading process produces high autocorrelation peaks (Figure 1.20). Sidelobes of the received signal will either add constructively or destructively, but the autocorrelation peaks will not be affected. The overall result

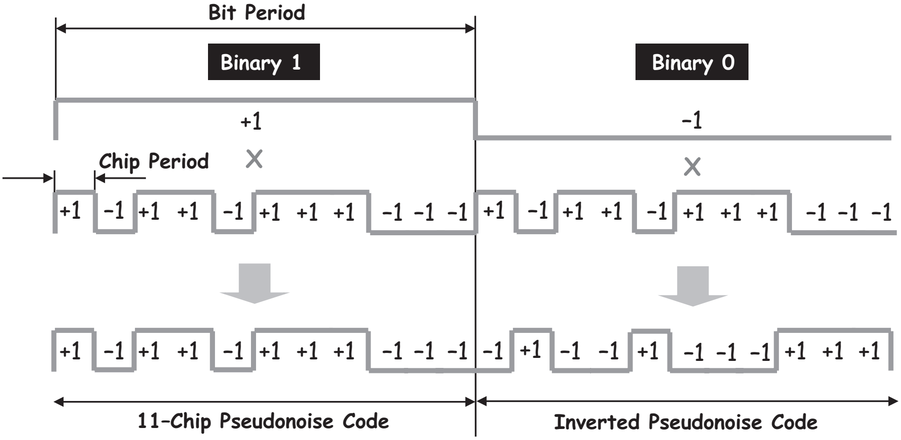


Figure 1.18 Combining a higher-speed PN code with lower-speed digital data.

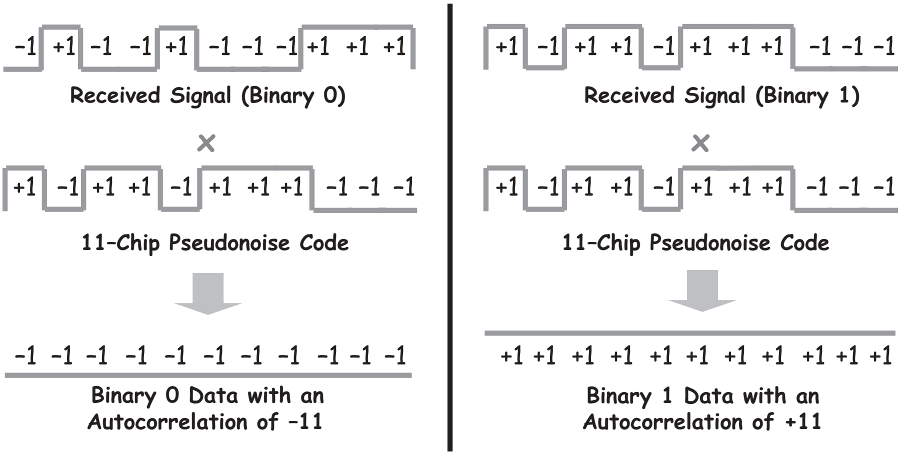


Figure 1.19 PN code is synchronized at receiver to recover lower-speed digital data.

is a single peak within each bit interval. The energy of noise and narrowband interference that may have been added during the transmission are spread and hence suppressed by the PN code (Figure 1.21).

Clearly, DSSS systems take up more bandwidth due to the spreading. To achieve higher aggregate efficiency, different PN codes can be used, thereby allowing concurrent transmissions to take place. This gives rise to CDMA. Code division refers to the fact that transmission with orthogonal PN codes may overlap in time

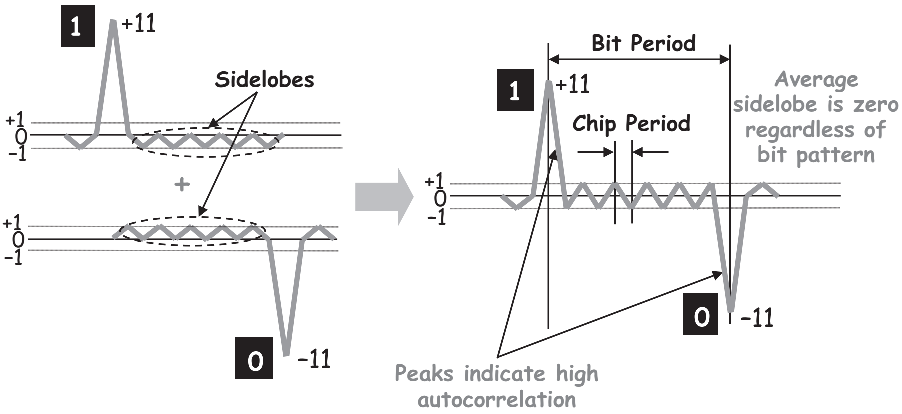


Figure 1.20 Autocorrelation and sidelobes in despread digital data.

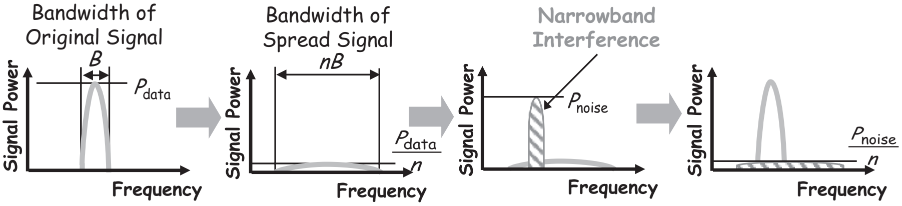


Figure 1.21 Interference mitigation using DSSS.

with little or no effect on each other. Different users transmit using unique codes. For a receiver tuned to the code of one transmission, other signals (using other PN codes) appear as background noise. In the despreading process, this noise will be suppressed by the processing gain. A random access CDMA system requires complex receivers that can synchronize or demodulate all PN codes. This is because the codes may not be orthogonal in the presence of multipath delay spread, leading to multiuser interference that ultimately limits the capacity of the system. A long PN code supports more concurrent transmissions. However, for a fixed channel bandwidth, the rate of each transmission is reduced. CDMA systems are often used to estimate the distance between the transmitter and receiver in a process called ranging. DSSS wireless LANs utilize the same PN code and therefore do not have a set of codes available as is required for CDMA operation. A single code enables information to be broadcast easily. In addition, the code can be made shorter, thereby increasing data bandwidth.

Like narrowband systems, multipath, shadow fading, and interference impose limits on the coverage region and on the number of users in spread spectrum systems. For example, if a 100-Mchip/s chip rate is selected, then multipath components must be separated by at least 10 ns in order for these components to be resolved. Hence, the appropriate area of coverage should have a propagation delay greater than

10 ns, or, equivalently, a minimum span of 3×10^8 m/s × 10 ns or 3 m. In general, a bigger coverage area is preferred for spread spectrum communications since this results in more resolvable multipaths. However, the opposite is true for narrowband systems. Furthermore, polarization antenna diversity becomes more effective since there is significant coupling between the vertical and horizontal polarization directions when the coverage area is large.

1.5.5 Error Control

There are two general types of error control in wireless communications. Forward error correction (FEC) techniques employ error control codes (e.g., convolutional codes and block codes) that can detect with high probability the error bit's location and correct the error. Due to the additional bit overheads, FEC tends to degrade data throughput (i.e., the usable bit rate available to applications). FEC may also require long bit interleaving intervals to randomize and spread the error bits so that they can be corrected easily. This increases the transmission delay. Code rates of 1/2, 2/3, 3/4, 5/6, are common. A code rate of 1/2 implies that for every 1 data bit, an additional FEC bit is added and so 2 bits are transmitted. Thus, a code rate of 1/2 results in a 50% reduction in data throughput. A high code rate of 5/6 is more efficient than a low code rate of 1/2 but demands a higher SNR and hence, a shorter operating range or higher transmit power. Low code rates are usually used with QPSK because high-order modulations are not effective with low rate codes. Thus, a more efficient code rate is normally chosen with a high-order modulation. For example, a code rate of 5/6 can be chosen with 64-QAM to send a group of 6 bits (representing one data symbol). The coded bits may be repeated (which further reduces coding efficiency) or punctured to remove some of the parity bits.

A convolutional code is an example of FEC that can be implemented with a constraint length of x. In this case, x memory registers are needed, each holding 1 input bit. The encoder contains n modulo-2 adders and n polynomial generators, one for each adder. Using the generators and the existing values in the remaining registers, the encoder outputs n bits. Convolutional codes can be decoded using several algorithms. For small values of x, the Viterbi algorithm provides maximum likelihood performance and is highly parallelizable. Puncturing is often used with the Viterbi algorithm to improve coding efficiency. Larger values for x are more practically decoded with sequential decoding algorithms (e.g., Fano algorithm). Unlike Viterbi decoding, sequential decoding is not maximum likelihood. However, since its complexity increases only slightly with x, this allows the use of stronger long constraint-length codes. Turbo codes are a class of high-performance FEC codes. While normal FEC codes typically achieve a channel capacity of 3 dB below the theoretical limit (i.e., Shannon bound), Turbo codes may potentially achieve a better performance of 0.5 dB below that limit. Two encoders with an interleaver are employed (Figure 1.22).

Low-density parity-check (LDPC) coding is another example of FEC. It is a class of linear block codes whose parity-check matrix contains only a few 1's in comparison with the number of 0's. This requires the number of 1's in each row or column of the matrix to be much less than the dimensions of the matrix. Like Turbo

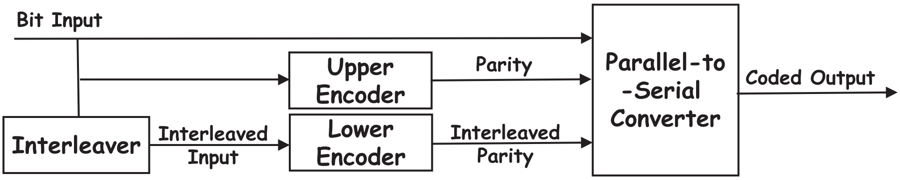


Figure 1.22 Turbo coding.

codes, LDPC codes offer near-capacity performance for many channels. However, constructing good-performing LDPC codes can be challenging, especially in implementation, which requires a high degree of parallelism to shorten the processing time. Theoretically, iterative decoding algorithms of sparse LDPC codes perform close to the optimal maximum likelihood decoder and better than most Turbo codes, but are complex to encode and decode. As such, LDPC has been offered as an optional feature in many wireless standards.

In automatic repeat request (ARQ), the receiver employs simpler error detection codes (which incur less bit overheads than FEC) to detect bit errors in the received packets and then request the transmitter to resend any error packet. Thus, ARQ may achieve a packet error rate of zero. The simplest ARQ method is the Stop-and-Wait ARQ. In this case, each packet transmitted by the sender is accompanied by an acknowledgment (ACK). If the ACK is not received after a predetermined interval (i.e., a timeout), the packet is retransmitted by the sender. Instead of a single packet, other more complex ARQ methods (e.g., selective repeat ARQ) employ ACKs that deal with a group of packets. In general, ARQ schemes are simpler and more flexible to implement than FEC but may suffer from variable delay. Many wireless standards employ a combination of FEC and ARQ, but these methods operate independently, and hence may not give the best performance. To this end, hybrid ARQ (HARQ) has been developed to balance and optimize the tradeoff in FEC overheads and ARQ retransmissions.

Since FEC, ARQ, or HARQ may not correct packet losses caused by buffer overflow at the receiver, packet retransmission by higher network layers (e.g., TCP) or error concealment by the application (e.g., video frame copy, image pixel interpolation) may still be needed.

1.6 MULTIPLE ANTENNA SYSTEMS

Wireless communication is characterized by random propagation conditions commonly referred to as fading channels. A fading channel can be modeled as a linear time-varying channel whose impulse response is a random process. Reliable communication over a fading channel can be achieved through the use of diversity techniques. Since the channel response at any given time and frequency (i.e., channel dimension or "degree of freedom") is a random variable, it is essential that each bit of information is spread over several dimensions. Diversity is traditionally achieved by coding in the time or frequency domains. Because time duration and frequency bandwidth are expensive resources, the number of channel degrees of freedom for diversity can be increased in the space domain, by adding spatially separated antennas at the receiver and the transmitter.

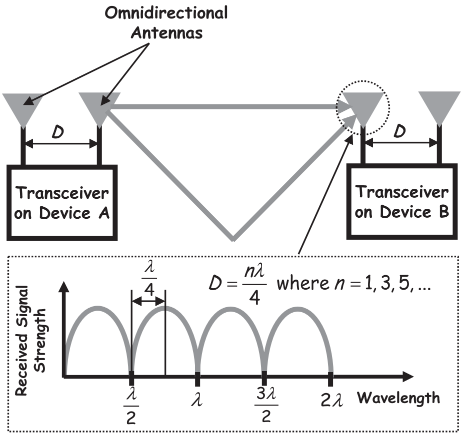
Under poor propagating conditions, the channel can be changed using antenna diversity. The basic principle for using multiple antennas is that if one antenna receives a deeply faded signal, the other antenna(s) receive only a slightly faded version of the signal. Since the angle of departure of the signal from the transmit antenna is typically narrower than the angle of arrival at the receiver, this gives the receiver the possibility of recovering different signals even if the antennas are closely spaced. Macroscopic (spatial) diversity is important when the fading is slow (long-term). It is often the best approach because it allows receivers to choose independent channels. Among the various forms of spatial diversity, selection combining is the simplest to implement since the receiver chooses to process only one channel—the channel with the highest SNR. Maximal ratio combining provides the largest SNR improvement because the signals from all available channels are weighted and then summed. This requires coherent modulation since the phase of each channel must be taken into account before the combining can occur. Microscopic diversity using two frequencies at the same (collocated) antenna site can help reduce short-term fading. Polarization diversity can also be used to select the best channel at a particular location. For example, circular polarized directional antennas, when used in LOS channels, can provide much lower delay spread than linear polarized antennas with similar directionality.

1.6.1 Receive Diversity versus Transmit Diversity

In multiantenna receive diversity, signal combining techniques are employed at the BS or AP to improve the UL performance (i.e., transmission from handset to BS or AP). The simplest method involves choosing the best signal among two or more antennas. Two antennas separated by an odd multiple of a quarter of a wavelength is sufficient to cause almost independent fades at the receiving antennas. Alternatively, different antenna polarization can be employed to keep the antenna dimensions small. Such diversity techniques are effective in overcoming multipath fading without any bandwidth penalty, and no additional transmit power is required. The overall impact is to reduce the fade margin and enhance interference rejection. It is difficult to implement receive diversity on the DL (i.e., transmission from BS or AP to handset). This is due to the size and battery power limitations of the handset. Mobility causes further problems. Thus, new multiantenna systems focus on transmit diversity where multiple antennas at the BS or AP transmit simultaneous data streams on the DL to handset.

1.6.2 Switched Antenna Receive Diversity

Switched antenna receive diversity is the simplest spatial diversity technique involving the use of two or more receive antennas at the user device. To avoid concurrent long-term fading at the collocated receive antennas, the antennas are separated by a



If $n = 1$ and $f = 5$ GHz, then $\lambda = 6$ cm, $D = 1.5$ cm
If $n = 1$ and $f = 2.4$ GHz, then $\lambda = 12.5$ cm, $D \approx 3$ cm

Figure 1.23 Switched antenna diversity.

distance that is a multiple of a quarter of the operating wavelength (Figure 1.23). At any time, only one of the two internal antennas is actively transmitting or receiving. If the receive SNR for the active antenna becomes poor, the antenna is deactivated and the other internal antenna becomes active. A 5 dBi external antenna can also be connected to improve the range and provide more coverage. In this case, the internal antennas are turned off, which removes receive diversity. The different antenna spacings at 2.4 and 5 GHz imply that two different radios may be needed to operate in these bands.

1.6.3 Multiple Input Multiple Output Systems

Unlike single antenna systems, which are also known as single input single output (SISO) systems, multiple input multiple output (MIMO) systems employ multiple antennas to transmit and receive simultaneous data streams on the same frequency channel. Unlike CDMA, MIMO is spectrally efficient even when one user is transmitting since it allows concurrent transmission of data streams without spreading the radio spectrum. When combined with OFDM, MIMO systems can realize impressive spectral efficiencies. For example, up to 15 bits/s/Hz can be realized for the 802.11n wireless LAN standard. This is in contrast to prior standards, such as 802.11b (0.5 bit/s/Hz) and 802.11a/g (2.7 bit/s/Hz), which are all based on single-antenna transmission. Because no switching of transmit antennas is required in MIMO, this may lead to faster response times. Some MIMO configurations are highly recommended for indoor environments due to the rich multipath signal reflections from walls and structures. Such reflections may not always be present in outdoor environments.

Reducing the MIMO antenna size is a key challenge for small personal handsets, such as smartphones. These devices may restrict the achievable benefits of MIMO due to limited space to separate the antennas in the microwave frequency bands. On the other hand, MIMO can be beneficial to handsets because it provides robustness or capacity without the need for a second radio transceiver (using a different frequency band) or additional frequency channels (using the same frequency band). This, in turn, reduces battery power consumption and handset cost. In addition, MIMO's greater reliability or higher rate minimizes the time to transmit or retransmit packets. Thus, the radio is turned on for less time when transmitting a given amount of data, conserving additional battery power. This benefit can also be achieved using broader channel bandwidths and higher-order modulation schemes.

Spatial diversity is available if the individual paths from the transmit to receive antennas fade more or less independently. This implies the probability of all signal paths fading at the same time at the receiver is significantly smaller than the probability of a single path experiencing fading (Figure 1.24). Spatial demultiplexing at the receive antennas separate the interfering streams. This improves signal processing capabilities at the same time because the overall MIMO processed signal exhibits less instances of deep fading than a single received signal. However, joint processing of multiple spatial streams may be required, and this increases receiver complexity and power consumption. The coupling relationship between signals received at the antennas (i.e., path correlation) determines the difficulty in recovering the individual data streams.

Using MIMO, the overall signal power is split equally among the transmit antennas. Accurate channel knowledge may be needed. The receiver decoding complexity is a critical factor in determining the actual performance. By selecting a

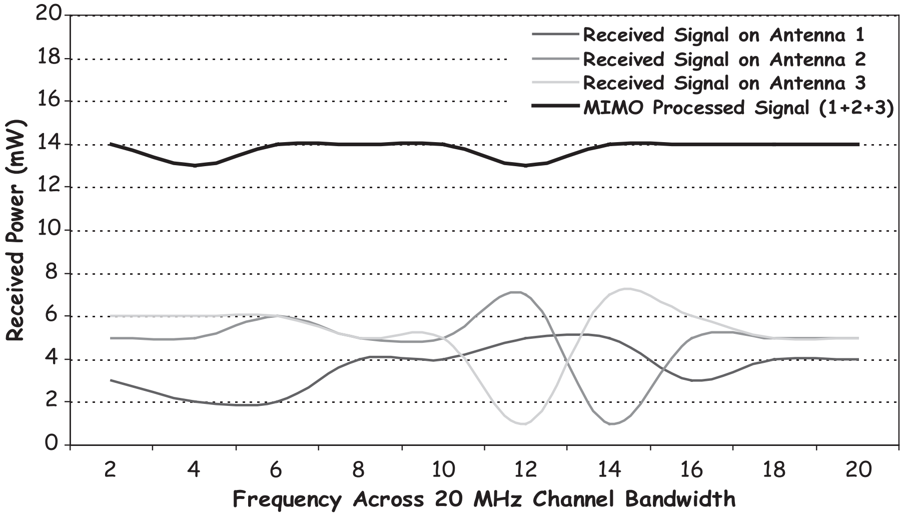


Figure 1.24 MIMO antenna diversity.

smaller set of receive antennas, fewer receiver chains are required. Depending on the number of receive antennas chosen, this may be done at the expense of a small sacrifice in capacity compared with the case when the full set of receive antennas are employed. Antenna selection is an exhaustive process of searching for the best pair of transmit and receive antenna combination that either gives the best SNR/range (for diversity) or spectral efficiency/capacity (for spatial multiplexing).

The number of parallel spatial streams (also known as layers or codewords) is determined by the rank of the channel matrix (H). The rank of H indicates the number of linearly independent rows or columns. If the rank of a channel is r, at most r, spatial streams can be transmitted successfully. More generally, the rank of the MIMO channel matrix is governed by:

$$\mathrm{Rank}(H) \le \min(N_\mathrm{T}, N_\mathrm{R}), \tag{1.3}$$

where N_T and N_R are the number of transmit and receive antennas, respectively. Clearly, the spectral efficiency and channel capacity improve in proportion to the number of antennas used at both the transmitter and receiver. The pioneering research of Foschini [2] and Telatar [3] shows that in theory, MIMO capacity can increase linearly with min (N_T, N_R) or the number of antennas. This is in contrast to the logarithmic increase in capacity with SNR for SISO systems. If H is not full rank (i.e., rank deficient), some of its columns become linearly dependent (i.e., some columns can be expressed as a linear combination of other columns in the matrix). The information in these columns becomes redundant and does not contribute to the overall channel capacity. Thus, a rank-deficient MIMO channel matrix leads to low spectral efficiency and hence, low capacity. This is usually the case when signals arriving at the receiver are highly correlated in the fading characteristics. The correlation between the signal paths can be introduced by a small antenna spacing at the transmitter and receiver, a lack of scattering, or a strong LOS component. However, low fading correlation in the received signals does not guarantee high spectral efficiency. In spite of the low correlation, there are channels that still exhibit poor rank properties and hence, low capacity. Proper selection of a set of transmit antennas can avoid antennas that lead to linearly dependent columns. By redistributing transmit power to other antennas, a full rank channel matrix can be obtained.

Rich RF signal scattering ensures full MIMO spectral efficiency because independent paths are commonly present in such a multipath environment. The key is to efficiently map data symbols to various signals transmitted by individual antenna elements using a code matrix. In general, the rank and determinant of the code matrix determines diversity gain and coding gain, respectively. The diversity gain affects the bit error probability at high SNR and should therefore be maximized before coding gain.

1.6.4 Spatial Multiplexing

In spatial multiplexing, different data streams are sent simultaneously on different antennas to achieve a higher data rate. It is also known as Bell Laboratories layered space–time (BLAST) coding. Each antenna transmits an independently modulated

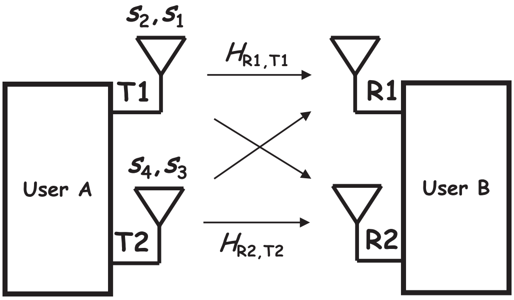


Figure 1.25 2 × 2 spatial multiplexing.

stream using a common channel bandwidth. Stream mapping is used if the number of data streams is less than the number of transmit antennas. The spatially multiplexed data streams are separated by the receive antennas using interference cancellation. Hence, the number of receive antennas should not be less than the number of transmit antennas. A 2 × 2 spatial multiplexing MIMO system is shown in Figure 1.25. This can be generalized to an $n \times m$ MIMO system. Data symbols s_1, s_2, s_3, and s_4 are mapped into a PSK or QAM constellation and applied independently to each OFDM subcarrier. The symbols are buffered to form the code matrix X that attempts to maximize the coding gain and channel capacity (data rate). In this case, the rate of X is 2 because four different symbols are transmitted in two symbol periods.

$$X = \begin{bmatrix} s_1 & s_2 \\ s_3 & s_4 \end{bmatrix}. \tag{1.4}$$

The general governing equation of matrices is given as

$$Y = HX + N, \tag{1.5}$$

where

Y = received signals, H = channel response

X = transmitted signals, N = channel noise.

If $H_{R1,T1} = 0.8$, $H_{R2,T1} = 0.2$, $H_{R1,T2} = 0.1$, $H_{R2,T2} = -0.8$, the channel matrix H for the 2 × 2 MIMO system becomes

$$H = \begin{bmatrix} 0.8 & 0.1 \\ 0.2 & -0.8 \end{bmatrix}.$$

In this case, the cross-coupled (self-interfering) signals are $H_{R2,T1}$ and $H_{R1,T2}$. A negative value for H indicates a signal phase reversal.

Full spatial diversity is not achieved with spatial multiplexing. The reduced diversity gain is due to a lack of redundancy across the antennas. This degrades the bit error performance and reduces the net data throughput at low SNR. In addition, the maximum data rate becomes limited due to path correlation, which increases as the number of transmit antennas increases. Finally, maintaining a consistent MCS is a challenge due to differing SNRs experienced by the receive antennas. Nonlinear receivers (e.g., maximum likelihood detectors) are complex to implement but may be needed to achieve better performance. These receivers require signals to be decoded jointly, allowing signal degradation to be overcome by decorrelating the streams using the channel knowledge. However, in the presence of an unknown interferer, the channel model may remain unknown to the receiver. Hence, instead of decorrelation across antennas, joint processing may actually amplify the detrimental impact of interference.

1.6.5 Space–Time Coding

Space–time coding (STC) was first introduced by Tarokh [4] from AT&T Research Labs in 1998. As the name implies, STC makes use of space (multiple antennas) and time domains in encoding/decoding data symbols. It combines channel coding (temporal diversity) with multiple antennas at the transmitter (spatial diversity) to improve transmission reliability and range. In general, STC optimizes diversity and range, whereas spatial multiplexing optimizes capacity. However, both methods require concurrent antenna transmissions in close proximity, as well as channel estimation at the transmitter or receiver. Recall that in spatial multiplexing, collocated antennas from the same device transmit different data symbols concurrently to increase the data rate. However, this is done at the expense of self-interference, which can be far more significant than weakened signal reflections resulting from multipath. Therefore, it is more difficult to recover the signal constellation in a spatially multiplexed MIMO system than a SISO system since there is a high probability of signal coupling in the data streams transmitted by antennas from the same device. Hence, spatial multiplexing demands better SNR than SISO to maintain the same BER.

In STC, replica data symbols among multiple antenna signals are transmitted to maximize spatial diversity gain and minimize bit error rate. This redundancy improves the robustness and SNR/operating range even in fading scenarios, but may achieve a lower data throughput than a comparable spatial multiplexing system. Unlike spatial multiplexing, multiple antennas at the receiver are optional. This is an advantage in asymmetric links where high DL rates are crucial. It also simplifies receiver design with the use of fewer antennas. The high DL rates are achieved using higher-order modulations. This is in contrast to spatial multiplexing, which improves data rates using independent spatial streams. Well-constructed space–time codes achieve a fair amount of coding gain in addition to diversity gain. These codes are useful if the channel is unknown to the transmitter. They include space–time block codes, Trellis codes, and Turbo codes. If space–time block codes are used, then this gives rise to space–time block coding (STBC).

1.6.6 Alamouti Space–Time Coding

The Alamouti STC is a special case of STC using two transmit and one receive antennas (2 × 1). It is an example of a multiple input single output system (MISO)

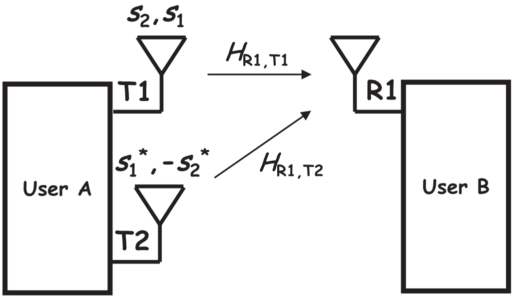


Figure 1.26 Alamouti STC.

that operates on pairs of consecutive OFDM symbols. In the first symbol interval, two OFDM symbols (namely s_1 from T1 and $-s_2^*$ from T2) are simultaneously transmitted from two antennas, where * denotes complex conjugation corresponding to a 180° reversal in phase. In the next symbol interval, signal s_2 is transmitted from T1, and signal s_1^* is transmitted from T2. This is illustrated in Figure 1.26. The code matrix becomes

$$X = \begin{bmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{bmatrix}. \tag{1.6}$$

Since two data symbols s_1 and s_2 are transmitted in two symbol periods t_1 and t_2, there is no increase in capacity, and the rate is 1. The redundant symbols are added in space and time to improve diversity. Note that $XX^T = \alpha I_{2\times 2}$, where X^T is transpose of X, α is a constant, and $I_{2\times 2}$ is a 2×2 identity matrix. The receiver buffers the received symbols (r_1, r_2) for two symbol periods, which are given by:

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} = \begin{bmatrix} s_1 h_1 + s_2 h_2 + n_1 \\ -s_2^* h_1 + s_1^* h_2 + n_2 \end{bmatrix}. \tag{1.7}$$

The receiver estimates the channel responses h_1 and h_2 via channel sounding. These responses correspond to $H_{R1,T1}$ and $H_{R1,T2}$, respectively, in Figure 1.26. Data preambles can be transmitted on alternate OFDM subcarriers from the two antennas. For example, the even subcarriers can be transmitted on antenna 1, whereas the odd subcarriers can be transmitted on antenna 2. Taking conjugates for r_2 (this operation does not introduce any correlation for the noise entities), we obtain:

$$\begin{bmatrix} r_1 \\ r_2^* \end{bmatrix} = \begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2^* \end{bmatrix}. \tag{1.8}$$

Thus, the recovered symbols s_1, s_2 are decoupled and can be decoded independently. This is an important feature in spatial stream recovery since the two spatial streams may encounter significantly different levels of signal attenuation and interference as they propagate through the wireless medium.

The Almouti STC achieves excellent performance even in highly correlated channels and is usable in the mobile mode. This is because the Alamouti STC employs orthogonal signals that work with a variety of channels. In addition, because only a single receive antenna is required, the best MCS can be chosen based on the prevailing SNR. If two receive antennas are used and placed a quarter of a wavelength apart, the independent fading on each antenna ensures maximum diversity due to the redundancy introduced by the Alamouti STC. Unlike spatial multiplexing, which requires joint multistream processing, single-stream decoding using linear processing is possible with the Alamouti STC receiver, resulting in a single receive chain. However, channel estimates are still needed at the receiver.

1.6.7 Beamforming MIMO Antenna Arrays

Conventional SISO and MIMO antenna types are designed to radiate laterally (not up or down) and this may not always be optimal in many deployments since the BS or AP is located at an elevated position compared with the user device in order to improve the signal coverage. For example, wireless APs are normally mounted above ceiling panels in enterprises. The use of active narrow-beam antennas offers the possibility of improved capacity at long distances and low power operation at short distances. Beamforming antenna arrays employ directional antennas that attempt to create LOS paths to the receivers so that the impact of multipath can be minimized. Dielectric lens and patch arrays on soft substrates can be used to produce low-cost antennas that are nearly omnidirectional in the azimuth direction but with narrow beams in elevation. Such antennas radiate energy only in a particular direction, providing gain along the intended direction and attenuation in the undesired directions. In doing so, a wireless coverage area can be broken down into smaller areas, each serviced by a directional antenna. This is equivalent to increasing the number of communication links without employing additional bandwidth, which increases system capacity. If only a single link is needed, the use of a directional antenna may reduce the transmit power and interference significantly, thus minimizing the overall energy consumption. Selection algorithms at the antennas assess the received signal strength and quality from different paths, giving the system a high probability of finding a path that is not corrupted by fading or interference. In addition, as can be seen in Figure 1.27, beamforming antennas are less dependent on correlated channels because a single channel estimate is sufficient for each link. This is in contrast to STC and spatial multiplexing, which require two or more channel estimates per link.

The use of beamforming antenna arrays is a promising technique to increase cell coverage and improve cell edge spectral efficiencies. Wireless links become implicitly distinct like in wired networks and can be physically separated and replicated using the same channel bandwidth. These arrays also provide a limited form of security by ensuring that the signal is directed at the legitimate station. Each pair of transmit–receive antennas may also use a different frequency channel to boost rates without causing self-interference. The antenna arrays can be made adaptive by using pilot signals on the UL and DL. When employed by the BS, such antennas can be used to track the locations of mobile users. Beamforming is normally applied on the DL and not on the UL. Since the antenna from the handset needs to be aimed at the correct antenna from the BS, it is difficult for beamforming systems to allow

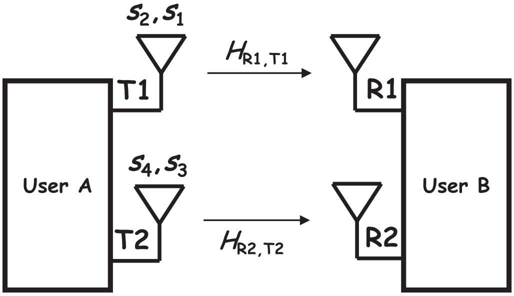


Figure 1.27 Beamforming antenna arrays.

new users to join a wireless cell or perform a handoff (or handover) when moving to a new cell. Thus, many handheld devices employ omnidirectional antennas for new user transmission. This implies that an access protocol is needed to resolve contention on the UL. A major drawback of beamforming is the so-called flashlight effect, where the measured channel quality becomes inaccurate due to the interaction of simultaneous active beams from neighboring arrays using the same frequency channel. Coordinated beamforming and switching has been suggested as a possible solution to this problem.

Beamforming antennas require more overheads in point-to-multipoint communications, since all arrays have to send the same information individually. For example, the same information has to be repeated periodically to all recipients or a specific group of recipients in broadcast or multicast operations. Although no additional bandwidth is required for this operation, the overall transmit power may be increased compared with unicast transmission. The need for more transmit power to service multiple links simultaneously is sometimes offset by the higher antenna gain that is available in the beamforming antennas. As the antenna gain is increased, the beamwidth becomes narrower. Very narrow antenna beamwidths—those associated with an antenna gain of 10 dBi or more—require automated antenna alignment. Beam-steering arrays can point the antenna beams in specific directions without manual intervention by dynamically adjusting the input voltage to each array to adaptively control the signal radiation angle. However, a beamforming protocol is needed to allow devices to discover each other, establish connections by aligning the beams, and coordinate operation in an efficient and interoperable manner. Such a protocol is critical to the performance of higher frequency systems (e.g., 60 GHz systems) since the beamwidths tend to be much narrower than lower frequency systems (e.g., 5 GHz systems) due to the higher gain. Once connected, antenna settings can be refined to maximize transmit and receive gains. As channels change, adjustments to the antenna settings can be made to optimize performance. This allows the highest data rates to be achieved, even with time-varying channels, such as those seen by mobile devices.

1.6.8 Downlink MIMO Architectures

Figure 1.28 illustrates two codeword options in DL MIMO implementation. In single codeword (SCW) or vertically encoded MIMO, only one encoder block or layer is

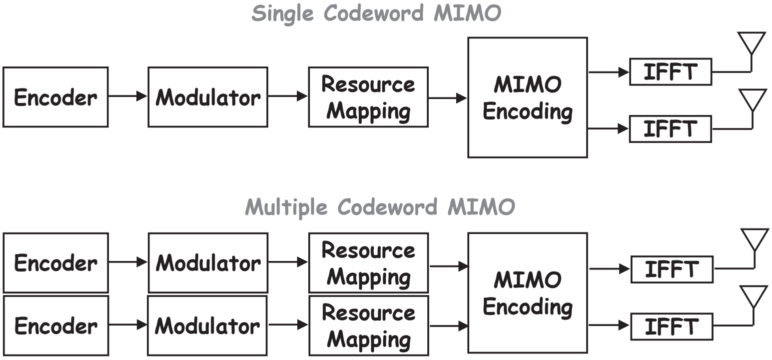


Figure 1.28 Single and multiple codeword MIMO.

required. Channel quality information (CQI) is used to indicate the quality of the DL channel as experienced by the user handset. For SCW, a single data stream is required, but spatial diversity is possible. Multiple codeword (MCW) or horizontally encoded MIMO employs multiple encoders or layers. In this case, CQI for multiple data streams are required. In addition, codeword-based successive interference cancellation to cancel out multiuser interference is needed at the receiver. For the same receiver type, the performance of the two MIMO implementations is almost the same.

1.6.9 Open-Loop and Closed-Loop MIMO

With open-loop transmit diversity, channel knowledge cannot be acquired at the transmitter due to the absence of an UL feedback path. Only CQI is available. Thus, open-loop MIMO is well suited for deployments that require high mobility speeds because UL feedback is restricted in these deployments. On the other hand, closed-loop transmit diversity allows channel state feedback to be sent by the handset when directed by the BS. Closed-loop MIMO requires CQI and channel state information (CSI) to be measured and sent back by the handset to the BS. This information is applied to a precoding procedure at the transmitter so that the BS can decide how best to modify the transmission, if needed. In this case, a robust signal format that uniquely identifies each transmitter is required before the signals are separated and demodulated. CSI is derived from channel sounding measurements. Examples of CSI are the channel matrix, the transmit rank indicator (RI), which indicates the number of useful spatial streams that can be spatially multiplexed based on the current channel conditions as experienced by the handset, and the precoding matrix indicator (PMI), which allows a handset to report to the BS its preferred precoding vector (normally selected from a codebook). With the additional information, closed-loop MIMO offers gains that may be comparable with using beamforming or adaptive MCS (Figure 1.29).

The handset can send quantized channel parameters to the BS to reduce the overheads. The PMI overhead can be measured by the feedback rate (in bits/ms/user) and the number of common DL pilots. Alternatively, both handset and BS can

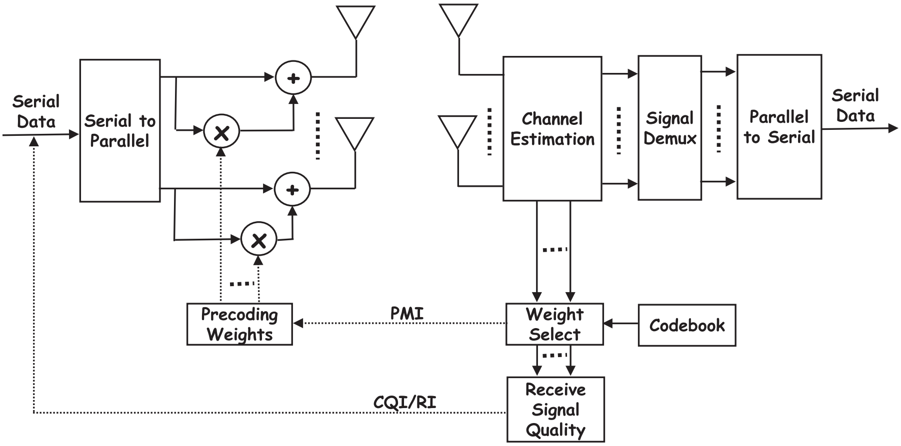


Figure 1.29 Precoding and channel quality measurement in MIMO.

predefine a set of codebooks, each containing an indexed set of coding matrices. For each channel rank, an index or value based on the codebook can be used, much like selecting a MCS from a table of MCSs. Thus, codebook-based closed-loop MIMO reduces the amount of feedback. Recall that the channel rank determines the number of simultaneous spatial streams that can be supported and should not exceed the number of transmit antennas.

1.6.10 Single-User and Multiuser MIMO

In single-user MIMO (SU-MIMO), only one user is communicating with the BS or AP. In multiuser MIMO (MU-MIMO), also known as collaborative or virtual MIMO, the BS or AP may communicate simultaneously with multiple users, thereby achieving higher aggregate rates compared with the single user case. User handsets listen to the pilot from the BS, which schedules the transmissions from different users to occur at the same time using the same channels. The overall network throughput is increased without the need for two or more transmitters at the user handset. However, the throughput of each user is not increased.

MU-MIMO is especially bandwidth efficient for the DL transmission due to the higher aggregation of traffic at the BS or AP. In addition, the BS or AP can efficiently allocate bandwidth resources based on channel feedback provided by the handset (e.g., transmit rank, PMI, CQI). Joint precoding can be centralized at the BS or AP, and cooperation between handsets is not required. MU-MIMO on the UL is challenging since the amount of cross-coupling needs to be estimated by the transmitting handsets, which may not be collocated. In addition, the BS or AP needs to estimate the channel responses of all independent spatial streams. Employing orthogonal UL transmissions (e.g., using the time or frequency domains) is a possible solution. A special case of such a solution is MU-SISO, where user handsets employ only a single antenna for transmission (Figure 1.30).

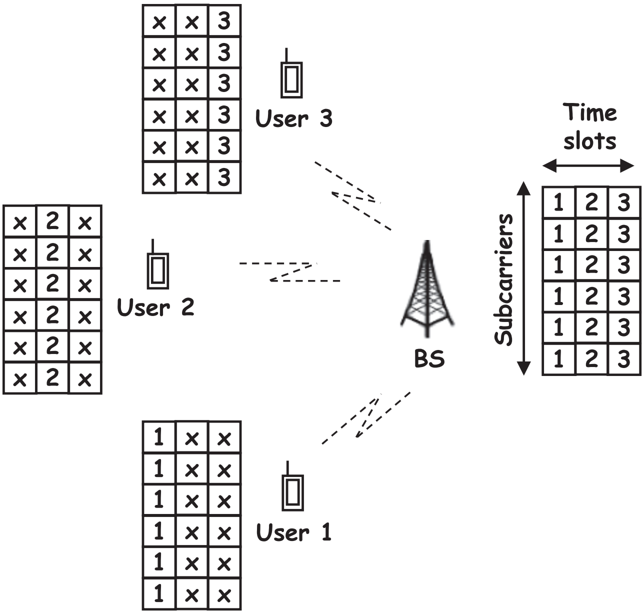


Figure 1.30 MU-SISO using nonoverlapping time slots.

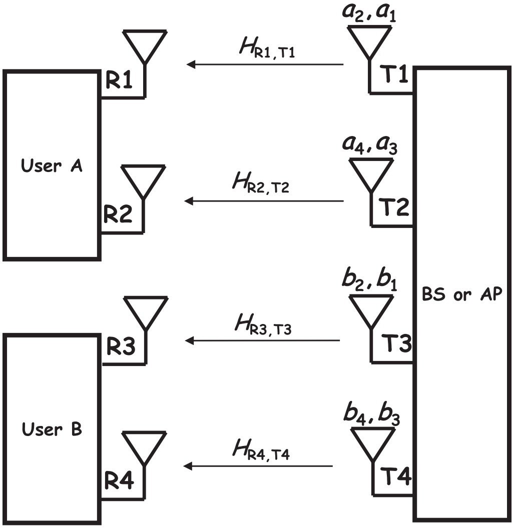


Figure 1.31 Downlink MU-MIMO transmission using beamforming antennas.

Since it is difficult to select the preferred spatial stream for different users using open-loop MIMO, closed-loop MIMO is recommended for MU-MIMO. Beamforming antenna arrays are capable of providing space division multiple access (SDMA) in MU-MIMO, where transmissions from multiple users can be divided in the spatial domain, and cross-coupling is significantly reduced. Figure 1.31 shows the operation of MU-MIMO using beamforming that is applied in the DL direction. Unlike DL beamforming, CSI may not be needed for UL beamforming. Note that MU-SISO results if only one transmit antenna is directed at the user handset.

1.7 INTERFERENCE

Interference is inherent in wireless transmission because it exists even when a single radio source is transmitting. Mitigating interference is a challenging problem in wireless networks because the transmitted signals cannot be shielded like in wired networks. In addition, the signal waveforms and transmission characteristics of the interfering sources may be unknown. Although transmit power control allows the radiation power to be utilized efficiently, this can be difficult to maintain in a fading or highly mobile environment. There are several forms of radio interference arising from known and unknown sources. The first category comprises self-interference caused by a single source (e.g., multipath propagation and multiantenna transmission) and multiuser interference caused by multiple sources sending information at the same time. In the latter case, because all users employ a known wireless transmission method, an access protocol can often be designed to resolve contention among the users, thereby transforming the multiuser interference into a single self-interfering source. The second category of interference may be generated randomly by radars (e.g., military and weather), malicious jammers, microwave ovens, and others. For licensed band operation (e.g., cellular networks), only known forms of interference may exist. For unlicensed band operation (e.g., wireless LANs), all interference types may exist. The carrier to interference plus noise ratio (CINR) is normally employed to characterize the channel quality in the presence of interference. CINR measures the fidelity of the desired carrier signal in the presence of interference, whereas SNR measures the fidelity of the demodulated signal. Since emerging 4G wireless systems are interference limited, CINR is widely used.

Unlike CDMA systems, where receivers attempt to cancel known interfering sources, 802.11 wireless LANs employ carrier sensing to avoid known and unknown interference sources and minimize contention. In this case, packet transmission is deferred when interference is detected. Unnecessary deferment may occur due to oversensitive sensing (e.g., sensing the transmission of a remote user), resulting in longer delays. In addition, the difficulty in achieving equal sensing and transmit ranges makes frequency reuse challenging for such networks.

1.7.1 Spatial Frequency Reuse

TDMA cellular users achieve orthogonality by transmitting at different times using different time slots assigned by the BS. An important feature of TDMA cellular systems is that the same set of frequency channels is reused after a far-enough distance (Figure 1.32). Frequency planning is therefore required, where a frequency channel is assigned to a BS in a specific cell or wireless coverage area. The bandwidth within a cell may be further subdivided into subbands, each served by a sector of the antenna at the BS. Hence, a hexagonal (honeycomb) cell is normally used to

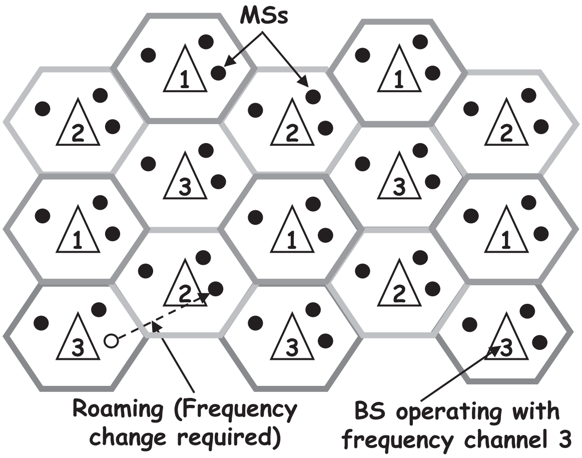


Figure 1.32 Three-cell frequency reuse in a sectorized TDMA cellular network.

represent six subbands in a single cell. When the user crosses a radio cell boundary, the assigned subband changes. Such spatial reutilization of bandwidth achieves a higher overall spectrum utilization by allowing multiple transmissions to take place simultaneously in different locations without causing excessive interference. A radio cell operating on one frequency channel is protected from the mutual interference arising in reused frequency channels by a surrounding ring of adjacent cells operating at other frequency channels. Ultimately, it is this mutual interference that limits the performance of TDMA cellular systems. In CDMA systems, the concept of frequency reuse translates to code reuse, where the same code is assigned to two users if they are far apart from one another (Figure 1.33).

A similar concept can be applied to indoor wireless LANs. However, because unlicensed frequency bands are employed in these networks, frequency planning is more complex and tedious due to the presence of known and unknown interfering sources that may change over time. With legacy 802.11 APs, manual site surveys are often conducted to calibrate the frequency channels according to the level of interference in each coverage area. However, newer 802.11 APs incorporate an ability to automatically select the best frequency channel based on channel measurement reports, thereby enabling plug-and-play operation. The channel may be switched dynamically as needed. The 802.11 user device will select the channel as dictated by the AP. In practice, channels that are spaced furthest apart (e.g., channels 1, 6, and 11 in the 2.4 GHz band) are used to reduce adjacent channel interference (caused by signal leakage from a neighboring channel).

The data rate selected by the AP (corresponding to a specific MCS) is typically based on the SNR of the wireless link at the receiver. A higher-order modulation, such as 64-QAM, is chosen if the SNR is high. The rate is continuously adapted to changing link conditions. If MIMO processing using spatial multiplexing is employed, rate selection is further complicated by the fact that multiple receive antennas on the same device may experience different SNRs. Since a single rate must be selected by all spatial streams originating from the same device, this implies

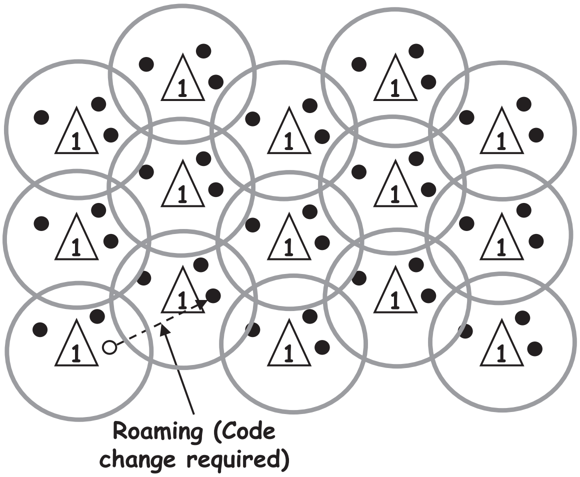


Figure 1.33 Spatial frequency reuse in a CDMA system.

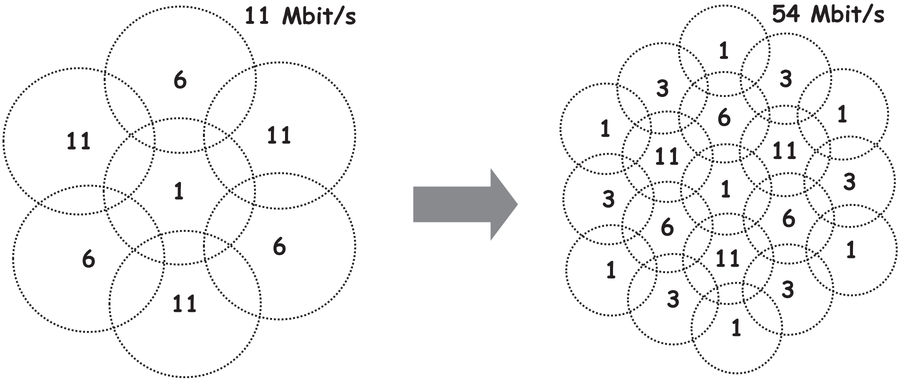


Figure 1.34 More 802.11 APs are needed to cover the same area with higher rates.

that the chosen rate is dictated by the lowest SNR from one of the receive antennas. Thus, the tradeoff between rate versus range is an important consideration in wireless LAN deployment. High data rates may only be relevant for small coverage areas. More 802.11 APs are therefore needed with denser coverage areas to carry more traffic (Figure 1.34). This implies greater cost, complexity, and interference with frequency reuse. In addition, more wires are needed to connect the APs.

1.7.2 Cochannel Interference

A common type of interference encountered in systems that employ frequency reuse is cochannel interference (CCI), which is caused by BSs or APs operating on

TABLE 1.4 CCI Mitigation Techniques

TDMA	CDMA	OFDMA
Control reuse patterns	Control spreading gain	Hard frequency reuse
Implement loose power control	Implement tight power control	Fractional frequency reuse
Implement guard times	Implement antenna arrays	Soft frequency reuse
Implement error control codes		
Implement antenna arrays		

the same channel. TDMA cellular users experience CCI interference from time-overlapping slots that originate from users transmitting using the same reuse frequencies in different radio cells. CCI can be minimized through proper frequency planning, but may not be completely eliminated. It is a major problem in direct sequence CDMA systems because the frequency channels are not spatially reused. In this case, CCI is a combination of multiaccess interference (MAI) due to transmission from multiple users, as well as ISI arising from multipath. The overall interference, which increases as the number of transmitting users increases, severely limits the capacity of these systems. It is important to realize that the performance of a spread spectrum receiver can vary widely even though the number of transmitting users is fixed. If the interfering users are located close to the receiver, MAI can be very large due to the higher receive power. Conversely, if the interfering users are far from the receiver, then MAI can be very small. This is known as the near/far effect, which can be fading induced. Thus, the geographic location of the receiver also influences the amount of MAI, and this changes as users move to different locations. Several CCI mitigation techniques commonly employed by TDMA, direct sequence CDMA, and OFDMA systems are listed in Table 1.4.

1.7.3 Multiuser Interference

Since MAI in direct sequence CDMA systems is highly structured, where spreading codes used by all users are known, interference cancellation techniques can be exploited. Examples include successive interference cancellation (SIC) and parallel interference cancellation (PIC) receivers that employ multiuser detection (MUD). These suboptimal linear detectors (usually implemented using a zero-forcing decorrelator) can completely eliminate MAI. They are an alternative to power control techniques that are used to overcome the near/far effect and channel fading, but the computational complexity may increase considerably with the number of users, especially when transmissions are conducted asynchronously. In addition to mitigating the near/far effect, MUD has the more fundamental potential of raising capacity by canceling multiuser interference. SIC appears to be simpler and requires less hardware than PIC. However, the processing delay presents the biggest drawback to SIC, since only a single user bit is decoded at each stage, and, thus, it takes at least x bit times to decode all users for each bit. In the case of PIC, it takes s bit-times (where s is the number of stages) to decode all users for each bit, and s is usually smaller than x. While increasing s improves the BER performance to a level similar to SIC, this may be achieved at the expense of excessive hardware complexity and processing delay. Both SIC and PIC schemes achieve better performance than the conventional detector. At the same time, they are simpler and less complex than the optimum detector. The processing speed of the decorrelator may limit the number of possible cancellations.

1.8 MOBILITY AND HANDOFF

Tracking a mobile user in a cellular environment creates additional complexity since it requires rules for handoff (or handover) and roaming. In addition, one must also cope with rapid fluctuations in the received signal power and traffic load variations that may change with user location and time. Handoff refers to the process of changing communication links so that acceptable channel quality and uninterrupted service can be maintained when users move across radio cell boundaries. Clearly, a handoff between two adjacent radio cells will change the number of active connections in each cell, and this in turn affects the traffic conditions and interference level in the cells. An effective handoff scheme must take into consideration three key factors:

- Propagation conditions
- Traffic load
- Switching and processing requirements.

It requires a neighboring BS to have a free channel with good signal quality and that the switchover be completed before any significant deterioration to the existing link occurs. The handoff process essentially comprises two stages namely,

- Channel quality evaluation and handoff initiation
- Allocation of bandwidth resources.

1.8.1 Intercell versus Intracell Handoff

When the user crosses two adjacent radio cells, an intercell handoff is required so that an acceptable channel quality for the connection can be maintained. Sometimes, an intracell handoff is desirable when the link with the serving BS is affected by excessive interference, while another link with the same BS (using a different antenna sector) can provide better quality.

1.8.2 Mobile-Initiated versus Network-Initiated Handoff

Network-controlled handoff (NCHO) has been widely used in first-generation analog cellular systems (e.g., AMPS). With NCHO, the channel quality is monitored only by BSs. The handoff decision is made under the centralized control of a mobile telephone switching office (MTSO). Typically, NCHO algorithms have handoff network delays in the order of several seconds, and have relatively infrequent updates of the channel quality from the alternate BSs. Mobile-assisted handoff (MAHO) is a decentralized strategy used in several digital cellular systems. The channel quality is measured by both the serving BS and the user. In TDMA systems, the user measures the signal strength during the intervals when it is not allocated a time slot. Channel quality measurements of alternate BSs are performed by the user. Link measurements at the serving BS are relayed to the user. The handoff decision is made by the user. The MAHO achieves the lowest delay (about 100 ms) and exhibits high reliability.

1.8.3 Forward versus Backward Handoff

Handoffs differ in the way a connection is transferred to a new link. Backward handoff algorithms initiate the handoff process through the currently serving BS. No access to the new link is made until resources have been allocated. The advantage of backward algorithms is that the signaling information is transmitted through an existing radio link, and, therefore, the establishment of a new signaling channel is not required during the initial stages of the handoff process. The disadvantage is that the algorithm may fail in conditions where the channel quality with the serving BS is rapidly deteriorating. This type of handoff is used mostly in TDMA cellular systems. Forward handoff algorithms activate the handoff process via a channel with the target (alternate) BS without relying on the current BS. Although the handoff process is faster, this is achieved at the expense of a reduction in handoff reliability. Handoffs can also be hard or soft. Hard handoffs release the radio link with the current BS at the same time when the new link with the target BS is established. Such handoffs are used in many TDMA cellular systems. Soft handoffs maintain radio links with at least two BSs. A link is dropped only when the signal level falls below a certain threshold. Soft handoffs are used in CDMA cellular systems.

1.9 CHANNEL ASSIGNMENT STRATEGIES

The assignment of frequency channels to BSs and to users is a fundamental operation of a mobile communication network. A classification of channel assignment schemes is shown in Figure 1.35. In fixed channel assignment, channels are assigned to cells for a relatively long period of time. This method obviously results in poor utilization of the available bandwidth when traffic patterns are bursty and change over time. Dynamic channel assignment represents the opposite extreme where channels are assigned to radio cells only when required. Such techniques can adapt to traffic load changes in real time but suffer from increased network management overhead. Between the extremes of fixed and dynamic channel assignment lie flexible channel assignment, channel borrowing, and hybrid channel assignment. Flexible channel assignment is essentially fixed assignment that is altered regularly according to predicted changes in the traffic load. Borrowing strategies can also be considered to be a variant of fixed channel assignment where channels not in use in their allocated cell can be temporarily transferred to congested cells on a

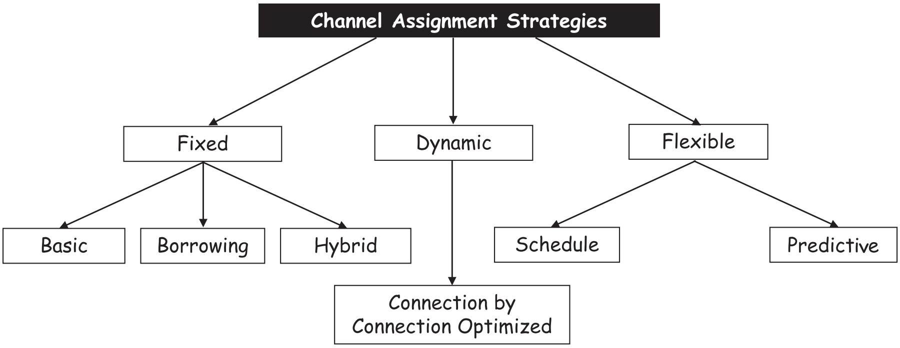


Figure 1.35 Channel assignment classification.

connection-by-connection basis. Hybrid assignment allocates a fraction of the channels according to fixed assignment and the rest according to dynamic assignment.

1.9.1 Medium Access Control Protocols

For wireless networks, sharing of bandwidth is essential because radio spectrum is not only expensive but also inherently limited. For example, Verizon recently spent \$3.6 billion to purchase just 20 MHz of the advanced wireless services (AWS) spectrum licenses from major U.S. cable companies to expand its 4G wireless services. This is in contrast to wired networks, where bandwidth can be increased arbitrarily by adding extra cables. However, the broadcast nature of the wireless link poses a difficult problem for multiuser access in that the success of a transmission is no longer independent of other transmissions. To make a transmission successful, interference must be avoided or at least controlled. Otherwise, multiple simultaneous transmissions may lead to collisions and corrupted signals. This sharing task is made harder when disparate traffic types created by multimedia applications are to be transported across the network. A multiple access or MAC protocol is required to resolve access contention among users and transform a broadcast wireless network into a logical point-to-point network. The domains that contention resolution can be achieved include time, frequency, code, space, or some combination. Note that in general, packet reception consumes less resources than packet transmission because there is no access contention in packet reception.

Defining a MAC protocol essentially implies the specification of a set of rules to be followed by each member of the user population in order to share a common bandwidth resource in a cooperative manner. The overall communications channel can be divided into subchannels, which are then assigned to contending users. Typically, there are more users than the available subchannels, but only a fraction of all users have packets to transmit at any given time. The central problem, therefore, is to locate the users with data to send in order for these users to share the channel efficiently. In general, MAC protocols can be categorized under contention, reservation, polling, and fixed allocation (Figure 1.36). A comparison of these protocols is shown in Table 1.5. At one extreme, where no control is enforced, two or more users

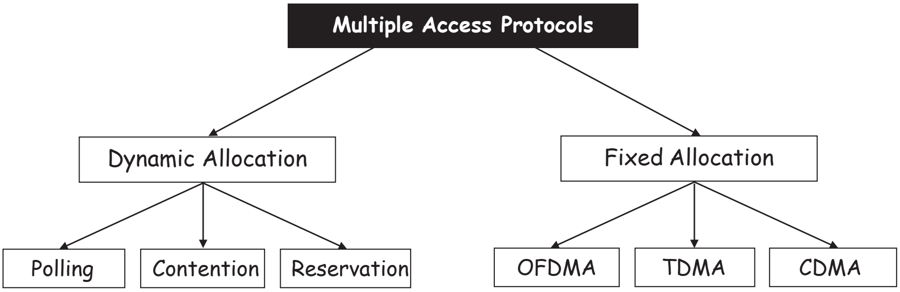


Figure 1.36 Classification of MAC protocols.

TABLE 1.5 Comparison of MAC Protocols

MAC protocol	Collisions	Control overhead	Idle time
Contention	Yes	No	No
Reservation	No	Yes	No
Polling	No	Yes	No
Fixed allocation	No	No	Yes

may transmit at the same time and conflicts may occur. These uncontrolled (contention) schemes are very easy to implement, but pay the price in the form of wasted bandwidth due to collisions. The other extreme is represented by a rigid system of static control (fixed allocation), where each user is permanently assigned a portion of the total bandwidth for its exclusive use. Such control places hard limits on the number of users sharing a given bandwidth. The class of dynamic control protocols (reservation and polling) dedicates a small part of the bandwidth for control, and this control information is used to determine the identity of the users with data to transmit. While the performance of contention and reservation techniques is dependent on the combined traffic from all users in the network, the performance of static allocation and polling schemes is strongly influenced by the traffic requirements of each individual user.

The multiple access capability of spread spectrum systems is distinctly different from narrowband systems. Not only can simultaneous transmissions be tolerated, the number of such transmissions should be large in order to justify the bandwidth spreading and achieve high network capacity. Thus, as long as different receivers are involved, MAC protocols in spread spectrum systems are normally designed to have multiple transmissions taking place at the same time. In this case, interference may sometimes dominate over noise as an error-producing mechanism.

1.9.2 Signal Duplexing Techniques

The duplexing scheme is usually described along with a particular MAC scheme. Half-duplex transmission is usually employed in radio transceivers due to cost and implementation considerations. For example, 802.11 transmission is half-duplex because either the AP is transmitting or one of the users within its coverage area is transmitting. Two half-duplex transceivers can be employed to achieve full-duplex transmission, each operating on an independent frequency band, one for the transmission and the other for reception. This introduces technical difficulties (e.g., saturation of the receiving antenna by the transmitter) and additional cost (e.g., separate radio hardware needed for each frequency band). An alternative is to employ duplexing methods, such as frequency division duplex (FDD) and time division duplex (TDD). These methods are commonly employed in cellular standards. The amount of spectrum required for both FDD and TDD is similar. The difference lies in the fact that FDD employs two bands of spectrum separated by a certain minimum bandwidth (guard band), while TDD requires only one band of frequencies for UL and DL transmissions. TDD has an advantage here, since it may be easier to find a single band of unassigned frequencies than two bands of frequencies separated by the required bandwidth.

TDD transceivers are half-duplex because the UL and DL bursts alternate on the same frequency channel but do not occur simultaneously. This is achieved by allocating a set of time slots to the UL and DL frames, with the ratio determining the relative UL and DL bandwidth requirements. This separation produces a delay, which is inherent in any time division techniques (e.g., TDMA) that employ a fixed frame structure. The DL frame is always transmitted first so that users at different locations can perform ranging operations on the UL to compensate for different signal propagation delays. The UL frame is transmitted after the DL frame. Sufficient turnaround time must be provided to allow the half-duplex transceivers to switch from transmit to receive and vice-versa.

In FDD, the UL and DL frames operate on nonoverlapping frequency channels, and can therefore transmit simultaneously (i.e., UL/DL frames can be coincident in time). FDD requires some frequency separation between UL/DL channel pairs to mitigate self-interference, but unlike TDD, no time gaps are required. Both TDD/FDD alternatives support adaptive burst profiles where MCS options may be dynamically assigned on a burst-by-burst basis. An important advantage of TDD over FDD is the ability to support asymmetric traffic loads since transmission time in TDD can be apportioned flexibly between the UL and DL. FDD systems may transmit at a longer range than TDD systems (or equivalently, requires lower transmit power). FDD also avoids the turnaround delay associated with TDD. Unlike TDD, a FDD BS may not use the UL channel information to enhance DL transmission.

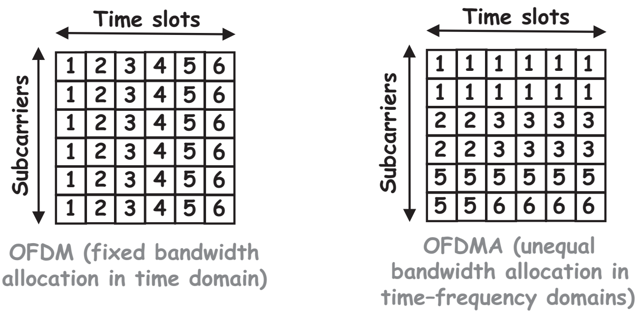
Since FDD employs different frequencies in the forward and reverse directions of transmission, diversity antennas have to be employed at both the BS and the user. TDD, however, benefits from antenna diversity without the need for multiple antennas. This is attributed to the fact that for a given frequency, the signal attenuation in a radio channel is typically reciprocal. For example, a TDD system can have the BS select the best signal from its antenna when receiving a specific signal from a user. When the BS next transmits to that user, it can employ the same antenna with the same power budget. Note that while the channel is reciprocal in attenuation, this may not apply to the interference. Thus, dual antennas may still be needed for a TDD system.

Another antenna-related design consideration when selecting a duplex scheme is whether a duplexer is required. A duplexer adds weight and cost to a radio transceiver, and can place a limit on the minimum size of a handset. This is because dual channels imply more complex receivers than a single-channel system, where all receivers operate at the same frequency. TDD is a burst mode transmission scheme. During the transmission, the receiver is deactivated. Thus, TDD systems are capable of providing bidirectional antenna diversity gain without employing duplexers. However, in terms of equipment utilization, a TDD transceiver effectively remains idle half of the time. Although duplexers are employed by most FDD systems, they are not required in FDD systems employing TDMA since the transmit and receive time slots occur at different times. A simple RF switch performs the function of the duplexer, but is less complex, smaller in size, and cheaper. Such a switch connects the antenna to the transmitter when a transmit burst is required and to the receiver for the incoming signal.

Since FDD uses different frequencies for each direction of transmission, interference is not possible even if the timing on the two frequencies is not synchronized. However, on each TDD link, precise synchronization is required. Otherwise, overlapping transmit and receive bursts will result in a reduction of overall system capacity. Each user is synchronized to a common clock time reference, allowing transmissions by each user to arrive at the intended receiver at a time agreed upon by all users. Ranging estimates of the forward plus reverse propagation times are also required to allow each handset to compensate for differences between its distance to the BS. This is fundamental for establishing new connections, since the handset can normally detect the BS signal (transmitted at higher power and from an elevated location), but the BS may not be able to detect the signal from the power-limited handset.

1.9.3 Orthogonal Frequency Division Multiple Access

Orthogonal frequency division multiple access (OFDMA) has become a key enabling air interface for 4G wireless standards. Like OFDM, OFDMA is mitigates multipath fading and achieves efficient spectral utilization. Unlike OFDM, where all subcarriers within the channel are assigned to same user, OFDMA allows BSs to assign a subset of subcarriers to different users for a predetermined amount of time (Figure 1.37). This procedure results in more granular bandwidth allocation and can be used to optimize the total system throughput under specific constraints (e.g., packet delay, transmit power, etc.). In addition, user contention on the UL is avoided, whereas a separate MAC protocol is required with OFDM. For instance, the 802.11 standard employs OFDM together with carrier sense multiple access (CSMA) to resolve access contention, which is not required in OFDMA. Like CDMA, an OFDMA system achieves a frequency reuse of 1. In this case, different sets of subcarriers may be assigned to different users within a cell or to the same user moving across cell boundaries. Thus, all wireless cells employ the same frequency band but may employ different sets of subcarriers, much like CDMA networks using the same frequency band but different spreading codes. This means that unlike conventional TDMA cellular deployment that requires different frequency bands for neighboring



OFDM (fixed bandwidth allocation in time domain)

OFDMA (unequal bandwidth allocation in time–frequency domains)

Figure 1.37 OFDM and OFDMA bandwidth allocation.

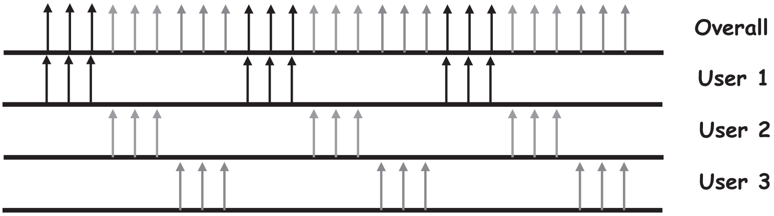


Figure 1.38 Grouped subcarrier allocation.

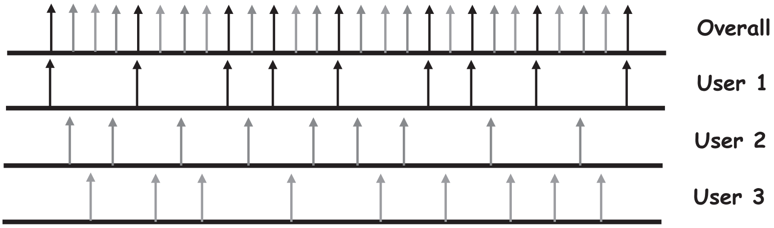


Figure 1.39 Permutated (distributed) subcarrier allocation.

cells, OFDMA allows overlapped cells to share the same frequency resource, which leads to a more flexible deployment, since frequency planning is not required. More importantly, fast handoffs in high-mobility communications can be achieved using the same frequency band in adjacent cells.

There are two general methods where OFDMA subcarriers can be allocated to different users. These methods provide frequency diversity that minimizes multipath interference. Figure 1.38 shows the grouped subcarrier allocation. Subcarriers may interact at the cell edge where coverage areas overlap. In this case, a group of subcarriers may collide. To further improve frequency diversity, permutated (distributed) subcarrier allocation allows non-contiguous subcarriers to vary from cell to cell (Figure 1.39). In this case, one or more subcarriers may collide at the cell edge.

OFDMA minimizes the transmit power from the user handset. Alternatively, it can allow more transmit power per subcarrier, thereby improving UL range performance. OFDMA requires a lower training overhead and subcarrier allocation can be dynamic (i.e., can vary in each data burst). OFDMA is less helpful on the DL because the power is shared, and the power per subcarrier from the BS typically remains constant. Time and frequency alignment is crucial in OFDMA due to the need to maintain orthogonal signals from different users.

OFDMA networks need to manage CCI. In particular, handsets located at the cell edge can be prone such interference. A solution is to selectively allocate subcarriers, rates, and power to the handsets depending on their location in the cell. There are three major frequency reuse patterns for mitigating intercell interference: hard frequency reuse, fractional frequency reuse (FFR), and soft frequency reuse. Hard frequency reuse splits the overall bandwidth into a number of distinct subbands so that neighboring cells transmit on noninterfering subbands. FFR splits the given bandwidth into inner and outer parts. It allocates the inner part to users located close to the BS with reduced transmit power. Thus, a frequency reuse factor of 1 is possible for the inner part. For users closer to the cell edge, a fraction of the outer part is employed with a frequency reuse factor greater than 1. With soft frequency reuse, the overall bandwidth is shared by all BSs (i.e., a reuse factor of 1), but for each subcarrier, the BSs are restricted to a specified power limit. Although hard frequency reuse is simple to implement, it suffers from reduced spectral efficiency. On the other hand, soft frequency reuse achieves high spectral efficiency but requires the coordination of multiple BSs, as well as accurate channel reports and interference measurements at the cell edge. FFR is a compromise between hard and soft frequency reuse.

1.10 PERFORMANCE EVALUATION OF WIRELESS NETWORKS

We present several measurement studies to illustrate the performance of single-antenna and multiantenna 802.11 wireless LANs. The studies are based on a single wireless connection between an AP and a laptop, both capable of supporting up to two spatial antenna streams. Figure 1.40 shows that at short distances, the rate for 802.11n is about 2.4 times greater than 802.11a/g. This is because 802.11n employs spatially multiplexed MIMO as opposed to SISO in 802.11a/g. However, at longer distances, the rates for 802.11a/g (54 Mbit/s maximum rate) are actually higher than 802.11n (maximum rate is set to 130 Mbit/s for this experiment). In all cases, the same channel bandwidth of 20 MHz is chosen. Thus, the gains for using multiple spatial streams may be relevant for short distances. At longer distances, single antenna systems achieve better throughput performance. The key factor that degrades MIMO performance at longer distances is the need for link adaptation when independent antenna branches collocated on the same device experience varying signal levels.

While mobility issues assume lesser importance in wireless LANs than in cellular systems, adjusting the MCS to accommodate different operating distances is a difficult challenge due to varying channel and interference conditions. As shown

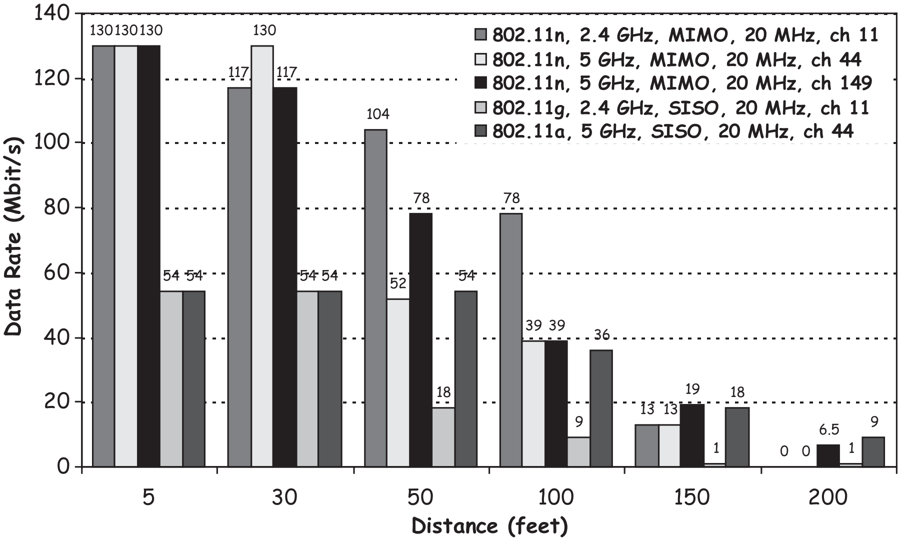


Figure 1.40 PHY data rate versus range (enterprise environment).

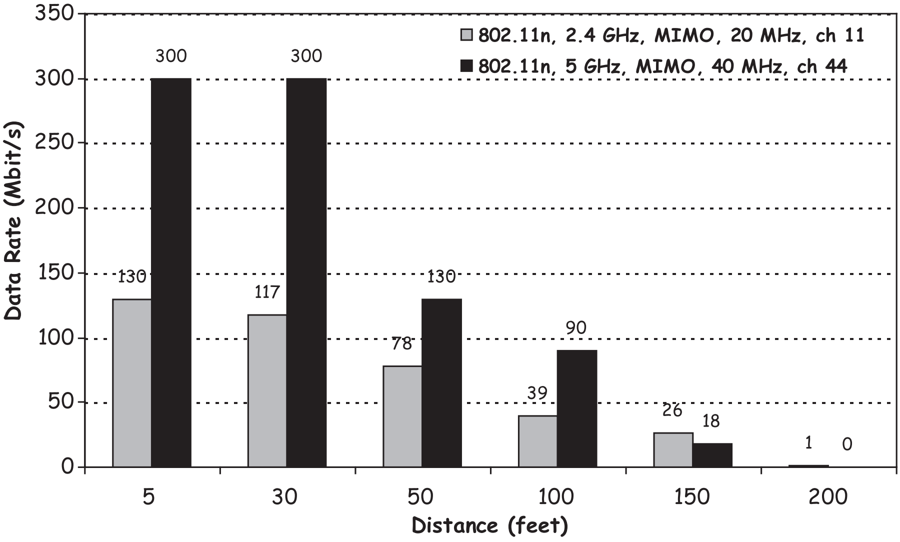


Figure 1.41 PHY data rate versus range (home environment).

in Figure 1.40, the rate of a 5 GHz radio can be better or worse than a 2.4 GHz radio due to different levels of interference in a dense office environment. However, a 2.4 GHz radio typically achieves a longer range than a 5 GHz radio for the same transmit power. This is exemplified in Figure 1.41, which shows the range performance in a typical single-family home. In this case, despite using a wider channel bandwidth of 40 MHz, the operational range for the 5 GHz radio is shorter than the

2.4 GHz radio. Similarly, the use of MIMO may only be effective at short distances (e.g., at 100 ft or less). A higher transmit power can improve range (e.g., using channel 149 instead of channel 44 in the 5 GHz band).

1.10.1 Impact of Link Adaptation

The use of a broad and dynamic range of transmission rates within a single coverage area and frequency channel may severely impact network performance. For instance, a few users transmitting continuously at low rates at the fringe areas of reception may degrade the overall data throughput of high-rate users located closer to the AP. The problem exists in both forward and reverse directions of transmission, and is due to data packets from high rate users frequently waiting for lower rate packets to complete transmission. This is caused by the native 802.11 MAC protocol, which is based on CSMA, where user devices or APs can initiate packet transmission whenever the wireless medium is sensed to be idle. Similarly, a mixture of low rate legacy SISO devices and new high rate MIMO devices operating within the same distance and channel may also lead to this problem. As an example, suppose a low-rate user is transmitting 500-byte packets at 1 Mbit/s (4 ms packet transmission time), whereas a high rate user is transmitting 500-byte packets at 130 Mbit/s (0.031 ms packet transmission time). The packets are transmitted continuously, which is typical of video streaming and peer-to-peer file download applications. If the packet transmission alternates between the high- and low-rate users, then the actual data rate for both users is 4000/(4 + 0.031) or 0.99 Mbit/s, which is very close to 1 Mbit/s, but over 130 times slower than 130 Mbit/s. One solution is to allocate additional airtime for the high rate users to transmit or receive a larger number of packets. This can be achieved by preallocating the transmission time for all users in the same coverage area using time slots. Another solution is to allow users located at the fringe areas to transmit at a higher power in order to maintain a high SNR for supporting a higher rate. However, this high-power option may be done at the expense of greater CCI and multipath interference. Yet another solution is to implement MU-MIMO, where different spatial streams are directed at users transmitting at different rates.

1.10.2 Impact of Higher Layers

While the PHY data rate is a key consideration, what really counts is the actual data throughput, which takes into account the overheads associated with the lower layers (i.e., PHY and MAC). The transmission control protocol (TCP) is a bidirectional network protocol that incorporates congestion and flow control using a sliding window mechanism to minimize network congestion. In doing so, device buffer overflows leading to packet loss are avoided. Thus, the end-to-end TCP throughput is highly dependent on how fast the window rotates. The maximum TCP throughput is affected by the roundtrip propagation delay and the congestion window size. The maximum throughput can be estimated by the ratio of the maximum window size over the roundtrip time. For example, if the maximum window size is 65,535 bytes (or 64 KB) and the roundtrip time is 0.5 second, then the maximum TCP throughput becomes limited to about 1 Mbit/s. The roundtrip time is dependent on the span of the network. Clearly, it is important to match the maximum TCP throughput with the available PHY and MAC rates. Unlike TCP, the user datagram protocol (UDP) is not subjected to such a throughput limit, since it is a unidirectional protocol.

Since TCP or UDP are designed for wired networks, there are several performance issues to consider when applying these protocols on wireless networks. This is because wireless networks encounter higher error rates, lower bandwidth, and more frequent signal outages than wired networks. For example, the performance of TCP over a wireless link may be affected by TCP's congestion avoidance algorithm, which is controlled by the sender. The algorithm is optimized for networks that experience low packet loss. Hence, TCP makes the implicit assumption that packet losses are a result of network congestion. However, pauses during handoffs (when users move among different wireless coverage areas) can be perceived as periods of heavy losses by the transport layer, causing retransmission timeouts. In addition, excessive retransmissions over the wireless link can lead to retransmission timeouts at the TCP layer. In both cases, TCP reacts by drastically reducing the current transmission rate. First, TCP reduces the transmit window size to restrict the amount of data flowing through the network. Second, it activates the slow-start mechanism that limits the increase in the data rate to the same rate that ACKs are received. The overall effects of the slow start algorithm is to increase the time required to reach the maximum throughput, both at the beginning of a TCP session and whenever a data segment is lost or corrupted. Finally, TCP resets the retransmission timer to a backoff interval that doubles with each consecutive timeout. These measures are effective in restricting the impact of congestion on the network. However, TCP takes a long time to recover from a transmission rate reduction, resulting in severe throughput degradation. Fortunately, such problems can be mitigated using a lossless link layer that provides local retransmissions using per-packet ACKs. In this case, a transmitter resends a packet after a timeout if it does not receive an ACK from the receiver. The 802.11 standard adopts this approach so that the wireless link becomes loss-free with a reduced effective bandwidth. As a result, most of the packet losses seen by the TCP sender are caused by network congestion due to buffer overflows at the user handset, AP, or other interconnected devices along the network path. Like TCP, UDP can also operate reliably over wireless networks that employ per-packet ACKs at the link layer, including 802.11 networks.

The hypertext transfer protocol (HTTP), based on the native TCP standard, is one of the most enduring network protocols of all time, providing a bidirectional mechanism to support Web browsing and more recently, high-quality audio/video transport. In addition, HTTP has been widely used to enable remote management and configuration of different Internet access devices, such as modems, routers, gateways, set-top boxes, and voice-over-IP (VoIP) phones. Figure 1.42 shows that the HTTP throughput can be reduced by as much as 50% of the PHY rates when streaming a 1080p high-definition (HD) video. The PHY rates are selected based on the received signal strength and not on the bandwidth demanded by the application. While the overheads of the 802.11n PHY and MAC can be substantial, optimizing the HTTP window size (and hence the HTTP throughout) is also important in ensuring that the PHY rate utilization is maximized. Figure 1.43 shows the impact of MAC layer ACKs on the HTTP throughput. As can be seen, when local packet ACKs are disabled, more end-to-end retransmissions will be handled by the HTTP layer.

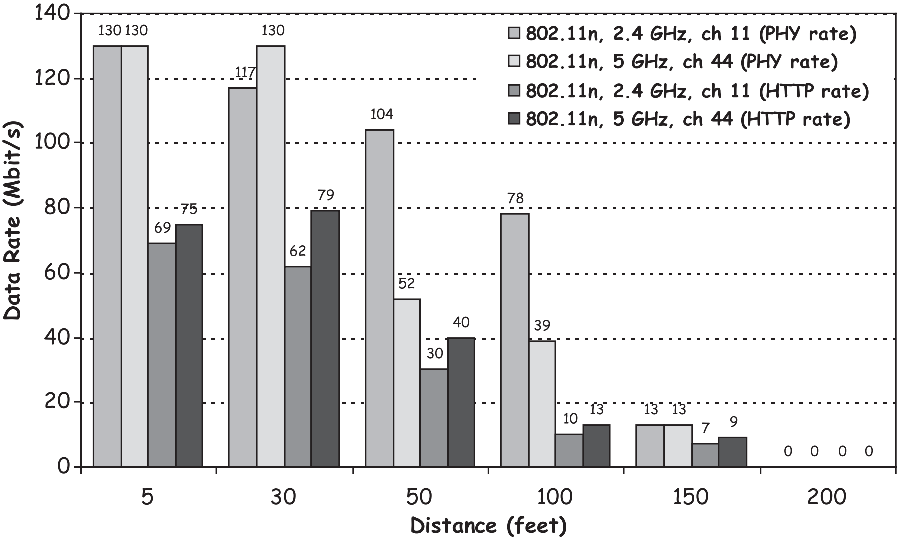


Figure 1.42 PHY and HTTP rates (MIMO).

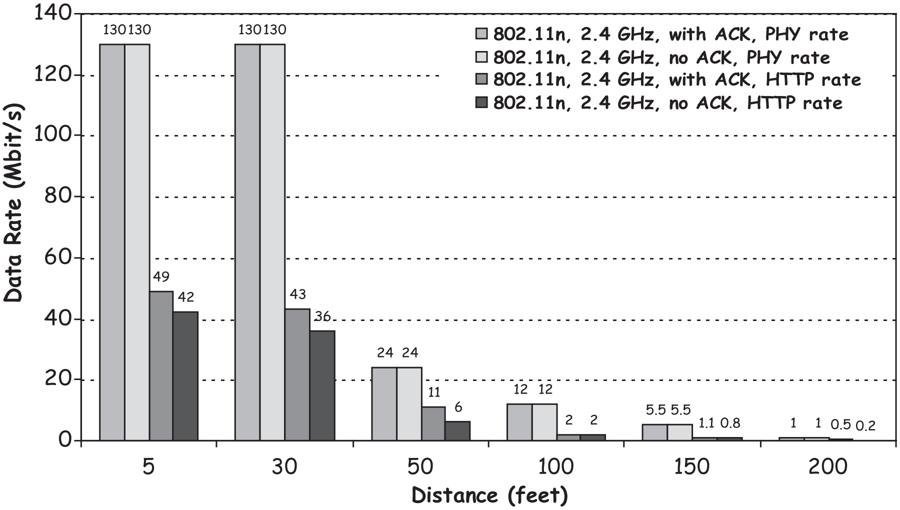


Figure 1.43 Impact of per-packet ACK on HTTP throughput.

This in turn reduces the transmit congestion window size and hence the available throughput.

1.10.3 Impact of Number of Antennas

The number of spatial streams supported by a device normally dictates the minimum number of antennas. For example, a three-stream device requires three or more antennas. The number of transmit antennas typically does not exceed the number of spatial streams. However, if there are more antennas than the number of spatial streams, the best antennas (e.g., antennas placed furthest apart) can be chosen among the available set of antennas to enhance the performance of transmission and reception. Suppose an 802.11n AP employs up to three spatially multiplexed streams, but the user device can only demultiplex up to two spatial streams. In this case, the maximum data rate of the wireless system will be limited to two streams. However, because there are three available antennas at the AP, it may select the two best antennas to transmit to the user device. Similarly, the AP may choose the two best antennas to receive transmissions from the device. Thus, the system performance can be better than an 802.11n AP using only two antennas.

1.10.4 Impact of Centralized Control

In 802.11 networks, data packets are transmitted asynchronously at the highest available data rate by the AP or user device whenever an idle wireless medium is detected. The AP provides a wireless interface for the user device to connect to the wired network but does not control the transmission from the device (i.e., APs and user devices are peers). In longer-range wireless access networks, including cellular networks, user handset transmission is tightly controlled by the BS. These networks also incur a greater signal propagation delay than 802.11 wireless LANs due to the longer operating range. Since TCP requires ACKs, both UL and DL transmissions are needed. Consider a file download application. The rate of TCP packets received by a user handset on the DL is inversely proportional to the request-grant cycle (RGC) duration (i.e., the turnaround time to receive a bandwidth grant from the BS) plus the TCP ACK transmit time. Suppose the RGC duration is between 3 and 5 ms, and each 40-byte UL TCP ACK takes a further 1.5 ms to send. Thus, the maximum packet rate in packets per second (pps) for the UL ranges from 154 to 222 pps, or equivalently, 49 to 71 Kbit/s. Assuming the maximum-length (1500-byte) TCP data packets on the DL, the same packet rate gives a maximum DL rate of between 1.848 and 2.664 Mbit/s. Thus, even if the UL rate is low, the DL can still achieve high TCP rates. The highest bandwidth utilization on a 1 Mbit/s UL and 10 Mbit/s DL becomes 7.1% and 26.64%, respectively. This implies that up to about five simultaneous TCP connections (running at the peak rate) can be supported on the DL. However, if the UL rate slows down to between 49 and 71 Kbit/s, then only 1 maximum rate DL TCP connection can be supported. Note that since UDP does not require ACKs, it requires only a unidirectional connection, and hence, the maximum data rate for the DL can be higher. For example, assuming the same RGC duration (3–5 ms) and a file download application, the highest UDP rate for the DL ranges from 2.4 to 4 Mbit/s, which corresponds to 200–333 pps.

1.11 OUTDOOR DEPLOYMENT CONSIDERATIONS

The very high frequency (VHF) band ranges from 30 to 300 MHz, and the ultra-high frequency (UHF) band ranges from 300 MHz to 3 GHz. These bands are usually too high a frequency for ionospheric refraction or reflection of the radio signal back

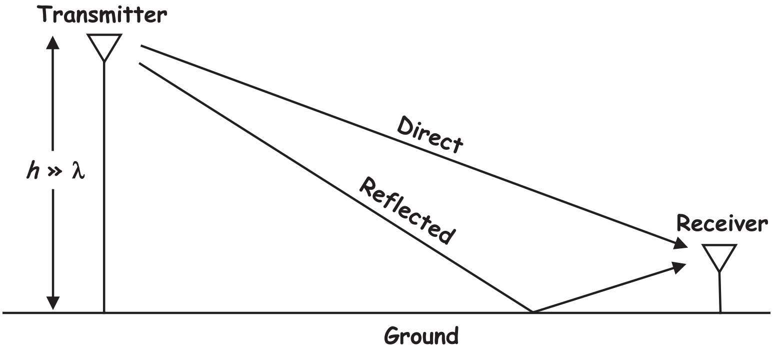


Figure 1.44 LOS signal propagation in VHF/UHF bands.

to earth. Therefore, LOS signal propagation is more appropriate for these bands. In this case, the transmitting antenna is elevated several wavelengths or more above the ground level, and the received signal is a summation of the direct and ground-reflected components (Figure 1.44). The available bandwidth is good for high-quality FM radio and TV channels, but the range of roughly 100 km restricts its use to local coverage. Propagation problems in these bands include reflection from ground and buildings, tropospheric refraction and scattering, diffraction over hilltops and buildings, and multipath scattering caused by buildings and trees. In urban settings, more physical obstacles are usually encountered, leading to signal reflections, whereas in rural areas, a larger cell coverage can be expected.

The earth bulges 12 ft every 18 miles (125 ft with buildings and trees). This becomes worse for shorter wavelength (or higher frequency) operation. To compensate for the Earth's curvature, 60% of the center of the first Fresnel zone (essentially an ellipsoid corresponding to the antenna radiation pattern with a circular aperture) should be kept free of obstacles for LOS operation. Signals traveling within the first Fresnel zone will add constructively at the receiver. Obstacles normally occur below the LOS path in outdoor transmission. Professional installation of outdoor antennas is recommended for LOS operation. Mobility may cause large-scale signal fading and the Doppler effect. The following sections provide simulation studies to illustrate the impact of fixed and mobile path loss models (based on Reference 5), single and multicarrier operation, and the modulation and operating frequency on wireless network performance.

1.11.1 Fixed Access Path Loss Model

The free space model can be represented as

$$P_l = -20\log\left(\frac{\lambda}{4\pi d}\right), \tag{1.9}$$

where

d = distance between users (m)

λ = wavelength of radio signal (c/f_c).

The suburban model can be represented as

$$P_l = 20\log\left(\frac{4\pi\times 100}{\lambda}\right)+10\left(a-bh_1+\frac{c}{h_1}\right)\log\left(\frac{d}{100}\right)+6\log\left(\frac{f_c}{2\times 10^9}\right)-20\log\left(\frac{h_2}{2}\right), \tag{1.10}$$

where

h_1 = BS height

h_2 = Handset height

a, b, c = Terrain dependent parameters.

Typical terrain values for fixed access are:

Terrain A (hilly-high tree density): $a = 4.6$, $b = 0.0075$, $c = 12.6$

Terrain B (moderate): $a = 4.0$, $b = 0.0065$, $c = 17.1$

Terrain C (flat-moderate tree density): $a = 3.6$, $b = 0.005$, $c = 20.0$.

As can be seen from Figure 1.45, for the same transmit power, bandwidth, and data rate, the suburban B model incurs higher signal attenuation than the free space model, significantly reducing the operating range. Single carrier transmission is used in this case.

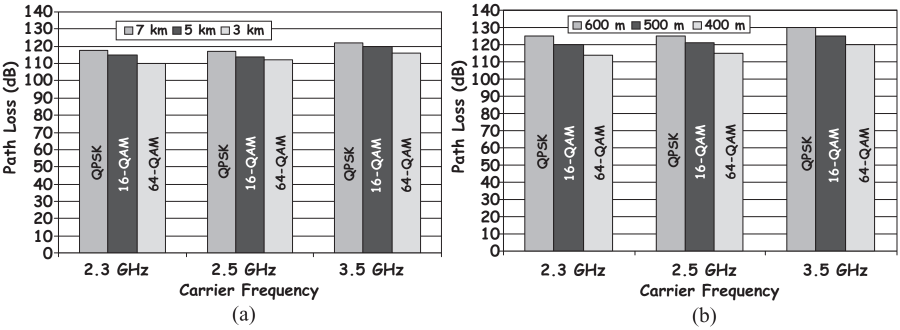
1.11.2 Mobile Access Path Loss Models

The pedestrian model can be represented as

$$P_l = 40\log\left(\frac{d}{1000}\right)+30\log\left(\frac{f_c}{10^6}\right)+49. \tag{1.11}$$

The vehicular model can be represented as

$$P_l = 40\times(1-0.04h_1)+\log\left(\frac{d}{10^3}\right)-18\log(h_1)+21\log\left(\frac{f_c}{10^6}\right)+80. \tag{1.12}$$



(a) (b)

Figure 1.45 Path loss for free space and suburban models. (a) Free space. (b) Suburban B.

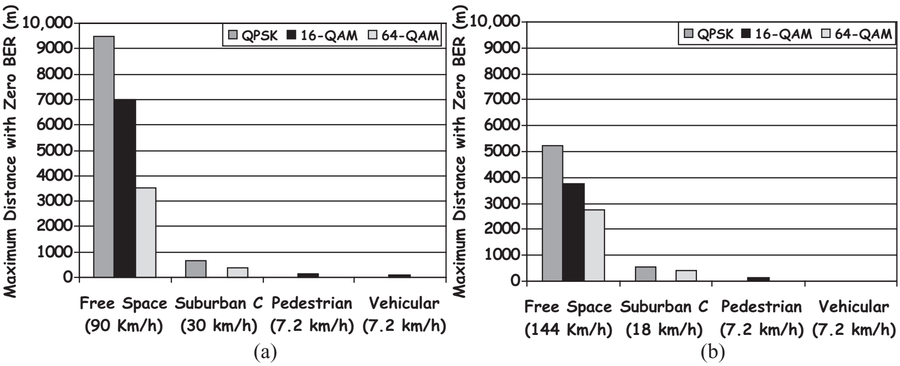


Figure 1.46 Maximum distance performance for different channel models. (a) Path loss only. (b) Path loss plus multipath model.

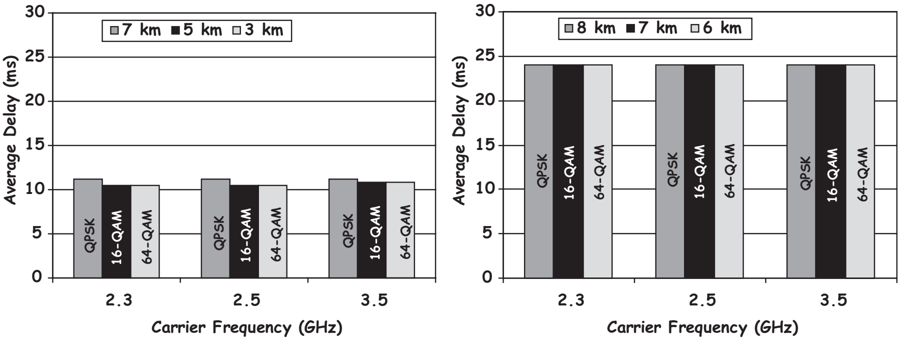
Figure 1.46 shows the performance of the mobile access models when compared with the fixed access models. The performance of these models is also evaluated in the presence of a multipath channel model. In this case, there is a fixed transmitter communicating with a stationary or moving station. The configurations are OFDM transmission, 2.5 GHz carrier frequency, 20 MHz bandwidth, and 0.5 W transmit power. The suburban C path loss model is chosen. The multipath channel model has an rms delay spread of 400 ns, giving a coherence bandwidth of 2.5 MHz. The Doppler spread at 2 m/s, and 2.5 GHz is 16.67 Hz. The peak throughput is 0.8 Mbit/s for the vehicular channel model and 1.136 Mbit/s for all other channel models. As expected, the maximum distance for error-free transmission using the vehicular model is much shorter than the rest, due to the higher signal attenuation. When multipath is taken into account, the performance of all models degrades.

1.11.3 Single Carrier and Multicarrier OFDM Comparison

The delay performance of single carrier and multicarrier transmission is shown in Figure 1.47. The free space path loss model is employed with a 20 MHz bandwidth, 0.5 W transmit power, and a peak throughput of 1.136 Mbit/s is maintained for all frequencies and distances. As can be seen, multicarrier transmission consistently incurs a longer delay (more than double) when compared with single-carrier transmission.

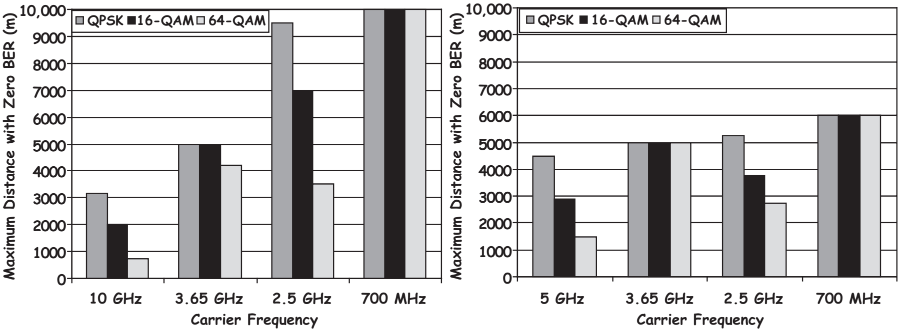
1.11.4 Impact of Modulation and Operating Frequency

Figure 1.48 shows the maximum operating range for error-free transmission improves with a more robust modulation, such as QPSK. For example, QPSK consistently performs better than 64-QAM. In this case, the operating bandwidth is 20 MHz bandwidth with a 0.5 W transmit power. The peak throughput is maintained at 1.136 Mbit/s. The free space path loss model is used. The single carrier mode performs better than OFDM mode due to the absence of a multipath model. However,



(a) (b)

Figure 1.47 Single-carrier and multicarrier delay performance. (a) Single carrier. (b) OFDM.



(a) (b)

Figure 1.48 Single-carrier and multicarrier delay performance. (a) Single carrier. (b) OFDM.

using a lower operating frequency, such as the 700 MHz band, removes the differences in the range performance regardless of modulation.

REFERENCES

[1] Apple Computer, et al., *Report and Recommendations of the Millimeter Wave Communications Working Group to the Federal Communications Commission*, 1996, http://transition.fcc.gov/oet/dockets/et94-124/etiquette.pdf.

[2] G. J. Foschini, "Layered Space-Time Architecture for Wireless Communication in a Fading Environment when using Multi-element Antennas," *Bell Labs Technical Journal*, pp. 41–59, Autumn 1996.

[3] I. E. Telatar, "Capacity of Multi-antenna Gaussian Channels," *European Transactions on Telecommunications*, Vol. 10, pp. 585–595, 1999.

[4] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-Time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction," *IEEE Transactions on Information Theory*, Vol. 44, pp. 744–765, 1998.

[5] OPNET WiMAX Specialized Model, 2007, http://www.opnet.com/WiMAX.

HOMEWORK PROBLEMS

1.1. Why are equally spaced subcarriers used in OFDM? Which consideration is more important when increasing the number of OFDM subcarriers: to increase the data rate or to mitigate a larger delay spread?

1.2. Which of these metrics are dependent on link conditions: CINR, SNR, or BER? Evaluate the pros and cons of using the SNR and signal-to-interference-noise ratio (SINR) for estimating signal quality versus using the absolute value of the signal power. Consider single and multiple antenna systems. Can interference be included as "noise" in SNR, or should it be measured separately as in SINR? Since spread spectrum systems suppress interference, is SINR relevant for such systems? How can one measure SNR or channel quality without including interference? Hint: Will it be possible to measure the channel impulse response of known and unknown interference sources?

1.3. The Rician channel matrix can be represented by the following equation, where H_d is the direct LOS component channel matrix, H_s is the scattered component channel matrix (i.e., the Rayleigh channel matrix), P_d is the power of the direct component, and P_s is the power of the scattered component.

$$H = \sqrt{\frac{K}{K+1}}H_\mathrm{d} + \sqrt{\frac{1}{K+1}}H_\mathrm{s} \quad \text{where } K = \frac{P_\mathrm{d}}{P_\mathrm{s}}.$$

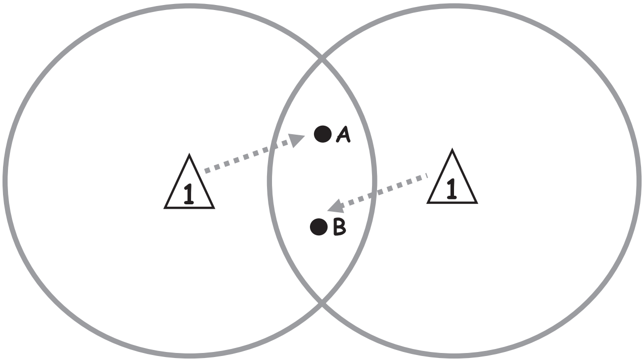
If P_d and P_s are normalized, and $P_\mathrm{d} + P_\mathrm{s} = 1$, derive a simplified equation for H based on P_s. Explain whether Rician fading (with a strong LOS path) occurs more frequently in an indoor environment than in an outdoor environment. How will such fading impact the performance of STC, spatial multiplexing, and beamforming MIMO systems? Which of these systems are more suited for non-LOS deployments?

1.4. Can carrier sensing be used to evaluate channel quality? If the transmission from one 802.11 user interferes with another 802.11 user, can the transmission be classified as an unknown interfering source since the content of the information generated by the source is random and unknown? Is carrier sensing needed for devices operating with beamforming antennas and 60 GHz frequencies? How will carrier sensing performance be affected if a station chooses channel 1 and another station uses channels 1 and 2? If two or more transmit antennas are used on the same device, must carrier sensing be performed on all antennas?

1.5. It can be shown that transmitting the shortest Ethernet packet (i.e., 64 bytes) achieves the lowest throughput because this incurs the most processing overheads compared with transmitting longer packets. Explain why this is so and whether it applies to both wired and wireless networks. The maximum Ethernet packet length for wireless networks is typically restricted to 500 bytes, whereas wired Ethernet networks are limited to 1500-byte packets.

1.6. Different equalizer types have their individual merits, but the important considerations are how much power they consume and the amount of training overheads. These considerations are related to the complexity in design. A key reason why OFDM is used in many wireless standards is to reduce the symbol rate to mitigate ISI, thereby simplifying the design of equalizers, possibly eliminating them in some instances. Recently, there has been interest in frequency domain equalizers. Single-carrier OFDM (SC-OFDM) is normally used in conjunction with frequency domain equalization. Frequency domain equalization in a SC-OFDM system is the frequency domain analog of a conventional linear time domain equalizer. For channels with severe delay spread, it is simpler than the corresponding time domain equalizer because of the FFT operations and the simple channel inversion operation. This technique can achieve better performance in bit error rate performance and PAPR compared with conventional OFDM. Explain how SC-OFDM can be used on the UL transmission to simplify the design of the user handset and shift the complexity of signal processing to the BS.

1.7. 4G wireless systems are designed to support a frequency reuse of 1 to enable the efficient use of valuable radio spectrum. However, this gives rise to more interference at the cell edge, where MSs and BSs must employ higher transmit power. Consider a point to multipoint network, with one BS managing several MSs. Transmissions between the BS and MSs are realized by means of OFDMA. Orthogonal subcarrier allocation on the UL will eradicate intracell interference between different MSs. However, since MSs A and B in the overlapped region can receive signals from two adjacent BSs, subcarriers may collide when the two BSs transmit at the same time on the DL. This is because a BS may not coordinate the allocation of subcarriers for a neighboring cell, which uses the same set of subcarriers. This reduces the success rate of handoffs and compromises the data throughput of the system. In addition, packet retransmissions will require more power from the user device and create higher CCI. Consider a new subcarrier allocation scheme to resolve collisions in the overlapped cell region. When there is a collision, the MSs will select the next subchannel with half the number of subcarriers allocated by the BS for the current subchannel. The subcarriers in the subchannel are selected in a random fashion. Explain whether this approach reduces the collision probability of the retransmission and the waiting time.



1.8. OFDM employs FEC to correct symbol errors at the receiver. The FEC is embedded in each transmitted packet, and this constitutes a recurring overhead. Consider a different method of FEC, where a dedicated FEC packet is transmitted after a group of *N* packets. Do you think this method will incur less overheads? What are the pros and cons? Can such a method be used to replace the need for data compression and reduce storage requirements? How can you modify this method to service real-time traffic involving voice and video?

1.9. In a wired LAN, only physically connected stations can hear each other. These stations are therefore authorized stations. However, this may not apply to wireless LANs. If a wireless LAN is connected to a wired LAN, how will this compromise the security of the wired LAN? Evaluate whether it is easier (and faster) to detect a denial of service (DoS) attack on a wired network than on a wireless network. Such attacks can be initiated by malicious users constantly transmitting information on the network.

	t_1	t_2	t_3	t_4
T1	s_1	s_2	s_5	s_6
T2	$-s_2^*$	s_1^*	$-s_7^*$	$-s_8^*$
T3	s_3	s_4	s_7	s_8
T4	$-s_4^*$	s_3^*	s_5^*	s_6^*

Figure A Code matrix for four transmit antennas.

	t_1	t_2	t_3	t_4
T1	s_1	s_2	0	0
T2	$-s_2^*$	s_1^*	0	0
T3	0	0	s_3	s_4
T4	0	0	$-s_4^*$	s_3^*

Figure B Code matrix for four transmit antennas.

	t_1	t_2
T1	s_1	s_2
T2	$-s_2^*$	s_1^*
T3	s_3	s_4
T4	s_5	s_6

Figure C Code matrix for four transmit antennas.

	t_1	t_2
T1	s_1	s_2
T2	$-s_2^*$	s_1^*
T3	s_3	s_4
T4	$-s_4^*$	s_3^*

Figure D Code matrix for four transmit antennas.

	t_1	t_2
T1	s_1	s_2
T2	$-s_2^*$	s_1^*
T3	s_3	s_4

Figure E Code matrix for three transmit antennas.

1.10. (a) The code matrices in Figures A and B employ four transmit antennas with four symbol periods. "0" indicates no transmission. Find the rates. (b) The code matrices in Figures C and D employ four transmit antennas with two symbol periods. Find the rates. (c) The code matrix in Figure E employs three transmit antennas with two symbol periods. Find the rate. From the code matrices of Figures A–E, list the code matrices that allow all the received symbols to be decoded independently.

1.11. Suppose a 3 × 3 MIMO system (three transmitters) achieves a similar performance as a 4 × 2 MIMO system (four transmitters), which system would you choose? What are the advantages of combining MIMO with multiband OFDM where different antennas employ different frequency channels? What antenna spacing would you choose for STC, spatial multiplexing, and beamforming MIMO: 1/4 or 1/2 wavelength spacing?

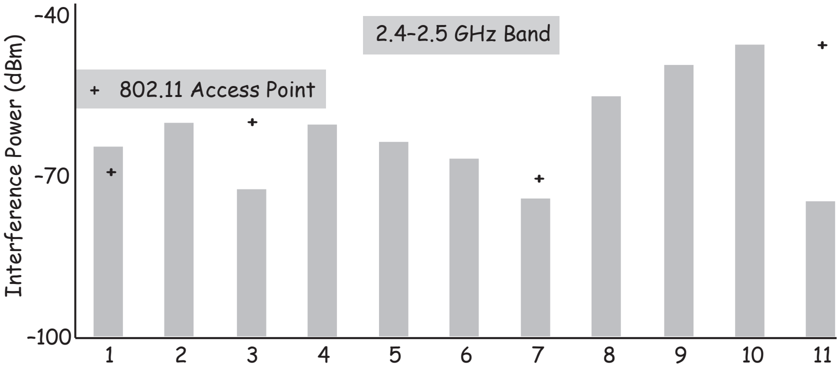
1.12. Consider an Alamouti STC system with two transmit and two receive antennas. The second receive antenna is used for additional diversity. Express the receive symbols (r_1, r_2) as a function of the transmitted symbols (s_1, s_2). Repeat for the 2 × 2 spatial multiplexing and 2 × 2 beamforming MIMO systems. Evaluate the number of channel estimates for 2 × 2 spatial multiplexing, STC, and beamforming. If $h_1 = h_2$, derive the receive symbols for the three MIMO types. What is the rank of the channel matrix? Note that this corresponds to the case of correlated channels. Repeat for the case when the channels are reciprocal. Is antenna beamforming useful for 2 × 2 and 2 × 1 Alamouti STC? Will MU-MIMO perform better using spatial multiplexing, STC, or beamforming? How can Alamouti STC be employed in UL and DL MU-MIMO?

1.13. More robust modulations (e.g., QPSK, 16-QAM) tend to achieve the full spatial multiplexing gain almost all the time. However, 64-QAM tends to achieve less. Is this related to the operating range, with longer-range lower-order modulations providing a richer multipath scattering environment? Or is this due to receive antennas experiencing different SNRs? Are beamforming systems affected by this problem? Should unequal transmit power be allocated to the spatial streams? Strictly speaking, in MIMO, the total transmit power across all antennas should be the same as an equivalent SISO transmission.

1.14. What are the advantages and disadvantages for deploying a 2.4 GHz network with four surrounding cells instead of six in first ring? Assume all cells have the same area of coverage.

1.15. Identify the wireless cells in Figure 1.34 that will experience cochannel or adjacent interference in the 2.4 GHz band. Assume that the transmit and receive ranges are confined within each circle (i.e., transmitter power and receiver sensitivity are about the same). How will the interference pattern change if the transmit range is two times shorter than the receive range? In practice, the sensing range is longer than the transmit range because the receiver need not decode a signal when sensing the presence of interfering signals. What are the drawbacks for having a more sensitive sensing range? Note that as users roam between cells served by different APs, they will not be affected by the sensitivity of the carrier sensing mechanism since the relative signal strengths are evaluated, rather than the absolute signal values. Smaller cells may lead to less contention within each cell. Evaluate the advantages and disadvantages of using small cells with packet forwarding in each cell to improve the coverage area and network capacity.

1.16. In the figure below, analyze whether cochannel and/or adjacent channel interference are present in each 20 MHz channel of the 2.4 GHz band. For 802.11 APs operating on channels 3, 7, and 11, the operating transmit power is above the interference power. However, the AP on channel 1 is operating below the interference power. Rearrange the APs (without changing the transmit power) such that all APs operate at an optimum level above the interference power. The operating channels remain the same: 1, 3, 7, and 11.



1.17. Show that 25 MHz is the minimum bandwidth to separate two 20 MHz channels in the 2.4 GHz band. For this bandwidth separation, show that carrier 6 is closest to carrier 1. Repeat for the case when 30 MHz is the minimum separation and confirm that carriers 1 and 7 are the closest. Determine the minimum separation to separate two 40 MHz channels and identify the appropriate carriers.

1.18. Are MIMO systems interference or noise-limited systems? How would you quantify the performance of MIMO systems using E_b/N_0? How will MIMO perform in an environment with a strong LOS signal, such as in an outdoor environment? Joint processing at the receiver for multiple spatial streams together with unknown interfering sources may have a detrimental impact. Explain whether channel sounding procedures can model the interference.

1.19. By implementing a MAC protocol to control channel access, only one user is transmitting at any one time in a wireless cell. Therefore, multiuser interference is reduced, although it may not be completely eliminated because of possible CCI. In this case, if there is a rich scattering environment with no LOS, STC MIMO performs well. Explain whether STC MIMO will perform well when a LOS path and rich scattering are both present.

1.20. Given an interference-prone environment in the 2.4 GHz band, which MIMO method will you use with OFDM: STC or spatial multiplexing? Will either method achieve higher data rates and/or better range performance than an OFDM system with no MIMO capability? Both OFDM and spread spectrum systems are able to mitigate the impact of fading. OFDM achieves this by using multiple subcarriers. Spread spectrum systems employ a wider signal bandwidth so that fading affects only a small portion of the bandwidth. In addition, spread spectrum systems are able to spread the power of unknown interference sources over a wide bandwidth, making them appear as low-power background noise. Like OFDM, the effective data symbol rate is reduced due to the need for bandwidth spreading. Explain whether OFDM is able to mitigate unknown interference sources, such as spurious emissions from microwave ovens in the 2.4 GHz band. Can MIMO systems help reduce the impact of these interference sources? What can you conclude about the effectiveness of MIMO-OFDM in mitigating frequency-selective fading, multiuser interference (including CCI), and unknown interference sources?

1.21. For wireless access networks covering several miles, are licensed bands preferred over unlicensed bands? For wireless communications in a large convention hall or an airport terminal with many active devices, would you expect a licensed spectrum, low bandwidth channel (e.g., 5 MHz channel in the 2.3 GHz band) to provide a higher bit rate than an unlicensed spectrum, higher bandwidth channel (e.g., 20 MHz channel in the 2.4 GHz band)? Evaluate whether carrier sensing and MIMO (requiring channel estimates) are useful for operation in these bands. How will the performance of open and closed loop MIMO systems differ for licensed and unlicensed bands?

1.22. When spatially multiplexing two data streams over two transmit antennas, would you recommend the use of one common channel accessible by both antennas (e.g., a single 40 MHz channel) or two independent channels (e.g., two 20 MHz channels), each channel assigned to an antenna? If one of the two channels experiences low SNR, whereas the other channel maintains high SNR, how will this impact the spatial multiplexing gains? Will it be better off to transmit a single stream on a single channel using one antenna? Suppose, instead of low SNR, one of the channels is rank deficient. Will single antenna transmission be more desirable in this case? Note that these antennas are collocated on the same device. Explain whether the flashlight effect will occur in a spatially multiplexed MIMO system.

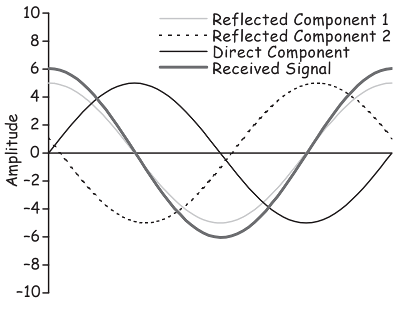
1.23. Which system performs better: using OFDM to service two different users simultaneously with two 20 MHz channels or using OFDMA to service these two users asynchronously using a 40 MHz channel? Can OFDMA achieve better performance than MU-MIMO? Justify your reasoning in terms of the number of simultaneous users supported, multipath mitigation, and the achievable data rate or throughput. Hint: OFDMA may allow all channels of say the 2.4 or 5 GHz bands to be used simultaneously.

1.24. Beamforming enhances the bit rate performance, energy efficiency, and security of the spatial streams, and can help locate and identify user devices. Beamforming also allows the rate and quality of each stream to be adapted to link conditions even in multicast operations. Can beamforming be used to reduce multiuser interference in MU-MIMO with concurrent transmissions arising from the BS and directed to different users? Will beamforming be useful for the UL transmission? If beamforming is used on the DL but not on the UL, how will this impact the overall performance on the UL? If beamforming is used on the UL, is a MAC protocol still needed? Note that if beamforming is possible on the UL, acknowledgment or channel feedback sent by the handsets can be used to improve link performance. This is in contrast to wireless systems that employ broadcast antennas to send information to multiple devices simultaneously. In this case, acknowledgments or channel estimates that are concurrently sent by two or more handsets will result in collisions.

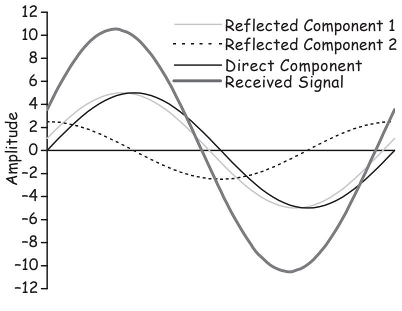
1.25. Compare the use of SIC in MU-MIMO and CDMA systems. Are CDMA systems likely to support more concurrent transmissions than MU-MIMO systems? Since there is no bandwidth spreading in MU-MIMO, will this imply that it may be more prone to interference than CDMA systems?

1.26. 60 GHz systems experience high signal attenuation but virtually no multipath or CCI. These systems employ channel bandwidths that are more than a 100 times wider than 2.4 and 5 GHz systems (2.16 GHz vs. 20 MHz). However, a wider bandwidth may potentially imply more interference sources, which may not be modeled accurately. Would you recommend a single antenna system with rate fallback or a two-antenna system with no rate fallback? If the two-antenna system is chosen, identify the specific type of MIMO system (i.e., STC, spatial multiplexing, or beamforming). Would you consider spatial multiplexing and beamforming to be equivalent at 60 GHz? Repeat for the case when the 700 MHz band is employed.

1.27. Multipath propagation is not always a bad phenomenon since it reduces the probability of complete phase cancellation at the receiver. This diversity property is sometimes exploited by radio receivers to improve reception. However, it can also result in higher amplitude variation in the received signal. These effects are shown in the figure below. For simplicity, it is assumed that the received signal comprises the sum of the direct LOS signal and the reflected signals (each with a random amplitude and phase). Whether the sum of the signal components cancel or reinforce each other strongly depends on the difference in their phase angles. If this phase difference is near 180°, then the net result is a deep fade in the received signal. As can be seen, these fades are separated by about half a wavelength. At 2.4 GHz, this is approximately 0.06 m. If a user moves at 10 km/h (pedestrian speed) in this frequency band, fades can be expected every 22 ms. Clearly, the rate of fading is proportional to the velocity of the user motion. Compare the multipath phenomenon with the use of subcarriers in OFDM. What can you say about the use of more subcarriers per channel? How will this impact the use of multiple antennas? Note that in all cases, the frequency of the received signal remains unaffected by the phase or amplitude distortions of the reflected signals.



(a) Total phase cancellation avoided



(b) Constructive interference enhances signal power

1.28. For a home networking environment where network connectivity is required between a wireless router and multiple consumer devices, such as HDTV, game console, and smartphone, would you recommend the use of STC, spatial multiplexing, or beamforming MIMO antennas? Justify your answer. What are the benefits of employing MU-MIMO in such an environment?

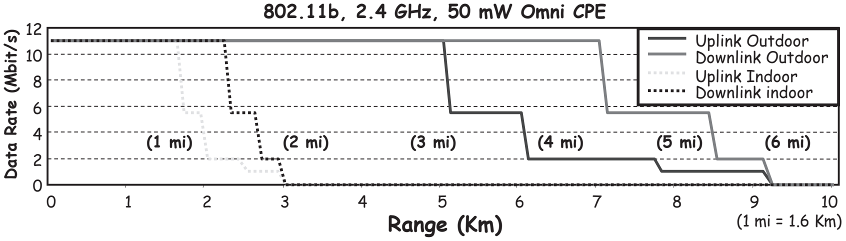
1.29. First-generation wireless LAN products employ the 900 MHz ISM unlicensed band, which ranges from 902 to 928 MHz. The allowed power output is up to 1 W. FCC section 15.249 allows 50 mV/m of electrical field strength at 3 m, which corresponds to an EIRP of −1.23 dBm. The harmonics limit is one hundredth of the fundamental level, 500 μV/m, corresponding to an EIRP of –41.23 dBm. Evaluate the pros and cons of using this band compared with the 2.4 and 5 GHz bands.

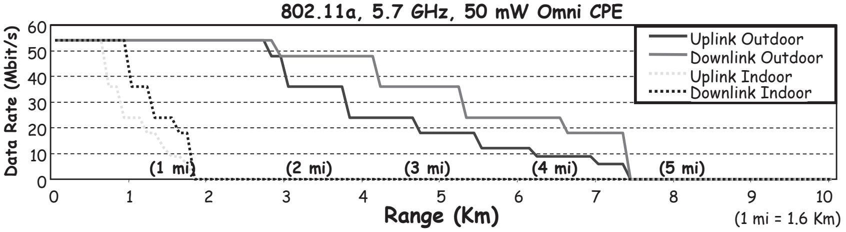
1.30. Consider a MIMO system where the transmit and receive antennas are chosen based on the receive power level. How will such a system perform compare with a system that chooses the antennas based on the channel response? Consider another system whether the number of transmitting antennas is always less than the total number of available antennas at the transmitter. This minimizes the number of receiver chains to be implemented. The transmitting antennas are selected such that the SNR is maximized at the receiver. How will such a system perform compared with a system where all available antennas at the transmitter are employed, regardless of SNR? Explain whether switched antenna diversity at the receiver will perform better than the case when a single antenna is transmitting at the source and both receive antennas at the destination are turned on and the individual signals from both antennas are weighted (in a manner that is proportional to the signal strength or amplitude) and then added. This latter process, which amplifies the stronger signal and attenuates the weaker signal, is also known as maximum ratio combining. It is an example of a single input multiple output (SIMO) system.

1.31. The table below shows the product specifications of a legacy 802.11 wireless LANs. Clearly, a longer operating range leads to a higher delay spread, which reduces the data rate. What is the connection between the bit error rate and the packet error rate? Can a higher receiver sensitivity reduce the transmit power?

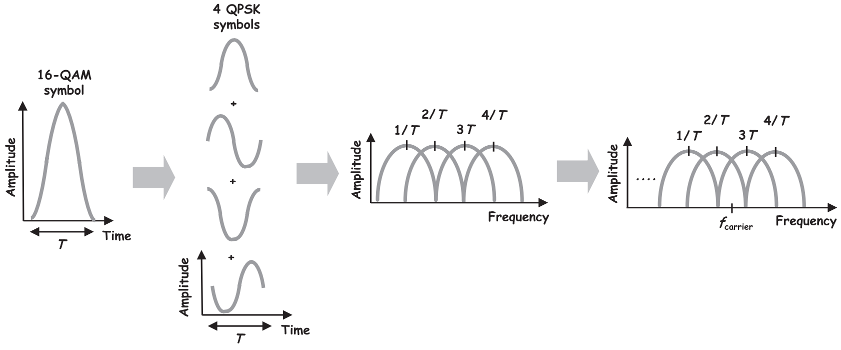
Bit error rate	$<10^{-8}$
Range in an open environment (100 bytes of user data)	1200 ft (2 Mbit/s)
	1400 ft (1 Mbit/s)
Receiver sensitivity	–90 dBm (2 Mbit/s)
	–93 dBm (1 Mbit/s)
Delay spread for a packet error rate of <1%	400 ns (2 Mbit/s)
	500 ns (1 Mbit/s)
Transmit power	15 dBm (32 mW)
Power consumption	Doze mode (9 mA)
	Receive mode (230 mA)
	Transmit mode (330 mA)

1.32. The figure below shows the range performance of a phased antenna array 802.11 switch operating in the 2.4 and 5 GHz bands. The customer premise equipment (CPE) uses a regular omnidirectional antenna and communicates with a central switch that can be located several miles away. Such a switch is useful for reducing the number of APs in office buildings and campus deployments. The antennas of the switch focus multiple narrow beams on the user devices. Since the same transmit power is used for the UL and DL, explain the differences in the operating range for the indoor and outdoor environments. Is carrier sensing useful for these systems? Why is the DL performance better than the UL in some cases but achieves the same performance at the maximum and shorter ranges? How would the performance of such a system differ from the case when the CPE uses beamforming arrays?





1.33. In the following multicarrier transmission method, a symbol (16-QAM) is split into four subsymbols (QPSK). The symbol interval remains constant. Comment on the bandwidth requirements of this scheme. Suppose four 16-QAM symbols are transmitted instead, using a symbol interval that is four times longer than the original symbol period (i.e., symbol rate is reduced four times). Comment on the bandwidth requirements of this scheme. What can you conclude on the tradeoff(s) between the number of subcarriers (representing lower speed symbols) and the spacing between the subcarriers?



1.34. Beamforming antenna arrays may be combined with spatial multiplexing to achieve better range while maintaining high capacity. Explain whether STC, including Alamouti STC, can be implemented using beamforming antennas, and describe the potential benefits and drawbacks. Note that in this case, the number of receive antennas can be less than the number of transmit antennas. However, independent fading on the two paths may not be sufficient.

1.35. HomePlug power line communications technology employs 84 OFDM subcarriers ranging from 4 to 21 MHz. It enhances the OFDM basic functionality by dynamically comparing each subcarrier with the current characteristics of the power line. Specific subcarriers that experience high attenuation or line noise are dropped. Since noise spikes can be random and can occur at any time, FEC is required to reconstruct any bits that may be corrupted. Explain whether a similar OFDM method can be adopted for wireless networks.

1.36. The public Internet is a lossy network where any router along the path can drop packets if it becomes congested. Depending on the distance of the path, losses of 2–3% are typical, although over 10% can also be experienced. Consider the following FEC scheme that adds an error recovery packet to *N* packets that are sent, minimizing retransmissions and reducing the associated delay. The FEC packet contains information that can be used to reconstruct the missing packets. The method can made adaptive to minimize overheads. For example, if no losses are detected for a connection, FEC is disabled. Explain whether such a method will perform well in an error-prone wireless link.

1.37. A lot of Web content is text-based and therefore highly compressible. This will reduce the number of packets that need to be transported from the application to the user device, which will reduce the response time. All popular Web browsers support compression. Through the use of the accept encoding request in the HTTP header, the client can indicate it can receive compressed content. The compression algorithm (e.g., gzip and deflate) can be specified in the same header. Compare the effectiveness of such a method with the packet-level FEC method (previous problem) when applied to an error-prone wireless link.

1.38. OFDMA is useful for resolving contention on the UL channel. Since the DL is a broadcast channel from a single source (i.e., the BS), discuss the merits of using OFDM versus OFDMA on the DL.

1.39. Beamforming antenna arrays may extend the coverage of wireless systems. However, a broadcast synchronization channel (also known as beacon or pilot channel) is normally needed to ensure all handsets in different locations are synchronized to the BS or AP. The synchronization channel requires omnidirectional transmission so that new and existing users can easily connect with the BS or AP. However, if the user moves out of synchronization range, the connection is lost even if the data channel (using beamforming) is within range of the BS or AP. Explain how this problem can be solved to improve the coverage performance of a wireless system using beamforming antennas. Suppose the 60 GHz band is used for data transmission. Discuss the merits of using a 2.4 or 5 GHz channel for synchronization.

1.40. Energy efficiency determines the sustainability of wireless communications. Design a wireless network that will consume the least power to achieve reasonable data rate and operating range. Justify your solution in terms of the antenna method used, the frequency band, the possibility of imperfect estimation of time-varying channel statistics, and the modulation and error control methods. Consider the tradeoffs in using smaller cells, which require more cells to serve the same number of users in the same coverage area, hence requiring more packet transmissions (possibly accompanied by more handoffs) compared with an equivalent single cell operation. Although deployment cost may increase with more cells, lower power BSs, APs, and handsets can be used.

1.41. Layer-free networks enable flexibility in accessing and responding to a variety of measured wireless channel and transmission statistics. Explain whether such an approach will lead to incompatibility issues in wireless communications (one of the major objectives of the Open Systems Interconnect (OSI) layered model is to standardize the structure of network layers so that transmitting and receiving devices from different vendors can seamlessly communicate). Justify your reasoning from the perspective of licensed and unlicensed frequency operation. Explain whether the CSMA protocol can be considered as a layer-free protocol since it uses interference information to transmit packets. Would you classify pure broadcast networks (e.g., satellite, GPS, and FM radio) as layer-free networks too?

1.42. Would you consider it odd that many CDMA cellular networks are deployed using licensed bands and yet interference cancellation is required in the CDMA receivers? Explain whether interference cancellation will work well in unlicensed frequency bands where unknown interference sources may be present. Is CSMA useful when used in conjunction with CDMA in unlicensed band operation?

1.43. Explain whether interference avoidance multiple access methods (e.g., CSMA) perform better or worse than interference driven methods (e.g., CDMA). Justify your answer in terms of the level of channel reliability (e.g., poor or good BER) and the level of packet losses (e.g., buffer overflow caused by excessive retransmissions).

1.44. Cellular networks are better at supporting mobile users but suffer from low efficiency due to the need to deal with unpredictable changes in the SNR resulting from fast movements. These SNR variations make it difficult to select the appropriate MCS. Describe the flaw(s) in this argument.

1.45. In a multipath environment, explain whether STC MIMO is spectrally more efficient than a spread spectrum system employing a single spreading code. Justify your answer for the single user case. Suppose a CDMA wireless handset uses different spreading codes to support concurrent transmissions from multiple antennas. In this case, how will the spectral efficiency compare with an equivalent STC system using the same number of transmit antennas as the number of spreading codes? Note that although both STC MIMO and spread spectrum systems generally perform well in a multipath environment, there is an upper limit to the interference they can tolerate.

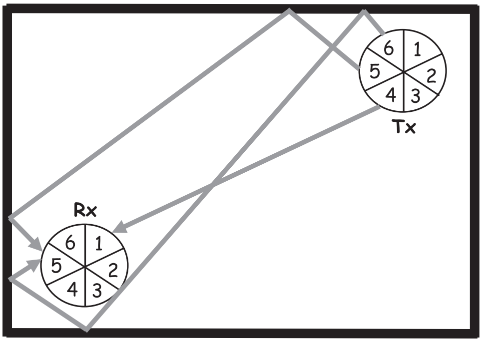
1.46. Explain whether STC and spatial multiplexing can be implemented in the UL of a MU-MIMO system where two handsets are transmitting simultaneously, each equipped with a single antenna. Thus, each transmitter resides on a different device, whereas the receiver resides on the same device, which is a prerequisite for MIMO communications. The advantage of such a system is the lower transmit complexity than a single device transmitting using two antennas. However, reference pilot signals from different transmitters have to be orthogonal (unique) in order to recover the channel coefficients, but precoding is not needed for the pilot signals. Will such a system lead to lower transmit power and more efficient energy usage for each transmitting handset?

1.47. The effects of multipath change with distance and operating frequency. Optical infrared systems employ baseband or unmodulated wireless transmission. Explain whether these systems will suffer from multipath fading.

1.48. In general, the effect of delay spread is to cause the smearing of individual symbols in the case where the symbol rate is sufficiently low, or to further cause time-dispersive fading and ISI if the symbol rate is high. ISI is a form of self-interference that increases the error rate in digital transmission, an impairment that cannot be overcome simply by improving the SNR. This is because increasing the signal power in turn increases the self-interference. For a data rate of 2 Mbit/s with 2 bits/symbol (i.e., using QPSK), the symbol rate is 1 Msymbol/s. Compute the rms delay spread that will cause adjacent symbols to overlap by 0.1 symbol. Note that at higher symbol rates or larger delay spreads, the difference in delay among the various signal reflections arriving at the receiver can be a significant fraction of the symbol interval. Normally, a delay spread of more than half a symbol interval results in indistinguishable symbols and a sharp rise in the error rate.

1.49. TCP is designed to operate over a variety of networks using an “hourglass” network architecture, serving different PHY and MAC layers (e.g., low- and high-speed networks, wireless and wireline networks) over a common IP framework. TCP file downloads perform better over links with a low bandwidth-delay product. Thus, a higher download rate demands a lower delay. TCP is also sensitive to packet loss. Hence loss rates may be restricted to 10^{-5} to 10^{-7}. This is in contrast to VoIP flows that can tolerate delays of about 100 ms and packet losses of up to 1%. Explain whether the poor performance of TCP over wireless networks is due to the lower rates and/or the higher error rates. One of the remedies to recover from lost data segments quickly is to incorporate fast retransmit and fast recovery into the TCP protocol. The random early discard is a method that forces faster retransmits rather than timeouts. It selectively penalizes high throughput TCP sessions. Explain whether this method will work well with wireless networks. Repeat for the case when the sender is made aware of the existence of wireless links so that it knows some packet losses are not due to congestion. The sender can then avoid invoking congestion control on noncongestion-related losses. The buffer (queue) length is sometimes used as a metric to indicate network congestion. Explain whether this metric applies to a network that comprises an interconnection of wireline and wireless networks.

1.50. An early wireless LAN product employs sectorized horn antennas operating in the 18–19 GHz band (see figure below). The overall transmission coverage is split into six sectors, each covering a 60° segment. A similar method is widely adopted in cellular systems where each sector of the BS is used to service a number of subscribers. How does omnidirectional MU-MIMO transmission and beamforming compare with this approach? Explain your reasoning in terms of the number of simultaneous stations supported, the effectiveness of multipath mitigation, the need for intracell handoff, and the achievable data rate.



Antenna Selection Table

Rx \ Tx	1	2	3	4	5	6
1				Best		
2						
3						
4						
5					Better	Good
6						

1.51. Many 802.11 APs allow the user to select a channel based on current interference conditions. For example, the user may evaluate active channels employed by several nearby APs and choose the channel that is least likely to encounter interference. Other routers go one step further and allow the user to fix the data rate (i.e., the MCS). Discuss the consequences of fixing the channel and data rate in an 802.11 wireless LAN. Note that fixing the MCS is effectively cutting off SNR feedback from the receiving device, much like UDP transmission. Since 802.11 employs asynchronous transmission at the maximum available rate with no centralized control, will TCP rates become similar to UDP rates?

1.52. In general, a more efficient code rate is normally chosen by a higher-order modulation. Since a higher-order modulation requires a higher SNR, will this imply that a more efficient code rate (e.g., 5/6) is more effective in correcting errors than a less efficient code rate (e.g., 1/2)?

1.53. A European hotel operates a 2.4 GHz 802.11 network using channels 1, 6, and 11 to serve international guests. It is suffering from high packet loss (5–35%), high packet latency (5–6 seconds), and spotty coverage. Evaluate the following solutions to improve the performance and coverage of the system:

- Switching some of the APs from 2.4 to 5 GHz band
- Employing four channels (i.e., 1, 5, 9, and 13) in the 2.4 GHz band
- Placing a smartphone near an AP with good signal strength and then relaying the 802.11 network connection over Bluetooth (also operating at 2.4 GHz) to the laptop
- Decreasing the AP receiver sensitivity (by reducing the coverage area from large to small)
- Increasing the minimum data rate from 1 to 2 Mbit/s
- Decreasing the transmit power from 20 dBm (100 mW) to 10 dBm (10 mW).

1.54. A big advantage of wireless networks operating at 5 GHz or lower is the ability to operate indoors as well as outdoors. Some measurements were conducted using the same hardware configurations to compare the performance of 802.11 in indoor and outdoor deployments. A wireless data rate of 60 Mbit/s at 300 m was obtained for an outdoor network, whereas a lower rate of 34 Mbit/s at a shorter distance of 93 m was obtained for an indoor network. Explain the substantial difference in the rate and range for the outdoor and indoor networks.

1.55. Etiquette transmission allows incompatible systems to co-exist. For example, the 802.11, Bluetooth, and Zigbee wireless standards may all co-exist in the 2.4 GHz band by sensing the wireless medium before data transmission. Are standards in etiquette transmission required to allow these systems to interoperate?

1.56. Having a reasonable span for a network is crucial for the success of any aspiring network technology. Some lessons can be learned from the approach taken by the wired Gigabit Ethernet Task Group more than 10 years ago. When the group realized that increasing the Ethernet clock rate by tenfold (which reduces the time needed to transmit an Ethernet frame by a factor of 10) directly shrinks the network span from 200 m (100Base-T) to 20 m, the IEEE 802.3z committee developed a mechanism to preserve the 200 m (660 ft) span. Evaluate the options for increasing the range performance of 60 GHz wireless systems without compromising the high data rate that is available in these systems.

1.57. Link adaptation using a set of MCSs may lead to fluctuations in wireless data rates and performance. Describe an antenna technology that can be used to maintain a fixed MCS even as the operating range or fading conditions change.

1.58. In a MU-MIMO network, is it possible for different users to transmit on the same time–frequency resource? If this is possible, spectrum efficiency is enhanced with more degrees of freedom than the case when a single user transmits using multiple spatial streams. In a cellular network, large number of users may potentially send data at the same time. How can the BS partition the users into different groups to reduce processing requirements, and how can spatial interference be reduced with precoding?

1.59. Cellular systems employ ranging operations to estimate the distance between the user handset and BS. To do this, the time of arrival or the power of the signal can be measured. Which metric is more accurate? Since more powerful error control is typically used for longer distance transmission, how will this impact ranging accuracy?

1.60. In general, for a 2 × 2 antenna system, 4 channel estimates are required for spatial multiplexing—2 for Alamouti STC and 1 for each beamforming link. Explain how these channel estimates affect the performance of the system. What is the channel rank for the Alamouti STC? ■