Chapter 1

Introduction to Radio Systems

Because radio systems have fundamental characteristics that distinguish them from their wired equivalents, this chapter provides an introduction to the various radio technologies relevant to the IP design engineer. The concepts discussed provide a foundation for further comparisons of the competing mobile radio access systems for supporting mobile broadband services and expanding into IP design for mobile networks.

Many excellent texts concentrate on the detail of mobile radio propagation, and so this chapter will not attempt to cover radio frequency propagation in detail; rather, it is intended to provide a basic understanding of the various radio technologies and concepts used in realizing mobile radio systems. In particular, this chapter provides an insight into how the characteristics of the radio network impact the performance of IP applications running over the top of mobile networks—characteristics that differentiate wireless networks from their fixed network equivalents.

In an ideal world, radio systems would be able to provide ubiquitous coverage with seamless mobility across different access systems; high-capacity, always-on systems with low latency and jitter; and wireless connectivity to IP hosts with extended battery life, thus ensuring that all IP applications could run seamlessly over the top of any access network. Unfortunately, real-world constraints mean that the radio engineer faces tough compromises when designing a network: balancing coverage against capacity, latency against throughput, and performance against battery life. This chapter introduces the basics of mobile radio design and addresses how these tradeoffs ultimately impact the performance of IP applications delivered over mobile radio networks. In Chapter 2, “Cellular Access Systems,” and Chapter 3, “All-IP Access Systems,” we use these concepts to differentiate the competing mobile access systems defined for supporting mobile broadband service offerings.

Spectrum

Radio Frequency Spectrum is a key distinguishing factor used to compare alternative mobile radio systems. Radio spectrum for communications ranges from approximately 30 Hz (termed Extremely Low Frequency [ELF]) to above 100 GHz (termed Extremely High Frequency [EHF]). Because of its capability to provide very wide area coverage and penetrate sea water, ELF has been used for global systems for providing low-rate submarine communications. EHF, on the other hand, can be used for Line-of-Sight (LoS) microwave communications. Table 1-1 shows the complete range of radio frequency spectrum used in communication systems and provides some examples of spectrum use.

Table 1-1 *Radio Frequency Spectrum*

Band Name	Frequency Range	Example Communication Use
Extremely Low Frequency	3–30 Hz	Submarine communications
Super Low Frequency	30–300 Hz	Submarine communications
Ultra Low Frequency	300–3,000 Hz	Underground communications
Very Low Frequency	3–30 kHz	Navigation
Low Frequency	30–300 kHz	AM broadcasting
Medium Frequency	300–3,000 kHz	AM broadcasting
High Frequency	3–30 MHz	Shortwave broadcast; amateur radio
Very High Frequency	30–300 MHz	Private mobile radio; FM and television broadcasting
Ultra High Frequency	300–3,000 MHz	Television broadcasting, cellular radio, and wireless LANs
Super High Frequency	3–30 GHz	Wireless LANs; point-to-point and point-to-multipoint microwave
Extremely High Frequency	30–300 GHz	Point-to-point microwave

Table 1-1 highlights how the characteristics of the different bands of the radio spectrum vary. In general, the lower the frequency, the better the range (for example, in the extreme case, a single ELF transmitter is able to cover the entire planet), but the bandwidths available are limited (for example, the same ELF systems typically provided a global system with total system capacity below 50 bps). Conversely, EHF systems can provide incredible capacity, but they incur significant attenuation by atmospheric effects due, for example, to extreme humidity, rain, or molecular absorption, and thus are prone to significant losses in non-Line-of-Sight (LoS) deployments.

In between these extremes is the "sweet spot" for the radio spectrum for conventional mobile systems, with the Ultra High Frequency (UHF) band ranging from 300 MHz to 3 GHz and providing what many consider to be the best compromise between usable bandwidths and propagation characteristics required for wide area coverage.

As a consequence, the UHF spectrum is a scarce resource with many competing users. In order to rationalize spectrum usage, the Radio Communications Sector of the International Telecommunications Union (ITU-R) has identified key bands that can preferably be used for International Mobile Telecommunications (IMT) operation. These IMT bands cover operation for the following:

- **IMT-2000 systems:** Covering legacy "3G" technologies, including Wide Band Code Division Multiple Access (WCDMA), cdma2000 1xrtt technologies, and most recently WiMAX.
- **IMT-enhanced systems:** Covering those systems offering improved mobile broadband services, including High-Speed Packet Access (HSPA) and EVolution-Data Only (EV-DO) technologies.
- **IMT-advanced systems:** Covering those systems offering very high-rate mobile broadband operation, including rates in excess of 1 Gbps to low-mobility users.

Note In Chapters 2 and 3, we provide more detail describing alternative mobile radio access systems; it will become apparent that none of the current competing systems, including WiMAX and Long-Term Evolution (LTE), meet the minimum requirements for IMT-advanced systems.

The spectrum for use by the IMT-2000 technologies was first identified by the ITU at the World Administrative Radio Conference (WARC) in 1992 and further augmented at the World Radiocommunication Conferences (WRC) in 2000 and 2007. Even when spectrum has been identified for use by IMT systems, it might not be available for sole use of mobile radio systems. However, the identification of spectrum by ITU-R, as illustrated in Table 1-2, provides equipment manufacturers with guidance on the range of frequency bands that are likely to be used in deploying IMT services, hopefully leading to economies of scale and consequential decrease in the overall cost of production of specialized IMT equipment.

Table 1-2 *IMT Spectrum Allocations*

Frequency Range	Regional Rules
450–470 MHz	All regions
610–790 MHz	Nine countries in Region 3 (Asia and Australasia): Bangladesh, China, Rep. of Korea, India, Japan, New Zealand, Papua New Guinea, Philippines, and Singapore

continues

Table 1-2 *IMT Spectrum Allocations (continued)*

Frequency Range	Regional Rules
698–790 MHz	Region 2 (Americas)
790–960 MHz	All regions
1,710–2,025 MHz	All regions
2,110–2,200 MHz	All regions
2,300–2,400 MHz	All regions
2,500–2,690 MHz	All regions
3,400–3,600 MHz	No global allocation, but over 80 administrations in Region 1 (Europe and Africa), plus nine in Region 3, including India, China, Japan, and Rep. of Korea

Even with the 885 MHz of spectrum allocated to IMT across all regions, as indicated in Table 1-2, the ITU has performed an analysis of the growing requirements for spectrum to address IMT deployments. ITU-R report M.2078 contains the results of that analysis, both for “legacy” systems in terms of pre-IMT systems, IMT-2000 systems, and IMT-enhanced systems that are already being deployed, as well as the spectrum that will be required for future IMT-advanced deployments. These results indicate that although the combined allocations of WARC-1992, WRC-2000, and WRC-2007 are sufficient for legacy deployments, the new IMT advanced systems are expected to require up to 420 MHz of additional spectrum to be allocated by year 2015 and up to 840 MHz of additional spectrum to be allocated by year 2020.

ITU-R Report M.2078 uses the service categorization defined in ITU-R Report M.2072, “World mobile telecommunication market forecast,” which includes seven service categories, as shown in Table 1-3, including services at speeds of up to 100 Mbps for *super high multimedia* services!

Table 1-3 *ITU Mobile Service Categorization*

Peak Bit Rate	Service Category
< 16 kbps	Speech
< 128 kbps	Multimedia messaging; low multimedia, low rate data
< 384 kbps	Medium multimedia
< 2 Mbps	High multimedia
< 10 Mbps	Very high multimedia
< 30 Mbps	Ultra high multimedia
< 100 Mbps	Super high multimedia

If the ITU estimates prove accurate, it is evident that future World Radiocommunication Conferences will be required to define increasing spectrum allocations for future IMT operations. These systems will be less telecommunications-focused and increasingly data-centric as capabilities evolve toward supporting super high multimedia service offerings.

Propagation

Because of its relative scarcity, mobile systems are required to re-use the allocated radio spectrum across a particular network of cell sites. Radio frequency signals need to propagate between the cell site antenna and the mobile wireless terminal. As the signals propagate, they exhibit a path loss as the emitted energy is dispersed over an increasing area. Estimating the path loss is critical in determining both the coverage provided by a single cell site and the bandwidth available to the IP services offered in that cell coverage area.

The benchmark of propagation loss is that of free space—in other words, the loss in a region that is free from all objects that might absorb or reflect the radio energy. Because the emitted energy from an isotropic antenna is dispersed over the surface of a sphere (with the transmitting antenna at the center of the sphere), the received energy is inversely proportional to the surface area of the sphere ($4 \pi r^2$, where r is the radius of the sphere), as illustrated in Figure 1-1. Using this approach, you can see that the free space path loss follows an inverse square law with changing distance from the antenna, r.

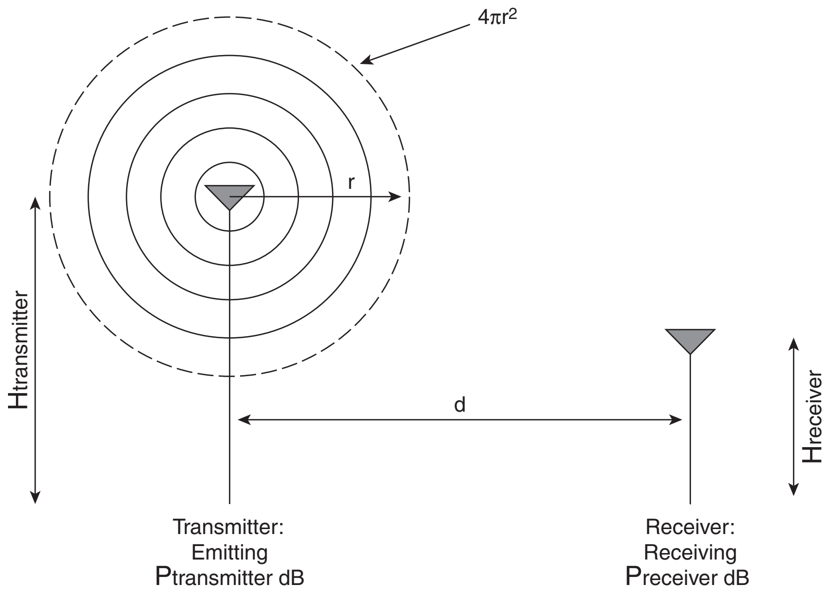

Figure 1-1 *Free Space Loss*

Radio propagation is often defined in logarithmic ratios, termed decibels (dB). When referring to power, a decibel is defined as follows:

$$X \text{ dB} = 10 \text{ Log}_{10}(X/X0)$$

Because path loss is an important quantity in defining the coverage and capacity of mobile radio systems, a useful unit used to compare different environments is to define the decrease in received power over a "decade," where a decade corresponds to an increase in the order of magnitude of distance—for example, when going from 1 mile to 10 miles, 10 kilometers to 100 kilometers, or 13 furlongs to 130 furlongs. Using such an approach and the inverse square law, you can see that the free space loss is equivalent to a path loss of 20 dB/decade; that is, the power received at 10 miles from a transmitter is 100 times less than the power received at 1 mile away from the same transmitter.

Outdoor Coverage

Unfortunately, cellular networks are not built in free space and instead need to accommodate reflections from the ground. Figure 1-2 shows such a model illustrating a direct path between transmitter and receiver as well as a ground reflection.

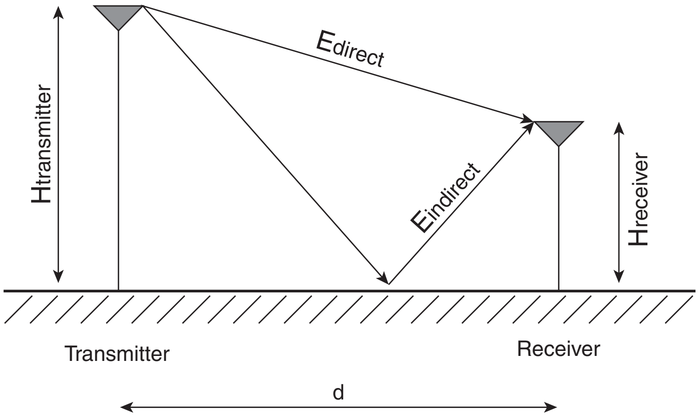

Figure 1-2 *Propagation with Ground Reflection*

The resulting estimate of the received power at the receiver is approximately equal to:

$$Preceiver = Ptransmitter * Gain * (Htransmitter * Hreceiver/ d^2)^2$$

where the *gain* corresponds to the antenna gains in the system and represents the directionality of the antennae radiation patterns when compared to an isotropic antenna.

Counterintuitively, the addition of the ground reflection produces an inverse fourth-power relationship between the power at the receiver and the distance, *d*, between the transmitter and receiver. Instead of the received power diminishing at a rate of 20 dB/decade as predicted by a free-space model, mobile communication systems frequently exhibit received power diminishing at a rate approaching 40 dB/decade as the receiver moves away from the transmitter.

Note The received power is also a function of the square of the product of transmit and receive antenna heights. This is why cellular antennas have conventionally been located on hilltops and raised high on towers, and if you are in a region of poor coverage, why it is often better to try receiving cell phone service on the top floor of a building.

Empirical modeling of radio propagation has been performed by Okumura[1] and Hata.[2] The typical urban Hata model defines the distance (*d*, in meters) related propagation loss as follows:

$$(44.9 - 6.55 \, \mathrm{Log}_{10} \, (h \; base\_station)) \, \mathrm{Log}_{10} \, (d)$$

Using an example 10-meter base station height, the empirical data indicates that the actual path loss should decrease at ~ 38 dB/decade.

This means that when moving from 1 kilometer to 10 kilometers away from a base station antenna, the signal will in fact decrease by a factor approaching 10,000. Although this allows for the scarce spectrum resources to be re-used as neighboring cell site emissions are rapidly attenuated, it also ensures that the cellular designers are constantly battling to provide improved coverage with lower path loss while limiting the number of cell sites required to cover a particular area.

Frequency-Dependent Propagation Loss

We can see from Table 1-1 that in general terms, lower frequencies propagate better than higher frequencies. This intrinsic property has resulted in different systems competing for the sub 1 GHz frequencies, which offer improved propagation characteristics compared with other IMT spectrum allocations.

For example, recent analysis of propagation in the UHF band has been performed,[3] indicating that there is a 26.7 dB increase in path loss when comparing IMT systems deployed in the IMT-defined 698–790 MHz band with those deployed in the 2,500–2,690 MHz band.

Given a typical macro-area propagation loss of 38 dB/decade, it is evident that a change in operating frequency from 698 MHz to 2,500 MHz needs to be compensated by decreasing the cell radius by a factor of $10^{(26.7/38)}$ or five times! The maximum throughput of an individual cell is bounded and will be reduced as the ratio between the wanted signal and the interfering signals decreases. In a lightly loaded system, the interfering signals will be low and hence lower frequencies can provide service over a large coverage area. We describe such deployments as being *coverage limited*, where the performance of the overall system is limited by the attenuation of the wanted signal.

However, as load increases, neighboring cells generate more interference and more cells will be needed to provide the required capacity. In such circumstances, the system becomes limited by interference. We describe such deployments as *capacity constrained*, where the rapid adoption of higher bandwidth IP services means that the maximum attainable cell radius is artificially reduced—for example, by reducing the maximum power—in order to support the required teletraffic density.

The same characteristic that allows a lower frequency signal to propagate over increased distances now also results in the increased effects of interference, as the emissions from neighboring cell sites are similarly attenuated to a lesser degree because of the lower frequency. Hence, the inherent advantages of operating at lower frequencies, which provided improved coverage in a lightly loaded system, diminish over time as the capacity increases.

When these characteristics are coupled with the fact that larger bandwidth allocations are often available at higher frequencies, it is evident that the optimum choice of frequency that delivers the lowest total cost for a specific radio system is a complex tradeoff.

Note The same analysis that indicated that 698 MHz systems had a 26.7 dB advantage in terms of path loss over 2,500 MHz systems also shows that as the 698 MHz systems become capacity constrained, they suffer deteriorating performance compared to the 2,500 MHz systems. For example, they fail to support coverage requirements exceeding 700 kbps/km^2, even in dense urban environments, compared to the 2,500 MHz systems that were able to support capacities in excess of 5 Mbps/km^2.

Fast Fading

Whereas the previous analysis concentrated on propagation in ideal free space or with simple two-ray models, the reality is that mobile radio systems operate with a variety of obstacles and reflections both between and around the base station and mobile terminal, as shown in Figure 1-3. The combination of the disparate propagation paths is called *multipath*—where the transmitted signal arrives at the receiver from various directions over a multiplicity of paths, with each individual path having its own electrical path length and degree of attenuation.

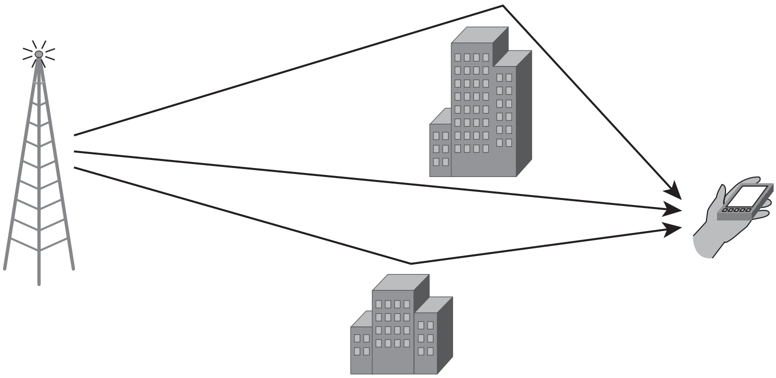

Figure 1-3 *Multipath Propagation*

The variation in path delay caused by the different reflections (termed *time dispersion*) creates distortion in the received signal. This distortion causes the signal from consecutive data intervals to interfere with each other at the receiver, a characteristic termed *Inter-Symbol Interference* (*ISI*). For example, in the Global System for Mobile (GSM) system, the data interval is ~3.9 μs, but operation needs to continue in extreme multipath environments where the time dispersion may approach 20 μs (corresponding to ~6 km difference in path length between the direct path and longest reflection), a scenario that may be experienced in extreme hilly locations with reflections from distant hilltops. In such circumstances, ISI means that at the receiver, a single data interval will be interfering with up to five successive data intervals.

Another consequence of the multipath reception is that different paths can combine constructively or destructively, which will lead to rapid changes in the received signal levels over time, termed *fast fading*, as shown in Figure 1-4. If each multipath component is independent of the others, the Probability Density Function (PDF) of the envelope of multipath components follows a Rayleigh distribution, and the rapid fluctuation in received signal level is termed *Rayleigh fading*. These short-term fluctuations of the Rayleigh fading envelope are superimposed on the long-term distance-related path loss, as defined by the preceding outdoor coverage model.

Just as fast fading generates time-varying changes in the received envelope, the same process can generate frequency-selective fades. This phenomenon creates nulls in the frequency response of the channel. Figure 1-5 illustrates a 15 dB null in the middle of the transmitted signal bandwidth, which may impact signal reception.

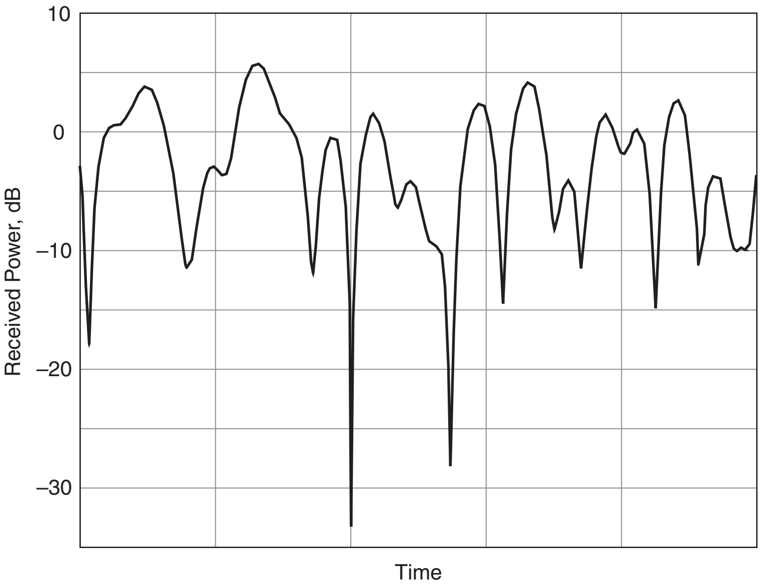

Figure 1-4 *Time-Varying Rayleigh Fading Envelope*

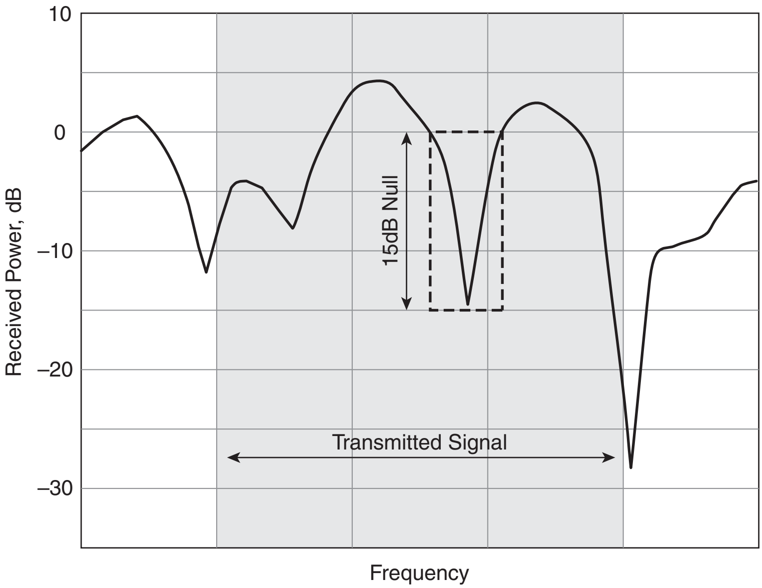

Figure 1-5 *Frequency Selective Fading*

Conversely, if the path lengths of the multipath components only vary slowly, as is normally the case with services delivered to a pedestrian user, flat fading can be experienced. This results in extended periods of time where destructive interference can produce particularly poor propagation conditions. Unless mitigation techniques are employed, the user of IP services may face extended periods where service is substantially degraded.

Shadowing and Building Penetration Loss

Although the first cellular radio systems were designed for vehicle-mounted systems, the mass adoption of cell phone technology has resulted in an increased usage by users located within buildings. JD Power (www.jdpower.com) reported that 2006 was the first year in which indoor wireless call transactions exceeded those made outside, and Telephia[4] reported that indoor viewing contributed 22% of total Mobile TV usage; this trend is set to continue.

Providing coverage to indoor users from outdoor base stations results in an additional Building Penetration Loss (BPL), as signals are required to propagate through the walls, floors, and windows of a building. The penetration loss experienced is dependent on the building materials, window coatings, building orientation, and transmission frequency. In order to provide for 90% in-building coverage, typical mean propagation losses have been shown[5] to be in the range of the following:

- 11 dB for a residential building with plywood wall construction
- 16 dB for a residential building with external brick type construction
- 30 dB for an office building with brick wall construction

BPL is one extreme example of the long-term statistical variation in the mean signal level termed *shadowing*. Because the statistical nature of such slow fluctuations results in the local mean having a Gaussian (or normal) distribution, such shadowing is often termed *log-normal fading*. The standard deviation of the log-normal distribution varies depending on the user's environment. A measurement campaign[6] conducted on behalf of the GSM Association (www. gsma.org; a trade body representing the GSM community) examining 3G coverage concluded that the log-normal standard deviation ranges from 10 dB in rural environments, to 12 dB in urban environments and 13–18 dB in indoor locations. This means that the radio engineer has to combat both the BPL, lowering the mean signal, and an increased spread of signal levels, associated with log-normal statistics, if IP services are to be offered to indoor users.

Historically, indoor coverage has typically been provided as a by-product of radio planning for good outside coverage. However, the increasing adoption of mobile broadband IP services by indoor users highlighted previously indicates that this will likely need to change. Indeed, a study for Signals Research Group (www.signalresearch.com) indicates that if a cellular network is planned with a 98% Probability of Coverage (PoC) to an outdoor voice user, the same network will be able to provide around a 70% probability of providing coverage for delivering a voice service to an indoor user. However, if the service is IP-based and requires a higher data rate (for example, 144 kbps or 384 kbps), the research indicates that the probability of indoor coverage falls to around 40% or 30%, respectively, in an urban environment, or to 26% to 16% in a rural environment, as shown in Table 1-4.

Table 1-4 *Probability of Coverage by Bearer Type*[7]

	Dense Urban	Urban	Suburban	Rural
Outdoor Probability of Coverage	98%	98%	98%	98%
Indoor Probability of Coverage				
12.2 kbps	67%	70%	72%	68%
64 kbps	50%	70%	72%	68%
144 kbps	41%	39%	34%	26%
384 kbps	31%	28%	22%	16%

As mobile services evolve from low-rate voice and short message-centric services toward higher-speed IP services, the challenge of providing reliable indoor service will become increasingly challenging. It may be likely that simply providing an increasing cell density in order to accommodate the 11 to 30 dB of building propagation loss is uneconomical and that the cellular service providers will be forced to adopt alternative approaches to providing reliable indoor mobile broadband IP services.

Note ITU-R Report M-2078 includes data on teletraffic density and macro-cellular coverage area per radio environment. In this report, Building Penetration Loss is assumed to decrease the urban cell coverage from 0.65 km^2 to 0.10 km^2, to decrease the suburban cell coverage from 0.65 km^2 to 0.15 km^2, and to decrease the rural cell coverage from 0.65 km^2 to 0.22 km^2.

Modulation

Modulation is the process of encoding information onto one or more radio frequency carriers. Digital modulation can encode information into the phase, frequency, or amplitude of a carrier, or use a combination of such techniques. One of the most common techniques of modulating a carrier is with Binary Phase Shift Keying (BPSK), where information is encoded by changing the phase of a reference carrier. Figure 1-6 shows an example of a BPSK waveform with modulated phase. Each encoded bit corresponds to either a phase shift of 0 degrees or a phase shift of 180 degrees. In this example, if a phase shift of 0 degrees represents the encoding of bit “0,” and a phase shift of 180 degrees represents the encoding of bit “1,” the waveform in Figure 1-6 encodes the seven-bit sequence 0100100.

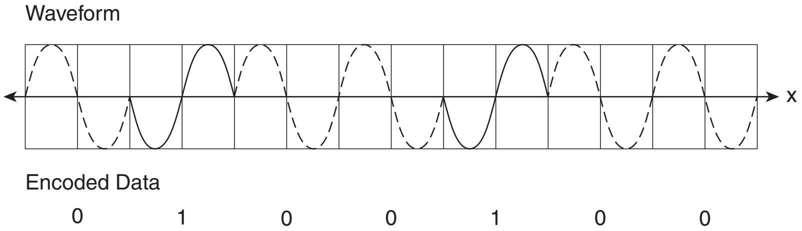

Figure 1-6 *Binary Phase Shift Keying Waveform*

Whereas BPSK encodes a single bit of information in each symbol, Quaternary Phase Shift Keying (QPSK) uses four possible phases to encode two bits of data per symbol. Figure 1-7 shows an example of a QPSK waveform:

- A phase shift of 0 degrees can be used to encode the two bits "00."
- A phase shift of 90 degrees can be used to encode the two bits "10."
- A phase shift of 180 degrees can be used to encode the two bits "11."
- A phase shift of 270 degrees can be used to encode the two bits "01."

Using this encoding scheme, the waveform in Figure 1-7 encodes the 14-bit sequence 00101101000111.

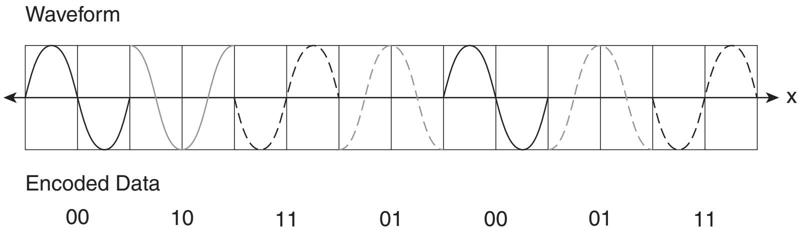

Figure 1-7 *Quaternary Phase Shift Keying Waveform*

Another way of representing the modulation schemes is using a constellation diagram, which displays the modulated waveform on a two-dimensional quadrature and amplitude axis. On a constellation diagram, BPSK and QPSK will be shown as the individual phases used to encode the information (two phases in the case of BPSK and four phases in the case of QPSK), as shown in Figure 1-8.

Obviously, because two bits of information are encoded in each symbol of QPSK compared with the one bit of information with BPSK, QPSK is able to transmit more information in the same time interval. One class of higher-order modulations makes use of a combination of Amplitude and Phase modulation and is termed Quadrature Amplitude Modulation (QAM). Figure 1-9 shows the constellation diagram for a 16-QAM signal, where now four bits of information are encoded in a single symbol.

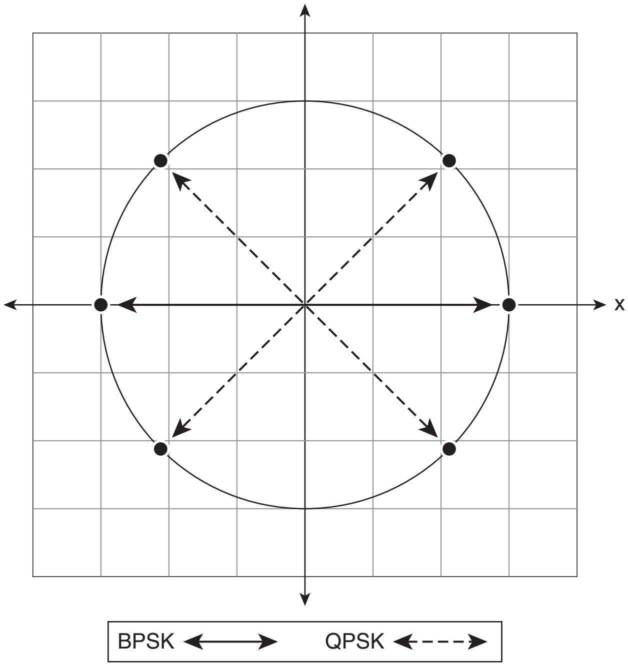

Figure 1-8 *Constellation Diagram for BPSK and QPSK*

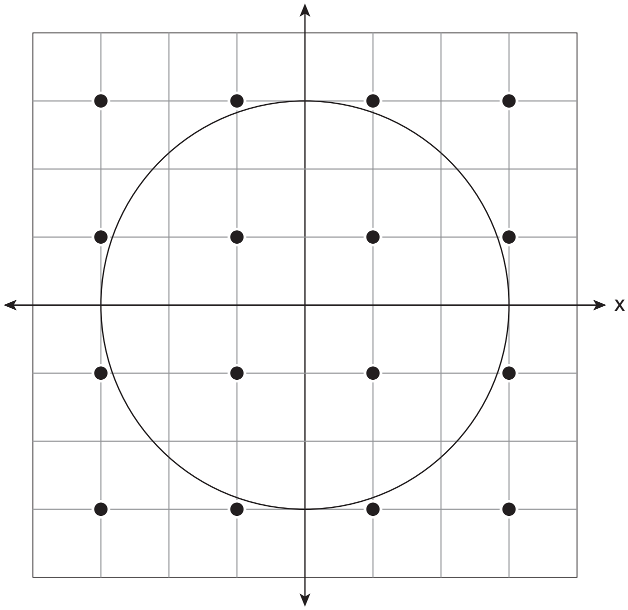

Figure 1-9 *Constellation Diagram for 16 QAM*

The more advanced mobile radio systems may include 64-QAM modulation, where six bits of information are encoded in each symbol, or six times the information encoded in the basic BPSK waveform. The IEEE 802.16 specification[8] used by the WiMAX system includes an option for 256-QAM modulation, where eight bits are encoded in each transmitted symbol.

The higher the order the modulation, the more data that can be encoded in a waveform and the greater the transmission speeds offered to the IP applications that utilize the mobile radio network. However, the higher-order modulation schemes are also more susceptible to impairments by noise and interference, leading to an increased probability of errored data being recovered at the receiver. As the distance between the mobile terminal and the base station increases, the Signal-to-Noise Ratio (SNR) or Carrier-to-Interference Ratio (CIR) will fall at around 38 dB/decade. Therefore, it is important to understand how the different modulation schemes perform in the presence of decreasing SNR ratios. A high-order modulation scheme that provides high speed throughput when the user is located in the vicinity of a base station actually performs worse in terms of the error-free throughput as the terminal moves away from the base station.

Figure 1-10 compares the Bit Error Rate (BER) of five different modulation schemes in the presence of Additive White Gaussian Noise (AWGN). Because a single symbol of 256-QAM modulation transmits eight bits of information compared with the single bit of information transmitted with a single symbol of BPSK modulation, the performances are normalized by plotting the BER against the ratio of Energy-per-bit (Eb) to Noise Power Spectral Density (No), or Eb/No. Eb/No is related to SNR by scaling for the bandwidth in which the SNR is measured, B, versus the data rate of the modulation scheme, R, where:

Eb/No = SNR * B/R

Consider a mobile radio system designed for a nominal BER operating point of 10^{-2}, and assume that the base station has a constant energy per symbol, independent of which order modulation is used. If the radio planner has set the transmit power such that at 1 kilometer from the base station, there is 16.5 dB of E_b/N_o at the 256-QAM demodulator, Figure 1-10 indicates that the BER operating point has been achieved. As the user moves away from the base station, the BER of the 256-QAM demodulator will degrade. If the user moves a further 459 m away from the base station to 1.459 kilometers away, using the Hata model, the carrier would be attenuated by an additional:

38 log10 (1.459) = 6.2 dB

If the modulation is unchanged, the E_b/N_o will be decreased to 10.3 dB and Figure 1-10 indicates that corresponding BER would degrade to over 0.07, or over seven times that at the nominal operating point.

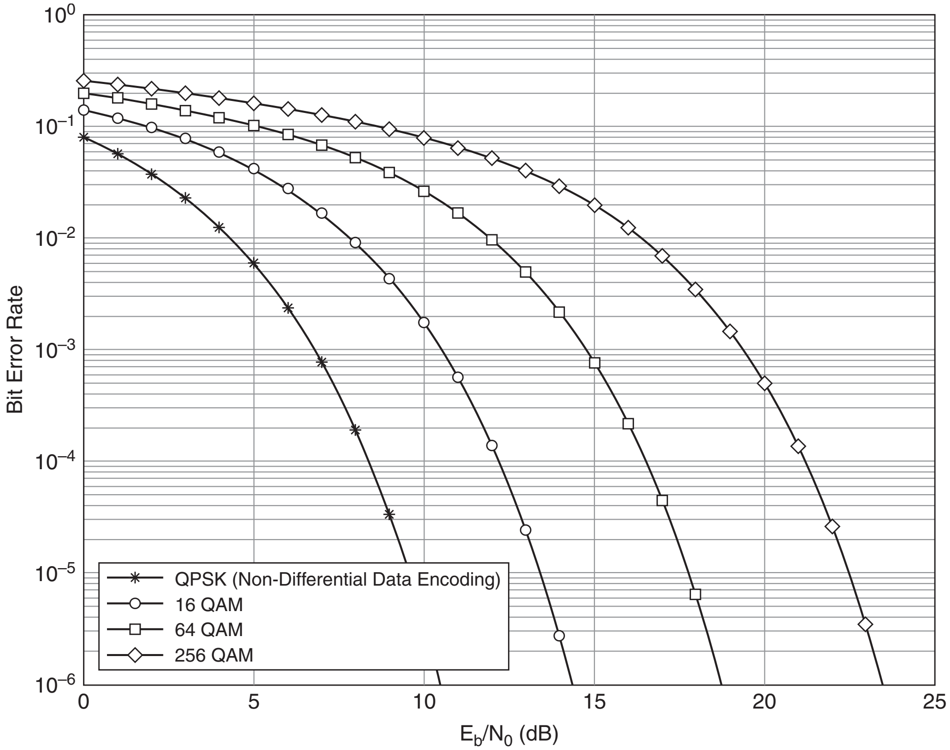

Figure 1-10 *BER Performance in AWGN of Different Modulation Schemes*

Now, if instead the modulation is decreased from 256 QAM to 64 QAM while keeping the symbol energy constant, since Eb/No is a normalized measure of the energy per symbol to noise spectral density ratio, there will a corresponding increase in E_b by a factor of:

$\mathrm{Log}_2\, M$

where M is the number of alternative modulation symbols.

Therefore, decreasing the order of modulation from 256 QAM to 64 QAM will increase the E_b/N_o by:

$\mathrm{Log}_2\, 256 / \mathrm{Log}_2\, 64 = 1.33$ or 1.24 dB

This means that by moving a further 459 meters away from the base station and switching modulation schemes, the E_b/N_o available at a 64-QAM demodulator is 11.5 dB, which, according to Figure 1-10, is sufficient for the nominal operating BER of around 10^{-2} to be maintained.

As the user moves further away from the base station, the path loss increases such that when the user is 2 kilometers away, the Hata model indicates that the 64-QAM BER would degrade to 0.07. The increase in path loss can be compensated by switching to 16-QAM modulation, such that the nominal BER is maintained at 0.01.

Finally, as the user approaches the edge of coverage at 2.75 kilometers away from the base station, the path loss will have increased by almost 17 dB compared to the nominal 256-QAM operating point at 1 kilometer. However, by switching to QPSK, the nominal BER is maintained.

Figure 1-11 summarizes the coverage by modulation type using the AWGN results of Figure 1-10 with the 38 dB/decade Hata model. The results highlight that even when the system is designed to operate with 256-QAM modulation, over half the cell is only able to receive data modulated with QPSK.

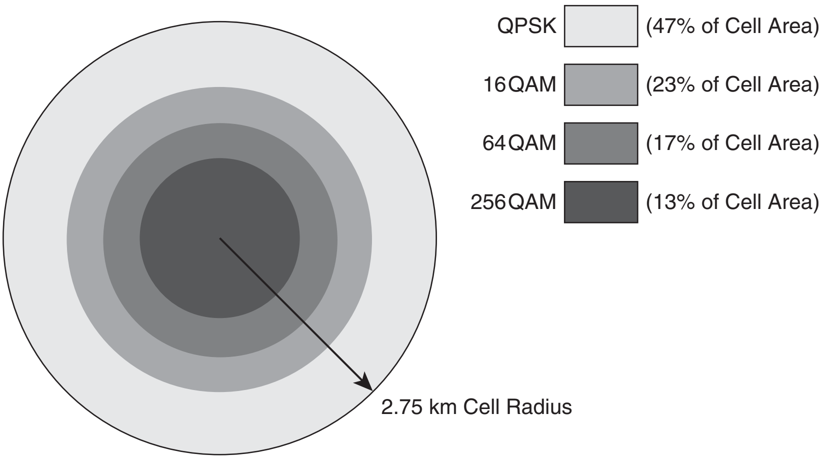

Figure 1-11 *Cell Coverage by Modulation Type for a Nominal BER of 0.01 in AWGN*

Unfortunately, by switching modulation schemes to keep the BER constant, the throughput achievable has been reduced by a factor of four from the original eight bits per symbol when operating at 256 QAM (when the use was in good coverage) to the two bits per symbol (when operating at the cell edge). In the downlink, this offset in throughput could be countered by applying a disproportionate amount of base station power to those users at the cell edge, resulting in increased interference in neighboring cells. Unfortunately, the same approach is often not possible in the uplink, where the device is peak power limited because of its power amplifier design. As a consequence, the IP design engineer needs to factor in the decrease in IP throughput able to be supported by those users located toward the cell edge.

The decrease in cell performance is further exacerbated by the impact of fading on the BER performance. The deep nulls characteristic of Rayleigh fading cause the BER to dramatically increase, as shown in Figure 1-12. The required Eb/No necessary to deliver the nominal un-coded BER increases by the order of 8–10 dB, decreasing the maximum achievable coverage of the cell in AWGN conditions from 2.7 kilometers to 1.5 kilometers.

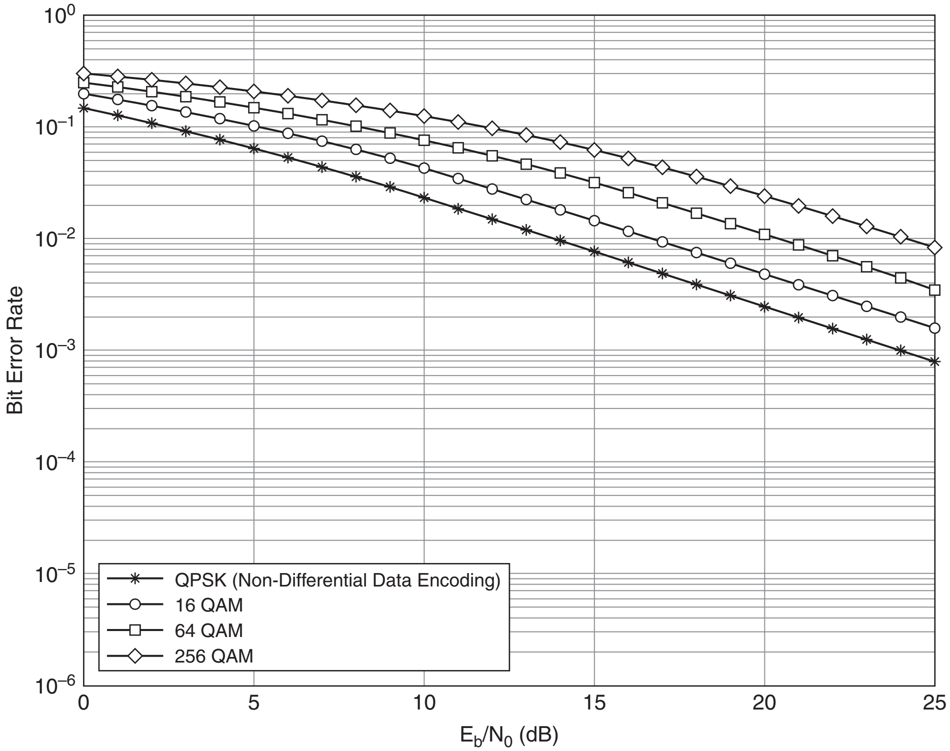

Figure 1-12 *BER Performance of Different Modulation Schemes in Rayleigh Fading*

Multiple Access Technologies

Although a simple modulated carrier may be sufficient for supporting a point-to-point communications link, mobile radio systems are characterized by their ability to support multiple users using a common communications resource. Multiple access techniques define how the communication resources are partitioned between the different users.

Figure 1-13 illustrates the various techniques for partitioning the radio resources, which are as follows:

- **Time Division Multiple Access (TDMA):** Each user is allocated a particular time interval when they have access to the communications resources.
- **Frequency Division Multiple Access (FDMA):** The communication resource is partitioned into separate carriers, and each user is allocated a subset of the overall frequencies when they have access to the communications resources.
- **Code Division Multiple Access (CDMA):** All users can use the same frequency at the same time but are isolated by a separate pseudo-random code that they use to spread their information.

- **Space Division Multiple Access (SDMA):** Isolation between communications resources is achieved using smart antenna techniques that generate directional antennas for each user.

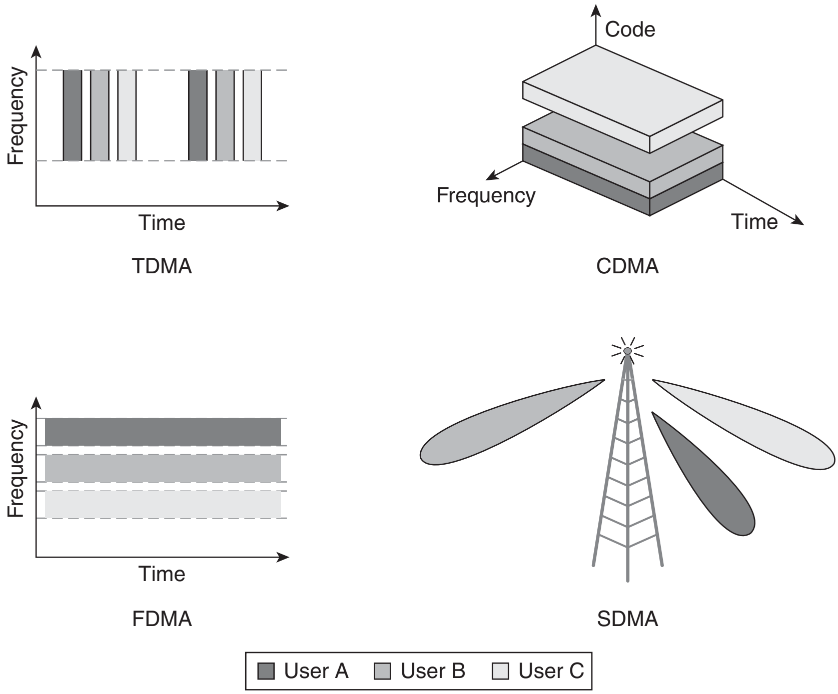

Figure 1-13 *Multiple Access Techniques*

Frequently, communications systems will use a combination of multiple access techniques; for example, the General Packet Radio Service (GPRS) system uses a combination of TDMA together with FDMA, and the High-Speed Downlink Packet Access (HSDPA) system uses a combination of CDMA together with TDMA.

Time Division Multiple Access

Time Division Multiple Access (TDMA) allows a number of users to share a communication resource by segmenting the resource into time slots and allocating each individual time slot to one user at a time. A communication resource may be dedicated to either transmission or reception, in which case two sets of frequencies are required to support bidirectional communications. Because of the frequency separation between transmit and receive resources, this technique is termed Frequency Division Duplex (FDD). The GSM system is one example of a communications system that uses TDMA with FDD.

Alternatively, a communication resource can be dedicated to transmission and reception, in which case there needs to be clear demarcation in time between those timeslots allocated to transmission and those allocated to reception. Because of the time separation between transmit and receive resources, this technique is termed Time Division Duplex (TDD). The Digital Enhanced Cordless Telecommunications (DECT) system is one example of a communication system that uses TDMA with TDD.

GSM uses a TDMA frame structure of 60/13 milliseconds (4.615 milliseconds) that is split into eight timeslots (TS0 to TS7) of length 0.577 milliseconds, as shown in Figure 1-14. Because timeslots cannot overlap, TDMA systems require the use of tight synchronization and guard periods, during which no transmissions occur. For example, Figure 1-14 shows that a GSM timeslot includes 8.25 bits of guard period out of a total timeslot duration of 156.25 bits, equivalent to over 5% of the overall communications resource.

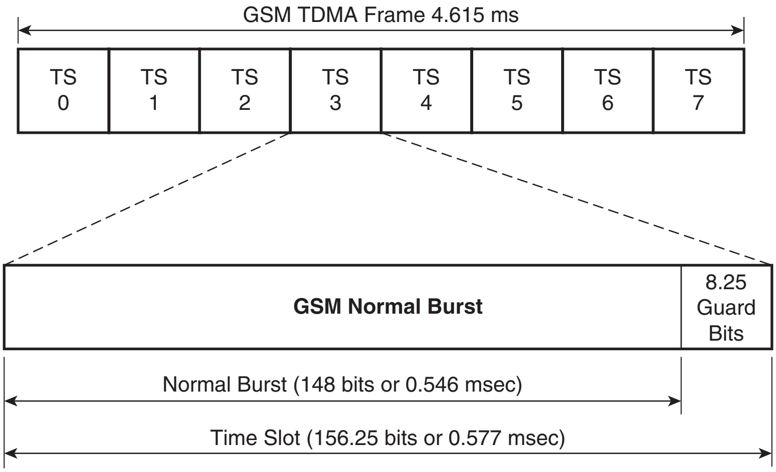

Figure 1-14 *TDMA Frame Structure*

When sending information, each user will be allocated one or more timeslots for transmitting and receiving information. In this way, users with more information to send can be allocated a larger percentage of the shared communications resource in order to support higher transfer speeds. Figure 1-15 highlights the multislot capability of the GSM system, where multiple timeslots can be allocated to a single user.

The figure shows the basic timeslot allocation of a single timeslot for receive (Rx), one for Transmit (Tx) and another timeslot for monitoring neighboring cells (Mn). The single timeslot operation is able to support 14.4-kbps service. Because GSM operates in FDD mode, a terminal can use a single radio, which can be switched between receive and transmit operation, allowing time to re-tune the oscillator frequency between bursts. The eight-slot TDMA frame allows different combinations of transmit and receive timeslots. The 2 + 2 configuration highlights how two timeslots can be allocated to reception and transmission, increasing the aggregate throughput to 28.8 kbps. Alternatively, the allocation may be asymmetrical with Figure 1-15 also showing a 3 + 1 multislot configuration that, assuming each slot can support 14.4 kbps, is able to support transmission speeds of 43.2 kbps in the downlink.

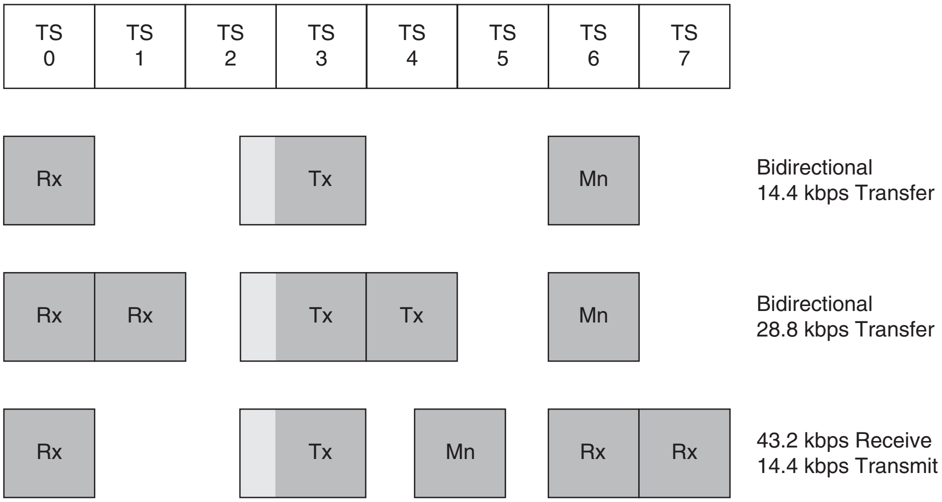

Figure 1-15 *TDMA Multislot Configuration*

Frequency Division Multiple Access

Frequency Division Multiple Access (FDMA) allows a number of users to share a communications resource by segmenting the resource into separate subcarriers and allocating individual subcarriers to one user at a time. In traditional mobile radio systems, a user was allocated a single subcarrier at any one time. For example, in the GSM system, a single user is allocated a 200-kHz channel during a particular time duration (corresponding to a TDMA frame). Because a large cell site will likely support multiple subcarriers, the same timeslot can be occupied by different users accessing using different subcarriers. Figure 1-16 shows three users simultaneously accessing via Timeslot 3, with each accessing via a different subcarrier.

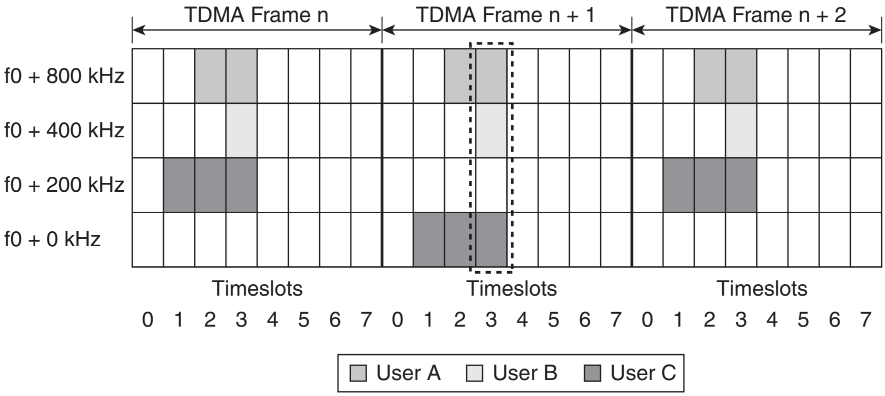

Figure 1-16 *Traditional FDMA Access*

Whereas traditional FDMA systems have operated in single-carrier mode, more recently FDMA techniques have been applied to multicarrier systems. In contrast to GSM, where a user is only allocated a single subcarrier at one particular time instant, a multicarrier FDMA user can be allocated multiple subcarriers, or tones, at a particular instant with each subcarrier being modulated independently. If the frequency separation between subcarriers is selected to be an integer of the modulating symbol rate, the tones are said to be orthogonal to each other, and the multiple access scheme is termed Orthogonal FDMA (OFDMA). With OFDMA, a receiver with optimum sampling and perfect frequency synchronization is able to reduce the Inter-Carrier Interference (ICI) to zero. Figure 1-17 shows the receiver outputs for five separate orthogonal tones over time. The figure shows that at the correct sampling instant, the contributions of the neighboring tones sum to zero and hence do not interfere with the reception of the wanted tone.

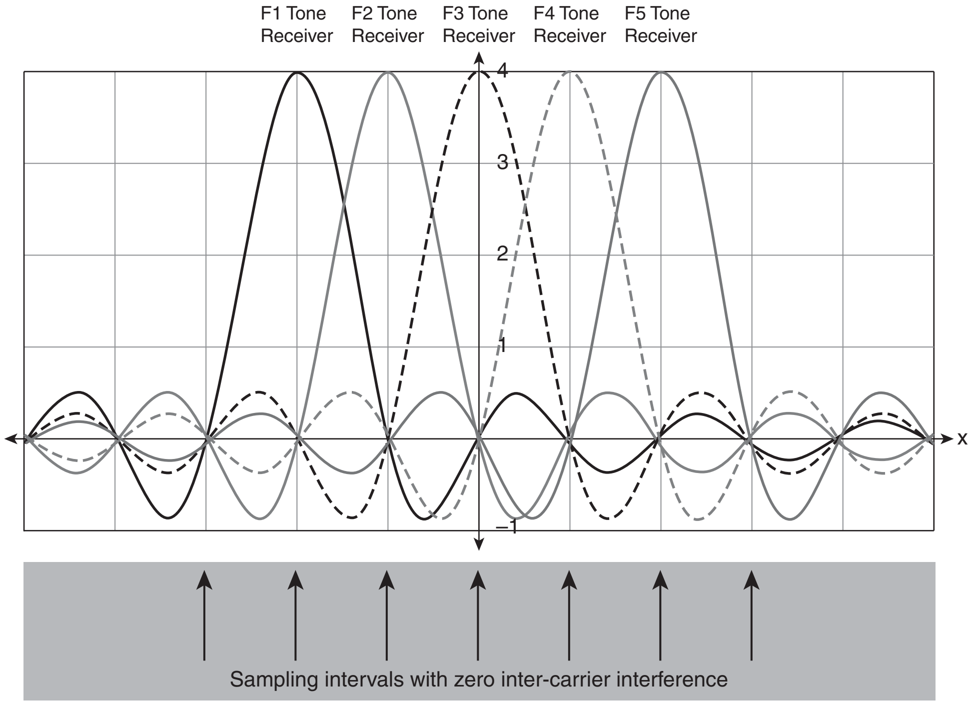

Figure 1-17 *Orthogonal Tone Separation in Multicarrier FDMA*

One of the advantages of OFDMA is that because multiple subcarriers are used, the overall symbol rate per subcarrier is reduced. For example, if 1,000 tones are used, the parallel tone symbol duration will be 1,000 greater than in a serial tone system, and this new subcarrier symbol duration can be significantly greater than the multipath duration, enabling OFDMA systems to offer robust mitigation of multipath effects.

Another advantage of OFDMA systems becomes apparent as the bandwidth/number of tones increases such that the channel experiences frequency selective fading. In such circumstances, the fading experienced by the tones at the lower end of the channel bandwidth will be de-correlated with the fading experienced by the tones at the upper end of the channel bandwidth, such that the effects of frequency selective fading are diminished.

Figure 1-18 shows an OFDMA system with 15 kHz subcarrier separation. The figure highlights how OFDMA enables users to be assigned multiple subcarriers in particular timeslots.

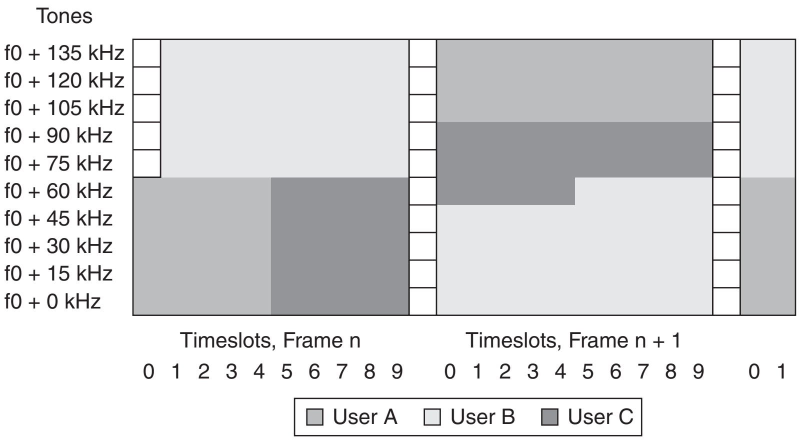

Figure 1-18 *Multiuser Operation in OFDMA*

Unfortunately, one of the disadvantages of OFDMA systems is that a handset transmitter needs to be able to transmit parallel tones simultaneously. The combination of these individual tones combine to generates a time waveform with a high Peak-to-Average Power Ratio (PAPR). A high PAPR places additional requirements on the linearity of the power amplifier. Linearity is normally increased by de-rating a particular amplifier, leading to decreased efficiency. Although inefficient or more expensive power amplifiers are not necessarily critical issues for the base station designer, they do pose challenges in the uplink where a decrease in terminal transmitter efficiency might decrease the maximum transmission time or lead to degraded uplink performance/coverage compared to the performance of single-carrier systems. In Chapter 3, we introduce techniques used by OFDMA-based mobile systems for reducing PAPR.

Code Division Multiple Access

Code Division Multiple Access (CDMA) is the key multiple access technology used in today's third-generation cellular systems, being the foundation for both 3GPP's WCDMA and 3GPP2's cdma2000 wireless systems and their respective evolutions to support mobile broadband data. CDMA operates by using special spreading codes to artificially increase the symbol rate of the information to a chip rate of the spreading code; the information to be transmitted is logically exclusive-ORed with the spreading code. (Exclusive-ORing is a logical function where precisely one input must be 1 [true] for the output to be 1 [true].) Figure 1-19 shows an example of spreading factor of 8 (where Td, the information bit duration of the original data, is 8 times Tc, the chip-timing duration).

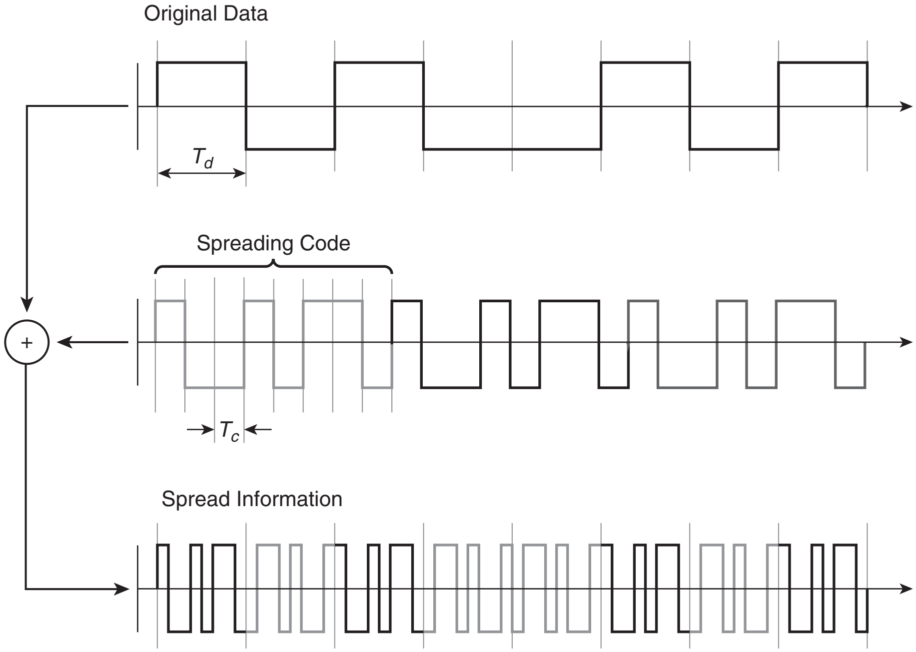

Figure 1-19 *Spread Spectrum Operation*

The spreading codes have special orthogonal properties that allow a particular user's information to be recovered from the composite waveform. Figure 1-20 shows an example of orthogonal spreading codes. Here, eight codes are shown with zero cross-correlation properties. The cross-correlation of two sequences is calculated using the vector product of two sequences.

When correctly synchronized, the vector product of sequence S1 and sequence S2 is equal to zero, as shown in Figure 1-21.

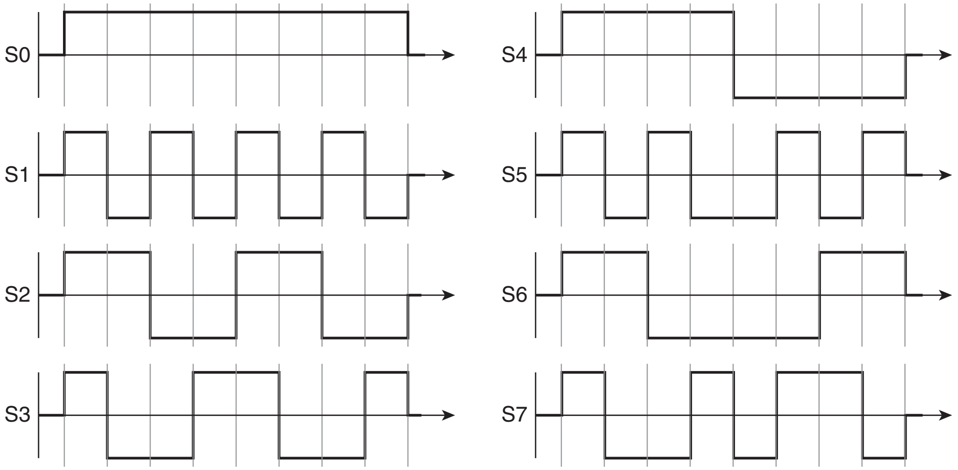

Figure 1-20 *Orthogonal Spreading Sequences*

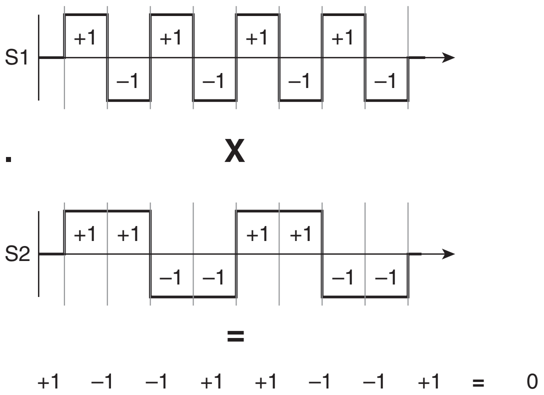

Figure 1-21 *Sequence Cross-Correlation Calculation*

Because of the good cross-correlation properties, each user can be allocated a separate spreading code and the receiver is still able to recover the original data from the combined waveform.

Instead of defining the use of fixed spreading codes in a system, Variable Spreading Factor (VSF) codes can be used. The different spreading factors result in different code lengths and allow different source rates to be effectively combined while still preserving orthogonality between the different users/sources. This is achieved by using an Orthogonal Variable Spreading Factor (OSVF) code, which can be represented by a code tree together with smart code allocation. Figure 1-22 shows a code tree that has codes up to 64 bits long.

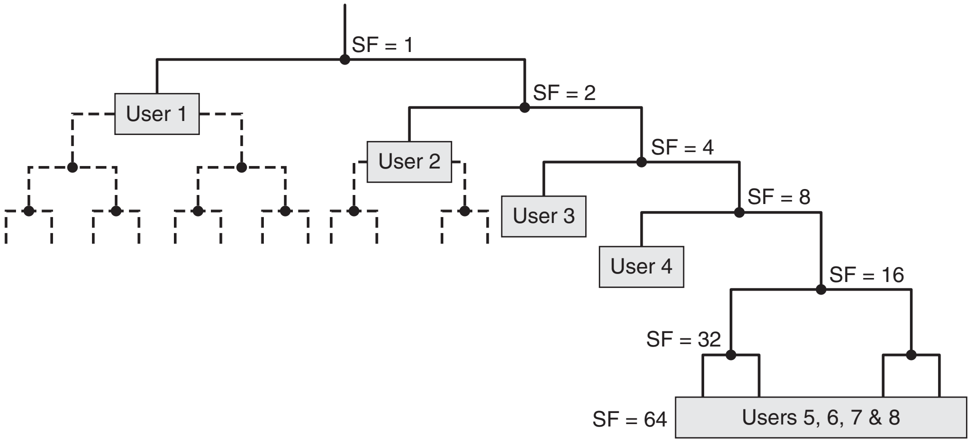

Figure 1-22 *Orthogonal Variable Spreading Factor Codes*

The code tree shows how eight users have been allocated different codes corresponding to different spreading factors. User 1 has been allocated a code with spreading factor 2 (SF=2). The rate at which User 1 can transmit data will be twice as fast as User 2, who has been allocated an SF=4 code. Similarly, User 2 can transmit data at a rate double User 3, who has been allocated an SF=8 code, and so on. As can be seen, the adoption of OVSF codes allows for the multiplexing of users with different IP data rate requirements. Note, however, that if the instantaneous data rate needs to change, as will frequently be the case during an IP session, corresponding functionality needs to be defined, which enables the spreading code to be updated in real time.

Although CDMA techniques have been used extensively in 3G cellular systems, they do have some unique challenges when it comes to supporting high-speed IP services, as follows:

- **Power Control:** In CDMA systems, one user's information signal is another user's noise source. If all CDMA users transmitted at the same power level, the base station will be able to decode those signals from those users near to the base station, whereas those signals from distant users will be perturbed by high levels of interference (the so-called near-far problem). In order to solve this, CDMA systems must include fast power control; for example, in UMTS, the power control is updated 1,500 times per second.[9] The efficiency of any power control system will decrease as the burstiness of traffic increases.

- **Low Spreading Factor Support:** Many of the advantages of CDMA are apparent only for high spreading factors. As the IP bandwidth requirements increase, the spreading factors decrease, leading to a decrease in CDMA efficiency.

Space Division Multiple Access

Space Division Multiple Access (SDMA) is an emerging technology for improving the throughput and capacity of mobile wireless systems. SDMA involves using advanced antenna systems and spatial signal processing to isolate the communications resources between multiple users, enabling those users to operate in the same channel simultaneously. Figure 1-23 shows how an eight-element base station is able to calculate the spatial signatures of the two users, effectively allowing re-use of the same frequency within the same cell.

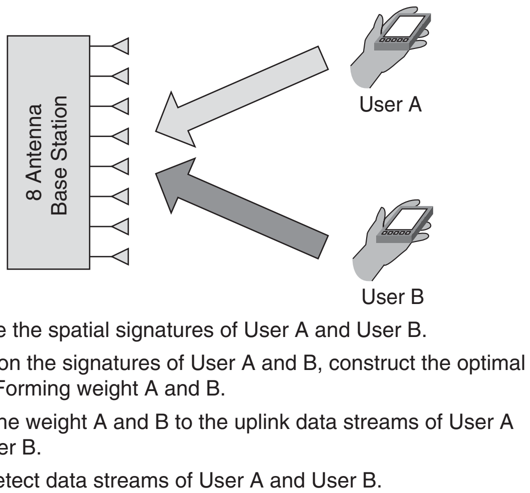

Figure 1-23 *Space Division Multiple Access*

It is anticipated that SDMA and Adaptive Antenna Techniques will be used for a variety of purposes in future mobile broadband systems, including the following:

- **Limiting co-channel interference:** The directional antenna patterns of SDMA decrease the downlink interference to other users in the same and neighboring cells, improving overall throughput.
- **Improving coverage:** The directional antenna gains of SDMA allow any Adaptive Modulation and Coding thresholds to be moved further away from the cell center, allowing more users to benefit from the throughput associated with higher-order modulation.
- **Stabilizing multipath:** The latest Multiple-Input Multiple-Output (MIMO) techniques take advantage of multipath propagation. SDMA techniques can be used to stabilize the multipath environment in outdoor environments, allowing an increasing percentage of macro-cellular users to benefit from MIMO gains.

SDMA techniques are not without their challenges. At the very least, conventional SDMA techniques rely on reciprocity between the uplink and the downlink channels. Although this is the case for Time Division Duplex (TDD) operation, where the same carrier is used for uplink and downlink transmission, the same is not true for Frequency Division Duplex (FDD), where there is a large frequency offset between the downlink channel and the uplink channel.

Combating Radio Impairments

This section introduces techniques that can be used to mitigate performance degradations, increasing the likelihood that IP applications can operate in such imperfect environments.

Forward Error-Correcting Codes

In the previous section, we described scenarios where adaptive modulation can be employed to minimize the effects of decreasing Eb/No as the user moves away from the center of a cell toward the cell edge. The fading analysis, however, has shown that when the channel suffers from Rayleigh fading, the resulting Bit Error Rates are significantly higher than those experienced in simple Additive White Gaussian Noise (AWGN) environments.

Forward Error-Correcting (FEC) codes add redundancy to the original information bits in a deterministic fashion, such that the receiver can correct and detect a certain subset of the errors received due to noise and interference impairments. FEC codes are characterized by their code rate, R, which indicates how much redundancy is added to the original information:

$$R = i/(i + r)$$

where i and r denote the number of information and check bits, respectively.

For example, a 1/3 rate FEC code transmits three coded bits for each information bit ($i=1$, $r=2$), and a 3/4 rate FEC code transmits four coded bits for every three original information bits ($i=3$, $r=1$).

The FEC operation decreases the probability that an IP packet will be received in error, reducing packet error rates and any delay associated with packet retransmission. However, the use of FEC coding effectively reduces the IP throughput, because the redundancy bits do not convey any new information. As a consequence, in order to enable systems with optimum throughput, systems that use adaptive modulation frequently define the use of adaptive FEC, which is a parallel technique where the amount of redundancy allocated to FEC coding is altered according to the prevailing channel conditions. The radio designer defines a range of FEC coding rates in order to accommodate different channel conditions. For example, in good propagation environments, the probability of a bit error will be low, so a high-rate code can be used (little additional redundancy added to the source information bits). Conversely, at the edge of a cell, even when using the lowest-order modulation, the probability of bit error might be high, so a low-rate code should be used.

The combination of adaptive modulation and adaptive FEC is termed Adaptive Modulation and Coding (AMC). Chapters 2 and 3 provide more details regarding the use of AMC in different wireless standards. As one example, 3GPP's High-Speed Downlink Packet Access (HSDPA) supports coding rates raging from 0.14 to 0.89.[10] When adaptive coding is used together with QPSK and 16-QAM operation, the nominal data rates supported can range from 68.5 kbps up to 12.779 Mbps, as shown in Table 1-5.

Table 1-5 *HSDPA Adaptive Modulation and Coding*

Modulation	Number of HS-DSCH*	Effective Code Rate	Instantaneous Data Rate
QPSK	1	0.14	0.07 Mbps
QPSK	1	0.27	0.13 Mbps
16 QAM	2	0.38	0.72 Mbps
16 QAM	5	0.65	3.10 Mbps
16 QAM	15	0.89	12.78 Mbps

*HS-DSCH = High-Speed Downlink Shared Channel

AMC is increasingly being adopted as a technique used to improve the overall throughput of next-generation mobile radio systems. This is important information for IP designers who need to accommodate transport over cellular systems; there might be a significant difference in the throughput available to those users in the good coverage toward the center of a cell and those users located at the edge of coverage.

Mitigating Multipath Effects

You can combat multipath effects using a variety of techniques, including the following:

- **Using channel equalization techniques:** Channel equalization involves estimating the impulse response or the multipath channel, and applying the inverse of this impulse so as to minimize the effects of Inter-Symbol Interference.
- **Using spread spectrum techniques:** Spread spectrum systems spread the signal energy over a wide band by decreasing the symbol duration by the processing gain of the system. The wide bandwidth makes spread spectrum systems more immune to narrow frequency nulls generated as a result of multipath fading. In CDMA systems, multiple correlator receivers, sometimes called *fingers*, are then used to independently decode the individual multipath components in the receiver. The contribution from all the correlator receivers are then combined (using a *Rake* receiver) in order to leverage the energy dispersed over different multipath components.

- **Using systems of parallel tone modems:** In OFDMA systems, the information to be transmitted is converted into a large number (*N*) of parallel streams. Each stream is used to modulate a narrow band subcarrier, with the overall transmission being comprised of *N* finely spaced subcarriers. The effective symbol rate of each of the subcarriers is decreased by a factor of *N*, resulting in the symbol duration being increased by a factor of *N*. If the new subcarrier symbol duration is significantly greater than the multipath delay spread, only a small guard period needs to be used in order to mitigate the effects of ISI and a simple receiver can be used to recover the original information.
- **Using frequency selective transmission:** In conventional systems, the frequency selective channel response can be combated by using frequency hopping. Frequency hopping systems use a pseudo-random hopping sequence to repeatedly change the carrier frequency of the transmitted modulated waveform. The pseudo-random sequence is known by the receiver, which synchronizes the sequence phase with the transmitter. If a slowly moving user is unfortunate enough to be positioned at a location of destructive interference for one frequency, there is a good probability that the location will not suffer destructive interference after the next frequency hop. In multicarrier systems, frequency selective multiple access can be achieved by smart allocation of tones. For an individual user, those tones that are exhibiting constructive interference are allocated in preference to those exhibiting destructive interference.
- **Interleaving:** The time-varying nature of Rayleigh fading, as shown previously in Figure 1-4, results in errors being generated in bursts corresponding to the nulls in the amplitude of the received signal. Unfortunately, Forward Error-Correcting (FEC) codes are not suited for correcting bursts of errors. Consequently, interleaving is a technique that is used to randomize the location of errors, ensuring that, even in the presence of multipath, the bit errors at the FEC decoder should be more uniformly distributed.

Note An unfortunate characteristic of interleaving is an increase in the end-to-end delay. For example, the GPRS system[11] interleaves one Radio Block over four consecutive TDMA frames. With one TDMA frame lasting 4.615 milliseconds, the use of the interleaver in GPRS to combat burst errors adds an additional 18 milliseconds to the packet transmission delay.

Radio systems frequently employ a range of techniques for mitigating multipath interference. For example, the GSM system includes a combination of equalization, frequency hopping, and interleaving in order to combat Rayleigh fading and Inter-Symbol Interference.

Complexity of Multipath Mitigation

As mobile radio systems move to increasingly supporting mobile broadband IP services, the overall peak throughput speeds required to be supported necessarily increase and the symbol durations correspondingly decrease. As data speeds increase, the use of a pure-spread spectrum Rake receiver technique is challenged, because the increase in speeds necessarily requires the spread spectrum processing gain to be decreased. The performance of the Rake receiver is dependent on a minimum processing gain, below which its performance becomes suboptimal, resulting in an error-floor that cannot be mitigated.

This has caused 3GPP to define the use of additional equalization techniques, previously avoided in 3G systems because of their computation complexity. Analysis[12] has shown that such advanced receivers can improve the performance of HSDPA in the pedestrian environment, increasing the average throughput from around 30 kbps to 90 kbps when the user is located at the edge of a cell and from around 700 kbps to 1,200 kbps when the user is located near the center of the cell.

The problem with relying on equalizers to combat the impairments generated by multipath propagation is their complexity, which rises rapidly with system bandwidth. Complexity comparisons have been performed by Van Nee and Prasad,[13] contrasting the implementation of a 24-MBps modem realized using a Gaussian Minimum Shift Keying (GMSK) serial modem (a type of frequency modulation used by the original GSM system) and an OFDM multicarrier implementation. The analysis estimates that in terms of multiplications per second, the equalization of the single-carrier GMSK system is ten times more complex than the Fast Fourier Transform (FFT) required for OFDMA reception. The OFDMA FFT complexity grows only slightly faster that the product of the bandwidth-delay spread product—that is, a function of the relative amount of Inter-Symbol Interference, compared to the equalizer-based solutions that have complexity growing at the square of the bandwidth-delay spread product.

As IP data rates increase, transmission bandwidths will increase further, and OFDMA's advantage over single carrier will more than likely continue to grow.

Smart Scheduling

Previous voice-centric mobile radio systems used deterministic scheduling; for example, these systems used a round-robin procedure, where the system would attempt to use a range of techniques to combat the multipath fading conditions experienced in the allocated scheduling interval. The move to data-centric mobile radio systems enables the radio design engineer to leverage the increased jitter tolerance of IP applications by defining the use of smart schedulers. Smart schedulers leverage the fact that when considering IP data transmissions in a multiuser system, some users may be momentarily disadvantaged by being located in a region of destructive multipath interference, whereas others will benefit from being located in a region of constructive multipath interference.

Techniques termed *multiuser diversity* are used to segment the resources on a time basis; during each time interval, the system is able to determine the instantaneous fading environment on a per-user basis. This information is used to decide which user to serve in each time interval. Compared to the deterministic round-robin scheduler, which would expend power to provide a degraded service through the deep fading nulls (for example, using a low-order modulation and high coding rate), the smart scheduler expends the same amount of power, offering higher throughput to a more advantaged user (for example, using a high-order modulation and low coding rate). Because of the time-varying nature of the multipath fading, the users who were momentarily disadvantaged will most likely find themselves being served at some later time when their multipath fading resulted in constructive interference. Figure 1-24 shows the scheduling to two users based on their respective instantaneous fading characteristics, showing how both users are served using a smart scheduler. Because transmission power is used more efficiently, the system throughput of the multiuser diversity smart-scheduling system increases over the case of the simple deterministic scheduler.

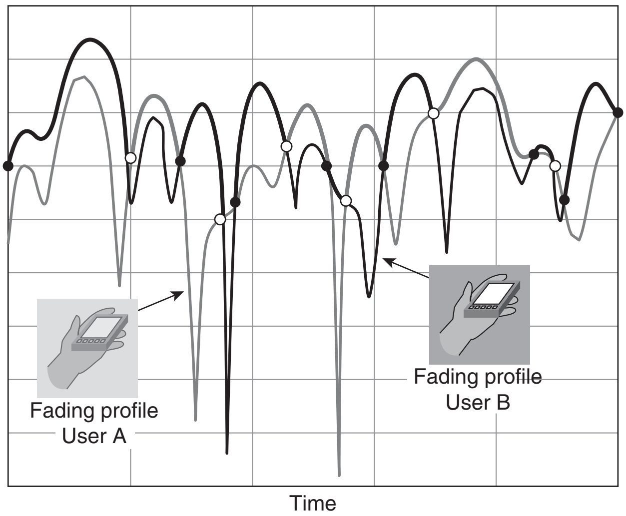

Figure 1-24 *Smart Scheduling Using Multiuser Diversity*

The scheduling algorithm defines how the shared radio resources are divided between the multiple users in a cell. These algorithms can be characterized by the scheduling metric (M_n), which is used to determine which of the n users to allocate radio resources to in the timing interval. If r_n is the instantaneous rate that a user n can support, a *Maximum C/I* scheduler will set the scheduling metric M_n equal to r_n, so as to dedicate resources to the user with the highest r_n value, as demonstrated previously in Figure 1-24. A *Maximum C/I* scheduler ensures that the throughput of the cell is maximized as resources are dedicated to a small subset of users in preferential coverage. Unfortunately, those users who are experiencing less-preferential propagation conditions (for example, those located at the edge of a cell) will be starved of resources.

In order to normalize the scheduling metric and remove the effects of wide-ranging propagation losses across a cell, the Proportional Fair (PF) scheduler has been defined, which uses a scheduling metric M_n equal to $r_n/Avg(R_n)$; R_n is the average data rate for user n. This then equates to scheduling users "at the top of their fades" but also ensures that those users in disadvantaged locations would have their scheduling priority increased over time as $Avg(R_n)$ slowly decreases.

Under optimum conditions and for best-effort traffic, HSDPA simulations have shown that the proportional fair scheduler offers up to 130% capacity improvements over round-robin scheduling. However, as the user speed increases, the fading environment becomes less temporally stable, and so the scheduler is less likely to be able to estimate the instantaneous r_n achievable during the subsequent scheduling instances and the benefits of PF scheduling diminish.

The adoption of smart-scheduling algorithms impact the average delay experienced by the transmitted IP packets. Typically those schedulers that optimize throughput result in degraded delay characteristics. In the extreme example of the Maximum C/I scheduler, certain users will be starved of resources, and the associated packet delay will continue to increase until the user changes location. Consequently, smart schedulers normally define a Discard Timer, which specifies the maximum queuing delay after which the scheduler will drop a packet.

Figure 1-25 illustrates the results of performed simulations[14] of the HSDPA queuing delay per TCP segment in a typical low-speed environment using a Proportional Fair scheduler. The results show that optimization for mobile broadband services requires careful tradeoffs between cell throughput and packet delay characteristics.

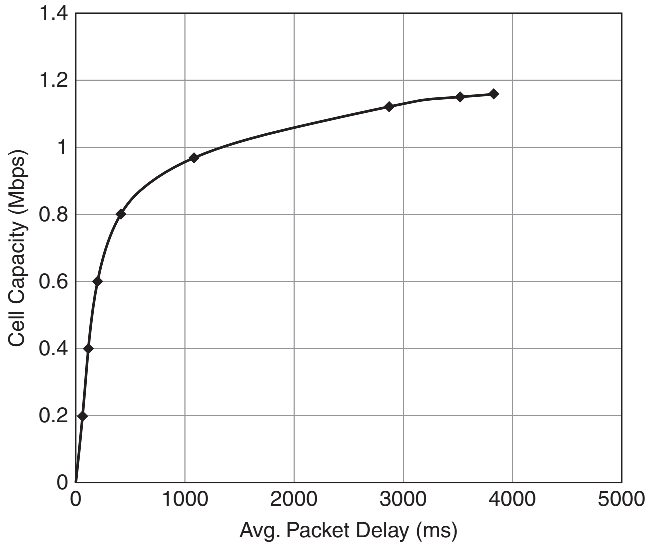

Figure 1-25 *HSDPA Proportional Fair Throughput versus Average Queuing Delay*

Automatic Repeat Request

Because degraded conditions might still generate a residual BER even after FEC decoding, additional techniques need to be applied in situations where error-free transmission is important. In conventional 3G data services, the Radio Link Control (RLC) layer has traditionally been the location where Non-Real-Time (NRT) services could receive error-correction capabilities. In UMTS, the RLC can operate in Acknowledged Mode (AM), which uses Automatic Repeat Request (ARQ) functionality. In RLC-AM operation (the default mode for packet-based services), error-detection capabilities together with a selective repeat operation for error correction are defined.

The centralization of the network-side RLC functionality in traditional mobile radio systems results in substantial retransmission delays between this and the peer RLC functionality in the radio terminal. These delays further degrade the performance of the RLC operation with the mean packet delay increasing as a function of BLock Error Rate (BLER). Typical mean packet delays due to UMTS RLC-AM operation can easily be shown to be in excess of 150 ms,[15] unfortunately leading to interactions with timers integrated into higher-layer protocols such as TCP. In Chapter 9, "Content and Services," we provide more detail of how TCP can be optimized for wireless operation.

To combat the excessive delays of centralized ARQ processing, more recent mobile broadband systems including High-Speed Downlink Packet Access (HSDPA), EVolution Data Only (EV-DO), and WiMAX have all adopted Hybrid-ARQ (H-ARQ), which integrates FEC functionality with the ARQ process.

Note In addition, to reduce the retransmission delays, the newly introduced H-ARQ functionality in these mobile broadband systems is typically located in the base station. Centralized ARQ operation may still operate to recover from H-ARQ failures.

The H-ARQ re-transmissions might be identical to the original transmission, called *chase combining*, or if the receiver has a sufficiently large memory buffer, the retransmission might instead provide additional redundancy/check bits, called *incremental redundancy*. If the original transmission used i information bits and r check bits, the first transmission would have a coding rate of $i/(i+r)$. Now if the re-transmission included r_{add} additional redundancy bits, the overall coding rate following the first re-transmission would be equal to $i/(i+r+r_{add})$ and the receiver would accordingly have a better probability of recovering the original data.

Simulations have shown that in certain conditions, incremental redundancy can decrease the average number of re-transmissions and hence the packet transfer delay compared to chase combining,[16]—for example, reducing the number of re-transmissions by 4 for low carrier-to-interference ratio conditions.

Diversity Combining

Diversity combining is used to combine the multiple signals received over different fading multipath components. The concept holds that if the fading experienced by the different components is independent, the chances of different signals received over diverse paths experiencing deep fades simultaneously is dramatically reduced; if the probability of a single Rayleigh fading signal experiencing a fade in excess of 20 dB is 1%, the probability that simultaneous fades in excess of 20 dB will occur on two independent signals is 0.01%. The diversity in a system is characterized by the number of independently fading diversity paths or branches, also known as the diversity order. There is a diminishing combining gain as the diversity order increases, where the greatest degree of improvement in performance is achieved by going from a single-branch system (no-diversity) to a second-order system with two diversity branches.

Note The Hybrid-ARQ combining techniques described in the previous section are examples of time-diversity. If the different re-transmissions are separated in time by a period greater than T_s, where $Ts = \lambda/2V$ and V is the user's velocity relative to the base station and λ is the wavelength of the carrier frequency, the signals will be de-correlated and diversity gains will be obtained.

Antenna diversity refers to the use of spatially separated antennas for generating the independent signals necessary for delivering diversity combining gains. The antenna separation required to generate uncorrelated signals has been shown to be equal to 0.4 λ at the mobile.[17] At the base station, measurements have indicated that horizontally spaced antennas need to be separated by 10–30λ for the correlation between antennas to be less than 0.7.[18]

Antenna diversity has been primarily used in the uplink on cellular systems. More recently, mobile receive diversity has been added to EV-DO and HSDPA systems, where it has been shown to assist in increasing data rates and sector capacity—for example, increasing the sector throughput from 1.24 Mbps with EV-DO Revision 0 to 1.5 Mbps with 2-Receive forward link diversity.[19]

Spatial Multiplexing

Many of the techniques described in the preceding sections concentrated on attempting to mitigate the degradations generated by the multipath mobile radio channel. *Spatial multiplexing* is a technique that looks to leverage the separate multipath components, using each as an independent channel to enable multiple data streams to be transmitted at the same frequency but over different spatial channels. Because the spatial multiplexing procedure requires multiple receive and transmit antennas to realize the multiplexing gains, the technique is often referred to as *Multiple-Input Multiple-Output (MIMO)*. Spatial multiplexing techniques have been adopted by IEEE 802.11n, IEEE 802.16e/WiMAX, HSDPA, EV-DO, and LTE radio systems.

Figure 1-26 shows an example of a 4x4 MIMO system, so called because of the four transmit and four receive antennas. In general, an array of N_{tx} transmitting antennas and N_{rx} receiving antennas can provide a spatial multiplexing order of SM_{order} = MIN (N_{tx}, N_{rx}), which allows SM_{order} parallel streams to be simultaneously transmitted, effectively increasing the spectral throughput by SM_{order}.

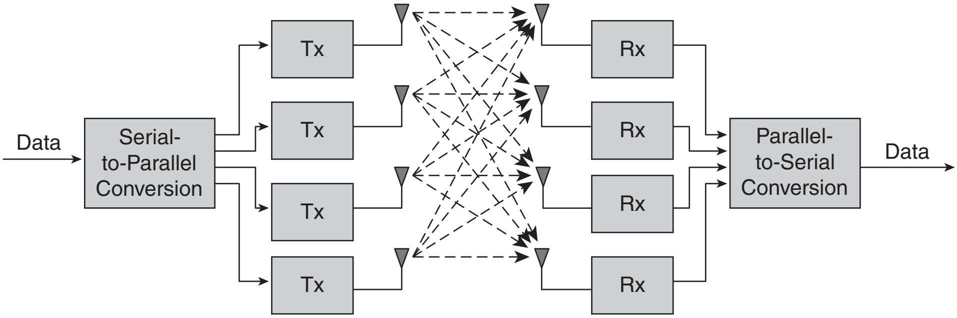

Figure 1-26 *4x4 MIMO Spatial Multiplexing System*

The capacity enhancements of spatial multiplexing requires a certain Signal-to-Interference-and-Noise Ratio (SINR) that enables sufficiently high SINRs on the parallel streams to achieve the overall increase in throughput. Simulations[20] have indicated that the SINR needs to be greater than 12 dB to achieve the gains promised by spatial multiplexing.

Note The very high throughput figures often used in marketing mobile broadband services that rely on MIMO operation will only be available to those users experiencing high SINR, and will thus typically be available to less than 10% of users in a particular cell. This, coupled with the use of AMC, will increase the disparity in mobile broadband services offered to those users in good coverage toward the center of a cell and those users at the edge of cell coverage.

Below these levels, the same MIMO concepts can be re-used to enhance coverage using Space Time Codes (STC), where instead of transmitting different data over parallel streams, a single stream of data is replicated and transmitted over the multiple antennas such that the transmit sequences from each antenna are orthogonal. Figure 1-27 shows an orthogonal Space Time matrix used for a rate-1 STC with two transmit antennas; the figure demonstrates how the matrix is used to map the streams of modulated symbols to the different antennas over time. The gains obtained using Space Time Coding with N_{tx}=2 transmit antennas and N_{rx} receive antennas is equivalent to providing a diversity order of $2N_{rx}$[21].

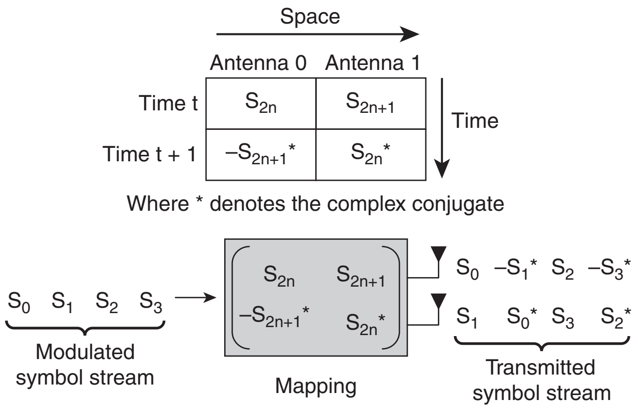

Figure 1-27 *Space Time Coding*

This leads to the definition of systems that use Space Time Coding techniques when the SINR is low to extend the reach of mobile radio systems and spatial multiplexing techniques when the SINR is high in order to increase the spectral efficiency and throughput delivered by mobile radio system.

Note WiMAX has defined its own MIMO nomenclature, specifying the use of 2x2 antenna configurations for both Space Time Coding, which in WiMAX is termed "Matrix-A," and spatial multiplexing, which in WiMAX is termed "Matrix-B."

Summary

The evolution of cellular systems from legacy car phone-centric systems, created to deliver narrow band voice services to outdoor users, toward systems optimized for the economical delivery of high-bandwidth IP services to users located both outdoors and indoors, creates significant challenges for the radio engineer. A level of background is necessary to understand the key compromises faced when trying to design a system, carefully balancing those characteristics that impact the performance of IP services supported by the latest mobile broadband systems.

Compared to legacy systems where service availability could be defined in binary terms (either voice service was available or not), the range of techniques used to increase the peak throughput of mobile broadband systems leads to an increasing disparity between services delivered to different users; the operation of Adaptive Modulation and Coding, Hybrid-ARQ, multiuser diversity, and spatial multiplexing are all techniques that preferentially benefit those users in good coverage conditions compared to those experiencing worse conditions.

The designer of IP services delivered over mobile networks will need to increasing adopt a more probabilistic definitions of service, where the throughput, latency, and jitter characteristics experienced by a mobile broadband user may vary considerably across the cell coverage area and between indoor- and outdoor-located users.

Endnotes

[1] Y. Okumura, E. Ohmori, T. Kawana, and K. Fukuda, "Field Strength and Its Variability in VHF and UHF Land Mobile Service," Review of the Electrical Communications Laboratory, Vol. 16, pp. 825–873, Sept.–Oct. 1968.

[2] M. Hata, "Empirical Formula for Propagation Loss in Land Mobile Radio," IEEE Trans. VT-29, pp. 317–325, August 1980.

[3] http://www.wimaxforum.org/technology/downloads/comparative_analysis_of_spectrum_alternatives_for_wimax_networks_with_scenarios_based_on_the_us_700_mhz_band.pdf.

[4] http://www.agbnielsen.co.nz/Common/Files/Downloads/TAMWise_03_2006_press.pdf.

[5] O.W. Ata, "In-building Penetration Loss Modeling and Measurement in Suburban, Urban, and Dense Urban Morphologies," IEEE Int. Symp. on Antennas and Propagation, pp. 779–782, July 2005.

[6] http://cnfrs.institut-telecom.fr/pages/pages_evenements/journees_scient/docs_journees_2008/GATI-JS08.pdf.

[7] http://www.signalsresearch.com/Mobile%20Broadband%20Wireless%20Aug%2007.pdf.

[8] IEEE Std 802.16-2004, "IEEE Standard for Local and Metropolitan Area Networks—Part 16: Air Interface for Fixed Broadband Access Systems," October 2004.

[9] 3GPP TS 25.214, "Physical Layer Procedures (FDD)."

[10] 3GPP TS 25.321, "Medium Access Control (MAC) Protocol Specification," RTS/TSGR-0225321v5d0.

[11] GSM 43.064, "Overall Description of the GPRS Radio Interface."

[12] A. Ghosh and R.A. Kobylinski, "Advanced Receiver Architectures for HSPA and Their Performance Benefits," Texas Wireless Symposium, pp. 49–53, 2005.

[13] R. Van Nee and R. Prasad, "OFDM for Wireless Multimedia Communications," Artech House, 2000.

[14] P.J.A. Gutierrez, "Packet Scheduling and Quality of Service in HSDPA", Ph.D. thesis, Aalborg University, 2003, http://kom.aau.dk/ADM/research/reports/PhDThesis_Pablo_Ameigeiras.pdf

[15] N. Enderle and X. Lagrange, "Radio Link Control-Acknowledged Mode Protocol Performance Modeling in UMTS," 4th Int. Workshop of Mobile Wireless Communications Network, pp. 332–336, 2002.

[16]Performance Comparison of Chase Combining and Incremental Redundancy in HSDPA, 3GPP TSG RAN1, TSGR1#17(00)1428, www.3gpp1.org/ftp/tsg_ran/ WG1_RL1/ TSGR1_17/Docs/PDFs/R1-00-1428.pdf.

[17]W.C.Y. Lee, "Mobile Communications Engineering," McGraw-Hill, 1982.

[18]L. Aydin, et al, "Reverse Link Capacity and Coverage Improvement for CDMA Cellular Systems Using Polarization and Spatial Diversity," IEEE Int. Conf. on Comms., Vol. 3, pp. 1887–1892, 2002.

[19]"Beyond 3G, Part II (the CDMA2000 & 1xEV-DO evolution)," Signalsahead, Vol. 2, No. 16, August 31, 2005, http://www.signalsresearch.com/Links/ Signals%20Ahead%200831.pdf.

[20]Na Wei, "MIMO Techniques for UTRA Long-Term Evolution," PhD Thesis, Aalborg University, 2007.

[21]S.A. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications," IEEE J. on Sel. Areas in Comms., Vol. 16, No. 8, pp. 1451–1458, October 1998.