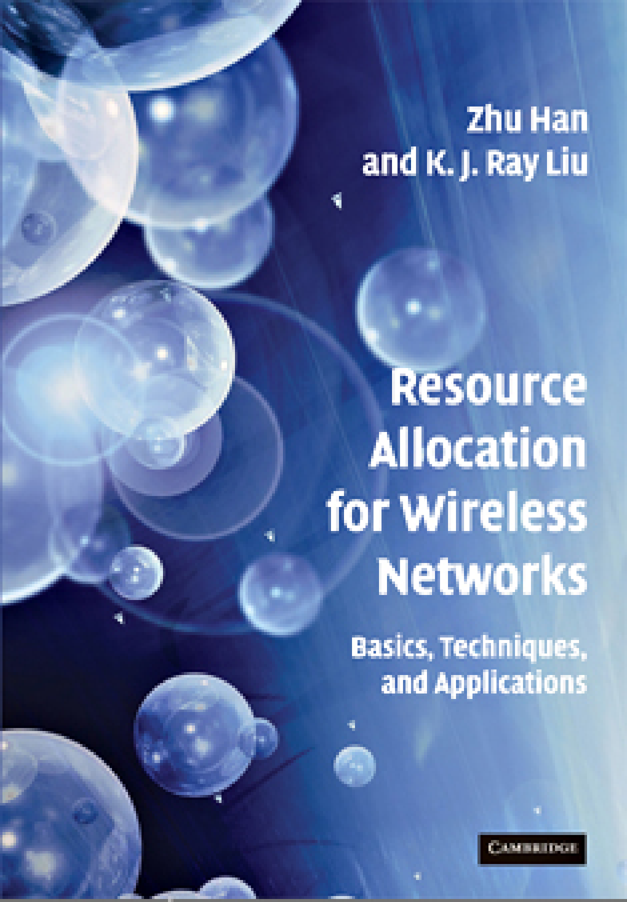


Zhu Han
and K. J. Ray Liu

Resource Allocation for Wireless Networks

Basics, Techniques, and Applications

CAMBRIDGE

CAMBRIDGE www.cambridge.org/9780521873857

This page intentionally left blank

Resource Allocation for Wireless Networks

Merging the fundamental principles of resource allocation with the state of the art in research and application examples, Han and Liu present a novel and comprehensive perspective for improving wireless system performance. Cross-layer multiuser optimization in wireless networks is described systematically. Starting from the basic principles, such as power control and multiple access, coverage moves to the optimization techniques for resource allocation, including formulation and analysis and game theory. Advanced topics, such as dynamic resource allocation and resource allocation in antenna-array processing and in cooperative, sensor, personal-area, and ultrawide-band networks, are then discussed. Unique in its scope, timeliness, and innovative author insights, this invaluable work will help graduate students and researchers understand the basics of wireless resource allocation while highlighting modern research topics and will help industrial engineers improve system optimization.

Zhu Han is currently an assistant professor in the Electrical and Computer Engineering Department at Boise State University, Idaho. In 2003, he was awarded his Ph.D. in electrical engineering from the University of Maryland, College Park. Zhu Han has also worked for a period in industry, as an R & D Engineer for JDSD. Dr. Han is PHY/MAC Symposium Vice Chair of the IEEE Wireless Communications and Networking Conference, 2008.

K. J. Ray Liu is a Distinguished Scholar–Teacher of the University of Maryland, College Park. Dr. Liu is the recipient of numerous honors and awards, including best paper awards from IEEE Signal Processing Society (twice), IEEE Vehicular Technology Society, and EURASIP, as well as recognitions from University of Maryland, including the university-level Invention of the Year Award and the college-level Poole and Kent Company Senior Faculty Teaching Award.

Resource Allocation for Wireless Networks

Basics, Techniques, and Applications

ZHU HAN
Boise State University, Idaho

K. J. RAY LIU
University of Maryland, College Park

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK
Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521873857

First published in print format 2008

ISBN-13 978-0-511-50037-4 eBook (Adobe Reader)

ISBN-13 978-0-521-87385-7 hardback

To my wife as well as my parents and sister in China.
– Zhu Han

To Lynne Liu.
– K. J. Ray Liu

Contents

Preface

Because of fading channels, user mobility, energy/power resources, and many other factors, cross-layer design and multiuser optimization are the keys to ensuring overall system performance of wireless networks. And resource allocation is one of the most important issues for implementing future wireless networks.

In the past decade, we have witnessed significant progress in the advance of resource allocation over wireless networks. It is not only an important research topic, but is also gradually becoming an integral teaching material for graduate-level networking courses.

Yet there are few books available to date that can serve such a purpose. Why? Because the field of resource allocation is such a versatile area that covers a broad range of issues, it is not easy to develop a comprehensive book to cover them all. For instance, resource allocation across various networking layers encounters different design constraints and parameters; different networking scenarios have different performance goals and service objectives; and different formulations of resource allocations need to employ different optimization tools.

To respond to the need of such a book for graduate students, researchers, and engineers, we try to tackle the difficulties by bringing together our research in resource allocation over the past decade and the basic material of resource allocation and optimization techniques to form the foundation of this book. Its intent is to serve either as a textbook for advanced graduate-level courses on networking or as a reference book for self-study by researchers and engineers.

This book covers three main parts. In Part I, the basic principles of resource allocation is discussed. Part II provides the background of optimization tools needed to conduct research and development in resource allocation. And in Part III, examples of advanced topics in resource allocation for different networking scenarios are the focus, to illustrate what one may encounter in different applications.

We would like to thank many for their contributions to some of the material presented in this book, including Thanongsak Himsoon, Ahmed Ibrahim, Zhu Ji, Andres Kwasinski, Xin Liu, Charles Pandana, Farrokh Rashid-Farrokhi, Javad Razavilar, Ahmed Sadek, Wipawee Siriwongpairat, Guan-Ming Su, Weifeng Su, Lee Swindlehurst, Beibei Wang, and Min Wu. We also would like to thank Greg Heinzman for his editing assistance.

Zhu Han
K. J. Ray Liu

1 Introduction

Over the past decade, there has been a significant advance in the design of wireless networks, ranging from physical-layer algorithm development, medium-access control (MAC) layer protocol design, to network- and system-level optimization. Many wireless standards have been proposed to suit the demands of various applications. Over time, researchers have come to the realization that, for wireless networks, because of fading channels, user mobility, energy/power resources, and many other factors, one cannot optimize wireless communication systems as has been traditionally done in wired networks, in which one can simply focus on and optimize each networking layer without paying much attention to the effects of other layers. For wireless networks, cross-layer optimization is a central issue to ensure overall system performance. Yet resource allocation is one of the most important issues for cross-layer optimization of wireless networks.

For instance, across different layers, one cannot design physical-layer coding, modulation, or equalization algorithms by assuming that the MAC layer issues are completely perfect, and vice versa. There are also user diversities—different users at different times and locations may suffer different channel conditions, and therefore may have different demands and capability. Fixing and allocating bandwidths and resources without considering such user diversity can simply waste system resources, and thus performances. In addition, in wireless networks there are space, time, and frequency diversities as well. Taking advantage of those diversities can significantly improve communication performance. All those factors contribute to the need of careful consideration of resource allocations.

We have witnessed the advance of resource allocation in recent years with tremendous progress. As one can imagine, because of the number of degrees of freedom and many different parameters, resource allocation is a broad issue covering a wide range of problems. Therefore, the optimization tools employed also vary a lot. Besides the commonly used convex optimization in communication system design, many resource allocation problems are nonlinear and nonconvex in nature. When it comes to channel allocation and scheduling, sometimes the problems become integer, combinatorial, or both. If one takes into account time-varying conditions, then the problem evolves into one of dynamic optimization. When cooperation among distributed and autonomous users is considered, game theory can be employed to find the optimal strategy and solution. It is fair to say that there is no single optimization tool available to solve all resource-allocation problems at once.

What makes resource allocation more challenging is that, in fact, when it comes to the applications, different wireless networks aim at different service goals, and therefore have different design specifications. One network can be severely energy sensitive and power constrained, whereas the other can be bandwidth limited or throughput hungry. In some situations, a network may have a high degree of mobility with opportunistic access, whereas in other cases a network has an ultrawide bandwidth to share with others but little mobility.

As such, different networks face different resource-allocation problems, different characteristics of problems employ different optimization techniques, and joint considerations of different layers encounter different constrained optimization issues.

This book aims at providing a comprehensive view to answer the preceding challenges in the hope of allowing readers to be able to practice and optimize the allocation of scant wireless resources over assorted wireless network scenarios. Given the nature of the topic, this book is interdisciplinary in that it contains concepts in signal processing, economics, decision theory, optimization, information theory, communications, and networking to address the issues in question. In addition, we try to provide innovative insight into the vertical integration of wireless networks through the consideration of cross-layer optimization.

The goals of this book are for readers to have a basic understanding of wireless resource-allocation problems, to be equipped with an adequate optimization background to conduct research on or design wireless networks, and to be well informed of state-of-the-art research developments. To achieve these goals, this book contains three parts:

- Part I: Basic Principles
  We will study the basic principles of resource allocation for multiple users to share the limited resources in wireless networks under different practical constraints. In addition to the explanation of the basic principles, we will also illustrate the limitations and trade-offs of different approaches.
- Part II: Optimization Techniques
  We will consider various optimization techniques that can be applied to wireless resource-allocation problems. These techniques will be categorized and then compared for their advantages and disadvantages. Some applications for different network scenarios are given as examples.
- Part III: Advanced Topics
  Through the use of some state-of-the-art design examples of different wireless networks, we will illustrate the wide varieties of topics and their potential future design directions. By considering the technical challenges of a variety of networks, we will show how to employ different techniques for different scenarios such as cellular networks, wireless local-area networks, ad hoc/sensor networks, ultrawide-band networks, and collaborative communication networks.

In Part I, the chapters that cover the basics of resource allocation are as follows:

2. Wireless Networks: An Introduction
   In this chapter, we first consider different wireless channel models such as large-scale propagation-loss models and small-scale propagation-loss models. Then, according to

the decreasing order of the coverage areas, we discuss four types of wireless networks: cellular networks, WiMax networks, WiFi networks, and wireless personal-area networks. In the wireless ad hoc networks without a network infrastructure, autonomous users should be able to establish the basic network functions in a distributed way. Wireless sensor networks can detect the events and transmit the information to the data-gathering point with a low consumption of energy. Finally, to cope with the limited spectrum, a cognitive radio can detect the spectrum hole and utilize the unused spectrum.

3. Power Control
   Power control is an effective resource-allocation method to combat fading channel and cochannel interference. The transmitted power is adjusted according to the channel condition so as to maintain the received signal quality. Power control is not a single user's problem, because a user's transmit power causes other users' interferences. We describe the basics of power control first. Then we classify the power-control schemes and discuss the centralized, distributed, and statistical schemes in details. Finally, code-division multiple-access power control is highlighted.
4. Rate Adaptation
   Rate adaptation is one of the most important resource-allocation issues, because the system can adapt the users' rates so that the limited radio resources can be efficiently utilized. In this chapter, we give an overview of the rate-adaptation system. Rate controls over different layers are discussed, such as source rate control, rate control for network/MAC layers, channel-coding rate control, and joint source–channel coding.
5. Multiple Access and Spectrum Access
   The multiple-access scheme is a general strategy to allocate limited resources, such as bandwidth and time, to guarantee the basic quality of services, improve the system performances, and reduce the cost for the network infrastructures. In this chapter, we first study some fixed multiple-access methods such as frequency-division, time-division and code-division multiple access. Then scheduling and random-access protocols are investigated. A third-generation multiple-access system is given as an example. Although multiple access considers the problem of allocating limited radio resources to multiple users, spectrum access decides whether an individual user can access a certain spectrum. We study channel allocation and opportunity spectrum access. Finally, handoff and admission control are illustrated.

In Part II, the chapters cover optimization techniques commonly used in resource allocation:

6. Optimization Formulation and Analysis
   In this chapter, we discuss how to formulate the wireless networking problem as a resource-allocation optimization issue. Specifically, we study what the resources are, what the parameters are, what the practical constraints are, and what the optimized performances across the different layers are. In addition, we address how to perform resource allocation in multiuser scenarios. The trade-offs between the different optimization goals and different users' interests are also investigated. The goal is to provide readers with a new perspective from the optimization point of view for wireless networking and resource-allocation problems.

7. Mathematical Programming
   If the optimization problem is to find the best objective function within a constrained feasible region, such a formulation is sometimes called a mathematical program. Many real-world and theoretical problems can be modeled within this general framework. In this chapter, we discuss the four major subfields of mathematical programming: linear programming, convex programming, nonlinear programming, and dynamic programming. Finally, a wireless resource-allocation example using programming is illustrated.
8. Integer/Combinatorial Optimization
   Discrete optimization is the problem in which the decision variables assume discrete values from a specified set. Combinatorial optimization problems, on the other hand, are problems of choosing the best combination out of all possible combinations. Most combinatorial problems can be formulated as integer programs. In wireless resource allocation, many variables have only integer values, such as the modulation rate, and other variables, such as channel allocation, have a combinatorial nature. Integer optimization is the process of finding one or more best (optimal) solutions in a well-defined discrete problem space. The major difficulty with these problems is that we do not have any optimality conditions to check whether a given (feasible) solution is optimal. We list several possible solutions such as relaxation and decomposition, enumeration, and cutting planes. Finally, a resource-allocation example is formulated and solved as a Knapsack problem.
9. Game Theory
   Game theory is a branch of applied mathematics that uses models to study interactions with formalized incentive structures ("games"). It studies the mathematical models of conflict and cooperation among intelligent and rational decision makers. "Rational" means that each individual's decision-making behavior is consistent with the maximization of subjective expected utility. "Intelligent" means that each individual understands everything about the structure of the situation, including the fact that others are intelligent, rational decision makers. In this chapter, we discuss four different types of games, namely, the noncooperative game, repeated game, cooperative game, and auction theory. The basic concepts are listed, and simple examples are illustrated. The goal is to let the readers understand the basic problems and basic approaches. As a result, we hope the readers can formulate the problems and find solutions in their research areas.

In Part III, we consider some network-aware advanced topics to illustrate the versatility of resource allocation:

10. Resource Allocation with Antenna-Array Processing
    For spatial diversity, transceivers employ antenna arrays and adjust their beam patterns such that they have good channel gain toward the desired directions, while the aggregate interference power is minimized at their output. Antenna-array processing techniques such as beamforming can be applied to receive and transmit multiple signals that are separated in space. Hence, multiple cochannel users can be supported in each cell to increase the capacity by exploring the spatial diversity. We

investigate two examples. First, joint power control, beamforming, and base-station assignment are studied. Second, if the channel information is not available, blind beamforming can be employed to control multiple users' power and beam pattern to achieve the desired link qualities.

11. Dynamic Resource Allocation
    A general strategy to combat detrimental effects, such as fading, is the dynamic allocation of resources such as transmitted power, modulation rates, channel assignment, and scheduling based on the channel conditions. Several design challenges need to be overcome: To optimize radio resource utilization, an important tradeoff exists between system performances and fairness among users. To satisfy the growing demands for heterogeneous applications of wireless networks, it is critical to deliver flexible, variable-rate services with high spectral efficiencies to provide a different quality of service. Finally, if the dynamics of channels is known, each user can calculate the optimal dynamic strategies to maximize the long-term benefits.
12. Game-Theoretic Approaches for Resource Allocation
    Some wireless networks, such as ad hoc networks, consist of a collection of radio transceivers without requiring centralized administration or a prearranged fixed network infrastructure. As a result, ensuring cooperation among selfish users becomes an important issue for designing wireless networks. Game theory is an effective method for analyzing and designing the distributed resource allocation. In this chapter, for noncooperative game theory, we study three examples for power control, multicell orthogonal frequency-division multiple-access channel allocation, and source–relay resource allocation for cooperative communications. For repeated game theory, we study a punishment-based approach for rate control and a self-learning-based approach for packet forwarding. Finally, for cooperative game theory, we use a negotiation-based approach for single-cell orthogonal frequency-division multiple-access resource allocation and opportunistic spectrum access for a cognitive radio.
13. Resource Allocation for Cooperative Networks
    Cooperative communications have gained attention as an emerging transmit strategy for future wireless networks. Cooperative communications efficiently take advantage of the broadcasting nature of wireless networks. The basic idea is that users or nodes in a wireless network share their information and transmit cooperatively as a virtual antenna array, thus providing diversity that can significantly improve system performance. In this chapter, we investigate the impact of cooperative communications on the design of different layers.
14. Ad Hoc/Sensor/Personal-Area Networks
    Over the past few decades, the increasing demands from military, national security, and commercial customers have been driving the large-scale deployment of ad hoc networks, sensor networks, and personal-area networks, which have no sophisticated infrastructures such as base stations. In these scenarios, the mobile users have to set up the network functionality on their own. For ad hoc networks, we investigate the connectivity problem. For sensor networks, we study how to prolong the lifetime.

Finally, for personal-area networks, we employ resource allocation to extend the coverage area.

15. Resource Allocation for Wireless Multimedia
    With the advancement of multimedia compression technology and wide deployment of wireless networks, there is an increasing demand especially for wireless multimedia communication services. To overcome many potential design challenges, dynamic resource allocation is a general strategy used to improve the overall system performance and ensure individual quality of service. Specifically, in this chapter, we consider two aspects of design issues: *cross-layer optimization* and *multiuser diversity*. We study how to optimally transmit multiuser multimedia streams, encoded by current and future multimedia codecs, over resource-limited wireless networks such as third-generation cellular systems, wireless local-area networks, fourth-generation cellular systems, and future wireless local-area networks and wireless metropolitan-area networks.

Part I

Basics Principles

2 Wireless Networks: An Introduction

2.1 Introduction

"Wireless network" refers to a telecommunications network whose interconnections between nodes is implemented without the use of wires. Wireless networks have seen unprecedent growth during the past few decades and will continuously evolve in the future. Seamless mobility and coverage ensure that various types of wireless connections can be made anytime, anywhere. In this chapter, we introduce some basic types of wireless networks and give the readers some preliminary backgrounds for the current state-of-the-art development.

Wireless networks use electromagnetic waves, such as radio waves, for carrying information. Therefore the performance is greatly influenced by randomly fluctuating wireless channels. To understand the channels, in Section 2.2, we will study the existing wireless channel models used for different network scenarios.

There are many existing wireless standards. We consider them according to the order of coverage area, and start with cellular wireless networks. The third-generation (3G) wireless cellular network standards have been enhanced to offer significantly increased performance for data and broadcast services through the introduction of high-speed downlink packet access, enhanced uplink, and multimedia broadcast multicast services. In Section 2.3, we provide an overview of the key elements and technologies. Specifically, we discuss WCDMA, CDMA2000, TD/S CDMA, and 4G and beyond.

WiMax, based on the IEEE 802.16 standard for a wireless metropolitan-area network (WMAN), is expected to enable true broadband speeds over wireless networks at a cost that enables mass-market adoption. WiMAX has the ability to deliver true broadband speeds and help make the vision of pervasive connectivity a reality. We discuss some techniques and the standard in Section 2.4.

A wireless local-area network (WLAN) is a network in which a mobile user can connect to a local-area network (LAN) through a wireless connection. The IEEE 802.11 group of standards specifies the technologies for a WLAN. Based on IEEE 802.11, WiFi is a brand originally licensed by the WiFi Alliance to describe WLAN technology. WiFi provides a low-cost and relatively simple way to gain high-speed access to the Internet. In Section 2.5, we study some specifications in IEEE 802.11 standards.

A wireless personal-area network (WPAN) is a personal-area network (PAN) for wireless interconnecting devices centered around an individual person's workspace. Typically, a WPAN uses a certain technology that permits communication within about

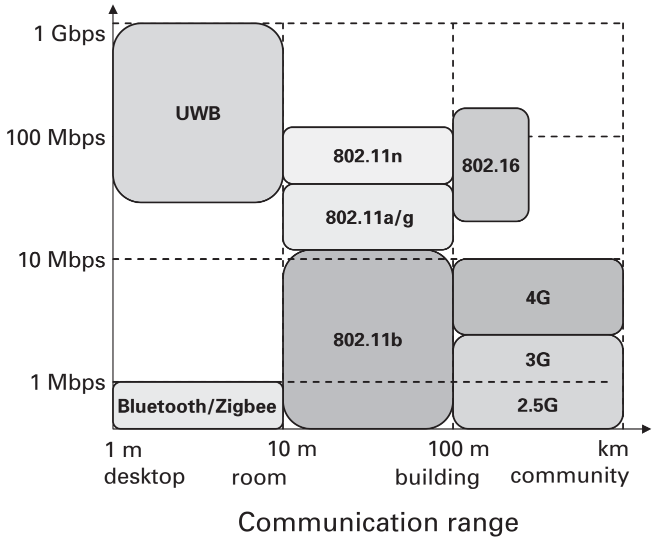


Figure 2.1 Standards comparison.

10 m; in other words, a very short range. IEEE 802.15 standards specify some technologies used in Bluetooth, Zigbee, and Ultra Wide Band. We investigate these technologies in Section 2.6.

We list different standards in Figure 2.1 for different communication rates and different communication ranges. Those standards will fit different needs of various applications, and we also discuss the techniques that can utilize multiple standards in different situations, so that a connection can be made anytime and anywhere.

Finally, in the last three sections in this chapter, we discuss some wireless networks without standards. Specifically, we study wireless ad hoc networks, wireless sensor networks, and cognitive radios, respectively. The motivations for deploying such networks, the design challenges to maintain basic functionality, and recent developments in real implementation are explained in detail.

2.2 Wireless Channel Models

Unlike the wired channels that are stationary and predictable, wireless channels are extremely random and hard to analyze. Models of wireless channels are one of the most difficult challenges for wireless network design. Wireless channel models can be classified as large-scale propagation models and small-scale propagation models, relative to the wavelength.

Large-scale models predict behavior averaged over distances much greater than the wavelength. The models are usually functions of distance and significant environmental features, and roughly frequency independent. The large-scale models are useful for modeling the range of a radio system and rough capacity planning. Some theoretical models (the first four) and experimental models (the rest) are listed as follows:

- Free-Space Model
  Path loss is a measure of attenuation based on only the distance from the transmitter to the receiver. The free-space model is valid only in the far field and only if there is no interference and obstruction. The received power $P_r(d)$ of the free-space model as a function of distance d can be written as

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2 L}, \tag{2.1}$$

  where P_t is the transmit power, G_t is the transmitter antenna gain, G_r is the receiver antenna gain, λ is the wavelength, and L is the system-loss factor not related to propagation. Path-loss models typically define a "close-in" point d_0 and reference other points from the point. The received power in decibel form can be written as

$$P_d(d)\ \text{dBm} = 10 \log \left[\frac{P_r(d_0)}{0.001\text{W}} \right] + 20 \log \left(\frac{d_0}{d} \right). \tag{2.2}$$

- Reflection Model
  Reflection is the change in direction of a wave front at an interface between two dissimilar media so that the wave front returns into the medium from which it originated. A radio propagation wave impinges on an object that is large compared with the wavelength, e.g., the surface of the Earth, buildings, and walls.

  A two-ray model is one of the most important reflection models for wireless channels. An example of reflection and the two-ray model is shown in Figure 2.2. The two-ray model considers a model in which the receiving antenna sees a direct-path signal as well as a signal reflected off the ground. Specular reflection, much like light off of a mirror, is assumed, and, to a very close approximation, the specular reflection arrives with strength equal to that of the direct-path signal. The reflected signal shows up with a delay relative to the direct-path signal and, as a consequence, may add constructively (in phase) or destructively (out of phase). The received power of the two-ray model can be written as

$$P_r = P_t G_t G_r \frac{h_t^2 h_r^2}{d^4}, \tag{2.3}$$

  where h_t and h_r are the transmitter height and receiver height, respectively.

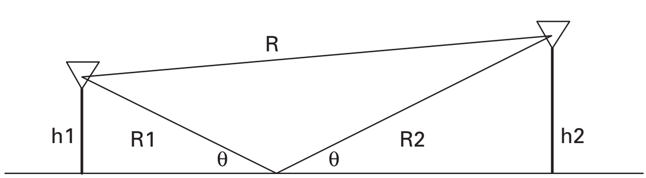


Figure 2.2 Reflection and two-ray model.

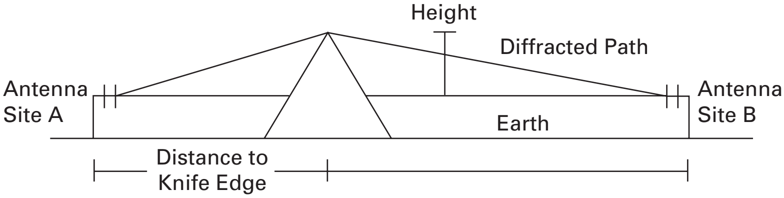


Figure 2.3 Diffraction and knife-edge model.

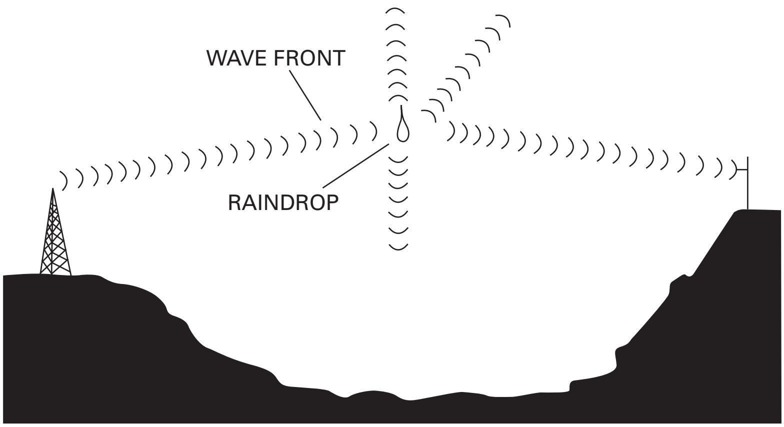


Figure 2.4 Scattering.

- Diffraction Model
  Diffraction occurs when the radio path between transmitter and receiver is obstructed by a surface with sharp irregular edges. Radio waves bend around the obstacle, even when a line of sight (LOS) does not exist. In Figure 2.3, we show a knife-edge diffraction model, in which the radio wave of the diffraction path from the knife edge and the radio wave of the LOS are combined together at the receiver. Similar to the reflection, the radio waves might add constructively or destructively.
- Scattering Model
  Scattering is a general physical process whereby the radio waves are forced to deviate from a straight trajectory by one or more localized nonuniformities in the medium through which it passes. In conventional use, this also includes deviation of reflected radiation from the angle predicted by the law of reflection. The obstructing objects are smaller than the wavelength of the propagation wave, e.g., foliage, street signs, and lamp posts. One scattering example is shown in Figure 2.4.
- Log-Scale Propagation Model and Log-Normal Shadowing Model
  From the experimental measurement, the received signal power decreases logarithmically with distance. However, because of the variety of different factors, the decreasing speed is very random. To characterize the mean and variance of this randomness, the log-scale propagation model and log-normal shadowing model are proposed, respectively.

The log-scale propagation model generalizes path loss to account for other environmental factors. The model chooses a distance d_0 in the far field and measures the path loss PL(d_0). The propagation path-loss factor α indicates the rate at which the path loss increases with the distance. The path loss of the log-scale propagation model is given by

$$\mathrm{PL}(d)\ (\mathrm{dB}) = \mathrm{PL}(d_0) + 10\alpha \log\left(\frac{d}{d_0}\right). \tag{2.4}$$

Shadowing occurs when objects block the LOS between the transmitter and the receiver. A simple statistical model can account for unpredictable "shadowing" as

$$\mathrm{PL}(d)\ (\mathrm{dB}) = \mathrm{PL}(d) + X_0, \tag{2.5}$$

where X_0 is a zero-mean Gaussian random variable with variance typically from 3 to 12. The propagation factor and the variance of log-normal shadowing are usually determined by experimental measurement.

- Outdoor Propagation Models
  In outdoor models, the terrain profile of a particular area needs to be taken into account for estimating the path loss. Most of the following models are based on a systematic interpretation of measurement data obtained in the service area. Some typical outdoor propagation models are the Longley–Rice model, ITU terrain model, Durkins model, Okumura model, Hatas model, PCS extension of the Hata model, Walfisch and Bertoni model, and wideband PCS microcell model.
- Indoor Propagation Models
  For indoor applications, the distances are much smaller than those of the outdoor models. The variability of the environment is much greater, and key variables are the layout of the building, construction materials, building type, and antenna location. In general, indoor channels may be classified either as LOS or obstruction with a varying degree of clutter. The losses between floors of a building are determined by the external dimensions and materials of the building, as well as the type of construction used to create the floors and the external surroundings. Some of the available indoor propagation models are the Ericsson multiple-breakpoint model, ITU model for indoor attenuation, log-distance path-loss model, attenuation-factor model, and Devasirvathamòs model.

Small-scale (fading) models describe signal variability on a scale of wavelength. In fading, multipath effects and Doppler effects dominate. The fading is frequency dependent and time variant. The focus is on modeling "fading," which is the rapid change in signal over a short distance or length of time.

Multipath fading is caused by interference between two or more versions of the transmitted signal, which arrive at slightly different times. Multipath fading causes rapid changes in signal strength over a small travel distance or time interval, random frequency modulation that is due to varying Doppler shifts on different multipath signals, and time dispersion caused by multipath propagation delays.

To measure the time dispersion of multiple paths, the power-delay profile and root mean square (RMS) are the most important parameters. Power-delay profiles are generally represented as plots of relative received power as functions of excess delay with respect to a fixed time-delay reference. The mean excess delay is the first moment of the power-delay profile and is defined as

$$\bar{\tau} = \frac{\sum_k a_k^2 \tau_k}{\sum_k a_k^2}, \tag{2.6}$$

where τ_k is the delay of the kth multipath and a_k is its corresponding amplitude. The RMS is the square root of the second central moment of the power-delay profile:

$$\sigma_\tau = \sqrt{\bar{\tau}^2 - (\tau)^2}, \tag{2.7}$$

where

$$\bar{\tau}^2 = \frac{\sum_k a_k^2 \tau_k^2}{\sum_k a_k^2}. \tag{2.8}$$

Typical values of RMS delay spread are of the order of microseconds in outdoor mobile radio channels and of the order of nanoseconds in indoor radio channels.

Analogous to the delay spread parameters in the time domain, coherent bandwidth is used to characterize the channel in the frequency domain. Coherent bandwidth is the range of frequencies over which two frequency components have a strong potential for amplitude correlation. If the frequency correlation between two multipaths is above 0.9, then the coherent bandwidth is

$$B_c = \frac{1}{50\sigma}. \tag{2.9}$$

If the correlation is above 0.5,

$$B_c = \frac{1}{5\sigma}. \tag{2.10}$$

Coherent bandwidth is a statistical measure of the range of frequencies over which the channel can be considered flat.

Delay spread and coherent bandwidth describe the time-dispersive nature of the channel in a local area. But they do not offer information about the time-varying nature of the channel caused by relative motion of transmitter and receiver. Next we define Doppler spread and coherence time, which describe the time-varying nature of the channel in a small-scale region.

Doppler frequency shift is due to the movement of the mobile users. Frequency shift is positive when a mobile moves toward the source; otherwise, the frequency shift is negative. In a multipath environment, the frequency shift for each ray may be different, leading to a spread of received frequencies. Doppler spread is defined as the maximum Doppler shift,

$$f_m = \frac{v}{\lambda}, \tag{2.11}$$

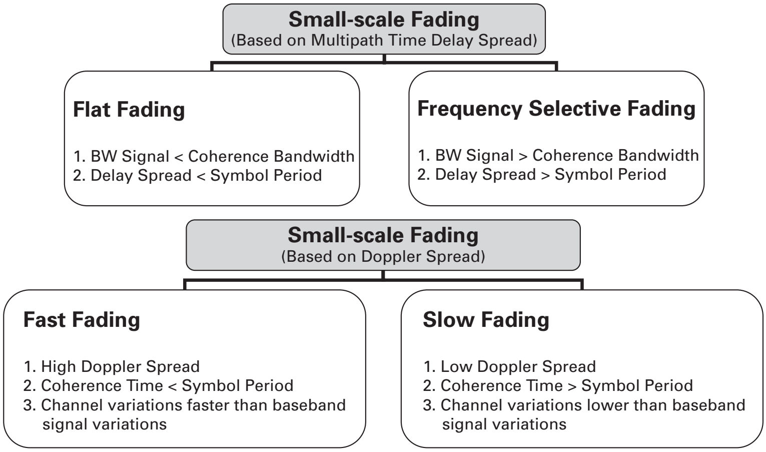


Figure 2.5 Classification of small fading.

where v is the mobile user's speed and λ is the wavelength. If we assume that signals arrive from all angles in the horizontal plane, the Doppler spectrum can be modeled as Clarke's model [247].

Coherence time is the time duration over which the channel impulse response is essentially invariant. Coherence time is defined as

$$T_c = \frac{C}{f_m}, \tag{2.12}$$

where C is a constant. The coherence time definition implies that two signals arriving with a time separation greater than T_C are affected differently by the channel. If the symbol period of the baseband signal (reciprocal of the baseband signal bandwidth) is greater than the coherence time, then the signal will distort, because the channel will change during the transmission of the signal.

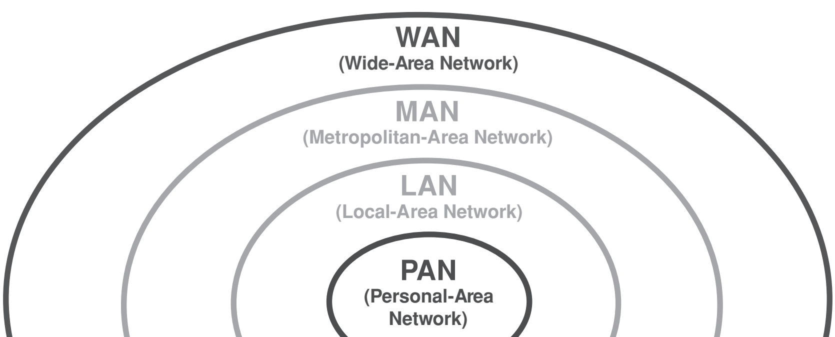
Based on the transmit signal's bandwidth and symbol period relatively to the multipath RMS and coherent bandwidth, small-scale fading can be classified as flat fading and frequency-selective fading. The classification means a bandlimited transmit signal sees a flat frequency channel or a frequency-selective channel. Based on coherence time that is due to Doppler spread, small-scale fading can be classified as fast fading and slow fading. The classification means that, during each signal symbol, the channel changes or does not change. The details are shown in Figure 2.5.

Multipath and Doppler effects describe the time and frequency characteristics of wireless channels. But further analysis is necessary for statistical characterization of the amplitudes. Rayleigh distributions describe the received signal envelope distribution for channels, in which all the components are non-LOS. Ricean distributions describe the received signal envelope distributions for channels in which one of the multipath components is the LOS component. Nakagami distributions are used to model dense scatters. Nakagami distributions can be reduced to Rayleigh distributions, but they give more control over the extent of the fading.

2.3 3G Cellular Networks and Beyond

Third-generation (3G) mobile communication systems based on wideband code-division multiple-access (WCDMA) and CDMA2000 (CDMA stands for code-division multiple-access) radio-access technologies have seen widespread deployment around the world. There are more than 160 3G systems in commercial operation in 75 countries with a total of more than 230 million 3G subscribers as of December 2005. The applications supported by these commercial systems range from circuit-switched services such as voice and video telephony to packet-switched services such as videostreaming, e-mail, and file transfer. As more packet-based applications are invented and put into service, the need increases for better support for different quality of service (QoS), higher spectral efficiency, and higher data rates for packet-switched services, in order to further enhance user experience while maintaining efficient use of system resources. This need has resulted in the evolution of 3G standards, as shown in Figure 2.6. For 3G cellular systems, there are two camps: the 3G Partnership Project (3GPP) [386] and the 3G Partnership Project 2 (3GPP2) [387], which is based on different second-generation (2G) technologies.

The development of 3G will follow a few key trends, and the evolution following these trends will continue as long as the physical limitations or backward compatibility



	PAN	LAN	MAN	WAN
Standards	Bluetooth/UWB 802.15.3	802.11	802.11 802.16 802.20	GSM, CDMA, Satellite
Speed	< 1 Mbps or < 480 Mbps	11 to 54 Mbps	10–100+ Mbps	10 Kbps–2 Mbps
Range	Short	Medium	Medium-Long	Long
Applications	Peer-to-Peer Device-to-Device	Enterprise Networks	Last Mile Access	Mobile Data Devices

Figure 2.6 Comparison of different wireless networks.

Table 2.1 Comparison of 3G standards

	CDMA2000	WCDMA	TD-SCDMA
Carrier bandwidth	1.25/3.75 MHz	5 MHz	1.6 MHz
Multiple access	DS/MC-CDMA	DS-CDMA	TDMA/DS-CDMA
Chip rate	1.2288/3.6864M chips/s	3.84M chips/s	1.28M chips/s
Frequency reuse	1	1	1
Channel coding	Convol./turbo code	Convol./turbo code	Convol./turbo code
Spreading code	Walsh, pseudo noise	OVSF	OVSF
Spreading factor	4-256	4-256	1,2,4,8,16
Data modulation	DL:QPSK;UL:BPSK	DL:QPSK;UL:BPSK	QPSK, 8-PSK
Frame length	5 ms, 20 ms	5 ms, 20 ms	10 ms
No. of slots/frame	16	16	7
Max. data rate	2.4 Mbps	2 Mbps	2 Mbps
Spectrum efficiency	1.0	0.4	1.25
Power control	Open/close 800 Hz	Open/close 1600 Hz	open/close 200 Hz
Receiver	Rake	Rake	MUD, Rake
Inter-BS timing	GPS synchronous	Asynch./synch.	Synchronous

requirements do not force the development to move from evolution to revolution. The key trends include the following:

- Voice services will also stay important in the foreseeable future, which means that capacity optimization for voice services will continue.
- Along with increasing use of Internet-protocol-based (IP-based) applications, the importance of data as well as simultaneous voice and data will increase.
- Increased need for data means that the efficiency of data services needs to be improved.
- When more and more attractive multimedia terminals emerge in the markets, the usage of such terminals will spread from office, homes, and airports to roads, and finally everywhere. This means that high-quality high-data-rate applications will be needed everywhere as well.
- When the volume of data increases, the cost per transmitted bit needs to decrease in order to make new services and applications affordable for everybody. The other current trend is that in the 3G evolution path very high data rates are achieved in hot spots with WLAN rather than with cellular-based standards.

In Table 2.1, we compare some of the technical parameters for the major 3G standards. In the following subsections, we discuss these standards in detail.

2.3.1 CDMA2000

The CDMA family of cellular networks grew out of work undertaken by Qualcomm, a California-based company. Working on direct-sequence spread-spectrum (DSSS) techniques, by using different spreading codes, a large number of users could occupy the

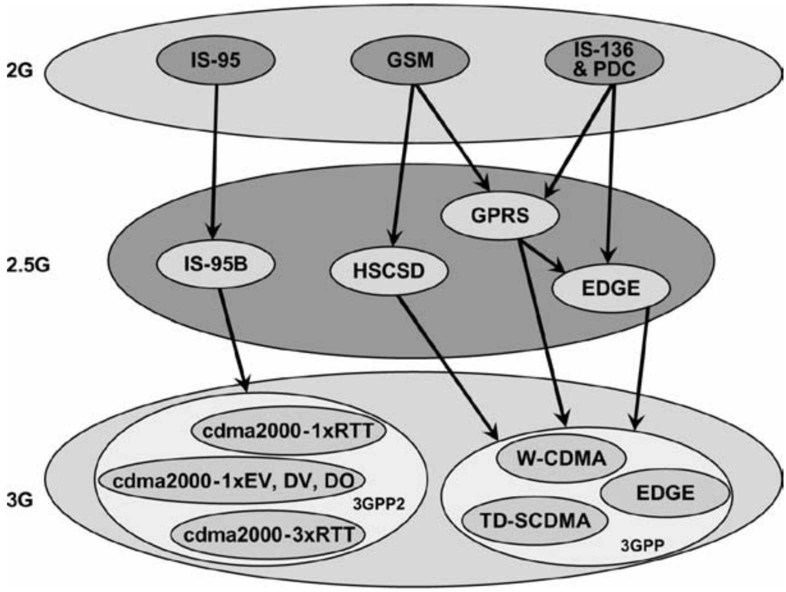


Figure 2.7 Evolution of 2G to 3G cellular networks [247].

same channel at the same time, which could provide a multiple-access scheme for cellular telecommunications. The first standard was the IS-95, and first network was launched in Hong Kong in 1996 under the brand name CDMAOne.

The CDMA system also has the following standards in its developmental stages, as shown in Figure 2.7: IS-95, IS-95A, IS-95B, CDMA2000(1x/EV-DO, 1xEV-DV, 1xRTT and 3xRTT). The first version of system standard IS-95 has never been launched for commercial purposes because of its prematurity. The IS-95A was originally applied for business and is still used widely nowadays. IS-95B was a short version because CDMA2000 standard was announced six months after it came out. The original IS-95A standard allowed for only circuit-switched data at 14.4 kbps, and IS-95B provided up to 64-kbps data rates as well as a number of additional services. A major improvement came later with the development of 3G services. The first 3G standard was known as CDMA2000 1x, which initially provided data rates up to 144 kbps. With further developments, the systems promise to allow a maximum data rate of 307 kbps.

In terms of the technology used, CDMAOne used a bandwidth of a 1.25-MHz channel so as to fit in with existing band plans and channel allocations. When using the system, different users are allocated different Walsh codes, the orthogonal spreading codes. Quadrature-phase-shift keying (QPSK) was used as a modulation form on the forward channel and offset QPSK (OQPSK) on the reverse channel. For CDMA2000, the Walsh code length was changed from 64 bits to 128 and new channels were introduced. Additionally, turbocodes were initially used for error correction.

The packet data MAC functions for IS-95 have only two states, active and dormant, as previously described. This simple approach works well for fairly low-speed data services with relatively low occupancy for any given user. However, this MAC model is inadequate to meet the aggressive requirements for very-high-speed data services, compared with many competing users in 3G systems. This is due to the excessive interference caused by idle users in the active state and the relatively long time and high system overhead required for transitioning from the dormant to the active state. To address these requirements, the CDMA2000 system incorporates a sophisticated MAC mechanism that includes two intermediate states (control hold state and suspended state) between the IS-95 active and dormant states.

To be backward compatible with IS-95 networks, the CDMA2000 radio interface retains many of the attributes of the IS-95 air interface design. In IS-95-B, higher data rates are provided through code aggregation. In CDMA2000, higher rates are achieved through either reduced spreading or multiple-code channels. In addition, there are a number of major enhancements in the CDMA2000 physical layer that facilitate advanced data services with higher rates and improved capacity. As a result, the upgrade cost of IS-95 to CDMA2000 is relatively low.

The designation “1xRTT” (1 times radio transmission technology) is used to identify the version of CDMA2000 radio technology that operates in a pair of 1.25-MHz radio channels (1×1.25 MHz, as opposed to 3×1.25 MHz in 3xRTT). 1xRTT almost doubles voice capacity over IS-95 networks. Although capable of higher data rates, most deployments limit the peak data rate to 144 kbps. Although 1xRTT officially qualifies as 3G technology, 1xRTT is considered by some to be a 2.5G (or sometimes 2.75G) technology. This has allowed it to be deployed in a 2G spectrum in some countries that limit 3G systems to certain bands. At this point, the development of CDMA2000 diverged. One development known as CDMA2000 1xEV-DO (data only) provided a data-only or data-optimized evolution, whereas another development known as CDMA2000 1xEV-DV (data and voice) provided data and voice evolutions.

- EV-DO

  CDMA2000 1xEV-DO was originally not on the development roadmap for CDMA2000. As a result, EV-DO was defined under IS-856 rather than under IS-2000 and carries data only at broadband speed. The first commercial CDMA2000 1xEV-DO network was deployed by a Korea-based company, SK Telecom, in January, 2002. The system is becoming more widespread with 30 networks live across Asia, America, and Europe and 37 more networks deployed in 2006.

  With CDMA2000 1x and EV-DO release 0 well established, further development is looking to provide greater data speeds, better spectral efficiency, and improved network manageability. The evolution continues through EV-DO. It aggregates multiple channels for higher peak rates and supports up to 20-MHz bandwidth with peak rates increasing linearly. In this way, CDMA2000 1xEV-DO would be able to compete with WLANs while maintaining the mobility.
- EV-DV

  The original roadmap for the development of CDMA2000 1x was to adopt the data and voice system, but it is unlikely to be deployed, although it could be capable of providing data at speeds up to 3.09 Mbits/s in the forward direction. To meet the requirements of 1xEV-DV, there are a number of new features to be implemented. These included the addition of new channels, an adaptive modulation and coding scheme, the addition of an automatic repeat-request (ARQ) to the physical layer, and cell switching. 1xEV-DV has been developed to exploit the delay tolerance and diversity of multiuser packet data traffic via numerous air interface innovations. It has been developed in a cooperative fashion with the support and effort of the companies involved in the 3GPP2.

2.3.2 WCDMA/UMTS

WCDMA was developed by NTT DoCoMo as the air interface for their 3G network, FOMA. Later NTT DoCoMo submitted the specification to the International Telecommunication Union (ITU) as a candidate for the international 3G standard known as IMT-2000. The ITU eventually accepted WCDMA as part of the IMT-2000 family of 3G standards, as an alternative to CDMA2000, EDGE (enhanced data rate for GSM evolution), and the short-range, Digital Enhanced Cordless Telecommunication (DECT) system. Later, WCDMA was selected as the air interface for the Universal Mobile Telecommunications System (UMTS), the 3G successor to the GSM (Global System for Mobile Communication). During the evolution from 2G to 3G for WCDMA, there are some 2.5G technologies (shown in Figure 2.7) as follows:

- HSCSD: High-speed circuit-switched data (HSCSD) is an enhancement to circuit-switched data, the original data transmission mechanism of the GSM mobile phone system. One innovation in HSCSD is to allow different error-correction methods to be used for data transfer. The other innovation in HSCSD is the ability to use multiple time slots at the same time. Using the maximum of four time slots, this can provide an increase in the maximum transfer rate of up to 57.6 kbps (4×14.4 kbps).
- GPRS: General packet radio service (GPRS) is a mobile data service available to users of GSM and IS-136 mobile phones. GPRS data transfer is typically charged per megabyte of transferred data, whereas data communication via traditional circuit switching is billed per minute of connection time, independently of whether the user actually has transferred data or has been in an idle state. 2G cellular systems combined with GPRS are often described as "2.5G," that is, a technology between the second generation (2G) and third generation (3G) of mobile telephony. It provides moderate-speed data transfer by using unused time-division multiple-access (TDMA) channels in, for example, the GSM system.
- EDGE: Enhanced data rates for GSM evolution (EDGE), or enhanced GPRS (EGPRS), is a digital mobile phone technology that allows for increased data transmission rate and improved data transmission reliability. It is generally classified as a 2.75G network technology. EDGE has been introduced into GSM networks around the world since 2003, initially in North America. It can be used for any packet-switched applications such as an Internet connection. High-speed data applications such as video services and other multimedia benefit from EGPRS's increased data capacity. Circuit-switched EDGE is a possible future development. The highest rate is 384 kbps.

The rapid widespread deployment of WCDMA and an increasing uptake of 3G services are raising expectations with regard to new services such as web surfing and file transfer. Release 6 of WCDMA brought support for broadcast services through multimedia broadcast multicast services (MBMS), enabling applications such as mobile TV. WCDMA has been evolving to meet the increasing demands for high-speed data access and broadcast services. These two types of services have different characteristics, which influence the design of the enhancements. For high-speed data access, data typically arrive in bursts, posing rapidly varying requirements on the amount of radio resources required. The transmission is typically bidirectional, and low delays are required for a good end-user experience. As the data are intended for a single user, feedback can be used to optimize the transmission parameters. Broadcast/multicast services carry data intended for multiple users. Consequently, user-specific adaptation of the transmission parameters is cumbersome, and diversity not requiring feedback is crucial. Because of the unidirectional nature of broadcasted data, the low delay for transmission is not as important as for high-speed data access.

Release 5 of WCDMA introduced improved support for downlink packet data, often referred to as high-speed downlink packet access (HSDPA). In release 6, finalized early in 2005, the packet data capabilities in the uplink (enhanced uplink) were improved. In WCDMA, the shared downlink resource consists of transmission power and channelization codes in node B (the base station), whereas in the uplink the shared radio resource is the interference at the base station. A key characteristic of HSDPA is the use of shared-channel transmission. This implies that a certain fraction of the total downlink radio resources available within a cell, channelization codes and transmission power, is seen as a common resource that is dynamically shared between users, primarily in the time domain. The use of shared-channel transmission in WCDMA implemented through the high-speed downlink shared channel (HSDSCH), enables the possibility of rapidly allocating a large amount of the downlink resources to a user when needed. Fast scheduling is used to control allocation of the shared resource among users on a rapid basis. Additionally, a fast hybrid ARQ with soft combining enables fast retransmission of erroneous data packets. To meet the requirement on low delays and rapid resource (re)allocation, the corresponding functionality must be located close to the air interface. In WCDMA this has been solved by locating the enhancements in the base station as part of additions to the MAC layer.

- Link Adaptation
  Link adaptation is implemented by adjusting the channel-coding rate, and selecting between QPSK and 16 quadrature amplitude modulation (QAM). Higher-order modulation, such as 16 QAM, makes more efficient use of bandwidth than QPSK, but requires greater received Eb/N0. Consequently, 16 QAM is mainly useful in advantageous channel conditions. In addition, the data rate also depends on the number of channelization codes assigned for HSDSCH transmission in a transmission time interval (TTI). The data rate is selected independently for each 2-ms TTI by node B, and the link-adaptation mechanism can therefore track rapid channel variations.

- Scheduling
  The scheduler is a key element and to a large extent determines the overall downlink performance, especially in a highly loaded network. A practical scheduler strategy exploits the short-term variations (e.g., there that are due to multipath fading and fast interference variations) while maintaining some degree of long-term fairness between the users.
- Hybrid ARQ
  The third key feature of HSDPA is a hybrid ARQ with soft combining, which allows the terminal to rapidly request retransmission of erroneously received transport blocks, essentially fine-tuning the effective code rate and compensating for errors made by the link-adaptation mechanism. The terminal attempts to decode each transport block it receives and reports to node B its success or failure 5 ms after the reception of the transport block.

2.3.3 TD-SCDMA

Transmit diversity (TD) is one of the key contributing technologies to defining the ITU-endorsed 3G systems WCDMA and CDMA2000. Spatial diversity is introduced into the signal by transmitting through multiple antennas. The antennas are spaced far enough apart that the signals emanating from them can be assumed to undergo independent fading. In addition to diversity gain, antenna gain can also be incorporated through channel-state feedback. This leads to the categorization of TD methods into open-loop and closed-loop methods. Several methods of transmit diversity in the forward link have been either under consideration or adopted for the various 3G standards.

China has fully embraced the remarkable growth and unprecedented penetration of mobile services, and has become the world's largest mobile cellular market. TD-SCDMA was proposed by the China Wireless Technology Standard (CWTS) Group in 1998, approved as one of the 3G standards by ITU in May 2000, and joined 3GPP in March 2001. China puts a major effort into advancing its mobile communication systems and facilitating its own technological development in this critical area. TD-SCDMA, a combination of TDD and synchronous CDMA, offers several unique advantages over its alternatives, WCDMA and CDMA2000, such as flexible spectrum allocation, low-cost implementation, and easier migration from GSMs.

For TD-SCDMA, the channel includes three carriers using a low-chip-rate mode of 1.28 Mchips/s that corresponds to a carrier bandwidth of 1.6 MHz. This helps provide high flexibility in spectrum usage and network design, especially in densely populated areas. In addition, each TDMA frame of 5-ms duration is divided into seven time slots, which can be flexibly assigned to either multiple users or a single user that might require multiple time slots. In addition to the TDMA/TDD principle, TD-SCDMA uses a different CDMA mode from that of CDMA2000 and WCDMA systems, in which TD-SCDMA limits the number of codes for each time slot to a maximum of 16. This helps to reduce multiple-access interference (MAI) and to increase system capacity. Relying on a combination of TDD and synchronous CDMA, TD-SCDMA offers a number of attractive features, including unpaired frequencies, suitability for IP services, and capability to support asymmetric services in uplink/downlink. In addition,

Table 2.2 Comparison of 3G and 4G

Parameter	3G	4G
Major requirement driving architecture	Voice-driven data add on	Data/voice over IP
Network architecture	Wide-area cell based	Hybrid with WiFi and WPAN
Speed	384 kbps–2 Mbps	20–10 Mbps
Frequency band	1.8–2.4 GHz	2–8 GHz
Bandwidth	1.25, 5, or 20 MHz	100 MHz
Switching design	Circuit and packet	Packet
Access	DS-CDMA	OFDM/MC-CDMA
FEC	Convolution/turbo code	Concatenated coding
Component design	Antenna, multiband adapter	Smart antennas, software radios

TD-SCDMA systems also incorporate some new or unique technologies such as joint detection, adaptive antenna-array processing, dynamic channel allocation, and baton handover.

2.3.4 4G and Beyond

Looking at developments in the Internet and applications, we clearly see that the complexity of the transferred content is rapidly increasing and will increase further in the future. Generally it can be said that the more bandwidth that is available, the more bandwidth applications will consume. To justify the need for a new air interface, goals need to be set high enough to ensure that the system will be able to serve us long into the future. A reasonable approach would be to aim at 100-Mbps full-mobility wide-area coverage and 1-Gbps low-mobility local-area coverage with a next-generation cellular system in about 2010 in standards fora. Also, the future application and service requirements will bring new requirements to the air interface and new emphasis on air interface design. One such issue, which already has had a strong impact on 3G revolution is the need to support IP and IP-based multimedia. If both technology and spectra to meet such requirements cannot be found, the whole discussion of 4G may become obsolete. In Table 2.2, we compare key parameters of 4G with those of 3G.

For 4G standards, it is worth mentioning Flash-OFDM (fast low-latency access with seamless handoff orthogonal frequency-division multiplexing), a system that is based on OFDM and also specifies higher protocol layers. It was developed and is marketed by Flarion, which was acquired by Qualcomm. Flash-OFDM has generated interest as a packet-switched cellular bearer. Flash-OFDM competes with GSM and 3G networks. Flarion system is the first truly IP-based broadband cellular network designed for data, and it outperforms 3G in all critical areas of performance. For example, the system is capable of sustaining 12 Mbps of throughput per cell in a three-carrier three-sector configuration and has peak user data rates up to 3 Mbps, full cellular mobility, less than 20 ms of latency, and full QoS.

Next-generation wireless involves the concept of a major move toward ubiquitous wireless communications systems and seamless high-quality wireless services. 4G mobile communications involve a mix of concepts and technologies in the making.

Some can be recognized as derived from 3G and are called evolutionary (e.g., evolutions of WCDMA and CDMA2000), whereas others involve new approaches to wireless mobile and are sometimes labeled revolutionary, like OFDM/ WCDMA. What is important, though, is the common understanding that technologies beyond 3G are of fundamental relevance in the movement toward a new wireless world that is a total convergence of wireless mobile and wireless access communications. Any of these terms are meant to signify fundamentally better wireless mobile communications in the future.

2.4 WiMAX Networks

Wireless metropolitan-area network (WMAN) technology is a relatively new field that was started in 1998. From that time a new standard has emerged to handle the implementation, IEEE 802.16. The equivalent of 802.16 in Europe is HIPERMAN. The WiMAX Forum is working to ensure that 802.16 and HIPERMAN interoperate seamlessly. This standard has helped to pave the way for WMAN technology globally and since its first inception has now received six expansions onto the standards. WMAN differs from other wireless technologies in that it is designed for a broader audience, such as a large corporation or an entire city. The overall scheme is illustrated in Figure 2.8.

There are two main applications of WiMAX today: Fixed WiMAX applications are point-to-multipoint enabling broadband access to homes and businesses, whereas mobile WiMAX offers the full mobility of cellular networks at true broadband speeds. Both fixed and mobile applications of WiMAX are engineered to help deliver ubiquitous, high-throughput broadband wireless services at a low cost.

Next, we break up how WMAN technology works by frequency range (top to bottom) and available options in each range.

- 10GHz-66GHz range:
  In this range, a "WirelessMAN-SC" technique is used, which employs a single carrier for modulation. For allowing multiple users onto the network, WMAN uses either time-division duplexing (TDD) or frequency-division duplexing (FDD). TDD allows for variable asymmetry in both uplink and downlink connections and acts just like

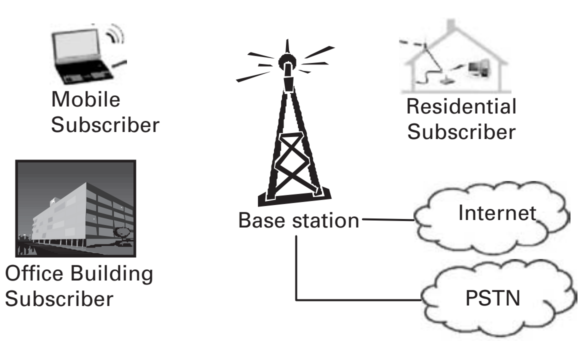


Figure 2.8 Basic implementation of a WMAN from the original IEEE 802.16 standard [372].

time-division multiplexing. TDD allows two users (in this case receiver/transmitter) with a two-way connection on a single frequency. FDD places the uplink and downlink signals on separate subbands and works best for symmetrical traffic (i.e., when the amount of uplink and downlink traffic is the same). For the base station, time-division multiplexing is utilized and for uplink, TDMA is employed.

- 2GHz-11GHz band:
  For the licensed 2–11-GHz band, there are three different ways for a WMAN to work. The first way is "WirelessMAN-SCa," which works the same way as it did in the 10–66-GHz range. The second way is "WirelessMAN-OFDM," which employs an OFDM method. This method is similar to a multiple-carrier modulation, in which it divides a high bit stream into several low bit streams across subcarriers that are orthogonal to each other. The last way is "WirelessMAN-OFDMA." OFDMA is a multiple-user version of OFDM, in which the subcarriers are broken down even further into subsets, and each subset is the representative of a different user.

  Another option when using the licensed 2–11-GHz band is the kind of acknowledgement system to use. For example, one of the acknowledgement systems that can be used in the 2–11-GHz band is ARQ. ARQ is an automatic repeat-request algorithm that has a transmitter send a packet. If the receiver retrieves the data with no error, it will send back an acknowledge (ACK) message saying it is ready for the next packet. If it sends a no-acknowledge (NACK) message or does not send a message, the transmitter will resend the data again.

  In the 2–11-GHz license-exempt band, WMAN employs the "WirelessHUMAN" method. The WirelessHUMAN (wireless high-speed unlicensed metropolitan-area network) method employs the same OFDM method that was previously discussed for WirelessMAN-OFDM.

The 802.16 MAC uses a scheduling algorithm for which the subscriber station needs to compete once (for initial entry into the network). After the competition, the subscriber station is allocated an access slot by the base station. The time slot can enlarge and contract, but remains assigned to the subscriber station, which means that other subscribers cannot use it. The 802.16 scheduling algorithm is stable under overload and oversubscription (unlike 802.11). It can also be more bandwidth efficient. The scheduling algorithm also allows the base station to control QoS parameters by balancing the time-slot assignments among the application needs of the subscriber stations. Moreover, the MAC layer is also in charge of protocol data unit (PDU) assembly and disassembly. A detailed illustration for different layer protocols of 802.16 is shown in Figure 2.9.

The operation standards for WMANs are regulated under IEEE standard 802.16 [377]. WMANs are allowed the operating frequency range of 10–66 GHz. With such a broad spectrum to work with, WMANs have the ability to transmit over previous wireless frequencies such as IEEE 802.11b/g, causing less interference with other wireless products. The only downside to using such high frequencies is that WMAN needs a LOS between the transmitters and receivers, much like a directional antenna. Using a LOS, however, will decrease multipath distortion, allowing higher bandwidths to be achieved,

Table 2.3 Comparison of 802.16 standards

Parameter	802.16	802.16a/802.16d	802.16e
Date	Dec. 2001	Jan. 2003/Q3 2004	Q3, 2004
Spectrum	10–66 GHz	<11 GHz	<6 GHZ
Channels	LOS only	Non-LOS	Non-LOS
Modulation	QPSK,16 QAM, 64 QAM	OFDM256, QPSK, 16 QAM, 64 QAM	Same as 802.11a
Mobility	Fixed	Fixed	Pedestrian mobility, regional roaming
Bandwidth	20.25.28 MHZ	1.25–20 MHz	Same as 802.16a
Throughput	Up to 75 Mbps	Up to 75 Mbps	Up to 30 Mbps
Cell radius	1–3 miles	3–5 miles	1–3 miles

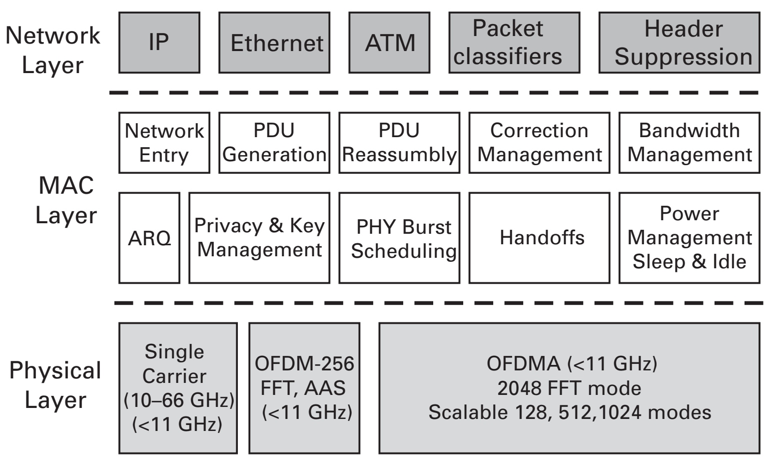


Figure 2.9 WMAX protocol stacks [383].

and can attain up to 75 Mbps for both uplink and downlink on a single channel [372]. Some extensions of 802.16 standards are listed as follows and in Table 2.3:

- IEEE 802.16a: The IEEE has developed 802.16a for use in licensed and license-exempt frequencies from 2 to 11 GHz. Most commercial interest in IEEE 802.16 is in these lower-frequency ranges. At the lower ranges, the signals can penetrate barriers and thus do not require a LOS between transceiver and antenna. This enables more flexible WiMAX implementations while maintaining the technologys data rate and transmission range. IEEE 802.16a supports mesh deployment, in which transceivers can pass a single communication on to other transceivers, thereby extending basic 802.16s transmission range.
- IEEE 802.16b: This extension increases the spectrum the technology can use in the 5- and 6-GHz frequency bands and provides QoS. WiMAX provides QoS to ensure priority transmission for real-time voice and video and to offer differentiated service levels for different traffic types.

- IEEE 802.16c: IEEE 802.16c represents a 10–66-GHz system profile that standardizes more details of the technology. This encourages more consistent implementation and, therefore, interoperability.
- IEEE 802.16d: IEEE 802.16d includes minor improvements and fixes to 802.16a. This extension also creates system profiles for compliance testing of 802.16a devices.
- IEEE 802.16e: This technology will standardize networking between carrier-fixed base stations and mobile devices, rather than just between base stations and fixed recipients. IEEE 802.16e would enable the high-speed signal handoffs necessary for communications with users moving at vehicular speeds.

In addition to IEEE 802.16, IEEE 802.20 (IEEE802.20) or the Mobile Broadband Wireless Access (MBWA) Working Group aims to prepare a formal specification for a packet-based air interface designed for IP-based services. The goal is to create an interface that will allow the creation of low-cost, always-on, and truly mobile broadband wireless networks, nicknamed Mobile-Fi. IEEE 802.20 will be specified according to a layered architecture, which is consistent with other IEEE 802 specifications. The scope of the working group consists of the physical (PHY), medium-access control (MAC), and logical-link control (LLC) layers. The air interface will operate in bands below 3.5 GHz and with a peak data rate of over 1 Mbps. The goals of 802.20 and 802.16e, the so-called "mobile WiMAX," are similar. A draft 802.20 specification was balloted and approved on January 18, 2006.

WiMAX can be viewed as "last-mile" connectivity at high data rates. This could result in lower pricing for both home and business customers as competition lowers prices. In areas without preexisting physical cable or telephone networks, WiMAX may be a feasible alternative for broadband access that has been economically unavailable. Prior to WiMAX, many operators were using proprietary fixed wireless technologies for broadband services. For this reason, WiMAX has its significant markets in rural areas and developing countries.

2.5 WiFi Networks

IEEE 802.11 denotes a set of WLAN standards developed by Working Group 11 of the IEEE LAN/MAN Standards Committee (IEEE 802). WiFi is a brand originally licensed by the WiFi Alliance to describe the underlying technology of WLAN based on the IEEE 802.11 specifications. It was developed to be used for mobile computing devices, such as laptops, in LANs, but is now increasingly used for more services, including Internet and VoIP (voiceover IP) phone access, gaming, and basic connectivity of consumer electronics, such as televisions, DVD players, or digital cameras.

In the PHY layer, 802.11b operates within the 2.4-GHz industrial, scientific, and medical (ISM) band. The original 802.11b defines data rates of 1 and 2 Mbps via radio waves using a frequency-hopping spread spectrum (FHSS) or a direct-sequence spread spectrum (DSSS). For FHSS, 2.4-GHz band is divided into 75 1-MHz subchannels. The sender and receiver agree on a hopping pattern, and data are sent over a sequence of the subchannels. Each conversation within the 802.11 network occurs over a different hopping pattern. Because of FCC regulations that restrict subchannel bandwidth to 1 MHz, FHSS techniques are limited to speeds of no higher than 2 Mbps. DSSS divides the 2.4-GHz band into 14 22-MHz channels. Adjacent channels overlap one another partially, with 3 of the 14 being completely nonoverlapping. The spreading code is the 11-bit Barker sequence. Binary-phase-shift keying (BPSK) and quadrature-phase-shift keying (QPSK) are used to provide different rates.

To increase the data rate to 5.5 Mbps and 11 Mpbs in the 802.11b standard, advanced coding technique, complementary code keying (CCK) is employed. A complementary code contains a pair of finite-bit sequences of equal length, such that the number of pairs of identical elements (1 or 0) with any given separation in one sequence is equal to the number of pairs of unlike elements having the same separation in the other sequence. A network using CCK can transfer more data per unit time for a given signal bandwidth than a network using the Barker code, because CCK makes more efficient use of the bit sequences. CCK consists of a set of 64 8-bit code words. The 5.5-Mbps rate uses CCK to encode 4 bits per carrier, whereas the 11-Mbps rate encodes 8 bits per carrier. Both speeds use QPSK as the modulation technique and signal at 1.375 Mps. Table 2.4 shows the differences rates for 802.11b.

802.11a adopts OFDM at 5.15–5.25 GHz, 5.25–5.35 GHz, and 5.725–5.825 GHz to support multiple data rates up to 54 Mbps. 802.11g utilizes the 2.4-GHz band with OFDM modulation and is also backward compatible with 802.11b. For OFDM, the fast Fourier transform (FFT) has 64 subcarriers. There are 48 data subcarriers and 4 carrier pilot subcarriers for a total of 52 nonzero subcarriers defined in IEEE 802.11a, plus 12 guard subcarriers. The IEEE 802.11a/g PHY layer provides eight PHY modes with different modulation schemes and different convolutional coding rates, and can offer various data rates. The configurations of these eight PHY modes are listed in Table 2.5.

To achieve higher data rates in the PHY layer, in January 2004, IEEE announced that it had formed a new 802.11 Task Group (TGn) to develop a new amendment to the 802.11 standard for WLANs. 802.11n builds on previous 802.11 standards by adding MIMO (multiple-input multiple-output). MIMO uses multiple transmitter and receiver antennas to allow for increased data throughput through spatial multiplexing and increased range by exploiting the spatial diversity. There are several proposal groups named TGnSync, WWiSE (short for "World-Wide Spectrum Efficiency"), and MITMOT ("MAC and MIMO Technologies for More Throughput"). All proposals occupy the 2.5-GHz frequency band with 20- or 40-MHz bandwidth so as to support the communication

Table 2.4 802.11b rates

Data rate	Code length	Modulation	Symbol rate	Bits/symbol
1 Mbps	11(DSSS)	BPSK	1 Mps	1
2 Mbps	11(DSSS)	QPSK	1 Mps	2
5.5 Mbps	8(CCK)	QPSK	1.375 Mps	4
11 Mbps	8(CCK)	QPSK	1.375 Mps	8

Table 2.5 PHY layer mode for IEEE 802.11a/g

Mode	Modulation	Channel coding	Data rate
1	BPSK	1/2	6 Mbps
2	BPSK	3/4	9 Mbps
3	QPSK	1/2	12 Mbps
4	QPSK	3/4	18 Mbps
5	16-QAM	1/2	24 Mbps
6	16-QAM	3/4	36 Mbps
7	64-QAM	2/3	48 Mbps
8	64-QAM	3/4	54 Mbps

Table 2.6 Comparison of 802.11 standards

Parameter	802.11b	802.11a/g	802.11n
Air Rate	11 Mbps	54 Mbps	200+ Mbps
MAC SAP Rate	5 Mbps	25 Mbps	100 Mbps
Range	30 m	30 m	50 m
Frequency	2.4 GHz	5.25,5.6,5.8 GHz/2.4 GHz	2.4 GHz
Bandwidth	20 MHz	20 MHz	20 or 40 MHz
Modulation	DSSS/CCK	DSSS/CCK/OFDM	DSSS/CCK/OFDM with MIMO
Special Streams	1	1	1,2,3,4

speed of more than 200 Mbps. 802.11n is backward compatible with 802.11b and 802.11g. In Table 2.6, we compare the parameters for the three 802.11 standards.

The IEEE 802.11 MAC protocol supports two kinds of access methods, namely, the distributed coordination function (DCF) and the point coordination function (PCF). In both mechanisms, only one user occupies all the bandwidth at each time slot. The PCF is based on polling, controlled by a point coordinator like an access point (AP), to communicate with a node listening and to see if the airwaves are free. The PCF seems to be implemented in only very few hardware devices as it is not part of the WiFi Alliance's interoperability standard.

In contrast, the DCF is an access mechanism using carrier-sense multiple access with collision avoidance (CSMA/CA). DCF mandates a station wishing to transmit to listen for the channel status for a DCF interframe space (DIFS) interval. If the channel is found busy during the DIFS interval, the station defers its transmission or proceeds otherwise. In a network in which a number of stations contend for the multiaccess channel, if multiple stations sense the channel is busy and defer their access, they will find that the channel is released virtually simultaneously and then try to seize the channel again at the same time. As a result, collisions may occur. To avoid such collisions, the DCF also specifies random backoff, which forces a station to defer its access to the channel for an extra period. The DCF also has an optional virtual carrier-sense mechanism that exchanges short request-to-send (RTS) and clear-to-send (CTS) frames between the source and destination stations before the long data frame is transmitted. The details of RTS/CTS will be given in a later chapter.

To take full advantage of the future market opportunity for WiFi, several key challenges must be overcome. In the following, we list some near-future design topics and their possible solutions.

- Security:
  Most concentration for WiFi is on free public access. However, eavesdroppers and hackers can take full advantage of WiFi systems. Currently, all 802.11a, b, and g devices support a WEP (wired equivalent privacy) encryption that has had flaws.

  IEEE 802.11i, also known as WiFi Protected Access 2 (WPA2), is an amendment to the 802.11 standard specifying security mechanisms for wireless networks. The 802.11i specification defines two classes of security algorithms: Robust Security Network Association (RSNA), and Pre-RSNA. Pre-RSNA security consists of WEP and 802.11 entity authentication. RSNA provides two data confidentiality protocols, called the temporal key integrity protocol (TKIP) and the countermode/CBC-MAC Protocol (CCMP). The RSNA establishment procedure includes 802.1X authentication and key management protocols. Beyond IEEE 802.11i, it is worth mentioning that WAPI (WLAN Authentication and Privacy Infrastructure) is a Chinese National Standard for WLANs (GB 15629.11-2003).
- Mobility:
  Mobility is an important attribute of wireless networks. Current WLAN standards provide mobility through roaming capabilities. IEEE 802.11p, also referred to as wireless access for the vehicular environment (WAVE), defines enhancements to 802.11 required to support intelligent transportation systems (ITSs) applications. This includes data exchange between high-speed vehicles and between the vehicles and the roadside infrastructure in the licensed ITS band of 5.9 GHz (5.85–5.925 GHz). 802.11p will be used as the groundwork for DSRC (Dedicated Short Range Communications), a U.S. Department of Transportation project, which will be emulated elsewhere, looking at vehicle-based communication networks, particularly for applications such as toll collection, vehicle safety services, and commerce transactions via cars. The ultimate vision is a nationwide network that enables communications between vehicles and roadside APs or other vehicles.
- QoS support:
  802.11e is the first wireless standard that spans home and business environments. It adds QoS features and multimedia support to the existing 802.11 wireless standards, while maintaining full backward compatibility with these standards. QoS and multimedia support are critical to wireless home networks where voice, video, and audio will be delivered. Broadband service providers view QoS and multimedia-capable home networks as an essential ingredient to offering residential customers video on demand, audio on demand, VoIP, and high-speed Internet access.

  802.11e introduces two enhancements, enhanced DCF (EDCF) and hybrid coordination function (HCF). In EDCF, a station with high-priority traffic waits a little less before it sends its packet, on average, than a station with low-priority traffic, so that high-priority traffic has a higher chance of being sent than low-priority traffic. In addition, each priority level is assigned a transmit opportunity (TXOP) that is a bounded time interval during which a station can send as many frames as possible. The HCF works more like PCF. With the PCF, QoS can be configured with great precision. QoS-enabled stations have the ability to request specific transmission parameters (data rate, jitter, etc.) that should allow advanced applications like VoIP and video streaming to work more effectively on a WiFi network.

- Integration of 3G and WLAN
  The 3G cellular networks and 802.11 WLANs possess complementary characteristics. 3G cellular networks promise to offer always-on, ubiquitous connectivity and mobility with relatively low data rates. 802.11 offers much higher data rates, comparable with those of the cellular networks, but can cover only smaller areas without mobility, suitable for hot-spot applications in hotels and airports. The performance and flexibility of wireless data services would be dramatically improved if users could seamlessly roam across the two networks. By offering integrated 802.11/3G services, 3G operators and Wireless Internet Service Providers (WISPs) can attract a wider user base and ultimately facilitate the ubiquitous introduction of high-speed wireless services. Users can also benefit from the enhanced performance and lower overall cost of such a combined service. For a network node changing the type of connectivity between 3G cellular phone and WLAN, the concept of vertical handoff will be discussed in a later chapter.

2.6 Wireless Personal-Area Networks

In this section, we first briefly review the 802.15 wireless personal-area standards. Then, for low data rate, we study the Bluetooth and Zigbee standards. Finally, we investigate the high-speed ultrawide-band standard.

A WPAN is a computer network used for wireless communication among devices [including telephones and personal digital assistants (PDAs)] close to one person. The devices may or may not belong to the person in question. The reach of a WPAN is typically a few meters. WPANs can be used for communication among the personal devices themselves (intrapersonal communication) or for connecting to a higher-level network and the Internet (an uplink). 802.15 is a communications specification that was approved in early 2002 by the IEEE Standards Association (IEEE-SA) for WPANs. Specifically, we list the following three substandards.

The IEEE Standard 802.15.1 was approved as a new standard for Bluetooth by the IEEE-SA Standards Board on 15 April 2002. The Bluetooth standard enables wireless communication between multiple electronic devices within 10 m of each other. Bluetooth devices are organized in piconets, which include one master device and up to seven slave devices. The Bluetooth devices communicate in the 2.4-GHz radio frequency (RF) band, enabling devices to communicate without LOS spacing, such as through walls or through a person's body. Bluetooth piconets utilize a FHSS in 79 1-MHz bands, reducing the likelihood of interference with other Bluetooth piconets.

802.15.3 is the IEEE standard for a high-data-rate WPAN designed to provide QoS for real-time distributions of multimedia content, like video and music. It is ideally suited for a home multimedia wireless network. The original standard uses a "traditional" carrier-based 2.4-GHz radio as the PHY layer. A follow-on standard, 802.15.3a, defines an alternative PHY layer; current candidate proposals are based on an ultrawide band (UWB), that will provide in excess of 110 Mbps at a 10-m distance and 480 Mbps at 2 m. This will allow applications requiring streaming of high-definition video between media servers and flat-screen high-definition (HD) monitors and extremely fast transfer of media files between media servers and portable media devices.

IEEE 802.15.4-2003 (Low-Rate WPAN) deals with a low data rate but a very long battery life (months or even years) and very low complexity. The first edition of the 802.15.4 standard was released in May 2003. In March 2004, after Task Group 4b was formed, Task Group 4 put itself in hibernation. The ZigBee set of high-level communication protocols is based on the specification produced by the IEEE 802.15.4 Task Group.

2.6.1 Bluetooth/Zigbee

Bluetooth is a standard for wireless communications that uses short-range radio frequencies to enable communication among multiple electronic devices. Bluetooth technology is envisioned as a replacement of the interconnection cables between personal devices such as notebook computers, cellular phones, PDAs, and digital cameras. Some typical applications are shown in Figure 2.10. If widely adopted, Bluetooth would enable a uniform interface for accessing data services. Thus calendars, address books, and business cards stored in personal devices could be automatically synchronized by use of push-button synchronization and proximity operation.

The Bluetooth Special Interest Group (SIG) was founded by Ericsson, IBM, Intel, Nokia, and Toshiba in February 1998 to develop an open specification for short-range wireless connectivity. The SIG offered all of the intellectual property explicitly included in the Bluetooth specification royalty-free to adopter members to facilitate the widespread acceptance of the technology. The SIG now includes thousands of companies. To use the intellectual property in the Bluetooth specification, adopter members must qualify any Bluetooth products they intend to bring to market through the Bluetooth qualification program. The Bluetooth qualification program includes radio and protocol conformance testing, profile conformance testing, and interoperability testing.

Bluetooth is a RF technology utilizing the unlicensed 2.4-GHz ISM band. Bluetooth enables wireless connections up to 10 m under standard transmitter power, and, because of the use of RFs, devices need not be within LOS of each other and may connect through walls or other nonmetal objects. In the active mode, Bluetooth devices typically consume 0.1 W of active power for class 1 with a range of 100 m, 2.5 mW for class 2 with a range of 10 m, and 1 mW for class 3 with a range of 1 m. The modulation technique utilized in Bluetooth technology is binary Gaussian frequency-shift keying,

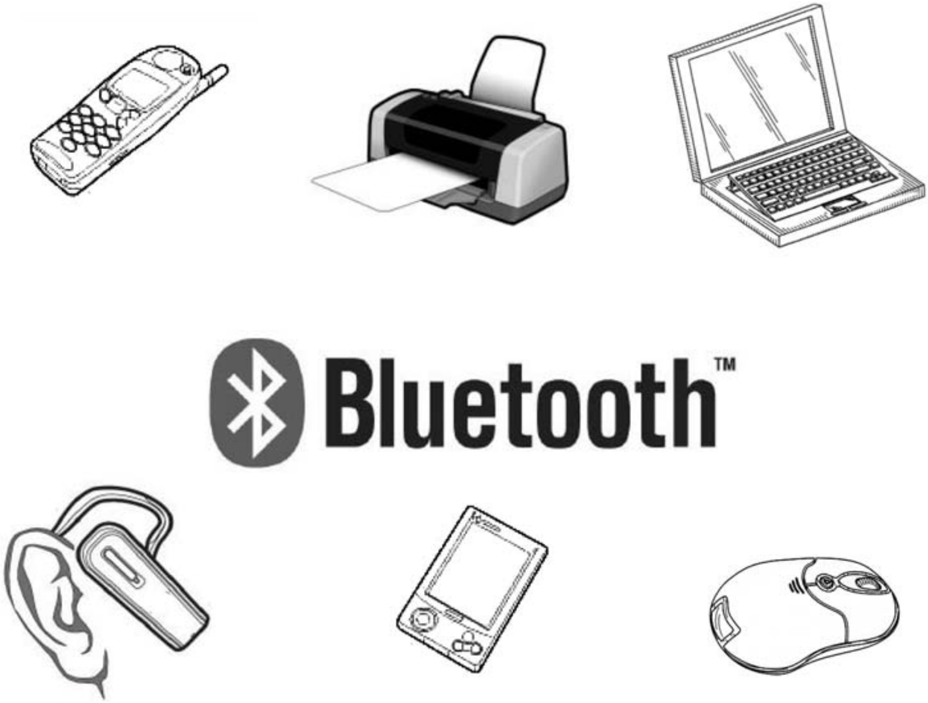


Figure 2.10 Typical Bluetooth applications.

and the baud rate is 1 Msymbol/s. Thus the bit time is 1 μs and the raw transmission speed is 1 Mbps.

The baseband signals used in Bluetooth devices, which are typically 1 MHz in bandwidth, cannot directly be transmitted on the wireless medium. Modulation of the 1-MHz baseband signals into the 2.4-GHz band is difficult to achieve in one step because complementary metal-oxide semiconductor (CMOS) transistors do not operate at these frequencies. Bluetooth radio devices solve this problem by modulating the baseband signal onto an intermediate frequency, such as 3 MHz, and then use a frequency mixer to increase the frequency of the signal to the 2.4-GHz band.

Because the unlicensed ISM band in which Bluetooth operates is often cluttered with signals from other devices, such as garage door openers, baby monitors, cordless phones, and microwave ovens, Bluetooth utilizes a FHSS for security and to avoid interference with the signals from other devices. Frequency hopping also allows multiple piconets to exist within range of each other with minimal interference. Frequency hopping typically involves generating a frequency-shift-keyed signal, and then shifting the frequency of the frequency-shift-keyed signal by an amount determined by a pseudo-noise code. The pseudo-noise code is random in that it appears to be unpredictable to an outsider, but it is generated by deterministic means. The pseudo-noise code is unique to the piconet and is determined by the master device.

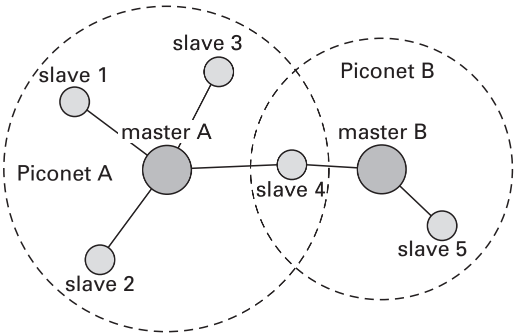


Figure 2.11 Piconets.

Bluetooth utilizes a slow hopping scheme, hopping in a pseudo-random fashion through 79 1-MHz channels. The frequency channels are located at $(2,402 + k)$ MHz, with $k = 0, 1, \ldots, 78$. A Bluetooth piconet hops through 1600 different frequencies per second. Each frequency hop corresponds to one slot, with each slot lasting 1/1600 = 625 μs. Each packet may be one, three, or five slots long. A frame consists of two packets, one packet being a transmit packet and the other packet being a receive packet.

A packet consists of an access code, a header, and a payload. The access code is 72 bits long and is used for clock synchronization, DC offset compensation, identification, and signaling. The header is 54 bits long and is used for addressing, identifying the packet type, controlling flow, sequencing to filter retransmitted packets, and verifying header integrity (ensuring that the header was not altered by another source). The payload is between zero and 2744 bits, depending on the type of packet. In packets that are one slot long, the payload is 240 bits long. In packets that are three slots long, the payload is 1500 bits long. In packets that are five slots long, the payload is 2744 bits long.

Each Bluetooth device includes a unique IEEE-type 48-bit address, called a Bluetooth device address, assigned to each Bluetooth device at manufacture, and a 28-bit clock. The clock ticks once every 312.5 μs, which corresponds to half the residence time in a frequency band when the radio hops at the rate of 1600 hops per second.

Bluetooth devices that are in communication with each other are organized into groups of two to eight devices called piconets, as shown in Figure 2.11. A piconet consists of a single master device and between one and seven slave devices. A device may belong to more than one piconet, but may be the master in no more than one piconet; thus a device may be a slave in two piconets or a master in one piconet and a slave in another piconet.

The slaves utilize the Bluetooth clock of the master to maintain time synchronization. The pseudo-random-hopping sequence is determined by the 48-bit Bluetooth device address of the master. The Bluetooth clock of the master clock determines the phase in the hopping pattern, thereby determining the particular frequency to be used at a particular time slot. Thus the communications channel in a particular piconet is fully identified by the master, and this communications channel serves to distinguish one piconet from another.

The master and the slaves alternate transmit opportunities according to a TDD scheme. According to this scheme, the master transmits on even-numbered time slots, as defined by the master's Bluetooth clock, while the slaves transmit on odd-numbered slots. A given slave may transmit only if the master has just transmitted to this slave.

To determine the presence and identities of other Bluetooth devices, Bluetooth devices engage in inquiry and page processes. The inquiry process is performed without knowledge of the identity or presence of other Bluetooth devices, whereas the paging process is performed with knowledge of the identity and presence of other Bluetooth devices.

During an inquiry process, a prospective master device makes its presence known by transmitting inquiry messages. Devices that are searching for inquiry messages respond with inquiry messages that contain their Bluetooth device addresses. After the master has acquired knowledge of the Bluetooth device address and presence of other Bluetooth devices within range, the master explicitly pages the other Bluetooth devices to join its piconet. Devices responding to the page will provide additional information, such as their clock phases, to the master.

Bluetooth devices have three low-power modes in which they reside when they are not in active communication. In sniff mode, a slave agrees with its master to listen for master transmissions periodically. In hold mode, a device agrees with another device in the piconet to remain silent for a given amount of time. A device that has gone into hold mode does not relinquish its temporary member address within the piconet. In park mode, a slave agrees with its master to park until further notice. In park mode, the slave device relinquishes its temporary member address within the piconet and periodically listens to transmissions from the master. The slave may be invited back to active communications by the master or may send a request to the master to be unparked.

Bluetooth devices typically provide link-layer security between any two Bluetooth radios. A challenge/response system, such as an E1 algorithm, is used for authentication. The authentication is based on a link key, which is a 128-bit shared secret between the two Bluetooth devices. The link key is generated by a challenge and response process between the two Bluetooth devices. Data sent between two Bluetooth devices may also be encrypted and may be ciphered with an E0 algorithm. An encryption key may be between 8 and 128 bits long and may be derived from the link key. The Bluetooth devices may use a configuration encryption key 0 to 16 bytes in length for key management and usage. The authentication and encryption keys may be generated with E2–E3 algorithms.

The specifications were formalized by the Bluetooth SIG. The SIG was formally announced on May 20, 1998. Today over 6000 companies worldwide are part of the SIG. It was established by Ericsson, Sony Ericsson, IBM, Intel, Toshiba and Nokia, and later joined by many other companies as associate or adopter members. Bluetooth is also known as IEEE 802.15.1 and the standards have the following versions:

- Bluetooth 1.0 and 1.0B
  Versions 1.0 and 1.0B had many problems, and the various manufacturers had great difficulties in making their products interoperable. 1.0 and 1.0B also had a mandatory Bluetooth Hardware Device Address (BD_ADDR) transmission in the handshaking

process, rendering anonymity impossible at a protocol level, which was a major setback for services planned to be used in Bluetooth environments.

- Bluetooth 1.1
  Many errors found in the 1.0B specifications were fixed. Support for nonencrypted channels and new features like the received signal-strength indicator (RSSI) were added.
- Bluetooth 1.2
  This version is backward compatible with 1.1 and the major enhancements include the adaptive frequency-hopping (AFH) spread spectrum, which improves resistance to RF interference by avoiding the use of crowded frequencies in the hopping sequence; higher transmission speeds in practice; extended synchronous connections (eSCOs), which improve the voice quality of audio links by allowing retransmissions of corrupted packets; host controller interface (HCI) support for three-wire universal asynchronous receiver/transmitter (UART), and HCI access to timing information for Bluetooth applications.
- Bluetooth 2.0
  This version is backward compatible with 1.x. The main enhancement is the introduction of an enhanced data rate (EDR) of 3.0 Mbps. This has the following effects: three times faster transmission speed up to 10 times in certain cases (up to 2.1 Mbps); 100-m range (depending on the class of the device); lower power consumption through a reduced duty cycle; simplification of multilink scenarios because of more available bandwidth; and further improved BER (bit error rate) performance.
- Bluetooth 2.1
  A draft version of the Bluetooth Core Specification Version 2.1 + EDR is now available from the Bluetooth website [384].

For a WPAN, besides Bluetooth technology, ZigBee is the name of a specification for a suite of high-level communication protocols using small, low-power digital radios based on the IEEE 802.15.4 standard for WPANs. ZigBee operates in the ISM radio bands; 868 MHz in Europe, 915 MHz in the United States, and 2.4 GHz in most jurisdictions worldwide. The technology is intended to be simpler and cheaper than other WPANs such as Bluetooth. The specification supports data transmission rates of up to 250 Kbps at a range of up to 30 m. ZigBee's technology is slower than 802.11b (11 Mbps) and Bluetooth (1 Mbps) but it consumes significantly less power.

2.6.2 Ultrawide Band

UWB is a technology for transmitting information spread over a large bandwidth (>500 MHz) that is able to share a spectrum with other users. In 2002, the FCC authorized the unlicensed use of UWB in 3.1 10.6 GHz. The intention is to provide an efficient use of scarce radio bandwidths while enabling high-data-rate personal-area network (PAN) wireless connectivity. Deliberations in the International Telecommunication Union Radiocommunication Sector (ITU-R) resulted in a report and recommendation on UWB in November of 2005.

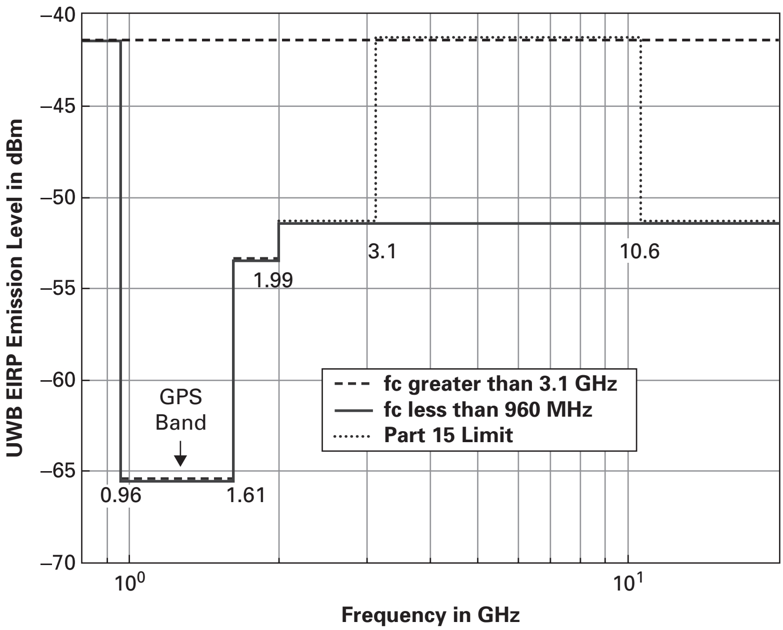


Figure 2.12 FCC UWB mask.

The FCC power spectral-density emission limit for UWB emitters operating in the UWB band is −41.3 dBm/MHz. This is the same limit that applies to unintentional emitters in the UWB band, the so-called Part 15 limit [371]. However, the emission limit for UWB emitters can be significantly lower (as low as −75 dBm/MHz) in other segments of the spectrum to prevent interference with other applications such as GPS. The FCC UWB spectrum mask is shown in Figure 2.12.

As in the IEEE 802.15.3a standard [77], the channel impulse response of the UWB can be modeled as the Saleh–Valenzuela (S-V) model [264]:

$$h(t) = \sigma^2 \sum_{c=0}^{C} \sum_{l=0}^{L} \alpha(c,l)\delta[t - T(c) - \tau(c,l)], \tag{2.13}$$

where σ^2 represents total multipath energy, $\alpha(c,l)$ is the gain of the lth multipath component in the cth cluster, $T(c)$ is the delay of the cth cluster, and $\tau(c,l)$ is the delay of the lth path in the cth cluster relative to the cluster-arrival time. The cluster arrivals and the path arrivals within each cluster are modeled as a Poisson distribution with rate Λ and rate λ (where $\lambda > \Lambda$), respectively. $\alpha(c,l)$ are modeled as zero-mean, complex Gaussian random variables with variances [77] $\Omega(c,l) = \mathrm{E}[|\alpha(c,l)|^2] = \Omega(0,0)\exp\left(-\frac{T(c)}{\Gamma} - \frac{\tau(c,l)}{\gamma}\right)$, where $\mathrm{E}[\cdot]$ is the expectation operation, $\Omega(0,0)$ is the mean energy of the first path of the first cluster, Γ is the cluster-decay factor, and γ is the ray-decay factor. The total energy contained in terms $\alpha(c,l)$ is normalized to unity, i.e., $\sum_{c=0}^{C}\sum_{l=0}^{L}\Omega(c,l) = 1$. The channel parameters corresponding are specified in [77] for four environment categories, CM1–CM4:

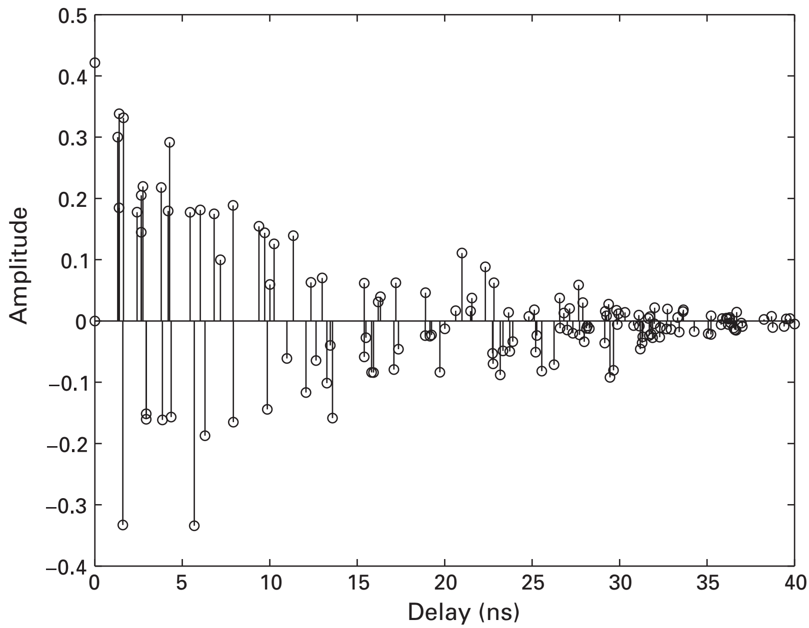


Figure 2.13 Typical UWB channel.

- Category CM1 (for Channel Model 1) consists of LOS paths in residences with $0 < d < 4$ m (RMS delay spread 5 ns);
- CM2 consists of NLOS paths in residences with $0 < d < 4$ m (RMS delay spread 8 ns);
- CM3 consists of NLOS paths in residences with 4 m $< d <$ 10 m (RMS delay spread 14 ns);
- CM4 consists of NLOS paths with extreme delay spreads (RMS delay spread 25 ns).

In Figure 2.13, we show a typical UWB channel response over time.

The ability of UWB technology to provide significantly high data rates within short ranges has made it an excellent alternative for Bluetooth for the PHY layer of the IEEE 802.15.3a standard for WPANs. However, as with 802.11 standards, two opposing groups of UWB developers are competing over the IEEE standard. The two competing technologies are single-band UWB and multiband UWB. The single-band technique, backed by Motorola/XtremeSpectrum, supports the idea of an impulse radio that occupies a wide spectrum. The multiband approach divides the available UWB frequency spectrum into multiple smaller and nonoverlapping bands with bandwidths greater than 500 MHz to obey the FCC's definition of UWB signals. The multiband approach is supported by several companies, including Staccato Communications, Intel, Texas Instruments, General Atomics, and Time Domain Corporation.

For the single-band UWB, the most popular proposal is the direct-sequence (DS)-UWB, which uses a combination of a single-carrier spread-spectrum design and wide coherent bandwidth. Unlike conventional wireless systems, which use narrowband-modulated carrier waves to transmit information, DS-UWB transmits data by pulses of

Table 2.7 Rate-Dependent Parameters

Data rate (Mbps)	Modulation	Coding rate	Conjugate symmetric inputs to IFFT	Time-spreading factor
53.3	QPSK	1/3	Yes	2
55	QPSK	11/32	Yes	2
80	QPSK	1/2	Yes	2
106.7	QPSK	1/3	No	2
110	QPSK	11/32	No	2
160	QPSK	1/2	No	2
200	QPSK	5/8	No	2
320	QPSK	1/2	No	1
400	QPSK	5/8	No	1
480	QPSK	3/4	No	1

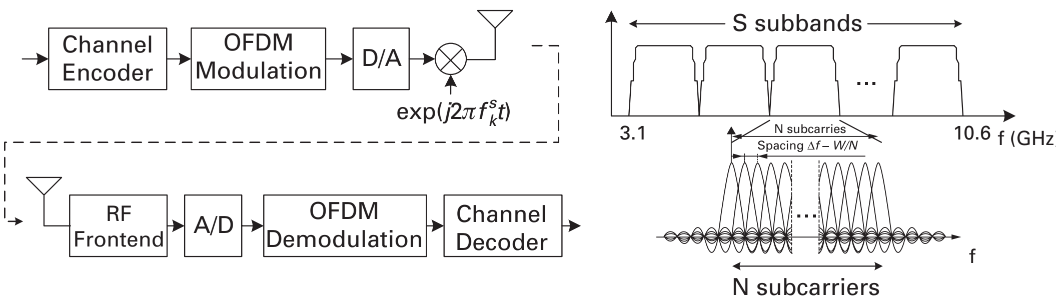


Figure 2.14 Multiband UWB system.

energy generated at very high rates: in excess of 10^9 pulses per second, providing support for data rates of 28, 55, 110, 220, 500, 660, and 1320 Mbps. A fixed UWB chip rate in conjunction with variable-length spreading codewords enables this scalable support.

For the multiband UWB, as shown in Figure 2.14, the available UWB spectrum, from 3.1 to 10.6 GHz, is divided into $S = 14$ subbands. Each subband occupies a bandwidth of at least 500 MHz in compliance with the FCC regulations. The UWB system employs OFDM with $N = 128$ subcarriers, which are modulated by QPSK. At each OFDM symbol period, the modulated symbol is transmitted over one of the S subbands. These symbols are time interleaved across subbands. Different bit rates are achieved by use of different channel-coding, frequency-spreading, or time-spreading rates. Frequency-domain spreading is obtained by choosing conjugate symmetric inputs to the IFFT (inverse fast Fourier transformation), whereas time-domain spreading is achieved by repeating the same information in an OFDM symbol on two different subbands [18]. The receiver combines the information transmitted via different times or frequencies to increase the signal-to-noise ratio (SNR) of received data.

As listed in Table 2.7, the multiband UWB system provides data rates ranging from 53.3 to 480 Mbps. For rates no higher than 80 Mbps, both time and frequency spreadings are performed, yielding an overall spreading gain of four. In the case of rates between 106.7 and 200 Mbps, only time-domain spreading is utilized, which results in an overall

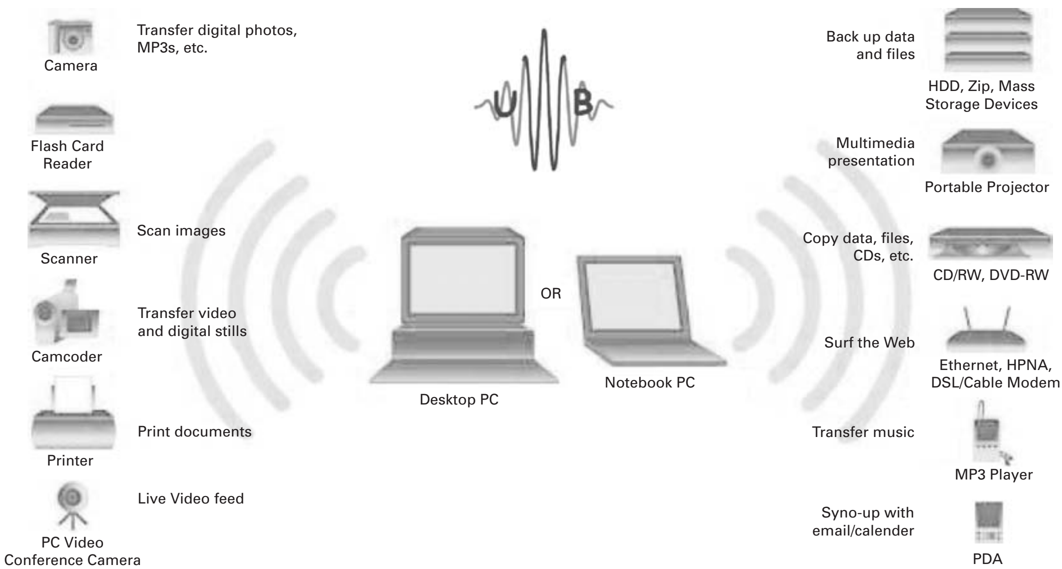


Figure 2.15 UWB applications.

spreading gain of two. The system with information rates higher than 200 Mbps exploits neither frequency nor time spreading, and its overall spreading gain is one. Forward-error-correction codes with coding rates of 1/3, 11/32, 1/2, 5/8, or 3/4 are employed to provide different channel protections with various information data rates.

Because of the extremely low emission levels currently allowed by regulatory agencies, UWB systems tend to be short range and high speed. High-data-rate UWBs can enable wireless monitors, the efficient transfer of data from digital camcorders, wireless printing of digital pictures from a camera without the need for an intervening personal computer, and the transfer of files among cell phone handsets and other handheld devices like personal digital audio and video players. Some applications are shown in Figure 2.15. UWB is also used in "see-through-the-wall" precision radar imaging technology, precision positioning and tracking (using distance measurements between radios), and precision time-of-arrival-based localization approaches.

2.7 Wireless Ad Hoc Networks

An ad hoc network is an autonomous collection of mobile users that communicate over bandwidth-constrained wireless links. The network is decentralized, in which all network activity, including discovering the topology and delivering messages, must be executed by the nodes themselves. Ad hoc networks need efficient distributed algorithms to determine network organization, link scheduling, and routing. For a special case of an ad hoc network, Mobile Ad Hoc Networks (MANETs), because the nodes are mobile, the network topology may change rapidly and unpredictably over time.

The first generation of ad hoc networks started about 1970, when Packet Radio Networks (PRNETs) were proposed by the U.S. Defense Advanced Research Projects Agency (DARPA) for multihop networks in a combat environment, and Areal Locations of Hazardous Atmospheres (ALOHA) was proposed in Hawaii for distributed channel access management. The second generation of ad hoc networks emerged in the 1980s, when the ad hoc network systems were further enhanced and implemented as a part of the Survivable Adaptive Radio Networks (SURAN) program. SURAN provided a packet-switched network to the mobile battlefield in an environment without infrastructure so as to be beneficial in improving the performance of radios by making them smaller, cheaper, and resilient to electronic attacks. In the 1990s, the concept of commercial ad hoc networks arrived with notebook computers and other feasible communications equipment. For example, the IEEE 802.11 subcommittee adopted the term "ad hoc networks."

The advantages of ad hoc networks are the easiness and speed of deployment, which are important requirements for military applications. For civil applications, ad hoc networks decrease dependence on expensive infrastructures. The set of applications for ad hoc networks is diverse, ranging from small, static networks that are constrained by power sources, to large-scale, mobile, highly dynamic networks. Some typical applications are PAN emergency operations such as policing and fire fighting, civilian environments such as taxi networks, and military use on the battlefields. One example of an ad hoc network is shown in Figure 2.16.

In contrast to the traditional wireless network with an infrastructure, an ad hoc network needs its own design requirements so as to be functional. We list some important aspects as follows:

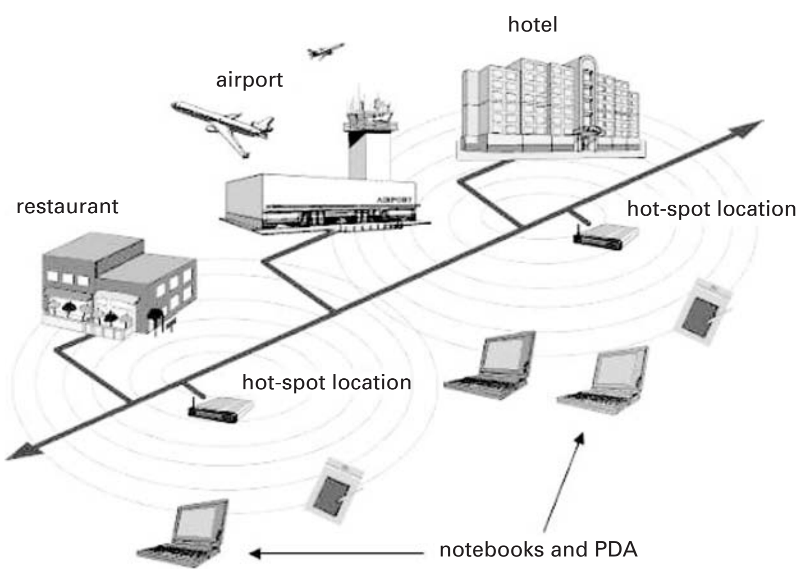


Figure 2.16 Ad hoc network example.

- Distributed Operation and Self-Organization
  No node in an ad hoc network can depend on a network in the background to support the basic functions like routing. Instead, these functions must be implemented and operated efficiently in a distributed manner. Moreover, in events such as topology changes that are due to mobility, the network can be self-organized to adapt to the changes.
  In addition, because the ad hoc nodes might belong to different authorities, they might not be necessary to cooperate to fulfill the network functions. However, this non-cooperation can cause severe network breakdown. Motivating distributed autonomous users is an important research and design topic. Traditionally, pricing anarchy is employed by use of the distributed control theory. Later in this book, we study how to explore the other methods, such as game theory, to motivate users' cooperative behaviors.
- Dynamic Routing
  For a MANET, the routing problem between any pair of nodes is challenging because of the mobility of the nodes. The optimal source-to-destination route is time variant. Moreover, compared with the traditional network in which the routing protocols are proactive, the ad hoc dynamic routing protocols are reactive. The routes are determined only when the source requests the transmission to the destination. There are two types of ad hoc dynamic routing protocols: table-driven routing protocols and source-initiated on-demand routing protocols.
  The table-driven routing protocols require each node to maintain one or more tables to store routing information. The protocols rely on an underlying routing table update mechanism that involves the constant propagation of routing information. Packets can be forwarded immediately because the routes are always available. However, this type of protocol causes substantial signaling traffic and power-consumption problems. Some protocols existing in the literature are destination-sequenced distance-vector routing [233], clusterhead gateway switch routing [46], and wireless routing protocols [212].
  Source-initiated on-demand routing creates routing only when desired by the source node. The disadvantage is that the packet at the source node must wait until a route can be discovered. But the advantage is that periodic route updates are not required. Some of the available routing protocols in the literature are ad hoc on-demand distance-vector routing [234], dynamic source routing [153], temporally ordered routing algorithm [230], associativity-based routing [309], and signal-stability-based adaptive routing protocol [59].
- Connectivity
  To achieve a connected ad hoc network, for any node there must be a multihop path to any other node. There are many types of connectivity definitions. In an undirected graph G, two vertices u and v are called connected if G contains a path from u to v. Otherwise, they are called disconnected. A graph is called connected if every pair of vertices in the graph is connected.
  One of the most adopted definitions is k-connectivity, which states that each node can at least connect to the rest of network if $k-1$ of its neighbor nodes are destroyed.