Wireless Communication Standards

A Study of IEEE 802.11™, 802.15™, and 802.16™

Todor Cooklev

IEEE
Published by
Standards Information Network
IEEE Press

INTEL-1020
U.S. Patent 9,265,043

Trademarks and Disclaimers

IEEE believes the information in this publication is accurate as of its publication date; such information is subject to change without notice. IEEE is not responsible for any inadvertent errors.

Library of Congress Cataloging-in-Publication Data

Cooklev, Todor, 1966-
Wireless communication standards : a study of IEEE 802.11, 802.15, and 802.16 / Todor Cooklev.
p. cm.
Includes bibliographical references and index.
ISBN 0-7381-4066-X
1. IEEE 802.11 (Standard) 2. wireless communication systems--Standards. I. Title.
TK5105.5668.C66 2004
621.382--dc22
2004054876

IEEE
3 Park Avenue, New York, NY 10016-5997, USA

Copyright © 2004 by IEEE
All rights reserved. Published August 2004. Printed in the United States of America.

No part of this publication may be reproduced in any form, in an electronic retrieval system, or otherwise, without the prior written permission of the publisher.

Yvette Ho Sang, Manager, Standards Publishing
Jennifer Longman, Managing Editor
Linda Sibilia, Cover Designer

INTEL-1020
U.S. Patent 9,265,043

IEEE Press/Standards Information Network publications are not consensus documents. Information contained in this and other works has been obtained from sources believed to be reliable, and reviewed by credible members of IEEE Technical Societies, Standards Committees, and/or Working Groups, and/or relevant technical organizations. Neither the IEEE nor its authors guarantee the accuracy or completeness of any information published herein, and neither the IEEE nor its authors shall be responsible for any errors, omissions, or damages arising out of the use of this information.

Likewise, while the author and publisher believe that the information and guidance given in this work serve as an enhancement to users, all parties must rely upon their own skill and judgement when making use of it. Neither the author nor the publisher assumes any liability to anyone for any loss or damage caused by any error or omission in the work, whether such error or omission is the result of negligence or any other cause. Any and all such liability is disclaimed.

This work is published with the understanding that the IEEE and its authors are supplying information through this publication, not attempting to render engineering or other professional services. If such services are required, the assistance of an appropriate professional should be sought. The IEEE is not responsible for the statements and opinions advanced in the publication.

Review Policy

The information contained in IEEE Press/Standards Information Network publications is reviewed and evaluated by peer reviewers of relevant IEEE Technical Societies, Standards Committees and/or Working Groups, and/or relevant technical organizations. The authors addressed all of the reviewers' comments to the satisfaction of both the IEEE Standards Information Network and those who served as peer reviewers for this document.

The quality of the presentation of information contained in this publication reflects not only the obvious efforts of the authors, but also the work of these peer reviewers. The IEEE Press acknowledges with appreciation their dedication and contribution of time and effort on behalf of the IEEE.

To order IEEE Press Publications, call 1-800-678-IEEE.

Print: ISBN 0-7381-4066-X SP1134

See other IEEE standards and standards-related product listings at: http://standards.ieee.org/

INTEL-1020
U.S. Patent 9,265,043

Trademarks

3Com is a registered trademark of 3Com Corporation (www.3com.com/).

Bluetooth is a registered trademark of Bluetooth SIG, Inc.(www.bluetooth.org/).

Cisco is a registered trademark of Cisco Systems, Inc. (www.cisco.com/).

Ericsson is the trademark or registered trademark of Telefonaktiebolaget L M Ericsson, Sweden (www.ericsson.com/).

Fujitsu is a registered trademark of Fujitsu Limited in Japan and/or other countries (www.fujitsu.com/).

IBM is a registered trademark of International Business Machines Corporation in the United States, other countries, or both (www.ibm.com/).

IEEE and 802 are registered trademarks of the IEEE (www.ieee.org/).

IEEE Standards designations are trademarks of the IEEE (www.ieee.org/).

Intel is a registered trademark of Intel Corporation (www.intel.com/).

Kodak is a trademark of Eastman Kodak Company (www.kodak.com/).

Lucent is a registered trademark of Lucent Technologies Incorporated (www.lucent.com/).

Microsoft is a registered trademark of Microsoft Corporation (www.microsoft.com/).

Motorola is a registered trademark of Motorola, Inc. (www.motorola.com/).

Nokia is a registered trademark of Nokia Corporation (www.nokia.com/).

SOMA Networks is a trademark of SOMA Networks, Inc. (www.somanetworks.com/).

Staccato Communications, Inc. is a registered trademark of Staccato Communications, Inc. (www.staccatocommunications.com/).

Texas Instruments is a trademark of Texas Instruments Incorporated (www.ti.com/).

Time Domain is a registered trademark of Time Domain (www.time-domain.com/).

Toshiba is a registered trademark of Toshiba Corporation (www.toshiba.com/).

Wi-Fi is a registered trademark of the Wi-Fi Alliance (www.wi-fi.org/).

WiMAX is a trademark of the WiMAX Forum (www.wimaxforum.org/).

XtremeSpectrum is a registered trademark of Motorola, Inc. (www.motorola.com/)

ZibBee Alliance is a trademark of Philips Corporation (www.zigbee.org/).

INTEL-1020
U.S. Patent 9,265,043

Dedication

To my wife Didi and my son Steven

INTEL-1020
U.S. Patent 9,265,043

Acknowledgment

A project like this book cannot be accomplished without the help of others, whom I would like to publicly acknowledge here. First, it has been personally and professionally rewarding to work with the many outstanding professionals within IEEE 802®. Second, I would like to acknowledge those I have worked closely with in industry over the years. I have gained substantially from interaction with former colleagues and friends Marcos Tzannes and Colin Lanzl. I want to also acknowledge former colleagues and friends Kurt Dobson, Dirk Ostermiller, Sy Prestwich, and Dr. Kevin Smart. I want to thank Dr. Gheorghe Berbecel for his friendship and the many stimulating discussions. I am also thankful to Victor Mozarowski for his friendship and discussions. I want to also acknowledge Dr. M. Sablatash for the interaction I have had with him over the years.

I want to thank also all of my friends. And last, but not least, I am thankful to my family, since they have made what I have achieved possible.

I am also grateful to Jennifer Longman from the IEEE Standards Information Network/IEEE Press for her work and above all for her patience during this two-year project.

I acknowledge Stefano Iachella for proof-reading the text.

I want to acknowledge specifically the anonymous reviewers from the IEEE for their constructive criticism of the completely inadequate first draft. They should be credited with all the improvements.

I am particularly indebted to the following people: Dr. Steve Halford, Globespan Virata, Palm Bay, FL; Dr. John Kowalski, Sharp Laboratories of America, Camas, WA; Dr. Martin Haenggi, University of Notre Dame, IN; and Dr. Sean Coffey, Texas Instruments, Santa Rosa, CA, for reviewing the manuscript and for their detailed comments.

INTEL-1020
U.S. Patent 9,265,043

Author

Todor Cooklev has been with San Francisco State University since 2002 where he teaches courses and conducts research in the area of wireless communications. During 1999–2002 he was with Aware, Inc. of Bedford, MA and Lafayette, CA. Prior to joining Aware he was with US Robotics, later 3Com Corporation, where he received the 3Com Inventor Award. In addition to wireless communications, he has worked in the area of voice band modems, DSL technology, and digital signal processing. He also has experience in consulting for government organizations and private technology and venture capital companies.

Dr. Cooklev has given a number of short courses, seminars, and invited talks and has also made contributions to the IEEE 802.11, 802.15, and 802.16 families of standards. He is a member of the IEEE Standards Association and has served the IEEE in a number of volunteer capacities. He is also on the Advisory Board of the IEEE International Conference on Telecommunications.

He received the Ph.D. degree in electrical engineering from Tokyo Institute of Technology, Tokyo, Japan, in 1995, and also received a NATO Science Fellowship Award in that same year. He is the inventor on several patents in the United States and author of a number of publications in the areas of communication systems, protocols, and signal processing.

INTEL-1020
U.S. Patent 9,265,043

Table of Contents

Preface xi

Acronyms and Abbreviations xiii

Chapter 1 Introduction 1
- Overview 1
- Government regulations 5
  - Recent regulatory changes 10
  - Future FCC directions 12
  - International Regulations 13
- Standardization bodies 17
- Wireless channels 24
- Introduction to cryptographic algorithms 32
- Design of a wireless communication standard 35
  - MAC requirements 36
  - PHY requirements 40
  - Sublayers of PHY and MAC 41

Chapter 2 The IEEE Standard for WLAN: IEEE 802.11 45
- Overview and architecture 45
- IEEE 802.11 security 49
  - Authentication 50
  - Privacy 58
- Interaccess Point Protocol (IAPP) 66
- Medium-access mechanism and real-time traffic over IEEE 802.11 70
  - Enhancements in IEEE 802.11e 81
- Power-saving mechanism 94
- IEEE 802.11 physical layers 96
  - DSSS PHY 97
  - IEEE 802.11b 98
  - IEEE 802.11a 103
  - IEEE 802.11g 113

INTEL-1020
U.S. Patent 9,265,043

WLAN installation 117
IEEE 802.11 technology and business trends 125

Chapter 3 Standards for Wireless Personal Area Networking (WPAN) . 133
Introduction 133
Comparing WPAN and WLAN 133
IEEE 802.15.1 135
IEEE 802.15.1 Physical layer 137
Baseband 139
LMP and L2CAP 162
Coexistence among wireless standards 164
Collaborative methods 168
Noncollaborative methods 170
Channel classification 174
PHY and MAC models 176
High-Rate WPAN 179
MAC layer 181
Physical layer for the 2.4 GHz ISM Band 195
Low-Rate WPAN 201
Overview and architecture 201
PHY Layer 204
IEEE 802.15.4 MAC 207
Coexistence issues involving IEEE 802.15.4 216
WPAN technology and business trends 218

Chapter 4 Air Interface for Fixed Broadband Wireless Access Systems . 225
MAC convergence sublayer 231
MAC common part sublayer 233
Network entry and initialization 238
Channel access and QoS 241
MAC security sublayer 248
Authentication 250
Data encryption with DES 255
MAC enhancements for 2–11 GHz operation 256
MAC enhancements for mesh systems 257
Advanced antenna systems (AAS) 261

INTEL-1020
U.S. Patent 9,265,043

Automatic repeat request (ARQ) 265
DFS for license-exempt operation 267
MAC enhancements for OFDM and OFDMA PHYs. 268
IEEE 802.16 physical layers 269
Physical layer for 10–66 GHz. 270
Physical layers for 2–11 GHz operation. 286
Coexistence 317
BWA business and technology trends 331

Chapter 5 Concluding Remarks 335

Bibliography 337

Glossary 351

Index 357

INTEL-1020
U.S. Patent 9,265,043

Preface

The purpose of this book is to discuss the design of the standards for wireless data communications. In particular, the book strives to answer the questions: How do these standards work? Why do they work the way they do? and How were they developed? The field of computer networking emerged with the advent of computers. The first mass-market data communication standard was Ethernet, standardized as IEEE 802.3™. Ethernet is a technology allowing computers to communicate over relatively short distances, such as offices, and campuses, and is the dominant wireline local-area networking technology (LAN). Ethernet requires special wiring such as CAT5 cables that are not available in homes and some buildings. The success and limitations of Ethernet, as well as the proliferation of laptops and mobile computing devices created significant market demand for wireless computer networks. The standards for broadband wireless data communication were developed in response to these market needs. These standards aim to provide integrated, packet-oriented transmission of data, voice, images, and video. These wireless technologies will have profound impact. In the near term, these wireless standards will create and expand markets for products and services. They will facilitate trade and commerce. In the long term, these wireless standards will lead to increased economic productivity and competitiveness. At present, the market for wireless data communications is one of the fastest growing communications markets, even as many other communication markets are in decline. As a result, even technology and service provider companies that have traditionally offered products and services for wired communication technology are beginning to offer products involving wireless communication technology. This confirms the shift from wired to wireless communication technology that is presently occurring. For many low-tech companies the use of wireless technology is new and results in substantially increased productivity. It is expected that the proliferation of laptops and, especially, of other mobile computing devices such as PDAs and home networking will continue to contribute to the growth in wireless data communications over the foreseeable future.

Almost all products and services currently being introduced are based on the standards developed by the IEEE 802 LAN/MAN Standards Committee. As a result many engineers and computer scientists, students, marketing specialists, and managers must educate themselves about these standards. This book is

INTEL-1020
U.S. Patent 9,265,043

designed to help practicing professionals learn about the standards. Graduate students doing research work in the area of wireless data communication technology will also benefit from using this book.

I have attempted to make the text accessible to a wide audience. The concepts are explained based on intuition and insights. Background in signal and system theory, modulation and coding, while helpful, is not particularly required. The book is suitable for novice readers who want to learn more about the subject, as well as for advanced readers who are working in wireless communications. The organization is as follows. In the introductory Chapter 1 the main requirements for the wireless standards (802.11™, 802.15™, and 802.16™) and the main technical difficulties that the wireless environment creates are discussed. The important issues of government regulations and the standardization process are also explained. Chapter 2 is devoted to the IEEE 802.11 standards for wireless local area networks, which historically were the first wireless data communication technology to be developed. The practical problem how to deploy wireless local area networks is also discussed. The IEEE 802.15 family of standards for wireless personal area networking is the subject of Chapter 3, and the IEEE 802.16 standard for broadband wireless access is the subject of Chapter 4. In general there are two ways to view any communications system: architectural and functional. The architectural approach emphasizes the logical divisions of the system and how they work together. The functional approach emphasizes the actual components. In each chapter, first the architecture is presented, followed by a description of the actual components.

It should be noted that the technical subject of this book is rapidly evolving. I have made an effort within the text to reflect not only the state of the wireless standards at publication time (2004), but to provide guidance as to what developments might happen in the future. It should be noted that the views expressed here represent solely my personal opinion. They do not necessarily coincide with the positions of either the IEEE 802 Committee, or the individual Working Groups or Task Groups.

INTEL-1020
U.S. Patent 9,265,043

Acronyms and Abbreviations

The following acronyms and abbreviations are used in this book:

2G	second generation
3-DES	triple data encryption standard
3G	third generation
AA	authorization agent
AAA	authentication, authorization, and accounting
AAD	additional authentication data
AAS	advanced antenna systems
AC	access category
ACK	acknowledgement
ACL	asynchronous connection-less or access control list
ACO	authenticated ciphering offset
ADSL	asymmetric digital subscriber line
AES	advanced encryption standard
AFH	adaptive frequency hopping
AGC	automatic gain control
AIFS	arbitration inter-frame space
AIFSN	arbitration inter-frame space number
AK	authorization key
AP	access point
APME	access point management entity
ARIB	Association of Radio Industries and Businesses
ARQ	automatic repeat request
AS	authentication server
ASTM	American Society for Testing and Materials
ATIM	ad-hoc traffic indication message
ATM	asynchronous transfer mode
AWMA	alternating wireless medium access
AWGN	additive white Gaussian noise

INTEL-1020
U.S. Patent 9,265,043

BB	baseband
BCC	binary convolutional code
BE	best effort
BER	bit error rate
BIFS	backoff inter-frame space
BPSK	binary phase shift keying
BTC	Block turbo code
BS	base station
BSS	basic service set
BSSID	basic service set identifier
BT	3 db bandwidth-bit duration product for Gaussian frequency shift keying
BTC	block turbo code
BW	bandwidth
BWA	broadband wireless access
CA	collision avoidance
CAT5	category 5
CAP	contention access period
CAZAC	constant amplitude zero autocorrelation
CBC	cipher block chaining
CBC-MAC	cipher block chaining-message authentication code
CBR	constant bit rate
CC	controlled contention
CCA	clear channel assessment
CCI	controlled-contention interval
CCITT	(French) International Telegraph and Telephone Consultative Committee (now ITU)
CCK	complementary code keying
CCM	counter + cipher block chaining-medium access control
CCMP	counter + cipher block chaining-medium access control protocol
CCS	common channel signaling
CD	collision detection or compact disc
CDMA	code-division multiple access

INTEL-1020
U.S. Patent 9,265,043

CEI	channel estimation interval
CEPT	European Conference of Postal and Telecommunications Administrations
CF	coordination function
CFB	contention-free burst
CFP	contention-free period
CFR	Code of Federal Regulations
CINR	carrier-to-interference-plus-noise-ratio
COF	ciphering offset number
CP	contention period or cyclic prefix
CRC	cyclic redundancy code
CS	convergence sublayer
CSMA/CA	carrier sense multiple access with collision avoidance
CTA	channel time allocation
CTC	convolutional turbo code
CTR	counter
CTS	clear to send
CVSD	continuously variable slope delta modulation
CW	contention window
DAMA	demand assigned multiple access
DBPSK	differential binary phase shift keying
DC	direct current
DCD	downlink channel descriptor
DCS	dynamic channel selection
DES	digital encryption standard
DEV	device
DFS	dynamic frequency selection
DFT	discrete Fourier transform
DHCP	dynamic host control protocol
DIAC	dedicated inquiry access code
DIFS	distributed inter-frame space
DL	downlink
DLP	direct link protocol

INTEL-1020
U.S. Patent 9,265,043

DQPSK	differential quadrature phase shift keying
DS	direct sequence
DSL	digital subscriber line
DSRC	dedicated short-range communication
DSSS	direct sequence spread-spectrum
DVD	digital video disc
EAP	extensible authentication protocol
EAPOL	extensible authentication protocol over local area network
EC	elliptic curve
ECB	electronic codebook mode
ECC	elliptic curve cryptography
ECDSA	elliptic curve digital signature algorithm
ECMQV	elliptic curve Menezes-Qu-Vanstone
EDCF	enhanced distributed coordination function
EIA	Electronic Industry Association
EIFS	extended inter-frame space
EIRP	effective isotropic radiated power
EN	European norm
ERC	European Radio Commission
ESS	extended service set
ETSI	European Telecommunication Standardization Institute
FCC	Federal Communications Commission
FCH	frame control header
FCS	frame check sequence
FDD	frequency-division duplexing
FEC	forward error correction
FFT	fast Fourier transform
FH	frequency hopping
FIPS	Federal Information Processing Standard
FTP	file transfer protocol
GF	Galois field
GFSK	Gaussian frequency shift keying
GIAC	general inquiry access code

INTEL-1020
U.S. Patent 9,265,043

GTS	guaranteed time slot
HEC	header error check
HC	hybrid coordinator
HCCA	hybrid coordination function controlled channel access
HCF	hybrid coordination function
HCI	host control interface
HDR	high data rate
H-FDD	half duplex frequency division
HIPERLAN	high performance radio local area network
HR	hop range
HUMAN	high-speed unlicensed metropolitan area network
HVAC	heating, ventilating, and air conditioning
I	inphase
IA	interference area
IAPP	inter-access point protocol
IBSS	independent basic service set
IDFT	inverse discrete Fourier transform
IEC	International Electrotechnical Commission
IETF	Internet Engineering Task Force
IEEE	Institute of Electrical and Electronic Engineers, Inc.
IF	intermediate frequency
IFFT	inverse fast Fourier transform
IFS	inter-frame space
IP	internet protocol
ISI	intersymbol interference
ISM	industrial, scientific, medical
ITU	International Telecommunications Union
JTC	joint technical committee
KEK	key encryption key
L2CAP	logical link control and adaptation protocol
LAN	local area network
LC	link controller
LFSR	linear feedback shift register

INTEL-1020
U.S. Patent 9,265,043

LLC	logical link control
LMP	link manager protocol
LMSC	Local and Metropolitan Area Network Standards Committee
LOS	line of sight
MAC	medium access control
MAN	metropolitan area network
MIB	management information base
MIC	message integrity code
MIFS	minimum inter-frame space
MIMO	multiple-input multiple-output
MLME	medium access control layer management entity
MMDS	multichannel multipoint distribution service
MP-MP	multipoint-to-multipoint
MSDU	medium access control service data unit
MWS	multimedia wireless systems
NAV	network allocation vector
NIST	National Institute of Standards and Technology
NLOS	non-line of sight
NOI	Notice of Inquiry
NPRM	Notice of Proposed Rulemaking
nrtPS	non-real-time polling service
NTIA	National Telecommunications and Information Administration
OCB	offset codebook mode
OFDM	orthogonal frequency division multiplexing
OFDMA	orthogonal frequency-division multiple access
OQPSK	offset quadrature phase shift keying
OSI	open system interconnection
PAN	personal area network
PBCC	packet binary convolutional coding
PBX	private branch exchange
PC	personal computer or point coordinator
PCF	point coordination function

INTEL-1020
U.S. Patent 9,265,043

PCM	pulse code modulation
PCS	personal communication service
PDU	protocol data unit
PER	packet error rate
PF	persistence factor
PHS	payload header suppression
PHY	physical layer
PIFS	point coordination function inter-frame space
PIN	personal identification number
PKM	privacy key management
PL	path loss
PLL	phase locked loop
PLME	physical layer management entity
PMD	physical medium dependent
PMK	pairwise master key
PMP	point-to-multipoint
PN	packet number or pseudo-noise
PNC	piconet coordinator
PPM	pulse position modulation
PS	power save
PSD	power spectral density
PTA	packet traffic arbitration
PSPS	piconet synchronized power save
PTA	packet traffic arbitration
PTK	pairwise transient key
Q	quadrature
QAM	quadrature amplitude modulation
QLRC	quality long retry count
QoS	quality of service
QPSK	quadrature phase shift keying
QSRC	quality short retry count
RADIUS	remote authentication dial-in user service
RF	radio frequency

INTEL-1020
U.S. Patent 9,265,043

RFC	request for comments
RIFS	retransmission inter-frame space
RLAN	radio local area network
RPE	radiation pattern envelopes
RRM	radio resource management
RS	Reed–Solomon or repeater station
RSA	Rivest, Shamir, and Adleman
RSSI	received signal strength indicator
rtPS	real-time polling service
RTS	request to send
RTTT	road transport and traffic telemetric
RX	receiver
SA	security association
SAP	service access point
SB	stuff bits
SC	single-carrier
SCO	synchronous connection-oriented
SDL	specification description language
SDP	services discovery protocol
SDR	software defined radio
SDU	service data unit
SIFS	Short inter-frame space
SIG	Special Interest Group
SINR	signal-to-interference-plus-noise-ratio
SLP	service location protocol
SNMP	simple network management protocol
SNR	signal-to-noise ratio
SME	station management entity
SP	service period
SPS	synchronous power save
STC	space-time coding
SS	Subscriber station
SSID	SS identifier

INTEL-1020
U.S. Patent 9,265,043

TBTT	targeted beacon transmission time
TC	traffic category
TCM	Trellis-code modulation
TCP	transmission control protocol
TCS	telephony control signaling
TDD	time-division duplexing
TDMA	time-division multiple access
TE	terminal equipment
TEK	traffic encryption key
TFTP	trivial file transfer protocol
TIM	traffic indication map
TKIP	temporary key integrity protocol
TPC	turbo product code or transmitter power control
TS	traffic stream
TSC	temporary key integrity protocol sequence number
TSPEC	traffic specification
TBTT	targeted beacon transmission time
TV	television
TX	transmitter
TXOP	transmission opportunity
UCD	uplink channel descriptor
UDP	user datagram protocol
UGS	unsolicited grant service
UL	uplink
UNII	unlicensed national information infrastructure
USB	universal serial bus
UW	unique word
UWB	ultra-wide band
VOIP	Voice over IP
WAN	wide-area network
WECA	Wireless Ethernet Compatibility Alliance (now Wi-Fi)
WEP	Wireless equivalent privacy
Wi-Fi	wireless fidelity

INTEL-1020
U.S. Patent 9,265,043

WLAN	wireless local area network
WPA	wireless protected access
WPAN	wireless personal area network
WRC	World Radio Congress
XOR	Exclusive "or"
XPD	cross-polar discrimination

INTEL-1020
U.S. Patent 9,265,043

Copyright © 2004 IEEE

Chapter 1 Introduction

People often take the view that standardization is the enemy of creativity. But I think that standards help make creativity possible—by allowing for the establishment of an infrastructure, which then leads to enormous entrepreneurialism, creativity, and competitiveness.

—Vint Cerf, *TCP/IP co-developer and Internet pioneer*

OVERVIEW

Wireless data communications technology is incessantly progressing from research to standardization to implementation. This book proposes that this process (i.e., the progress of communication technology) is governed by the following four main principles. First, Shannon's law:

$$C = W\log_2\left(1 + \frac{P}{N_0 W}\right) \qquad \text{Eq. 1–1}$$

where C is the capacity of a communications system in bits per second, P is the signal power in Watts, W is the signal bandwidth in Hertz, and N_0 is the one-sided noise power spectral density in Watts/Hz. Shannon's law says that if there is one transmitter and one receiver, the capacity of a communication channel depends linearly on the available bandwidth and logarithmically on the signal-to-noise ratio. In other words, it is more difficult to increase the capacity of a channel by increasing the signal-to-noise ratio than it is by increasing the available bandwidth.

The second fundamental principle is Moore's law, which states that the level of integration of integrated circuits will double every 18 months.

The third principle is that the value of a network is proportional to the square of the speed of connection.[1]

[1] This principle is sometimes attributed to R. Metcalfe, inventor of Ethernet and founder of the data networking company 3Com® Corporation.

INTEL-1020
U.S. Patent 9,265,043

The fourth main principle is that the value of a network is proportional to the square of the number of devices that can be connected. With respect to a particular device, the fourth principle says that the value of a device is proportional to the square of the number of devices with which it can communicate.

The place of Shannon's law among this set of main principles does not need to be justified. Communications technology, the wireless standards in this book included, has been trying to get closer to Shannon's capacity since Shannon's work in 1948. Moore's law is very well known in the area of integrated circuit design. Why is it also a communications law? It is a communications law because it provides the foundation to implement increasingly complex and powerful signal processing circuits. In this way, Moore's law effectively allows increasingly complex and powerful wireless devices to be implemented. Companies and businesses, of course, want constantly to increase the value of their networks and devices. While the first two principles address the question of what can be accomplished, the third and fourth principles address the question of how this increase in value can be accomplished—by increasing the rate of communication and by increasing the number of attached devices. The development of the IEEE 802® wireless data communication standards is a good manifestation of these fundamental laws of communications. The wireless standards provide the technology to interconnect a large number of devices at increasing data rates.

Before describing the standards for wireless data communications, the main characteristics of the ideal wireless technology should be identified. These characteristics include the following:

a) Ability to transport voice, audio, and video, in addition to computer data

b) Allowing devices with differing price, power consumption, and data rates to operate

c) Efficiently and dynamically allocating spectrum among the various networked devices

These requirements deserve to be discussed in greater detail.

Historically, much like the situation in wired networks, the development of wireless technology proceeded along two main paths—voice and data networks. Voice-oriented networks were developed first, followed by data-oriented networks. The main reason is business—the market for voice-oriented products

INTEL-1020
U.S. Patent 9,265,043

and services developed earlier. However, at present there are two notable long-term market trends. First, the market for data-oriented products and services is growing considerably faster than the market for voice-oriented products and services. Second, voice-oriented and data-oriented products and services are merging. As a result, voice-oriented networks have increased data capabilities, and modern data-oriented networks can efficiently transport voice. One consequence of these market trends is the increased competition between vendors of voice-oriented products and vendors of data-oriented products. Therefore, the ideal wireless technology must efficiently transport voice, audio, and video in addition to data. Voice, audio, and video have different requirements than data. For data, the most important parameter is throughput, but the delay (or latency) is not important. For voice and video, in addition to a minimum throughput, the signal delay is very important. Substantial progress has been made recently to satisfy this requirement. All wireless standards developed by the IEEE 802 Working Groups strive to transport voice, audio, and video, in addition to data.

The second and third requirements (item b and item c, previously) towards the ideal wireless technology are much more difficult to satisfy, despite many claims to the contrary. There even does not seem to be universal agreement about the meaning of the second and third requirements.

Almost every wireless technology supports multiple data rates. For example, IEEE Std 802.11b™ support data rates of 1, 2, 5.5, and 11 Mb/s, IEEE Std 802.11a™ supports data rates between 6 and 54 Mb/s, etc. However, IEEE Std 802.11a devices that can communicate only at 6 Mb/s would cost the same as devices that can communicate at 54 Mb/s. At present, wireless technologies provide data rate scalability, but they provide no cost scalability at the same time. It becomes apparent that there is no technology for wireless communication that can satisfy the second requirement.

The third requirement—frequency agility—is necessary because the electromagnetic spectrum is limited and shared with other wireless devices. As the number of different technologies increased, frequency agility became important. Recent standards for wireless data communication have some form of frequency agility. The protocols have features to detect whether a frequency band is used by other devices, and if so, to coordinate the relocation to another frequency band. Still, standards lack complete flexibility. Frequency agility means

INTEL-1020
U.S. Patent 9,265,043

truly dynamic spectrum allocation. Using the simple IEEE 802.11a example, a 6 Mb/s device will take as much bandwidth to communicate as a 54 Mb/s device.

It is obvious that existing technologies lack many of the characteristics of the ideal wireless technology. If a wireless technology had true scalability and frequency agility, then it would encompass all applications over short and long distances. The fact that such a technology does not exist has led to three main different wireless technologies—wireless local area networks (WLAN), wireless personal area networks (WPAN), and broadband wireless access (BWA) devices, also called wireless metropolitan area networks (WMAN).

The differences among WLAN, WPAN, and WMAN arise from the different design goals. WPAN devices are intended for very short-range communications—typically up to 10 m. In addition, they are used in small battery-powered devices and therefore must consume very little power. WLAN devices are intended for operation of up to 100 m and have intermediate power consumption requirements. Data rate for most WPAN devices is not as important as it is for WLAN devices, but battery life and cost are. As a result, WPAN devices achieve lower data rates and lower cost than WLAN devices. Some WPAN technologies being developed now aim for a different trade-off—very high data rates, but at very short distances. This trade-off is desirable to enable true "cable replacement." Both WPAN and WLAN devices work in unlicensed bands. WMAN devices are intended primarily to work over much larger distances—up to several kilometers. In addition, WMANs can work not only over unlicensed bands, but also over frequency bands that are licensed. WMAN devices are not intended to be used in battery-powered devices for extended periods of time.

The remainder of this chapter discusses government regulations relevant to the wireless data communication standards, followed by an overview of the standardization process. Because efficient design of wireless devices requires understanding of radio propagation, the characteristics of wireless channels are summarized in "Wireless channels" on page 24. An overview of the main technical characteristics of wireless standards is presented in "Design of a wireless communication standard" on page 35. The chapter concludes with an introduction to cryptographic systems which are used in current wireless communication standards. These standards are presented in subsequent chapters.

INTEL-1020
U.S. Patent 9,265,043

GOVERNMENT REGULATIONS

This section discusses why government regulations are necessary, what the relevant government regulations are in the United States and other major markets around the world, and how and in what direction these regulations are changing.

Arguably, spectrum is a precious resource in wireless communications. This situation is drastically different from wired communications where one can always run an extra wire, and spectrum is therefore essentially unlimited. The development of large markets for wireless products and services always depends, above all, on the availability of spectrum in appropriate frequency bands. The total radio spectrum comprises the region between 3 kHz and 300 GHz. Spectrum is of fundamental importance because from one frequency band to another, both radio propagation and the cost to build devices are different. In addition, antenna size depends on the frequency band of operation. In general, different frequency bands are appropriate for different systems. Furthermore, wireless devices are potentially subject to interference from other wireless devices that operate in the same frequency band. To maintain order, spectrum has been regulated by governmental organizations.

Another reason governments regulate the area of communications is the high cost of deploying equipment to provide nationwide coverage. In the U.S., the governmental organization that regulates spectrum for use by the Federal government is the National Telecommunications and Information Administration (NTIA), and the organization that regulates spectrum for commercial use is the Federal Communications Commission (FCC) [B55]. While the FCC rules officially govern operation only in the U.S., they are monitored by regulatory bodies all over the world, and even followed in varying degrees by many other nations in the Americas and the Pacific Rim.

In the past, communications services in all countries were provided by monopolies. In the last 20 years there has been a move away from monopoly in the United States, followed by Canada, and to a lesser extent in other parts of the world. In the United States, the Telecommunications Act of 1996 played an especially big role in opening the communication markets to competition.

INTEL-1020
U.S. Patent 9,265,043

In the area of wireless data communications, the FCC has the following two main responsibilities:

a) Allocates spectrum for wireless data communication devices and establishes regulations, which the devices operating in the allocated spectrum must satisfy.

b) Certifies that the equipment on the U.S. market satisfies the regulations.

The demand for spectrum is significant, especially for certain frequency bands that have attractive propagation characteristics and allow inexpensive devices to be manufactured. Spectrum has been allocated for use by the military, law enforcement, and other government agencies, radio amateurs, etc. Similar to other regulatory bodies around the world, the FCC can respond to the demand for spectrum primarily by steady migration toward higher frequency bands, and—whenever possible and economically justified—by restructuring the frequency band allocations. Higher frequency bands are wider and therefore can accommodate more services. Restructuring of the frequency band allocations is seldom used because it is associated with significant cost.

The FCC gives licenses for a specific purpose. They cannot be resold without FCC approval.

There are two main types of frequency bands as per FCC regulations: licensed and unlicensed. Only the license holders are allowed to operate in licensed bands. Since 1994, these licenses have been awarded by auctions. The goal of the auction is to award the license to those who will use it most effectively and not to maximize revenue for the government. Nevertheless, economic efficiency does translate into higher revenue, and licenses are awarded to the highest bidder. No license is required for operation in the unlicensed bands. In the United States the FCC regulations are under the 47th part of the Code of Federal Regulations (CFR).

There are several specific FCC rules of high interest to designers of IEEE 802 class systems. These sections are 15.35, 15.205, 15.209, 15.247, and 15.249.

FCC 15.35 gives the requirements for detector and averaging functions for certification measurements. FCC 15.205 documents the *restricted bands* where only spurious emissions are allowed, and where those must meet the general levels of 15.209. Above 1,000 MHz, averaging according to 15.35 may be used.

INTEL-1020
U.S. Patent 9,265,043

FCC 15.209 restricts the RF energy that electronic equipment may parasitically emit. The specific level of emissions is 200 μV/m at 3 m test range below 960 MHz, and 500 μV/m above. These field strengths are approximately equivalent to –49.2 and –41.2 dBm ERP, respectively.

This rule applies to the Industrial, Scientific, and Medical (ISM) bands, which are 902–928 MHz, 2400–2483.5 MHz, and 5725–5875 MHz.

FCC 15.247 is the primary category for U.S. operations of unlicensed equipment [B39]. This service category provides the potential for the highest performance of all the unlicensed service categories, allowing freedom from licensing, transmit powers up to 1 watt, and no limitations of application or transmit duty cycle. These requirements allow for carrier power up to 1 watt for direct sequence spread-spectrum systems in both 902–928 MHz and 2400–2483.5 MHz, if direct sequence processing gain is a minimum of 10 dB. However, effective May 30, 2002, the FCC eliminated the requirement for direct sequence spread-spectrum and replaced it by "digital modulation." Digital modulation was defined as the point when the combination of data rate, coding, and modulation method has a 6 dB bandwidth greater than 500 kHz and a maximum transmitted spectral density of less than +8 dBm/3 KHz. However, either direct sequence or frequency hopping spread spectrum may still be used to meet the requirements under 15.247 [B40].

In the same ISM bands of 902–928 MHz, 2400–2483.5 MHz, and 5725–5875 MHz, FCC 15.249 provides for narrowband (non–spread spectrum) operation of up to approximately 1 mW effective radiated power (if wideband, then 1 mW per 100 kHz below 1000 MHz and 1 mW/MHz above 1000MHz). Operation is allowed up to 50 mV rms of electric field strength from a transmitter at a 3-m test range. This is equivalent to 0.75 mW effective radiated power.

Other rules contained in CFR 47 provide allocations in the 260–470 MHz range (FCC 15.231), but because they are generally restricted on power, transmit duty cycle, and application, this band is primarily used for control and security applications (such as keyless entry). It is not of interest for high-speed wireless data communications.

The ISM bands are of particular interest to wireless data communications. These ISM bands were allocated to foster the development of new devices and service offerings that will stimulate economic development and the growth of new

INTEL-1020
U.S. Patent 9,265,043

industries. The ISM regulations have indeed been extremely successful and many ISM devices are now available on the market. Because the spectrum is unlicensed, many more devices are yet to come. There are several reasons why these frequency bands are very attractive for wireless data communications. At these frequencies, a transmitter with power of 1 W or less with omnidirectional antenna can provide coverage up to 300 m indoors and up to a half a mile outdoors. At such frequencies, the antenna size is about one inch. At the ISM frequencies, today's semiconductor technology allows wireless devices to be built with reasonable power consumption, size, and cost. As the cost to build wireless transceivers at higher frequencies decreases, migration to higher frequencies and wider bands can be expected in the future.

IEEE 802 equipment would most often be certified under the FCC 15.247 service category [B39], [B40].

Devices that operate in the ISM bands must share the spectrum not only with other wireless data communication devices, but with other unlicensed devices. For example, the 902–928 MHz band has been used for cordless phones, and an increasing number of cordless phones work in the 2400–2483.5 MHz band. All these devices must share these frequency bands not only among themselves, but also with other government wireless systems. Many of these government systems are radiolocation systems, which transmit at high power levels. The intention of the rules is to prevent interference from the unlicensed devices to licensed systems, but not the opposite. Licensed systems are considered primary users and can cause interference to unlicensed devices. It is the responsibility of the vendor to design the unlicensed system so that it can operate in the presence of interference. Every company is allowed to develop and sell products operating in the unlicensed bands, provided that these products satisfy the FCC Part 15.247 regulations. Unlicensed devices can use frequency-hopping spread-spectrum (FHSS), direct-sequence spread-spectrum (DSSS), and novel digital modulation technologies. Spread-spectrum systems have less capability to produce harmful interference. Spread-spectrum systems work by spreading the digital signal by a pseudo-random code. The spreading process reduces the power spectral density of the transmitted signal.

INTEL-1020
U.S. Patent 9,265,043

Direct-sequence spread-spectrum devices work by first using conventional modulation and then multiplying the signal with a chipping sequence. Generally speaking spreading rates of at least $N = 10$ chips per symbol are required to achieve $10\log_{10} N = 10$ dB processing gain. To allow digital modulation schemes to be used, such as orthogonal frequency division multiplexing (OFDM), recent regulations eliminated the processing gain requirement and, instead, specified only the maximum power density [B40].

The regulations before 2002 required direct-sequence spread-spectrum systems to have a processing gain of at least 10 dB. The 10 dB minimum was specified to ensure that the system is in fact spread spectrum.

Frequency-hopping systems work by changing or hopping the center frequency in a pseudorandom fashion. Therefore, frequency-hopping systems require a wider frequency band in which to hop. In the 902–928 MHz band, FH systems are allowed output power up to 1 W if they hop over 50 hopping channels of width 500 KHz, and 0.25 W if they hop over more than 24 but less than 50 channels. Frequency-hopping systems in the 2.4 GHz band are allowed output power of 30 dBm (or equivalently 1 W) if they hop over 75 channels 1MHz wide, and 0.125 W if they hop over more than 14 but less than 75 channels. In the 5.725–5.875 GHz band, FH systems are allowed output power up to 1 W. Note that the power in dBm is related to the power in megaWatts by

$$P_{dBm} = 10\log_{10} P_{mW}$$

and to obtain the power density, the signal power is divided by the occupied bandwidth. Allowing hopping to occur on less than the maximum number of channels is another recent modification of the rules. The new rules allow intelligent frequency-hopping systems to be built, which, by hopping on fewer channels, avoid other wireless systems operating in the same frequency band.

Until recently, only spread-spectrum systems were allowed, and digital modulation systems were not allowed in the ISM bands. The reason that the regulations in Part 15.247 were recently revised is that technological advances in digital modulation made possible systems that can cause less interference, much like spread-spectrum systems. At present in all three ISM bands digital modulation systems are allowed with output power levels of up to 1 W.

INTEL-1020
U.S. Patent 9,265,043

The power levels quoted previously are specified at the device-to-antenna connector and are valid for devices operating with antennas having gains of up to 6 dBi. If the antenna gain exceeds 6 dBi, the output power must be reduced. In the 902–928 MHz band, the output power must be reduced by 1 dB for every 1 dB of antenna gain above 6 dBi, and in the 2400–2483.5 MHz band, the output power must be reduced by 1 dB for every 3 dB of antenna gain above 6 dBi. For point-to-point systems in the 5.725–5.875 GHz band, highly directional antennas can be used with gains above 6 dBi without reduction of output power.

The ISM bands in the 900, 2400, and 5700 MHz regions are narrow. To enable high data rate wireless devices and services, in 1997 the FCC allocated a 300 MHz–wide block of frequencies for unlicensed use [B41], [B42]. These bands became known as unlicensed national information infrastructure (U-NII) bands and are from 5.15–5.25, 5.25–5.35, and 5.725–5.825 GHz. These bands are called lower, middle, and upper U-NII bands, respectively. The regulations, which all equipment operating in these bands must satisfy, are in Section 15.401 of the FCC regulations.

Recent regulatory changes

Very recently, important further developments took place. In a Notice of Proposed Rulemaking FCC 03-110 dated June 2003, the FCC amended Section 15.401. To harmonize US regulations with European regulations at the request of the Wi-Fi® Alliance, the FCC allocated additional 255 MHz of spectrum between 5.47 and 5.725 GHz. At the same time, the FCC took steps to protect licensed wireless services operating in this frequency band from possible interference from unlicensed wireless networks. The lower U-NII band is not allowed to be used outdoors, and the maximum EIRP that is allowed for the band of 5.15–5.25 GHz is 200 mW. For the bands of 5.25–5.35 GHz and 5.47–5.725 GHz, the maximum EIRP allowed is 1 W. These bands are shared with radiolocation, Space Research Service, and Earth Exploration Satellite Service. For the 5.725–5.825 GHz band, the maximum EIRP that is allowed is 4 W. These power levels are allowed with antennas that have up to 6 dBi gains. Devices may use antennas with higher gains, but the output power at the transmitter-to-antenna connector must be reduced by the same amount in decibels that the antenna gain exceeds 6 dBi. For fixed point-to-point systems operating in the upper U-NII band, antenna gains up to 23 dBi are allowed with no reduction of output power.

INTEL-1020
U.S. Patent 9,265,043

FCC regulations are constantly changing. The process for a FCC rule change is based on a petition from a private company, individual, or government organization. Depending on the scope of the intended change, the FCC will issue either a Notice of Inquiry (NOI), or a Notice of Proposed Rulemaking (NPRM), soliciting comments from all interested parties. Any individual, private business, or organization can provide comments and opinions. Once the FCC completes its decision making, the new rule is issued in a Report and Order document. This document not only lists the new rule, but discusses its rationale for adopting it.

Before vendors are allowed to sell products, they must demonstrate that their products comply with the relevant regulations. The FCC plays an important role in certifying all equipment. In the past, to help vendors obtain certification, the FCC issued documents to interpret its own rules. Applicants were required to test their equipment and submit a detailed report to the FCC. A subset of the report is kept in a database open to the public. The FCC interpretation database and grantee database provide useful information to equipment manufacturers. Tradition and common practice are also good guidelines.

Recently there have been notable changes to the role of the FCC in equipment certification. The reason is that the whole certification process used to take a couple of months, which is a significant length in a short product lifetime. Some companies have spent significant amounts of money to ensure that their products comply with the latest FCC regulations, only to find competing noncompliant products on the marketplace. The goals of the changes have been to make the process faster and allow innovative products to reach the market sooner, while still providing adequate protection to existing products and services. To achieve this goal, the FCC has been allowing private industry to develop the technical criteria necessary to protect existing wireless networks and services. The current FCC leadership generally feels that having private industry develop the technical criteria better serves the public interest, rather than having the FCC continue to do so. Ultimately, however, the responsibility rests with the FCC.

INTEL-1020
U.S. Patent 9,265,043

The government regulations in CFR 47 do not answer every question. Furthermore, the rules are not always clear. Some rules even may appear to contradict each other.

The Code of Federal Regulations 47 (CFR 47) [B39] is constantly changing in stride with the development of wireless markets and technologies. On one hand, advances in technology lead to changes in government regulations, and on the other hand, in the process of revising the government regulations the goal is to allow the introduction of new devices and services, without causing an unacceptably high level of interference to existing systems.

Among the recent FCC decisions, in February 2002 the Commission adopted landmark regulations that authorized the limited operation of ultrawide band (UWB) devices in the spectrum above 3.1 GHz. Based on these regulations, UWB has the potential to become a commercially successful technology and therefore there is a significant industry interest.[2]

Future FCC directions

Recently the FCC issued a Notice of Proposed Rulemaking to investigate the feasibility of expanding access for wireless broadband services in the frequency band of 2500–2690 MHz, currently occupied by instructional television. In connection with this initiative, the FCC is soliciting comments on the regulations necessary to promote competition and innovation. This development is considered to be significant because the frequency band is twice as wide as the 2.4 GHz ISM band, which is used by the very successful WLAN and WPAN technologies.

Another policy change contemplated by the FCC is to incorporate performance specifications on the receiver interference immunity. These performance guarantees state that interference is below a certain threshold. While the FCC does not yet plan on issuing regulations subjecting all receivers to mandatory standards, it plans to investigate whether market incentives and voluntary industry programs could be more effective for receiver immunity.

[2] In a Memorandum opinion and order issued in 2003, the Commission did not make any changes. In fact, it denied the requests for further restrictions on UWB devices and granted those requests that would not increase the interference of UWB devices, in the FCC's opinion.

INTEL-1020
U.S. Patent 9,265,043

In a Notice of Inquiry released in December 2002, the FCC invited comments on the feasibility of expanding the operation of unlicensed devices in the television broadcast spectrum at locations and times when the spectrum is not being used. Unlicensed devices can monitor the frequency spectrum and avoid the use of frequency bands that are used by other devices. The FCC is also exploring whether unlicensed devices should be allowed to operate in other bands such as the 3650–3700 MHz band, provided that they do not cause interference to licensed services. The FCC believes, and rightfully so, that expanding the available spectrum for unlicensed devices will lead to further development of new and innovative types of products for consumers and businesses.

The FCC assumed its present role in 1934, as a result of the 1934 Telecommunications Act of the U.S. Congress. The rationale of having regulations is, arguably, the scarcity of spectrum and potential for interference. However, some modern critics of the FCC argue that scarcity of spectrum is artificial and induced by regulations favoring the broadcast industry [B37]. Currently, there is talk about the FCC withdrawing entirely from the spectrum allocation task. Subsequently, the present spectrum allocation would need to be completely revamped. There are two specific ideas, not entirely in conflict, how can this change be accomplished. The first proposal, based on advances such as dynamic frequency selection (DFS) and transmitter power control (TPC), calls for allowing every intelligent wireless device access to any spectrum [B37]. Devices that cause interference can change their frequency and power of operation. The second proposal calls for all spectrum, like any other resource in the United States, to be available on the market, bought and sold freely by companies, organizations, and individuals, to achieve maximum economic efficiency [B37]. In this case, government organizations, such as defense and law enforcement, could purchase spectrum on the open market for their needs. Clearly, the idea to completely revamp the present spectrum allocations is revolutionary and most likely will not be implemented in the foreseeable future. However, in the long run this idea might become more attractive, especially with further advances in technology.

International Regulations

Canada provides for ISM bands and general operating modes that are almost identical to the United States [B88]. Canadian Radio Standards Specification

INTEL-1020
U.S. Patent 9,265,043

RSS-210 provides numerical requirements for narrowband operation in 902–928 MHz and 2400–2483.5 MHz and associated harmonic limits that are identical to FCC 15.249 operation. As of the time of this writing, the Canadian requirements for spread-spectrum operation given in RSS-210 are equivalent to the previous (before May 30, 2002) FCC 15.247 requirements. However, it is likely that Industry Canada [B56] will soon adopt the changes recently made in the U.S. regarding "digital modulation."

For general radio regulation in Europe, www.ero.dk provides documents and updates, including general description of recommended applications, frequencies, powers, and other specifications.

The more than 40 member nations of the European Conference of Postal and Telecommunications Administrations (CEPT) have established a fairly high degree of standardization throughout Europe on the operation of low-power radio equipment. Most disagreements concern allowed modes and transmit duty cycles that may be accounted for in software control, allowing the same hardware and technical standards to be used throughout Europe. The European Telecommunications Standard Institute (ETSI) develops technical standards for CEPT countries. Requirements are spread across multiple documents. ETSI EN 300 328-1, 328-2 [B35], [B36] are the European rules for spread spectrum systems, and ETSI EN 300 220-1 provides details on compliance certification [B34]. ETSI 300 328 is in general the formal governing document (though it may be overruled by the regulatory documents of a specific nation). For devices with less than 10 mW EIRP, either the rules of ETSI 300 328 section 5 or the rules of CEPT ERC Recommendation 70-03E Annex 1 may be used. For power levels between 10 mW and 100 mW in the 2400 MHz band, only ETSI 300 328 applies, and direct sequence or frequency-hopping spread spectrum must be used.

In Europe, the 433.05–434.79 MHz band segment is the common control and security band. It is primarily limited to these applications because of a general 10% duty cycle limit (see ERC 70-03E, downloadable from www.ero.dk). Europe does not offer an ISM band in the 902– 928 MHz range, but it does offer a limited band of 868–870 MHz (see ERC 70.03 [B32] for general rules).

INTEL-1020
U.S. Patent 9,265,043

Important differences in the European rules as compared to the FCC rules are provisions that go beyond preventing interference to other systems—a primary goal of the FCC rules. European rules also attempt to guarantee acceptable system performance.

The Electronic Communications Committee (ECC) within the CEPT is currently studying a number of changes to the short-range device rules. A broad description of some European changes proposed at the time of this writing is captured in [B28]. In Europe at the moment there is an initiative to expand the 868–870 MHz band to cover from 863–870 MHz for nonspecific short-range devices using spread spectrum with power levels up to 25 mW. Some relaxation of the duty cycle limits now applying in the 868–870 MHz range is being contemplated. The changes are generally friendlier to short-range device operation. The use of interference avoidance techniques such as frequency agility, dynamic channel assignment, and listen-before-transmit are encouraged.

The 2400–2483.5 MHz ISM band is also provided for in ERC 70-03 [B32] for short-range devices (any digital modulation form) and also in ETSI 300 328 [B35], [B36] for spread-spectrum devices with data rates equal to or greater than 250 kb/s. Note that European regulations are not unified. There are differences among individual countries.

In Japan, the government has chartered an industry organization called Association of Radio Industries and Businesses (ARIB) [B52] to perform certain quasi-governmental regulatory functions in supporting efficient use of the radio spectrum. The governing document is generally standard ARIB STD-T66 [B7]. This document seems incomplete, especially in the area of how to perform testing to ensure compliance with government regulations. It is also not clear from the standard exactly which specifications carry the force of law and which are recommendations. DSSS operation in the 2.4 GHz band is allowed with maximum output power of 10 mW/MHz, equivalent to 10 dBm/MHz. a minimum spreading factor of 5 is required. FH is allowed with maximum output power 3 mW/MHz or 4.77 dBm/MHz, and non–spread spectrum with 10 mW/MHz. Other requirements given in this standard include a maximum antenna gain of 2.14 dB at full power of 10 mW/MHz, though higher gain antennas can be used with a commensurate decrease in maximum transmit power.

INTEL-1020
U.S. Patent 9,265,043

In summary, the 2400–2483.5 MHz band is the only almost worldwide allocation of spectrum for unlicensed usage without any limitations on applications and transmit duty cycle. It provides up to 1 W transmit power in spread spectrum modes in the U.S., up to 100 mW in Europe, and up to 10 mW/MHz in Japan.

Note that spectrum allocation is only one part of the regulations. Spurious suppression requirements vary considerably not only by nation, but also by test methodology and the use of averaging. Because they are lower in frequency and thus accessible in lower capability integrated circuit processes, the European 868–870 MHz and the U.S. 902–928 MHz bands are special cases of regional bands that are still useful even though they are not allocated worldwide. Given that antennas remain approximately omnidirectional, they also provide larger antenna aperture than 2400 MHz. So at a given transmit power, data rate, receiver sensitivity, and reliability level, they will provide greater range.

Unfortunately, government regulations around the world are not identical (see Table 1–1). This creates problems especially in the area of wireless data communications, where devices can easily cross international boundaries. Users should make an effort to configure devices so that appropriate government regulations are always observed. Regulatory bodies around the world generally recognize that it is desirable for regulations to be identical and whenever possible make an effort to achieve this. Unfortunately, it does not seem possible to achieve complete harmonization of the regulations, because this would require the costly relocation of some existing wireless services to different frequency bands.

Table 1–1: Maximum output power in the ISM bands

Frequency band	Region	Maximum output power
2400–2483.5 MHz	USA	1000 mW
	Canada	1000 mW (with some limits on installation location)
	Japan	10 mW/MHz
	Europe	100 mW EIRP or 10 mW/MHz
902–928 MHz	USA and Canada	1000 mW
868–868.6 MHz	Europe	25 mW

INTEL-1020
U.S. Patent 9,265,043

STANDARDIZATION BODIES

The development of large markets for wireless communication services and equipment depends not only on the availability of sufficient spectrum in appropriate frequency bands, but also on the availability of standards. The main reason that large markets are not possible without standards is that big telecommunication service providers want to have the choice of buying equipment from multiple suppliers, rather than having to purchase equipment from a single company. In the absence of a standard, big service providers simply refrain from using the technology. The existence of standards also helps small technology companies to enter large markets by reducing risks associated with providing products such as semiconductor chipsets, software, firmware, etc. Standards can be viewed as voluntary agreements among technology and semiconductor companies, equipment manufacturers, and service providers. Until several years ago, there were no standards for wireless data communication; the total market was small, fragmented, and dominated by several proprietary technologies. The dramatic market growth is a direct result of the emergence of several standards, such as IEEE 802.11a, 802.11b, and 802.11g™. Standards allow equipment from different vendors to work together in a network. Standards create mass markets for equipment, which creates economy of scale for manufacturers. The economic effect of standards is even more significant than simply enabling mass markets. The market for standard-compliant products is characterized by significant competition, which results in lower prices. To counter this trend, many vendors are looking for ways to differentiate themselves in the marketplace, while still offering standard-compliant products. Overall, standards significantly stimulate creativity, innovation, and entrepreneurship.

Because standards enable mass markets, intellectual property required for the implementation of a standard is highly valuable. Accordingly, some of the most valuable patent claims are claims that cover a standard or parts of a standard. Therefore, the intellectual property policy of the standards-making body is very important. In industry consortia, the intellectual property behind their specification may or may not be available for licensing by companies that are not a member of the relevant consortium. This means that non-member companies that want to manufacture and sell products complying with a particular consortium's specification may have to pay license fees, which can create a barrier to market entry. In some cases, the intellectual property may not be

INTEL-1020
U.S. Patent 9,265,043

available for licensing at all. When the intellectual property is available, the conditions under which it can be licensed may create other barriers to market entry. For example, intellectual property can be made available on a reciprocal basis. The specification reflects the intellectual property position and the business models of the members of the consortia. Specifications produced by industry consortia are intended to maximize the revenue of the members of the consortium. The Bluetooth® SIG and HomeRF are industry consortiums that have recently produced specifications for wireless data communication.

With the standards revolution of the last five years came a plethora of standards-making bodies. In general, standards-making bodies are of two types: industry consortium and open organizations. Industry consortium type bodies are inherently closed. The companies that establish the consortium dictate the rules of the standards-making process. Membership in the consortium may or may not be open.

Open organizations, such as standards-development organizations, are designed to develop standards that do not favor any one company, but rather aim to benefit the public good. The IEEE is an open standardization body. Before an IEEE standard is adopted, companies that are known to have essential patents are asked to provide a patent letter of assurance stating that they will either not enforce any of their present or future patents that are required (essential) to implement either mandatory or optional portions of the standard against anyone wishing to comply with the standard, or that they will provide licenses, either without compensation or at reasonable rates, on a nondiscriminatory basis, with reasonable terms and conditions to those seeking to implement the standard. The meaning of "reasonable" is somewhat loosely defined. There is no guarantee from the IEEE that a standard does not violate the intellectual property of other parties, including parties unwilling to license their intellectual property on reasonable terms. Furthermore, only a court of law can accurately determine the meaning of "reasonable" and "nondiscriminatory."

INTEL-1020
U.S. Patent 9,265,043

Within the IEEE, standards for data communications are developed by IEEE 802 community, called the Local and Metropolitan Area Networks Standards Committee (LMSC) and sponsored by the IEEE Computer Society [B57]. IEEE 802 has the charter to develop and maintain global standards and recommended practices for computer communication. Some successful IEEE 802 standards are the IEEE 802.3™ or Ethernet standards [B91], IEEE 802.5™ or Token Ring standards and the IEEE 802.11™ or Wi-Fi standards. They all have been adopted by the ISO/IEC Joint Technical Committee 1 (JTC1) as International standards.

IEEE 802 was formed in February 1980 and has held three plenary meetings a year since then. It is legend that the 802 number was chosen to reflect the year and month it was created, however, this is untrue. The number 802 was simply the next number in the queue.

Although IEEE is rooted in the United States, it enjoys strong international participation at the IEEE meetings, and many standards produced by the IEEE are international standards. IEEE 802 consists of several working groups, which are organized around significant applications. The IEEE 802 standards deal with the physical and data link layers in the ISO open systems interconnection (OSI) reference model. IEEE 802 standards specify the data link layer in two sublayers, logical link control (LLC) and medium access control (MAC). The LLC is standardized in IEEE Std 802.2™ and is common to all IEEE 802 MACs.

Initially, work on wireless data communication took place in Working Group IEEE 802.4™, which worked on token-passing multiple-accessing method. In 1990 the Working Group reached the conclusion that token passing is not a suitable protocol for wireless local area networks (WLAN), and the IEEE 802.11 Working Group was established [B58].

Over the years other MAC and PHY working groups have been established. Some of these working groups did not publish a standard, and other groups became inactive after producing a standard. While not all standards have been successful, it must be noted that at present IEEE Std 802.11™ is the most commercially successful wireless standard in the world.

In 1999 two other wireless working groups were established—IEEE 802.15 [B59] devoted to WPAN and IEEE 802.16 [B60] for wireless broadband access. Membership in these working groups is by individual only and is open to anyone.

INTEL-1020
U.S. Patent 9,265,043

Company membership is not allowed. Only individuals vote. Membership includes representatives mainly from the United States, Canada, Europe, Japan, and Australia. Standards established by the working group first become IEEE standards, and then are submitted to the ISO/IEC for adoption as international standards. The commercial success of IEEE 802.11 and IEEE 802.15 and the work in IEEE 802.16 has made IEEE 802 recognized internationally as the main wireless standards-making body.

Just because an international standard exists does not mean that it will succeed in the marketplace. The success of a standard depends not only on the quality of the technology, but also on business and political reasons. Therefore, there has to a be certain alignment among interested companies on these business and political issues on which the success of the standard depends. In an industry consortium, this alignment is easier to achieve. To achieve this alignment for a standard produced by an open standardization body, companies often form other industrial organizations. Such organizations relevant to IEEE 802.11 and IEEE 802.16 are Wi-Fi [B61] and WiMAX™, respectively. The tasks of these organizations include testing to certify interoperability among products from different vendors and promoting the standard-compliant products in the marketplace.

IEEE 802.11, the Working Group for Wireless Local Area Networks (WLANs), is responsible for developing Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA)–based WLAN standards within IEEE 802. Since IEEE 802.11 was formed in July 1990, it has produced the ISO/IEC 8802-11:1999 (IEEE Std 802.11-1999) standard [B68] with several amendments (once referred to as "supplements"). Amendment IEEE 802.11b defines a physical layer achieving data rates of 11 megabits per second (11 Mb/s) in the 2400–2483.5 MHz ISM band [B70]. The amendment IEEE 802.11a is another physical layer for operation between 6 Mb/s and 54 Mb/s in the unlicensed bands above 5 GHz [B69]. IEEE 802.11b and IEEE 802.11a were adopted in 1999. Recently, however, the IEEE 802.11 Working Group has been working on several other amendments.

IEEE 802.11F™[B72], approved in 2003, is an amendment to the medium-access control layer that enables complex wireless networks to be built. IEEE 802.11e™ is another MAC amendment, which would allow higher Quality of Service [B71]. IEEE P802.11i™ is an amendment to ensure security [B75]. These amendments faced several technical challenges and will likely not be completed in 2003.

INTEL-1020
U.S. Patent 9,265,043

IEEE 802.11h™ is an amendment to IEEE 802.11a, which was completed in 2003 [B74]. IEEE 802.11g™ is yet another physical layer for data rates over 20 Mb/s in the 2400–2483.5 MHz band [B73]. It was also completed in 2003. Table 1–2 provides a summary of the completed and active projects within 802.11.

Table 1–2: 802.11 amendments

Number	Description
802.11a	Physical layer for the 5 GHz UNII bands, 6–54 Mb/s
802.11b	Physical layer for the 2.4 GHz ISM band, 5.5 and 11 Mb/s
802.11c	Supplement to support MAC bridge operation
802.11d	Specification for operation in different regulatory domains
802.11e	Enhancements for QoS (active)
802.11F	Interaccess point protocol
802.11g	Physical layer for operation in the 2.4 GHz ISM band
802.11h	Spectrum and power management enhancements to 802.11a
802.11i	Security enhancements (active)
802.11j	Enhancement to 802.11a for operation in 4.9–5.0 GHz in Japan (active)
802.11k	Radio resource management (active)
802.11m	Technical corrections and clarifications (active)
802.11n	High-throughput enhancements (active)

IEEE 802.15, the standards Working Group for Wireless Personal Area Networks, is responsible for developing standards for short distance wireless networks. IEEE 802.15 was formed in July 1999. The group has four projects and correspondingly four task groups. the first task group, IEEE 802.15.1™, produced a WPAN standard in the 2.4 GHz ISM band with PHY and MAC layers equivalent to the PHY and the MAC layers of Bluetooth [B78]. The success of both Bluetooth and

INTEL-1020
U.S. Patent 9,265,043

IEEE 802.11, which operate in the 2.4 ISM GHz, created interference problems between the two standards. Within IEEE 802.15, Task Group 2 was established to develop guidelines on how these two standards can coexist [B79]. The desire to achieve higher data rates led to the formation of Task Group 3 [B80]. The desire to develop WPAN devices with extended functionality and low data rates led to the formation of Task Group 4 [B81]. Table 1–3 summarizes the 802.15 projects.

Table 1–3: 802.15 projects

Number	Description
802.15.1	WPAN based on portions of Bluetooth v.1.1
802.15.1a	WPAN based on portions of Bluetooth v 1.2
802.15.2	Coexistence of WPAN with other systems in the 2.4 GHz band
802.15.3	High-rate WPAN
802.15.3a	Additional UWB physical layer for 802.15.3
802.15.4	Low-rate WPAN
802.15.4a	Additional physical layer for 802.15.4

IEEE 802.16, the Standards Working Group, is responsible for developing standards and recommended practices for Broadband Wireless Access Networks. IEEE 802.16 was formed in March 1999 [B60]. The group has several projects:

a) Air Interface (including a MAC and a PHY) for operation between 10 and 66 GHz

b) IEEE 802.16a™, an amendment specifying additional physical layers and appropriate MAC modifications for operation between 2 and 11 GHz, including licensed and unlicensed bands

c) IEEE Std 802.16.2™ and 802.16.2a™, which provide a recommended practice for coexistence

d) IEEE 802.16e, being developed at present to support mobile subscribers

INTEL-1020
U.S. Patent 9,265,043

In addition IEEE 802.16 has developed interoperability documentation, which is normally done outside of IEEE 802 standards (see Table 1–4).

Table 1–4: 802.16 projects

Number	Description
802.16	Fixed broadband wireless systems between 10 and 66 GHz
802.16a	Amendment for operation between 2 and 11 GHz
802.16c	Enhancement including system profiles between 10 and 66 GHz
802.16.2	Coexistence between 10 and 66 GHz
802.16/Conf01 **802.16/Conf02** **802.16/Conf03**	Test and conformance specifications (active)
802.16d	System profiles (active)
802.16e	Enhancement to support mobility (active)

Other related working groups have been established recently. The IEEE 802 standards committee decided that the monitoring and active participation in radio regulatory activities around the world are important and established a Radio Regulatory Technical Advisory Group—IEEE 802.18. One of the first priorities of IEEE 802.18 is to monitor the global allocation of a significant part of the 5 GHz band

As the number of wireless standards increased, the issues of coexistence among the different wireless standards became prominent. IEEE 802.19 is a Coexistence Technical Advisory Group (TAG), established in 2002. The responsibilities of IEEE 802.19 include helping other working groups such as IEEE 802.11, IEEE 802.15, and IEEE 802.16 address coexistence issues. IEEE 802.19 may develop coexistence documentation of its own.

INTEL-1020
U.S. Patent 9,265,043

Not all standards are commercially successful. Usually the lack of success is due to business reasons, rather than technology. At present, IEEE 802.11 and Bluetooth (IEEE 802.15.1) are the most successful wireless standards.

IEEE 802.20 is yet another recent working group. Its mission is to develop the MAC and PHY layers for mobile broadband wireless access systems. IEEE 802.20 was established in December 2002 and at present is in the process of accepting proposals.

The European Telecommunications Standardization Institute (ETSI) is another open standardization organization [B54]. Membership in ETSI is open only to companies and organizations, and not to individuals. Membership is open to companies from the entire world, and many American and Japanese companies participate. ETSI has produced two families of wireless standards called HiperLAN and HiperAccess, which are technologies for wireless LAN and wireless broadband access, much like IEEE 802.11 and IEEE 802.16.

On an international scale, a major international standardization body is the International Telecommunications Union (ITU), an agency of the United Nations. The ITU is responsible for communications standards and for treaty-based agreements on spectrum management. Its specific role in spectrum management is to minimize radio interference by establishing international rules standardizing the use of radio-frequency bands. It organizes and administers World Radio Congresses, which are held regularly to update the radio regulations.

WIRELESS CHANNELS

The first step in the design of a radio network is the characterization of the radio channel, because the performance of a wireless network—signal coverage and achievable data rate—heavily depends on the wireless communication channel. To help determine signal coverage and achievable data rate, a link budget calculation is performed. A block diagram of a radio system is shown in Figure 1–1.

INTEL-1020
U.S. Patent 9,265,043

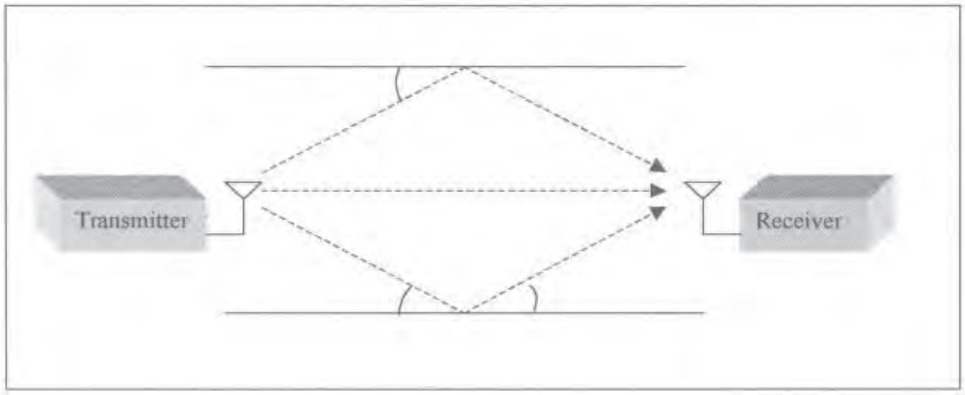

Figure 1–1: A general block diagram of a wireless transmitter and receiver

The transmitting antenna converts the radio-frequency (RF) signal into an electromagnetic wave. The effective isotropic radiated power (EIRP) is equal to the transmitted output power minus cable loss plus the gain the transmitting antenna:

$$\text{EIRP} = P_{out} + C_t + G_t \qquad \text{Eq. 1–2}$$

where P_{out} is transmitted output power in dBm, C_t is the cable loss in dB, and G_t is the gain of the transmitting antenna in dBi. The gain of an antenna is a measure of its directivity. It is the ratio of the radiation intensity in a given direction over the radiation intensity of an isotropic antenna. Isotropic antennas have gain of 0 dBi and radiate equally in all directions. The transmitted electromagnetic wave propagates in the wireless medium. It is intuitively clear that only a portion of the transmitted electromagnetic wave's energy will reach the receiver antenna. The most basic model of RF propagation is known as the "free space model." The free space propagation model is used to predict received signal strength when the transmitter and receiver have a line of sight path between them. The power of such a transmitted signal would degrade with an inverse square rule, or 20 decibels (dB) per decade. More specifically the received power will be given by the Friis free space equation:

$$P_r = \frac{P_t G_t G_r \lambda^2}{C_t C_r (4\pi)^2 d^n L} \qquad \text{Eq. 1–3}$$

INTEL-1020
U.S. Patent 9,265,043

where P_r and P_t are the received and transmitted power, respectively; G_t, G_r are the transmitter and receiver antenna gains; C_t and C_r are the cable losses at the transmitter and receiver; d is the distance between transmitter and receive;, and n is a number that depends on the wireless environment. In indoor environments, n can vary from 1.6 to 3.3, where values smaller than 2 can be achieved in corridors. The wavelength is $\lambda = c/f$, where c is the speed of light and f is the frequency of operation. The wavelength is 12.5 cm and 5.2 cm, correspondingly for the 2.4 GHz and 5.8 GHz bands. L represents a signal loss factor, including transmission losses, filter losses, and antenna losses. $L = 1$ signifies no losses. Assuming no cable losses, the free space path loss equation can be written as

$$P_r = \frac{P_0}{d^n} \qquad \text{Eq. 1–4}$$

where P_0 is the received power at a distance of 1 m and is equal to

$$P_0 = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 L} \qquad \text{Eq. 1–5}$$

Measured in decibels, we have

$$P_r = P_0 - 10n\log d \qquad \text{Eq. 1–6}$$

Assuming that $L = 1$, the power loss at 1 m is approximately 40 dB for 2.4 GHz and 47.5 dB for 5.7 GHz. In free space, n is equal to 2, producing $P_r = P_0 - 20\log d$, which explains where the 20 dB-per-decade rule comes from.

The free-space path loss model is valid for distances that are in the far field of the transmitting antenna, i.e., distances that exceed $2D^2/\lambda$, where D is the largest physical linear dimension of the antenna. Note also that the path loss is random according to a Gaussian (or normal) distribution, and the model provides only the average value. This effect is called log-normal shadowing. The standard deviation of this Gaussian distribution in the frequency bands of interest here is 8 dB for residential indoor applications and between 10 and 12 dB for indoor office applications [B122].

INTEL-1020
U.S. Patent 9,265,043

As an example, if the output power Pt is 30 dBm (or equivalently, 1000 mW) and the transmit and receive antennas have gains of –5 dBi each, and if the cable losses in the cables to the antennas are 1 dB, the EIRP is 30 – 5 – 1 = 24 dBm. For a transmission at 2.4 GHz, the loss at a distance of 3 km is

$$10\log\frac{0.125^2}{(4\pi)^2 3000^2} = 110\text{dB}$$

(assuming $L = 1$), and the received power would be $Pr = 24 - 110 - 5 - 1 = -92$ dBm. In this example, a receiver with sensitivity of at least –92 dBm is required. However, the performance of a wireless communication system is governed by two groups of factors. The first factor is the received power, determined by the transmitted power and the gains and losses on the link. The second factor includes the signal processing used to encode and modulate the signal. The power of the received signal cannot characterize the quality of the received signal. A measure of the quality of the received signal is the signal-to-noise ratio (SNR):

Receivers can operate if the power of the received signal is above a certain minimum value, called receiver sensitivity. Furthermore, because attenuation is random, the received signal power must exceed the receiver sensitivity by a certain margin.

$$\text{SNR} = \frac{P_r}{P_n} \qquad \text{Eq. 1–7}$$

where the noise power is $P_n = KT_{sys}B$, $K = 1.38 \times 10^{-23}$. J/K is Boltzmann's constant and B is the signal bandwidth in Hertz. T_{sys} measured in Kelvin (K) is the equivalent system noise temperature which represents all noise sources: external electromagnetic noise picked up by the antenna, noise generated by the antenna due to thermal emission by the antenna material, and noise generated internally by the receiver.[3] The external noise can be naturally occurring and unavoidable like atmospheric noise or man-made, like the interference from other wireless systems operating in the same frequency band. The biggest source of unavoidable noise is the thermal noise generated in the receiver electronics. The finite precision of the arithmetic operations in the receiver is another noise source. Because at the

[3] If T_{sys} is assumed to be 289 K, the noise power spectral density will be –204 dB/Hz. In addition, practical receivers will generate noise between 5 and 12 dB, and the finite-precision arithmetic operations will lead to an additional implementation loss of typically about 5 dB.

INTEL-1020
U.S. Patent 9,265,043

receiver the SNR must be above a certain threshold SNR_{th}, the path loss will limit the signal bandwidth (and therefore the data rate) according to

$$B \leq \frac{P_t G_t G_r \lambda^2}{SNR_{th}(4\pi)^2 L N_0 d^n} \quad \text{Eq. 1–8}$$

or the signal coverage,

$$d \leq \left[\frac{P_t G_t G_r \lambda^2}{SNR_{th}(4\pi)^2 L N_0 B}\right]^{1/n} \quad \text{Eq. 1–9}$$

where N_0 is the one-sided noise power spectral density.

This free-space model is accurate when there is a line of sight (LOS) between the transmitter and receiver. The line-of-sight condition can be more precisely defined in the following way. The optical LOS is the imaginary straight line between the two antennas. In wireless communications, it can be assumed that a clear LOS exists if a certain area around the optical LOS is free of obstacles. This area is called the Fresnel zone, and its radius is $R = \frac{1}{2}\sqrt{\lambda d}$. When the Fresnel zone is not clear of obstacles, there is not a line of sight, and the additional obstructions in the signal path will cause additional signal loss. The amount of additional signal loss will be random.

Table 1–5 shows the attenuation that some building materials can cause. When the signal path is known, all obstacles with their attenuation contributions can be added to the equation for a total path loss figure. It is this equation that is used when determining the signal strength at a particular point in many of the propagation models.

In addition to attenuation, the transmitted electromagnetic wave in Figure 1–1 will be reflected by walls, ceilings, and other objects if indoors, or buildings and terrain features if outdoors. Therefore, the signal will reach the receiver via multiple paths. This phenomenon is called multipath. If LOS is present, the LOS component arrives first at the receiver and is the strongest. The transmission delay of the LOS component is $\tau = d/c$, which is approximately $3d$ ns, or 3 ns delay per meter. In the absence of LOS, the signal on the first arriving path may not be the strongest. When the signal is reflected from ceilings, walls, etc., part of the signal energy is reflected and part of the signal energy will start propagating through the ceilings, walls, etc. Therefore, the signals on the different paths will have

INTEL-1020
U.S. Patent 9,265,043

Table 1–5: The average additional attenuation of some building materials at 2.4 GHz

Material	Additional Loss(dB)
Open space	0
Window	3
Floor/ceiling	12–15
Wall (wood)	10
Wall (metal and concrete)	15–20

different strengths. Because these paths are of different lengths, the arrival times of the signals on the different paths will be different, as shown in Figure 1–2.

Fading is the result of the randomness with which the multipath components add at the receiver. Fading depends on both the spatial position of the receiver and on the frequency of transmission.

The difference between the arrival times of the first and last paths is called delay spread. Each multipath component is characterized by amplitude and phase. The phases can be accurately modeled with uniform distribution between 0 and 360 degrees. The amplitudes have exponentially decaying Rayleigh distribution. The multipath phenomenon leads to fading—fluctuation in the received signal. Clearly, for some positions of the receiver, the multipath components will arrive with opposite phases and will cancel each other. At these positions, the receiver is in a null, and communication is not possible. Fading also depends on the frequency of transmission. For some frequencies, the multipath components will add destructively, resulting in deep fades or spectral nulls. The complex impulse response of the multipath channel can be represented as in:

$$\sigma_k^2 = \sigma_0^2 e^{-kT_s/T_{RMS}} \qquad \text{Eq. 1–10}$$

INTEL-1020
U.S. Patent 9,265,043

$$\sigma_0^2 = 1 - e^{-kT_s/T_{RMS}} \qquad \text{Eq. 1–11}$$

where $N\left(0, \frac{1}{2}\sigma_k^2\right)$ is a zero mean Gaussian random variable with variance $\frac{1}{2}\sigma_k^2$, and $\sigma_0^2 = 1 - e^{-kT_s/T_{RMS}}$ is chosen in such a way so that the condition $\Sigma\sigma_k^2 = 1$ is satisfied. The number of samples to be taken in the impulse response should ensure sufficient decay of the impulse response tail—for example, $k_{max} = 10T_{RMS}/T_s$. The radio channel is linear, and therefore is completely described by its impulse response. The goal of channel modeling is to estimate the impulse response. Note that the impulse response varies with the relative position of the receiver with respect to the transmitter and also varies with time.

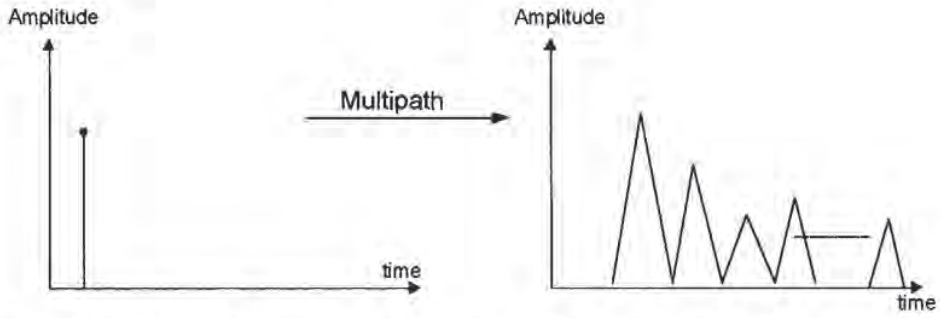

Figure 1–2: A transmitted pulse as a result of the multipath effect will appear to the receiver as multiple smeared impulses

When multiple pulses (or symbols) are transmitted they will start overlapping in time when they reach the receiver (Figure 1–3). In other words, the symbols will interfere with each other, or equivalently there will be intersymbol interference (ISI), as shown in Figure 1–3.

The root mean square (rms) delay spread is often used to characterize wireless channels. The rms delay spread and the symbol period are used to estimate the amount of ISI caused by multipath channel. Because the distances that the individual paths travel are smaller indoors, the rms delay spread for indoor communications is smaller than outdoors. The delay spread tends to increase as the distance between the antennas increases, in which case the path loss increases as well.

INTEL-1020
U.S. Patent 9,265,043

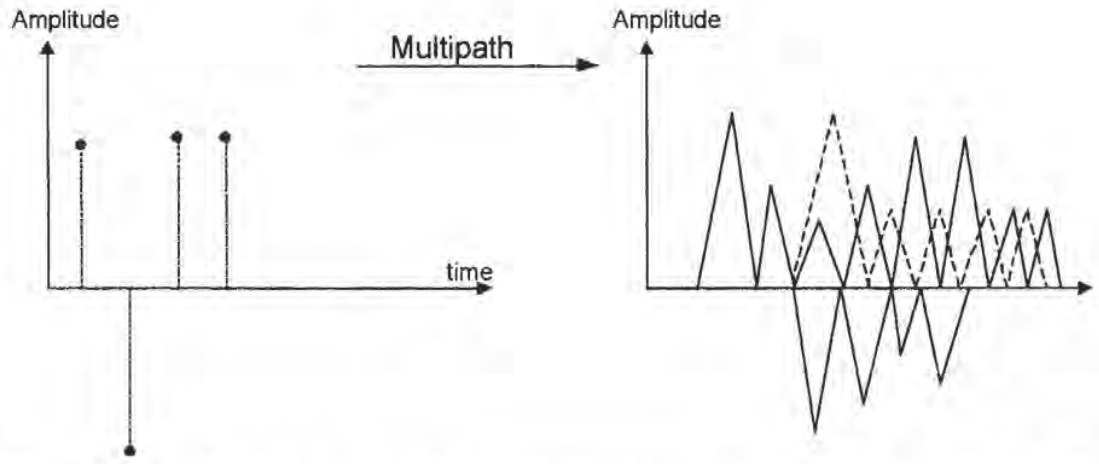

Figure 1–3: Multiple transmitted impulses as a result of the multipath will be received as shown on the right

The characteristics of reflected paths between the transmitter and receiver depends also on polarization, incidence angle, and the material's complex permittivity. Therefore, polarization and antenna radiation pattern can significantly affect indoor propagation characteristics. Directional antennas can reduce the rms delay spread compared with omnidirectional antennas. Circular polarization also can reduce the rms delay spread compared with linear polarization, independent of the frequency of operation. It is finally noted that reducing the rms delay spread is not always desirable, because some communication systems specifically take advantage of multipath effects.

Typical values for the rms delay spread are 50 ns in a home environment, between 100 ns and 150 ns for offices, and about 250 ns for large indoor spaces such as warehouses, airports, shopping malls, and convention centers.

Wireless communication channels clearly change with time as well. This change in general is a result of moving vehicles, objects, people, etc., as well as transmitter and/or receiver mobility. Mobility will lead to an effect called Doppler spread. Doppler spread is a measure of the spectral broadening caused by the channel time variation. The Doppler frequency is $f_D \leq \frac{v}{\lambda}$. For example, for operation at 5 GHz, mobility of 5 mph will lead to a Doppler frequency of 37 Hz, meaning that the channel will change at a rate of 37 Hz. Mobility of 60 mph will

INTEL-1020
U.S. Patent 9,265,043

result in a Doppler frequency of 444 Hz. However, when using high-speed wireless data communication products, users require limited mobility. As a result, it is both convenient and realistic to assume that the channel characteristics do not change within a packet, only from packet to packet. In the case of a laptop computer with a wireless network interface card, the LOS path can be shadowed by people moving in the vicinity of the laptop. Some measurements have observed fades of 10–15 dB in these cases [B122].

Multipath fading together with attenuation and noise limits the achievable signal coverage and data rate in high-speed wireless data communication. To mitigate the channel conditions, wireless standards use the following three methods:

a) Power management

b) Signal processing

c) Antenna diversity

Mobile and portable devices must be battery-powered. Therefore, low power consumption is crucial. The wireless data communication standards use sophisticated power management to reduce power consumption as much as possible. Receivers are generally implemented with equalizers to compensate for ISI caused by the multipath channel. Multicarrier modulation is another technology, very suitable for multipath channels. Advanced antenna systems are used in particular by broadband wireless access systems.

Wireless channels are described in considerably greater detail in [B116] and [B121].

INTRODUCTION TO CRYPTOGRAPHIC ALGORITHMS

The goal in this section is to introduce the specific cryptographic algorithms that are used in the IEEE 802 standards. Cryptographic algorithms are discussed in greater detail in [B132].

INTEL-1020
U.S. Patent 9,265,043

A message to be encrypted is called plaintext and the resulting encrypted message is called ciphertext. The objective of encryption is to be practically impossible for an unauthorized party to understand the contents of the ciphertext. The main types of cryptographic systems are secret-key and public-key systems.

In secret-key cryptographic system, the sender uses a secret key for the encryption. The intended receiver recovers the plaintext using the same key. It is the sharing of a secret that makes the communication secure. This method can be used in authentication. Suppose that two devices want to authenticate each other using this approach. Then one device can send a random number (the challenge) to the other device. The other device will respond with an encrypted version of the random number. The receiver will decrypt the number, and if the result is equal to the original random number, the authentication procedure is considered successful. If mutual authentication is desired, then the second device can issue its own challenge by sending its own random number. It is extremely important that the random numbers used in each challenge be different. Otherwise an eavesdropper can determine the key that is shared by observing the challenges and responses. The term nonce (number once) is used to describe the random numbers.

In public-key cryptography keys are not being shared. It relies on two different keys, a public key and a private key. Encryption is performed using a public key, and decryption is performed by using a private key. Privacy is provided, because only the holder of the specific private key can decrypt the message. Public-key systems can also be used for authentication. In this case, one device selects a nonce, encrypts it using the other device's public key, and transmits it to the other device. The other device decrypts the received message with its private key and responds with the originally selected nonce.

The Data Encryption Standard (DES) was adopted in 1977 by the National Bureau of Standards, now National Institute of Standards and Technology (NIST). DES is a shared-key cryptographic algorithm. In the encryption process, DES first divides the data into blocks of 64 bits. Each block is separately encrypted into a block of 64-bit ciphertext. The key is 56 bits. The encryption algorithm has 19 steps. Each step has a 64-bit input and produces a 64-bit output. The first step is a permutation independent of the key. The last step is a permutation that is the inverse of the initial permutation. The stage immediately before the last simply swaps the 32 bits

INTEL-1020
U.S. Patent 9,265,043

on the left with the 32 bits on the right. Each of the remaining 16 iterations performs the same processing, but uses a different key. The key at every step is generated from the key at the previous step by applying a permutation, circular rotation, and another permutation. Out of the 56 bits, only 48 bits are used for the key at every iteration.

At each iteration the following operation is performed: First, the 64-bit input is divided into two equal portions, L_{i-1} and R_{i-1}. The output L_i is equal to R_{i-1}. The right part of the output R_i is the result of a bitwise XOR of the left part of the input and a function of the key at the given iteration K_i, and the right part of the input, $R_i = L_{i-1} \oplus f(R_{i-1}, Ki)$. This is the electronic codebook (ECB) mode of DES.

ECB mode may not be secure when the structure of a message is known to an attacker, as is typical in communication standards. Improvements to ECB are possible. One way to improve the algorithm is at every step to XOR the current plaintext block with the preceding ciphertext block. The first plaintext block is XORed with an initialization vector (IV). This is the cipher block-chaining (CBC) mode. Another problem with DES is that the 56-bit key is not sufficiently long for protection against brute-force attacks. An improvement called triple DES mitigates this problem. Triple DES uses two keys and therefore the total key length becomes 112 bits. According to triple DES, ciphertext C is produced from plaintext P by

$$C = E_{K1}(D_{K2}(E_{K1}(P))) \qquad \text{Eq. 1–12}$$

The corresponding decryption algorithm is

$$P = D_{K1}(E_{K2}(D_{K1}(C))) \qquad \text{Eq. 1–13}$$

If the two keys are identical, triple DES reduces to single DES.

In 1997 NIST announced a public contest to select the successor of DES. In 2001 the Rijndael proposal by Belgian scientists Rijmen and Daemen was selected as the Advanced Encryption Standard (AES). The Rijndael algorithm encrypts 128-bit blocks of data with keys of 128, 192, and 256 bits long. It can be implemented in a very efficient manner.

INTEL-1020
U.S. Patent 9,265,043

While DES and Rijndael are shared-key algorithms, the RSA algorithm is the most popular public-key algorithm. It is named after its inventors Rivest, Shamir, and Adleman. It is based on the fact that while it is simple to find the product of two numbers, factorization is more complex. The public and private keys are generated first by choosing two large prime numbers p and q and multiplying them ($pq = n$). Then, a number e is found, which is relatively prime to $(p-1)(q-1)$, and a number d is found so that $de = 1 \bmod (p-1)(q-1)$. The public key is $\{e,n\}$, and the private key is $\{d,n\}$. This algorithm guarantees that for every number P smaller than n, $P^{de} = P \bmod n$. The encryption operation is $C = P^e \bmod n$. The decryption operation is $C^d = P \bmod n$.

Secret-key algorithms are computationally simpler. However, key management is complex. Public key algorithms are computationally intensive. Public-key algorithms simplify key management by requiring each user to have a secret private key and a public key that can be freely distributed. In practice, public-key algorithms require certificates to verify that a given public key corresponds to a certain user.

DESIGN OF A WIRELESS COMMUNICATION STANDARD

In the early days of computer networking, different vendors used systems with incompatible architectures. This required customers to stay with a single vendor. To remedy this situation, the International Standardization Organization (ISO) developed a seven-layer model for communication systems. This model is shown in Figure 1–4. The physical layer deals with the transfer of bits over an actual communication channel. The data link layer deals with the transfer of frames. The data link layer includes framing and address information, as well as flow control. The network layer provides for the transfer of packets across a communication network. It deals with finding a path in the network (routing) and handles congestion from temporary increases in traffic. The transport layer deals with the transfer of complete messages. The session layer controls the data exchange. The presentation layer handles the conversion of data from a machine-dependent form to a machine-independent form, and back. Finally, the purpose of the application layer is to provide services to the various applications.

The IEEE 802 standards deal with the lowest two layers in the ISO open systems interconnection (OSI) reference model: the physical and the data link layer

INTEL-1020
U.S. Patent 9,265,043

(Figure 1–4). Furthermore, IEEE 802 specifies the data link layer in two sublayers, logical link control (LLC) and medium access control (MAC). The LLC is situated above the MAC layer. It is standardized in IEEE 802.2 and is common to all IEEE 802 standards. The LLC provides two services to the next higher layer. These are the LLC data service and the LLC management service. The features of the LLC layer are packet segmentation and handshake. The network layer provides functionality to configure and maintain the topology of the network and to interface with the application through the API.

The different wireless standards developed by IEEE 802, like the other IEEE 802 standards, specify the MAC layer and the physical layer. In this section, the general requirements of these two layers are analyzed.

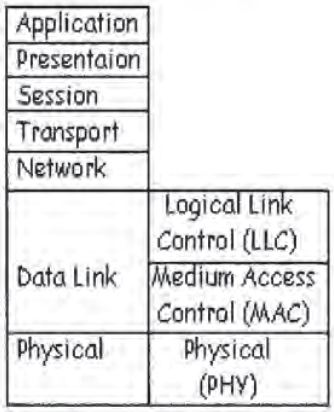

Figure 1–4: Mapping of ISO OSI to IEEE 802

MAC requirements

The access protocol defines the way the medium (the communication channel) is shared among multiple users. The medium in which wireless devices operate is a good example of a shared medium; here the medium is the space through which the radio waves propagate. The ultimate goal of a MAC for wireless communication is to allow a large group of otherwise uncoordinated users to efficiently use this shared medium. Therefore, the choice of a MAC protocol depends on the nature of the traffic and the performance demands of the users.

INTEL-1020
U.S. Patent 9,265,043

Traffic can be classified in two groups: periodic and bursty. When the interarrival variance between messages is very small, traffic is said to be periodic.[4] Signals such as voice and video generate periodic traffic. Periodic traffic requires a limit on the maximum end-to-end delay and the delay variation (jitter). Because the delay depends mainly on the time to assign channel access, the MAC design has a big influence on the signal delay. The data rate of a periodic traffic source is approximately constant. Therefore, the use of a dedicated, circuit-switched connection for periodic traffic is justified. Bursty traffic is characterized by messages of arbitrary length separated by intervals of random duration. Data communication (e.g., e-mail, Internet access, file sharing) in an office environment is an example of bursty traffic. The delay and jitter are not important for bursty traffic. A bursty traffic source leads to a data rate that varies considerably. The peak data rate is much higher than the average data rate. If enough capacity over a dedicated connection is provided to satisfy the peak demand channel, utilization will be poor. Packet-switched connections perform much better for bursty traffic. Because communication systems must support not only Internet access, but also voice and video, both traffic types must be handled. This requirement makes the design of a MAC a difficult task. In addition, in wireless networks, the wireless channel is also the only means to coordinate the stations in the network.

The law of large numbers applied to networking means that the combined requirements of a large number of users is equal to the sum of the average requirements of each user. In other words, although the peak demand for each user can be high, when there are many users, it is the average demand that matters because only a fraction of all users have data to transmit at any given time. Still, if several users attempt transmission simultaneously, the result will be a collision. The MAC must resolve these access contentions.

In general, access methods fall into three general categories: contention methods, polling methods, and time-division multiple-access (TDMA) methods. Contention protocols are also called carrier-sense multiple access (CSMA) or listen-before-talk. If a station has data to transmit, it senses the channel for a certain time period before transmitting. For each station, the length of this time period is random within a predefined interval. If the channel has been free during this entire time

[4] Note that this does not mean strictly periodic in the mathematical sense.

INTEL-1020
U.S. Patent 9,265,043

period, the station transmits. Otherwise, the station defers transmission and waits before again sensing the channel. In this way, the probability of collisions on the wireless medium is minimized, but is not zero. If two stations that have data to transmit choose time periods of equal lengths, collision will occur. In a CSMA with collision detection (CD), which is used in wired Ethernet (IEEE 802.3), the transmitting stations stop the transmission when a collision is detected. An acknowledgment is required for every transmission. Packets that are not acknowledged, due to collision or unsuccessful reception, are retransmitted. Retransmissions reduce the throughput. Under heavy traffic, the end-to-end delay performance of the network becomes poor because there are many collisions and retransmissions. The success of Ethernet demonstrated that contention systems are well suited to bursty traffic. However, contention systems have one fundamental disadvantage: there are no delay guarantees. As a result, purely contention systems are unsuitable for periodic traffic.

In slotted systems, such as TDMA, all users are effectively synchronized and have different time slots of certain duration assigned to them in periodic fashion. Obviously, this is very well suited for periodic traffic. For bursty traffic, however, channel capacity is wasted, because a station will have a time slot assigned to it even if it does not have data to transmit. A significant problem for TDMA protocols is selecting time slot duration or packet size. If the time slot durations are chosen to match the largest message lengths, shorter messages will not use the channel effectively. On the other hand, if shorter slot sizes are used, the delivery of longer messages will require several time slots. To complete the delivery of longer messages, more time will be required. The message size cannot be known in advance and is likely to change dynamically.

Polling, another type of access method, requires a central station. This central station controls the network by polling individual stations. Stations that have data to transmit will send it in response to a poll. Stations that have periodic data to transmit usually can request to be polled on a periodic basis. The central station typically maintains a global queue of requests. Polling is efficient in the sense that it achieves dynamic resource allocation. However, it has disadvantages as well. First, the overhead for maintaining the global queue can be high. The overhead incurred by the channel access mechanism depends on the number of users, unlike contention schemes. Second, all data must pass through the central station, even if

INTEL-1020
U.S. Patent 9,265,043

it is not destined for the central station. Third, polling is not suitable for wireless ad-hoc networks, which do not have a central station.

Access protocols can be classified as static or dynamic. The advantage of static allocation MACs is that each user is guaranteed a share of the network resource. The disadvantage is that resources cannot be transferred from one user to another, i.e., users that have no data to transmit are given the same resources as users with data to transmit. MACs with dynamic allocation try to provide network resources only to users that have data to transmit. Compared with other dynamic-allocation protocols, contention-based MACs are easier to implement. This ease comes from the fact that users can join or leave the network at any time. This advantage is the main reason why wired Ethernet is far more popular than the contention-free token ring network. Such a feature is of paramount importance in wireless networks where stations roam about freely. In general, for light-to-medium traffic loads, dynamic-allocation MACs perform better than static-allocation MACs. Under heavy traffic loads, the end-to-end delay that static-allocation MACs achieve is lower than that of dynamic-allocation MACs.

More detailed discussions of multiple-access methods can be found in the references [B9] and [B121].

Wireless communication protocols are affected by the mobility of stations. There are two types of stations to distinguish between—portable and mobile stations. A portable station is one that can be moved from one place to another place, but normally participates in a wireless network when it is in a fixed location. Fixed or slowly moving stations can be considered portable. Mobile stations can participate in a network while in motion.

Thus, the basic requirements of the MAC are:

a) The MAC protocol must be independent of the underlying physical layer.[5]

b) The access mechanism must be efficient for both bursty and periodic traffic.

c) The MAC must handle mobile users.

[5] In practical systems, it is not possible for the MAC to be completely independent of the physical layer. The MAC must have some knowledge of the underlying physical environment and certain parameters of the physical layer.

INTEL-1020
U.S. Patent 9,265,043

PHY requirements

The PHY is the first layer of the seven-layer OSI model and is responsible for transporting bits between adjacent systems over the air. The PHY performs two functions, depending on whether the device is in transmit or receive mode. In transmit mode, the PHY receives a bit stream from the MAC layer and performs signal processing operations such as error-correcting coding and modulation to convert the bit stream into an electric signal that is supplied to the antenna. In receive mode, the PHY receives the electric signal from the antenna. After signal processing operations such as demodulation and error-correcting decoding, the PHY converts the signal to a bit stream that is passed to the MAC. In addition to these two functions, all IEEE 802 standards require the physical layer to provide carrier-sense indication back to the MAC.

Error control is a form of diversity, similar to diversity transmission and reception.

The wireless channel typically leads to bit error rates much higher than for wired channels. Soft-decision decoding provides higher coding gain and is more important for wireless systems with multipath than in additive white Gaussian noise (AWGN) channels. In addition, errors frequently occur in bursts, coinciding with deep fades on the link. Therefore, error-control schemes are important. Three types of error control coding schemes are used:

a) Block codes

b) Convolutional codes

c) Automatic repeat request (ARQ) schemes

An important point here is that most error-correcting codes are designed to protect against random errors, but not burst errors. A technique that reduces the statistical dependence of errors is interleaving. With interleaving, the symbols contained in one code block are not transmitted in consecutive order, but instead are interspersed among other transmitted symbols so that a signal fade is less likely to impose a dense burst of errors on individual code block. If interleaving can be performed over a sufficiently long time, then errors in individual blocks can be made independent. In practical systems, the latency is limited, and this puts a limit on the interleaver time interval. After decoding to ultimately decide whether there is still an error, cyclic redundancy code (CRC) is performed.

INTEL-1020
U.S. Patent 9,265,043

The general requirements of the PHY are bandwidth efficiency and power efficiency. Because of the constant demand for higher data rates, bandwidth efficiency is becoming increasingly important. The importance of power efficiency varies among wireless networks. WLANs are typically used on portable terminals, which are battery-powered only for a limited time. These terminals are normally powered from the AC power sources. Thus, power efficiency is not very important for WLANs. WPANs, on the other hand, are used on mobile devices, which are battery-powered all the time. Power efficiency is of paramount concern. Class-C power amplifiers provide the highest power efficiency. However, they are highly nonlinear and require a constant-envelope modulation scheme. That is why frequency shift keying (FSK) and variations like GFSK are used in some wireless physical layers.

Power requirements can be divided into two categories: the power needed to operate the electronic circuits and that needed by the power amplifier to supply the antenna with an electric signal of a certain level. The power needed by the power amplifier translates directly into signal coverage. The power needed to operate the electronic circuits depends on the complexity of the signal processing algorithms and the implementation. Only the amount of power needed by the amplifier is usually considered.

PHY layers can be frequency hopping (FH), direct-sequence spread-spectrum (DSSS), or orthogonal frequency division multiplexing (OFDM). Frequency-hopping systems achieve a natural robustness against burst errors. If one carrier is in a fade, the next one is likely not to be in a fade.

Sublayers of PHY and MAC

Both the PHY and the MAC layers can be composed of sublayers. The interface between the MAC and the PHY, like the interface between any two protocol layers, is called service access point (SAP). These service access points are shown in Figure 1–5. A protocol data unit (PDU) is the data unit exchanged between peer entities of the same protocol layer. A service data unit (SDU) is defined as the data unit generated for the next lower layer (in the downward direction) or the data unit received from the previous lower layer (in the upward direction) (Figure 1–5).

INTEL-1020
U.S. Patent 9,265,043

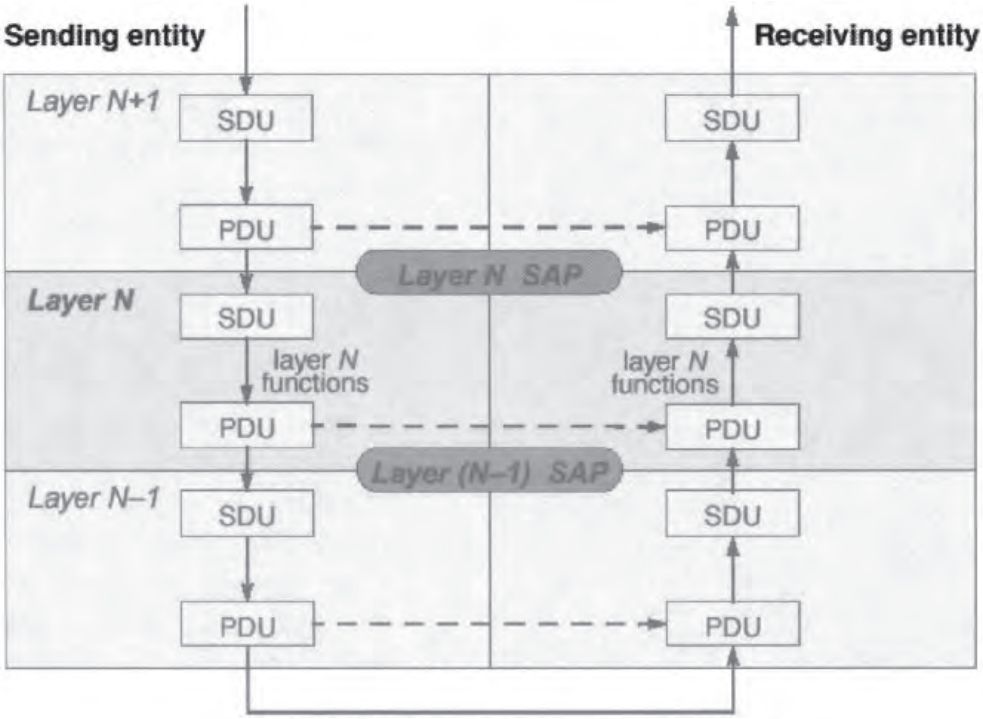

Figure 1–5: PDU and SDU in a protocol stack

The MAC can be considered as providing services to the layer above, and the PHY can be considered as providing services to the MAC layer. In a layered protocol system, the information flow across the boundaries between the layers can be defined in terms of primitives. These primitives represent different items of information and cause actions to take place. These primitives do not appear as such on the air interface; they serve to define more clearly the relations of the different layers. The "request" primitive is an initial request for service from a lower layer. The MAC sends the appropriate dynamic service request message (addition, change, or deletion) across the air link to the receiving MAC sublayer. The receiving MAC sublayer generates an "indicate" primitive to inform the peer convergence sublayer of the request. The peer convergence sublayer responds with a "response" to its MAC, causing it to respond to the initiating side MAC with the appropriate dynamic service response message. Finally, the MAC sublayer sends a "confirm" primitive to the requesting convergence sublayer. At any point along the way, the request may be rejected, terminating the protocol. There are special cases when it is not necessary to send a request over the air link

INTEL-1020
U.S. Patent 9,265,043

to the peer station. In these cases the "confirm" primitive is issued directly by the MAC on the originating side. Such cases may occur, for example, when the request does not depend on the other side. The use of these primitives to provide peer communication is shown in Figure 1–6.

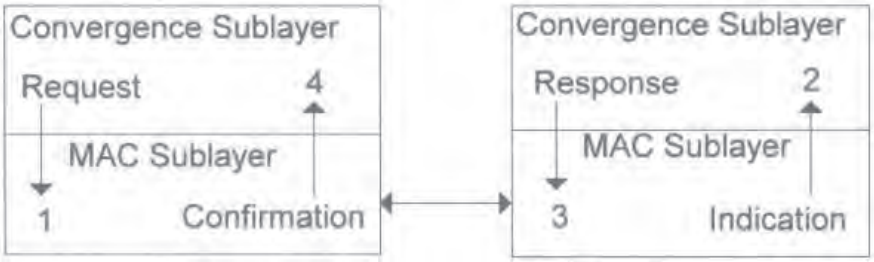

Figure 1–6: Use of primitives to request services

INTEL-1020
U.S. Patent 9,265,043

Copyright © 2004 IEEE

Chapter 2 The IEEE Standard for WLAN: IEEE 802.11

OVERVIEW AND ARCHITECTURE

Work on the first wireless data communications standard began in 1990 when IEEE 802.11 was started as a working group within IEEE 802. The success of Ethernet or IEEE 802.3 and the desire to have "wireless Ethernet" was the motivation for creating IEEE 802.11 as a working group. Wireless Ethernet was originally intended to provide connectivity where wiring was inadequate to support wired LANs. (Ethernet needs special cables called CAT5. CAT5 wiring is not available in older buildings and most homes, nor can it be inexpensively retrofitted.) Another motivation for forming an IEEE 802.11 working group was the allocation of the 2.4 GHz band for unlicensed devices by the FCC in 1986.

The general architecture of IEEE 802.11 is shown in Figure 2–1.

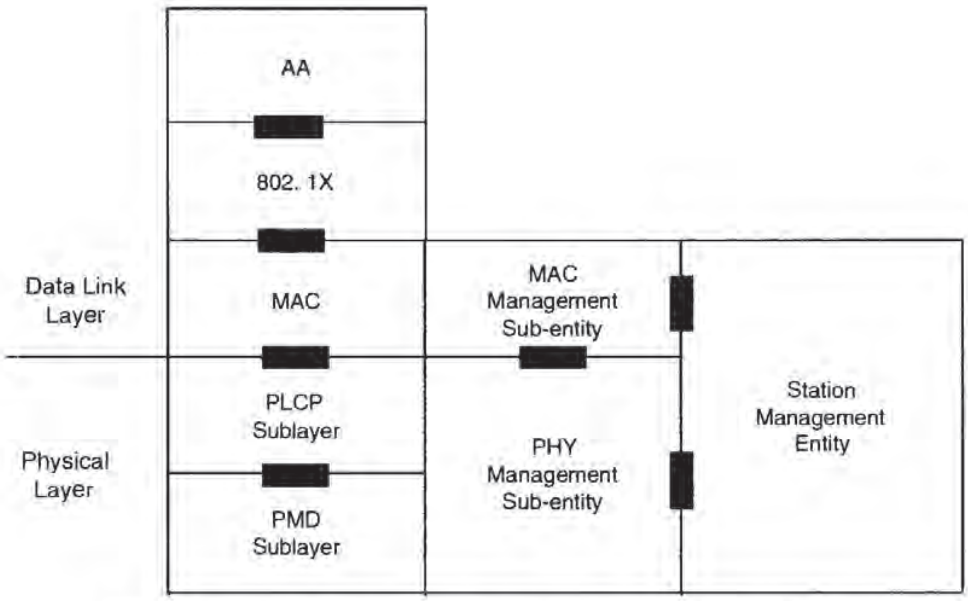

Figure 2–1: General architecture of IEEE 802.11

INTEL-1020
U.S. Patent 9,265,043

Figure 2–1 summarizes the architectural view, emphasizing the separation of the system into four major parts: the MAC of the data link layer, the PHY, IEEE 802.1X, and upper-layer authentication protocols. The data link layer consists of an IEEE 802.1X layer and a MAC sublayer. The physical layer consists of two sublayers: a physical layer convergence protocol (PLCP) sublayer and a physical medium-dependent (PMD) sublayer. The PHY and the MAC have management subentities, which communicate with the station management entity.

The IEEE 802.11 working group spent a lot of time on the medium-access protocol of IEEE 802.11 before deciding on carrier-sense multiple access with collision avoidance (CSMA/CA), similar to the CSMA with collision detection (CD) used in IEEE 802.3. The difference between the two is that in wireless communications, collision detection is not possible. A transmitting station cannot reliably detect collisions because the transmitted signal will be considerably stronger than the received signal. The main goal of the MAC is to make the upper layers unaware of the unreliable nature of the wireless channel. This is important because these protocols do not have the concept of mobility. The standard allows a mobile station to roam freely throughout a WLAN and appear to be stationary to the protocols above the MAC.

Recently the MAC of IEEE 802.11 was developed further. First, it was realized that the security mechanism in the original IEEE 802.11 [B68] was totally inadequate. As a response, part of the IEEE 802.11 group, called Task Group I or IEEE 802.11i, began working on an advanced security mechanism. Also, with the advent of multimedia communications requiring QoS, work began on providing QoS support. The corresponding task group is IEEE 802.11e. With the deployment of the first networks, it also became clear that the access points from different vendors must support a common protocol to communicate effectively. The result is IEEE 802.11f, the interaccess point protocol.

As noted previously, six physical layers are currently defined. Two reasons for this number are the constant need for a higher data rate and that IEEE 802.11 works in several different frequency bands. A frequency-hopping PHY and a direct-sequence spread-spectrum PHY, each capable of operation at 1 and 2 Mb/s, were envisioned in the original standard. In addition, an infrared PHY was also standardized. Before even the first standard was completed, the task group realized that the data rates were too low for the technology to succeed in the

INTEL-1020
U.S. Patent 9,265,043

marketplace. Thus, a new physical layer was created, resulting in IEEE Std 802.11b-1999 The new protocol is able to deliver data rate of 11 Mb/s. The modulation technology in IEEE 802.11b is known as complementary code keying (CCK) modulation. In addition, IEEE 802.11b contains an optional mode, known as packet binary convolutional coding (PBCC). The fifth physical layer, IEEE Std 802.11a, is motivated by the allocation of the UNII bands by the FCC in 1997. Yet another recent work, IEEE 802.11h, adds dynamic frequency selection (DFS) and transmitter power control (TPC) to IEEE 802.11a. The sixth physical layer, completed in 2003, is IEEE Std 802.11g. It provides the same data rate as IEEE 802.11a (up to 54 Mb/s) but works in the 2.4 GHz band. IEEE 802.11g was stimulated by the decision of the FCC to allow digital modulation in the 2.4 GHz band. Thus IEEE 802.11, like any standard in a growing industry, continues to be a work in progress.

The organization of this chapter is as follows. The remainder of this section is devoted to the general architecture of IEEE 802.11. Then the MAC of IEEE 802.11 is presented, followed by a discussion of the physical layers. The MAC presentation is organized as follows. The security mechanism is discussed on page 49, followed by a discussion of the interaccess point protocol on page 66. The section beginning on page 70 is devoted to the QoS mechanism. The section beginning on page 96 is devoted to the physical layers, mainly to IEEE 802.11b, IEEE 802.11a, and IEEE 802.11g. Then the installation of WLANs is discussed on page 117. At the end of the chapter on page 125, current technology and business trends are discussed briefly.

In a way, the goal of the wireless network is to extend the reach of the wired network or backbone.

According to IEEE 802.11, a logical device that participates in a network is referred to as a station. A station consists of a physical layer and a medium access control layer. The basic network is called a basic service set (BSS). One difference with wired LANs is that in wireless networks there are two distinct types of BSSs. The first one is an ad hoc or independent BSS (IBSS). Typical examples of IBSSs are networks formed by personal digital assistants, laptops, CD or DVD players, cell phones, etc. These networks are short lived. The second type of networks, denoted as just BSS, is distinguished by the presence of a special station called access point (AP). The access point allows the network to connect with another network, typically a wired network such as

INTEL-1020
U.S. Patent 9,265,043

Ethernet. The backbone network is typically wired, but can also be wireless. Note that the AP is simultaneously a member of two networks—the wireless BSS and the backbone network. In a BSS, client stations communicate only with the access point. Thus, if one station wants to communicate with another station all information must pass through the access point. 802.11 works this way because in most of the applications, client stations rarely communicate among themselves. A typical application is Internet access, where the client stations use the backbone network through the access point. A group of BSSs can be combined to form an extended service set (ESS). Figure 2–2 shows an ESS, consisting of two BSSs. A roaming station in an ESS needs a handoff protocol, which defines how the APs hand off connections for stations.

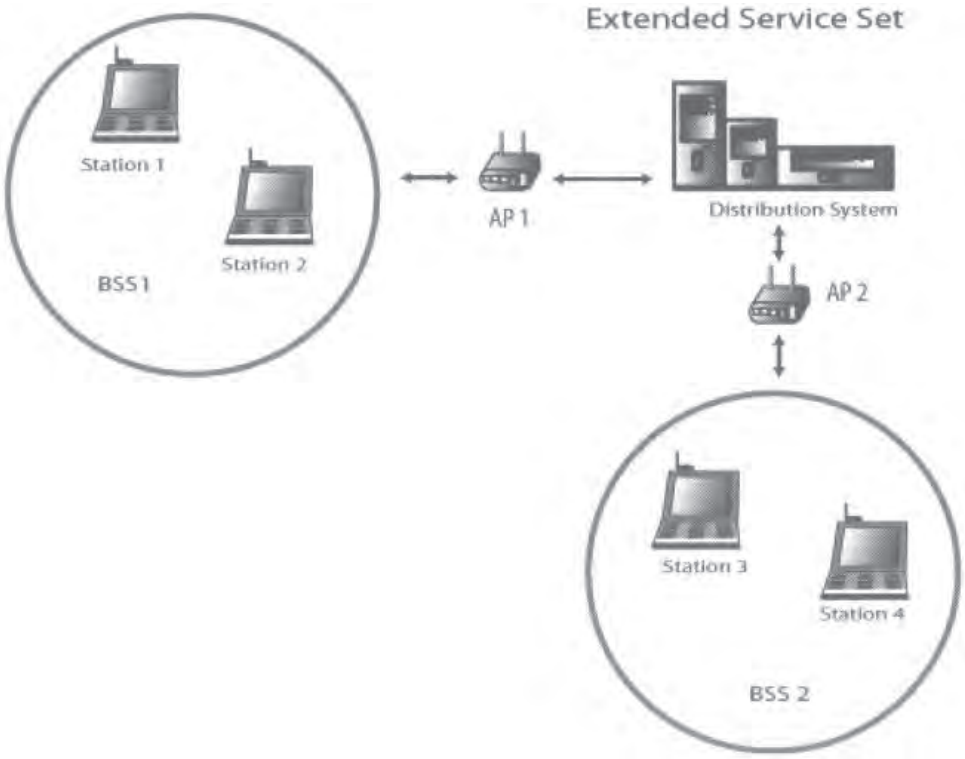

Figure 2–2: Extended service set and distribution system

The standard is designed to support both BSS and IBSS topologies. In a BSS, decision-making is centralized. In an IBSS, decision-making is distributed.

INTEL-1020
U.S. Patent 9,265,043

From the architecture of the IEEE 802.11 network it is clear that two groups of functions are called in the standard services. The first group, called station services, are part of every station. The station services are authentication, deauthentication, key distribution, data authentication, replay protection, privacy, and delivery of data. When there is an access point, another group of services, called distribution services, manages traffic. The distribution services are association, disassociation, distribution, integration, and reassociation. Association, disassociation, and distribution control access and provide data confidentiality. Distribution and integration support data delivery between stations.

The MAC layer exchanges three types of messages: data, management, and control. Management messages are used to support the services. Control messages are used to support the delivery of management and data messages.

IEEE 802.11 SECURITY

In this section, the security problem in wireless networks is discussed first, followed by a discussion of solutions offered by IEEE 802.11.

The problem of network security has attracted considerable attention, as networks are increasingly used in business transactions. The security threats that can arise in any network include the following:

a) Unauthorized users can get access to the network.

b) An unauthorized user can get access to the network by convincing the network that it is one of the legitimate users.

c) Information can be received and decoded by eavesdroppers. This information can be later replayed in its original or possibly modified form.

d) An attacker can flood the network with data, resulting in network overload and denial-of-service to authorized users.

e) An unauthorized user can be the man in the middle, convincing part of the network that it is a client (or station), and convincing the other part of the network that it is a server (or an access point).

Before stations can exchange data, a connection must be established. Scan, authentication, and association are necessary to establish a network. Scanning is

INTEL-1020
U.S. Patent 9,265,043

used to discover existing BSSs that are within range. APs periodically transmit beacon frames that contain timing and network information and can be used to discover BSSs. Immediately after discovering a network, a client cannot join. Before joining a BSS, a client must demonstrate, through authentication, that it is allowed to do so. (In a wired LAN, authentication is replaced by the physical connection to the network. In other words, in a wired LAN, if a station can physically connect to the network, it is authenticated.) All wireless networks require authentication, posing a formidable, but solvable security problem. The security mechanism that IEEE 802.11 uses consists of authentication, key distribution, replay protection, and privacy. The first security mechanism—authentication—is the protocol by which stations prove their identities. However, authentication only is not secure enough, because the wireless medium is still open to an attacker even after authentication. Unless a frame is cryptographically protected by a key tied to the original authentication, it is not possible to determine whether the frame was really sent by the association peer. Encryption and decryption form the privacy mechanism, which will be discussed on page 58. The functions that comprise the security mechanism are discussed next.

Authentication

Historically, IEEE 802.11 has defined three types of authentication: open system, shared key, and upper layer. The first two of these are called link-level authentication and were defined in the first IEEE 802.11 standard of 1997. According to open-system authentication, two messages are exchanged. The first message asserts identity and requests the identity of the other station. The second message returns the result of the authentication—success or failure. Open-system authentication is equivalent to trusting everyone, and therefore does not provide any security. Open-system authentication is sometimes referred to as a null authentication algorithm.

Shared-key authentication intends to do better. The goal of shared-key authentication is to authenticate only those stations that know a shared secret key (Figure 2–3). Shared-key authentication assumes that the shared key is delivered independently of the authentication process. Using encryption and decryption, a challenge text is encrypted at one station, transmitted, and decrypted at the other station. If the result is identical with the original challenge text, the stations do, indeed, have the same key, and authentication succeeds.

INTEL-1020
U.S. Patent 9,265,043

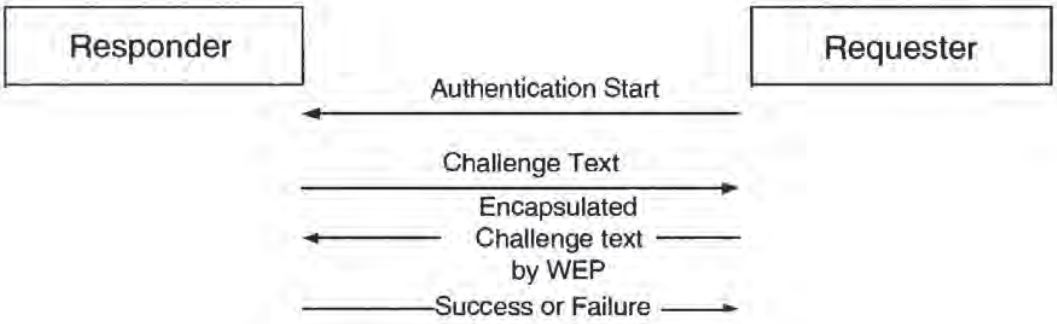

Figure 2–3: Shared-key authentication

Clearly, before shared-key authentication is performed, stations must have at least one common key. These common keys must be distributed through a separate secure channel. The first disadvantage of shared-key authentication is that it is only one-way authentication. At one time, this was acceptable because access points did not require authentication. This left the BSS open to the presence of rogue access points, which clearly compromised security. The second disadvantage of shared-key authentication is the use of static keys. There is no key management. Furthermore, keys are defined as 40-bit quantities, which even in 1997 was too short, but was the maximum key length permitted for export outside the U.S. and Canada by U.S. government regulations.

Shared-key authentication, however, has an even more serious disadvantage. It fails to achieve even its objective—to authenticate stations that possess the secret key. The reason is that it makes public all the information required to reconstruct the key. The conclusion is that shared-key authentication, much like open-system authentication, does not provide any security. The failure of shared-key authentication undermined considerably the trust in wireless networks security. The amendment to the standard, approved in June 2004, as IEEE Std 802.11i™ [B75], provides adequate security, although it may take some time to regain the public's trust in wireless security. To convey the message that wireless is indeed secure, companies have chosen a marketing name that the public can understand. Wireless networks that implement the novel security algorithms will be called WPA (wireless protected access). In this subsection, the security mechanism developed by IEEE P802.11i will be discussed.

INTEL-1020
U.S. Patent 9,265,043

Some background is necessary to understand the development of P802.11i. Because other standards organizations, e.g., the Internet Engineering Task Force (IETF) have worked on advanced security mechanisms, IEEE 802.11 decided to leverage that work. This avoids duplicating functions at the MAC sublayer that are already performed at higher layers. Thus, a robust security network utilizes non-802 protocols for its authentication and key management services. As shown in the IEEE 802.11 architecture in Figure 2–1, these protocols reside above the MAC layer. A robust security network requires three components, which have been developed outside of IEEE 802.11.

The first new component is an IEEE 802.1X Port [B67]. IEEE 802.1X Ports determine when to allow traffic across an IEEE 802.11 wireless link. They reside above the IEEE 802.11 MAC and all data traffic that flows through the MAC also passes through the IEEE 802.1X Port. The architecture of IEEE 802.11 includes an IEEE 802.1X port (Figure 2–1). IEEE 802.1X requires higher-level authentication and provides key management.

Before authenticating any supplicants, the authenticator and AS must authenticate each other. Therefore, it is assumed that a secure channel exists between them. It is important to note that the AS is a logical entity; in real implementations it may be convenient to be integrated into the same physical device as an AP.

The second new component is the Authentication Agent (AA). This component resides on top of the IEEE 802.1X Port (Figure 2–1) at each station and provides for authentication and key management. The Authentication Agent utilizes protocols above both the IEEE 802.1X and IEEE 802.11 layers to provide its services.

The third new component is the Authentication Server (AS). The AS is an entity that resides in the distribution system that participates in the authentication of all stations (including access points) in the Extended Service Set. The AS provides material that every station in the robust security network can use to authenticate every other station. The AS communicates with the AA on each station.

A robust security network, illustrated in Figure 2–4, is a network supporting upper-layer authentication based on IEEE 802.1X. A robust security network has the following characteristics:

INTEL-1020
U.S. Patent 9,265,043

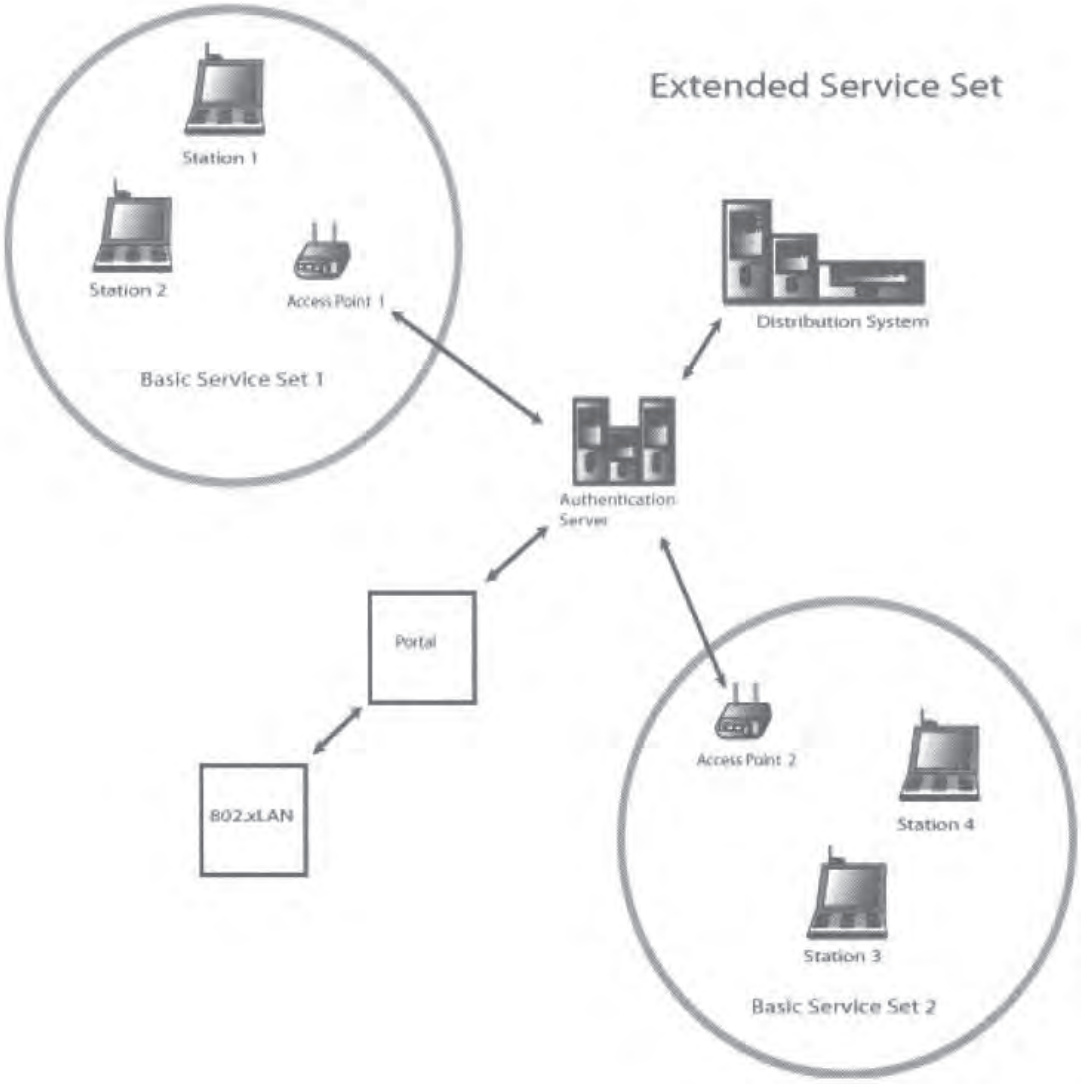

Figure 2–4: A robust security network

INTEL-1020
U.S. Patent 9,265,043

a) Mutual authentication between access points and stations.

b) Key management algorithms (Cryptographic keys age and must be refreshed).

c) Cryptographic key establishment.

d) An enhanced data encapsulation mechanism, called Counter-Cipher-block chaining Message authentication code (CCM), and optionally temporary key integrity protocol (TKIP).

First let's consider what is IEEE 802.1X and how it is used with 802.11. IEEE 802.1X "Port-Based Network Authentication" was originally designed for networks, in which undetected eavesdropping is not possible [B67]. However, the standard does not prohibit operation over shared-media networks. As IEEE 802.11 LANs increased in popularity, the need for a proper authentication and key management presented itself, and it was natural to want to leverage mechanisms already been defined in another IEEE 802 standard (802.1X).

IEEE 802.1X is a standard for port-based network access control. A port is a point at which a system attaches to a LAN. Through its ports a system can access services offered by other systems on the LAN, and also it can export services to other systems on the LAN. The system's ports are controlled to ensure that only authorized systems access its services. IEEE 802.1X-2001 is based on the Extensible Authentication Protocol (EAP)[6] over LANs, also known as EAPOL. EAPOL is used to exchange EAP messages. The EAP entities are supplicant, authenticator, and AS. These entities exchange EAP messages to authenticate the supplicant. A port of a system can be either an authenticator or a supplicant. The port configured to enforce authentication before allowing access to services that are accessible via that port adopts the authenticator role. The port configured to access the services offered by the authenticator's system adopts the supplicant role. The AS performs the authentication function necessary to check the credentials of the supplicant on behalf of the authenticator, and indicates whether or not the supplicant is authorized to access the authenticator's services. As can be seen from these descriptions, all three roles are necessary in order to complete an authentication exchange. Authentication is mutual. Any system can—and in a

[6] The EAP was originally designed to support authentication over the Point-to-Point Protocol (PPP), and is a product of the Internet Engineering Task Force (IETF).

INTEL-1020
U.S. Patent 9,265,043

secure network, must—authenticate any other system before communicating with it. Therefore, any system is capable of adopting the roles of supplicant or authenticator. An authenticator and an AS can be co-located within the same system, allowing that system to perform the authentication function without the need for communication with an external server.

Upper-layer authentication does not occur as a part of IEEE 802.11 authentications, so it does not appear in MAC-sublayer authentication management frames. Instead, IEEE 802.1X implements upper-layer authentication, by encapsulating an authentication scheme outside of IEEE 802.11. The upper-layer authentication mechanism is shown in Figure 2–5.

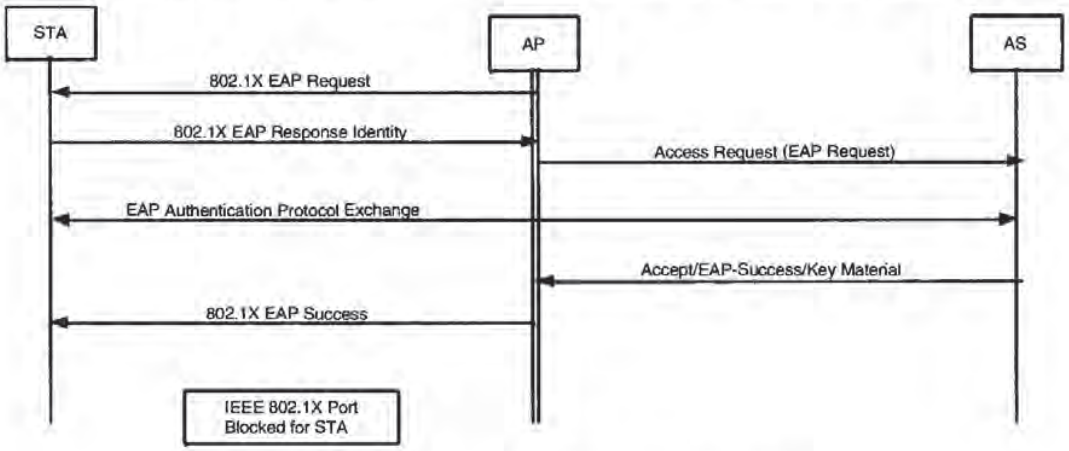

Figure 2–5: EAP Authentication

Because there are three approaches—open, shared key, or upper layer—before authentication, stations must determine which particular approach will be used. In a robust security network, only upper-layer authentication is used, and no authentication operates at the MAC sublayer itself. It should be noted that the choice of an acceptable authentication protocol is an issue for both access points and stations, because the goal of upper-layer authentication is mutual authentication between the access point and the station, not just authentication of the station to an access point. An AP that supports

Mixing the two authentication methods by performing link-level and upper-layer authentication simultaneously is not allowed. A station's attempt to use IEEE 802.11 authentication when upper-layer authentication has been negotiated should be considered as a service attack.

INTEL-1020
U.S. Patent 9,265,043

advanced security advertises its capability by asserting appropriate bits in the security subfield in its beacons and probe response messages. It may also advertise the explicit authentication methods that it supports. In the passive or active scanning process stations detect the capability of the AP and respond accordingly. Typically a client station will choose an authentication method from the list of supported authentication methods by the AP. Because the advanced security mechanism is new, it can only be supported by APs and stations manufactured in the future. Therefore, for some time there will be mixed networks in which some of the stations support robust security measures and some do not.

If an AP does not support advanced security (legacy AP), stations that support advanced security may use the IEEE 802.11 MAC authentication algorithm or may decline to associate. An AP that has robust security capabilities can support legacy stations. Legacy stations ignore the appropriate bits in the Beacons and Probe Responses and do not assert them in their own Association and Reassociation Requests. Similarly, legacy APs do not advertise robust security in their Beacons and Probe, Association, and Reassociation Responses. In this way, new equipment can identify and interwork with legacy equipment. Clear legacy stations can use the legacy MAC sublayer authentication (open or shared key) and then associate. Any station or access point might choose not to associate with another device for many reasons. Networks, in which some equipment supports advanced security algorithms and some does not, are called transition security networks. It must be emphasized that transition security networks are not secure, because compromising the security between devices capable of robust security and legacy devices equipment can compromise even the connections among devices capable of robust security. Only a network that allows robust security network associations is a robust security network

The security mechanisms depend also on the network architecture. In an IBSS, each pair of stations selects the security algorithms to be used, and each station must make its own authentication decision. In an ESS, the AP can enforce a uniform security model. In an ESS, the station initiates all associations. In an IBSS, a station must be prepared for other stations to initiate communications. Thus, a station in an IBSS can negotiate its desired security algorithms when it accepts an association initiated by another station, while in an ESS the AP always chooses the security suite being used. In a secure ESS, the AP may delegate the authentication decision to a physically separate AS, while in an IBSS each station

INTEL-1020
U.S. Patent 9,265,043

must make its own authentication decision regarding each peer. There is no architectural difference between the two; as in the IBSS case, every station implements its own AS.

It must be noted that IEEE 802.11i does not specify a mandatory authentication method. There are several upper-layer algorithms that can be used. EAPOL is one such protocol. Other protocols can also be used as long as they satisfy two requirements: (1) The authentication algorithm must perform mutual authentication, and (2) it must provide for key agreement based on the authentication. A practical requirement towards any algorithm is to be fast enough not to affect the overall QoS. For example, voice over IEEE 802.11 applications require authentication to be completed within approximately 20 ms.

At present there are several Extensible Authentication algorithms used in commercial products, offered by companies such as Cisco®, 3Com, Lucent®, and others. EAP-Cisco (aka LEAP) is password-based, EAP-TLS (Transport Security Layer) is certificates-based, and EAP-PEAP (Protected EAP) and EAP-TTLS (Tunneled TLS) are hybrid, based on certificates and password. The open architecture of the standard ensures that future algorithms can also be used.

Once authentication completes, IEEE 802.1X allows data traffic beyond IEEE 802.1X traffic and stations can associate. Association creates a logical connection between a station and the AP. Once association is established, the AP will buffer, deliver, or forward traffic to the station. A single IEEE 802.1X Port corresponds to one association, and each association maps to one IEEE 802.1X Port. The IEEE 802.1X Port does not permit general data traffic to pass between the station and the AP until after an authentication procedure completes. Either the station or the AP can use disassociation. A station sends a disassociation message when it is leaving the BSS. An AP may send a disassociation message to a station if it has no

Authentication can be terminated by deauthentication. To terminate an existing open or shared-key authentication, the deauthentication service within IEEE 802.11 is invoked. To terminate an upper-layer authentication, deauthentication is also an upper-layer function.

In an ESS using open or shared-key authentication, IEEE 802.11 authentication is a prerequisite for association. Hence, the act of deauthentication causes the station to be disassociated.

INTEL-1020
U.S. Patent 9,265,043

resources to handle this station. Disassociation makes the AP unreachable to the station, and vice versa. Reassociation is similar to association. A station roaming in an ESS uses reassociation to establish a connection with a new AP, identifying in the message the AP that it was previously associated with.

Note that although, logically, association takes place after authentication, in a network with higher-level security, stations are required to associate without any MAC sublayer authentication. In this way, the upper-layer authentication process is allowed to take place above the MAC layer. Authentication packets (contained in IEEE 802.11 MAC data frames) are passed via the IEEE 802.1X uncontrolled port. However, the association exists only for a period of time sufficient for authentication to take place. If authentication is not completed within that time, the station noticing the delay will disassociate. Nevertheless, this seems a weakness in the standard, because the 802.1X port cannot filter out unauthorized packets during this period of time. With this model, IEEE 802.1X, rather than the IEEE 802.11 MAC, makes decisions as to which packets are permitted onto the Distribution System. If the association or reassociation completes successfully with the selection of upper-layer authentication, the IEEE 802.11 MAC passes all data packets it receives from higher layers, delegating the filtering of any unauthorized traffic to IEEE 802.1X.

APs use the distribution service to forward data from a station in its BSS to another station within the BSS, or the ESS, or to a router for delivery outside the WLAN. Integration takes an IEEE 802.11 frame and recasts it as a frame for a different type of a network, such as Ethernet.

Privacy

Privacy (or encryption) is required to achieve security. IEEE 802.11 provides three cryptographic algorithms: Wired equivalent privacy (WEP), Temporal Key Integrity Protocol (TKIP), and Counter-Mode/CBC-MAC protocol (CCMP).

The cipher suite specified in the original IEEE 802.11 is called WEP. The encapsulation algorithm is illustrated in Figure 2–6. Encapsulation here means the way cryptographic data is constructed from plaintext data; decapsulation is the reverse process. The WEP algorithm is a form of stream cipher, in which plaintext is bit-wise XORed with a pseudorandom key sequence of equal length.

INTEL-1020
U.S. Patent 9,265,043

In Basic WEP, the per-packet key (also called a seed) is simply a concatenation of an initialization vector (IV) and a private encryption key. (Figure 2–6.) A new IV is selected for every packet, but the encryption key remains the same. Decryption (not shown) is simply the reverse of the process in Figure 2–6. The same algorithm can be used for keys of different lengths. WEP is only an encryption algorithm, it does not have data integrity keys. Note that shared-key authentication can be implemented only if WEP, which is optional, is implemented.

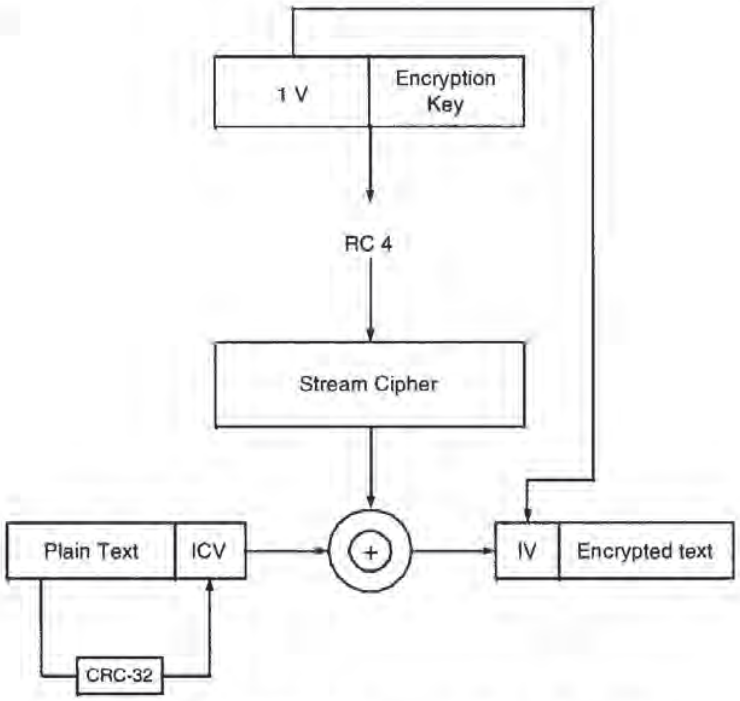

Figure 2–6: WEP Encapsulation algorithm

In basic WEP, not only is the 40-bit key short, but the key remains constant across several frames. Because the same plaintext encrypted with the same key always results in the same ciphertext, it seems possible that the key can be computed. Several researchers investigated this weakness of WEP in 2001 and confirmed that this assumption is correct. Because the number of possible keys is not large enough, a passive attacker can derive the WEP key by only listening to the network and performing statistical analysis. Another approach to derive the key, suitable for an active attack, is to solicit information from the network.

INTEL-1020
U.S. Patent 9,265,043

For example, an attacker can send a text to another station and then will listen waiting to see the ciphertext. The key is simply derived from the ciphertext and the original text. Another possible way is to capture a frame, change a few bits in the frame and resend the frame. This will likely result in a higher-layer protocol error. Again, the attacker waits for the error ciphertext to derive the key. The WEP algorithm provides security only against casual monitoring. It offers no protection against cryptographically sophisticated enemies. Note also that the WEP algorithm itself does not provide data integrity guarantees. It is possible to compute an Integrity Check Value (ICV) and embed ICV in the plaintext before encryption.

In IEEE 802.11 jargon the MIC is also affectionately called "Michael."

TKIP and CCMP are new cipher suites, developed by IEEE Std 802.11i (approved in June 2004). WEP and TKIP are based on the RC4 algorithm. CCMP is based on Advanced Encryption Standard (AES). A means is provided for stations to select the algorithm to be used for a given association. Encryption provides a data privacy function only. In particular, it does not afford any protection against data modification. To provide data authenticity requires the use of a data authenticity mechanism. The way to accomplish data authenticity is to compute a tag, called a Message Integrity Code (MIC), using a keyed cryptographic function. Without the MIC, it would be easy for an attacker to intercept a message, flip a few bits in the message, flip the corresponding bits in the ICV to get it to match, regenerate the CRC, and forward the packet on to the ultimate destination. Note that the attacker would not have to know the encryption key to carry out this attack, and the attacker does not need to decrypt the message to carry out the attack. These message-forgery attacks attracted considerable debate in the IEEE 802.11i Task Group. The MIC protects against that. This tag is transported with the data over an unprotected channel with the data it protects, and its value verified by the receiver, using the same key with the same cryptographic function. If the integrity check value is successful, the responder responds with "success." However, even if an attack fails, it still gives some information about the key. It becomes feasible to recover the key after many attacks against the key. Therefore, a MIC failure indicates an attack. After the first MIC failure, stations are required to disassociate and rekey. To limit the probability of a successful attack, any station that detects two MIC failures within 60 seconds must stop all communication for 60 seconds.

INTEL-1020
U.S. Patent 9,265,043

Notice that the cryptographic checksum is a one-way function in the sense that it is extremely difficult to find a message to produce a given checksum.

Stations using CCMP or TKIP can authenticate the data origin and can detect replays. Authenticating the data origin means that when receiving data from another station, the recipient can determine whether the data actually was sent by that station. The replay detection mechanism detects whether the frame is an unauthorized retransmission.

TKIP ensures strong encryption by including (1) dynamic key management, where the key for every packet is different, and (2) a 64-bit message integrity check (MIC).

A 64-bit MIC provides significant protection against message forgery. There are possible MICs and if we assume a message size of 1000 bits, there would be 2^{1000} possible messages. This means that $\frac{2^{1000}}{2^{64}} = 2^{936}$ messages produce the same MIC. This does not mean that it is easy to find a message with the same MIC. One needs to make 2^{64} attempts to find a message that will produce a given MIC. If it takes 1 μs to generate a message and compute its MIC, it will take over 10^{12} years to find a message that will produce a given checksum.

TKIP was designed for hardware devices supporting WEP only. IEEE 802.11 does not recommended using TKIP. TKIP modifies WEP as follows:

a) To defend against forgery attacks, there is a MIC.

b) There is a TKIP sequence counter (TSC). The receiver drops frames received out of order. This is a weak form of replay protection.

c) TKIP uses a cryptographic mixing function to combine the TSC, a temporal key, and TA into the cryptographic key.

The encapsulation and decapsulation mechanisms are shown in Figure 2–7 and Figure 2–8.

IEEE Std 802.11i requires using the Counter-Mode/CBC-MAC protocol (CCMP). It provides all four security services: authentication, confidentiality, integrity, and replay protection. CCM employs the AES encryption algorithm using the CCM mode of operation. The CCM mode combines counter mode (CTR) for confidentiality and Cipher Block Chaining Message Authentication Code (CBC-MAC) for authentication and integrity. (The AES is described in [B113].)

INTEL-1020
U.S. Patent 9,265,043

The AES is used with a 128-bit key. CCM is a generic mode that can be used with any block-oriented encryption algorithm. CCM needs two parameters: the size of the MIC M, and the length of the frame. CCMP supplies the values $M = 8$, meaning the MIC is eight bytes, and $L = 2$, meaning that the length of the frame can be represented as a two-byte number. CCM requires a fresh temporal key (TK) for every communication session and a unique number, called a nonce, for every frame. A nonce is a number that is pseudo-randomly generated and does not repeat within a frame. If a nonce repeats within a frame, all security protection is lost.

The CCMP encapsulation process, shown in Figure 2–9, consists of the following steps:

1) The packet number (PN) is incremented for every frame. The PN must never repeat for the same TK.
2) Additional authentication data (AAD) is computed from the MAC header. In this way the CCM algorithm provides integrity protection for the fields included in the AAD. MAC Header fields that may change when retransmitted are excluded when the AAD is computed.
3) The CCM nonce is constructed from the PN, A2, and the priority of the frame.
4) The CCMP header is computed.
5) The ciphertext and the MIC are formed from the TK, AAD, nonce, and plaintext data. CCM encryption needs four inputs: a 13-byte nonce, key, AAD, and frame body. The encapsulated frame is formed by concatenating the original MAC header, the CCMP header, the encrypted data, and the MIC.

INTEL-1020
U.S. Patent 9,265,043

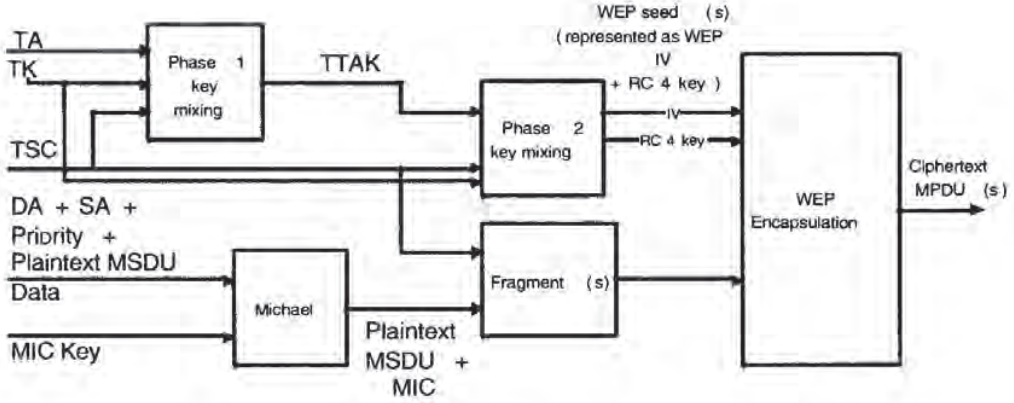

Figure 2–7: TKIP encapsulation diagram

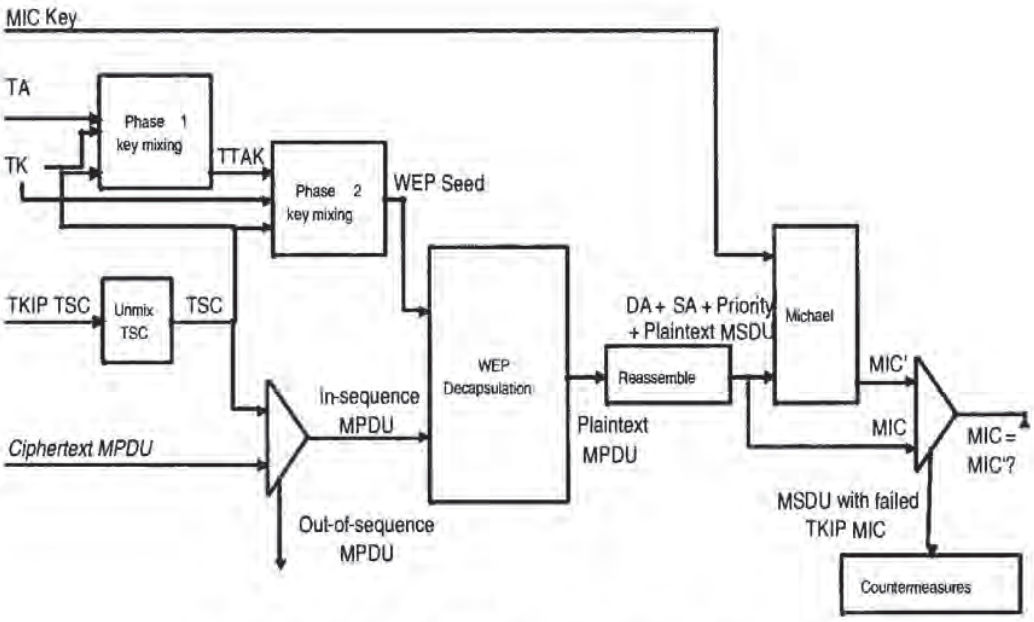

Figure 2–8: TKIP decapsulation diagram

INTEL-1020
U.S. Patent 9,265,043

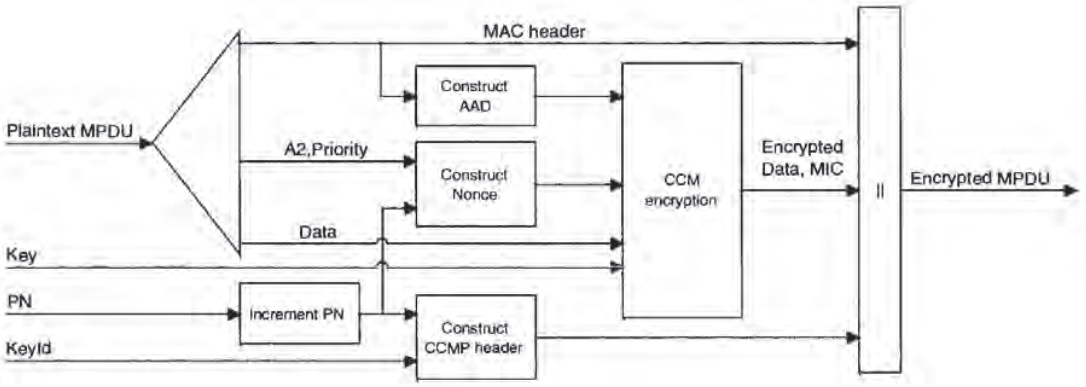

Figure 2–9: CCM encapsulation process

In connection with the traffic classes, introduced by the IEEE P802.11e QoS[7] mechanism, discussed in the following section, the receiver maintains a separate replay counter for each IEEE 802.11 traffic class. The PN recovered from a received frame is used to detect replayed frames. A replayed frame occurs when the PN extracted from a received frame is repeated or not greater than the current traffic class replay counter.

The decapsulation process (Figure 2–10) is based on the following steps:

1) The received encrypted frame is parsed to construct the AAD and nonce.
2) The MIC is extracted for use in the integrity check process.
3) The temporal key (TK), AAD, nonce, MIC, and ciphertext are used to recover the plaintext, as well as check the integrity of the AAD and the plaintext.
4) The receiver extracts the packet number to detect replayed frames. The decryption processing prevents unauthorized replay of frames by validating that the packet number is greater than the packet number maintained for the session. A frame is replayed when the packet number repeats.

[7] IEEE P802.11e is not yet an approved standard [B71].

INTEL-1020
U.S. Patent 9,265,043

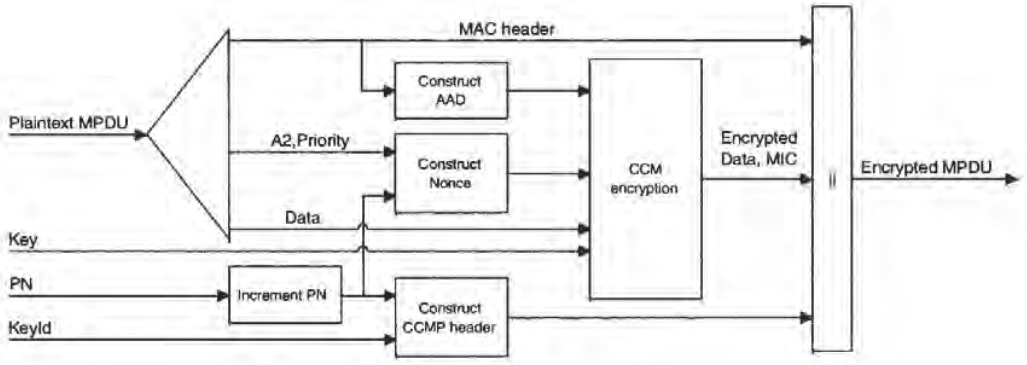

Figure 2–10: CCM decapsulation process

IEEE 802.11 uses the notion of a security association to describe secure operation. A secure association has a finite lifetime. As noted before, the quality of protection with any cryptographic algorithm degrades through key usage. Dynamic key management ensures that the cryptographic keys are fresh. Therefore, a fresh, never-before-used key is required whenever a new "session" begins, so that keys really cannot be used independently of some notion of a session. Fresh cryptographic keys are required for enhanced confidentiality, data authentication, and replay protection. The replay protection counter assumes that peers somehow synchronize a fresh key whenever they reinitialize their replay state, again to avoid having to maintain history of every data frame received.

In an ESS, after a station selects an AP with an appropriate SSID, it uses open-system authentication, followed by association. Negotiation of the security parameters takes place during this stage. Then mutual IEEE 802.1X authentication takes place, followed by key exchange. The AP and the supplicant use two key confirmation handshakes, the four-way handshake and group-key handshake, to indicate that the link has been secured by the keys so is safe to allow normal data traffic. The handshake is performed using EAPOL. In addition to the pairwise keys used for unicast frames, there are group keys used for broadcast messages. Once the keys are known to be established by both association peers, the MAC must discard any data received over the association that is unprotected by the encapsulation algorithm. Without this provision, it is trivial for an attacker to forge a valid message by simply sending a cleartext message.

INTEL-1020
U.S. Patent 9,265,043

As shown in Figure 2–5, the authentication process creates cryptographic keys shared between the supplicant and the AS. A fresh, never-before-used pairwise master key (PMK) is generated and distributed to the AP. The PMK is 256 bits. From the PMK, a Pairwise Transient Key (PTK) is generated. The PTK is partitioned into EAPOL-Key MIC and encryption keys, and temporal keys used by the MAC to protect unicast communication between the authenticator and the supplicant.

Note that roaming must be considered with the security policy. When a station roams in an ESS and uses reassociation, the station must remove the cryptographic key shared with the old AP. A roaming station can cache several keys for several APs, and in this case it will not have to do a complete IEEE 802.1X authorization when it reassociates with each.

The IEEE 802.11 Advanced Encryption Standard Privacy algorithm identifies unicast key "sessions" with IEEE 802.11 associations and uses (re)association messages for the synchronization function, and a random nonce exchange and key derivation to refresh keys. An IBSS desiring to use the Advanced Encryption Standard Privacy algorithm must therefore implement (re)association messages.

The Advanced Encryption Standard Privacy algorithm architecturally lies above the IEEE 802.11 retry function. This is required because data may be accepted by the local IEEE 802.11 implementation, but its acknowledgment may get lost in transit to the peer. If the Advanced Encryption Standard Privacy algorithm were to lie below the IEEE 802.11 MAC retry function, then it would be impossible to recover from this state because the replay protection function would discard all further retries.

INTERACCESS POINT PROTOCOL (IAPP)

Confining stations to a single BSS has little benefits. It is much better to allow stations to roam within an Extended Service Set (ESS), as illustrated in Figure 2–11.

To achieve roaming capabilities, the APs must communicate among themselves. The APs are connected via an abstract layer, called distribution system (DS). The protocol that is used across the DS is new and is called interaccess point protocol.

INTEL-1020
U.S. Patent 9,265,043

This protocol, defined in IEEE Std 802.11F, which was approved in June 2003, is the subject of this section.

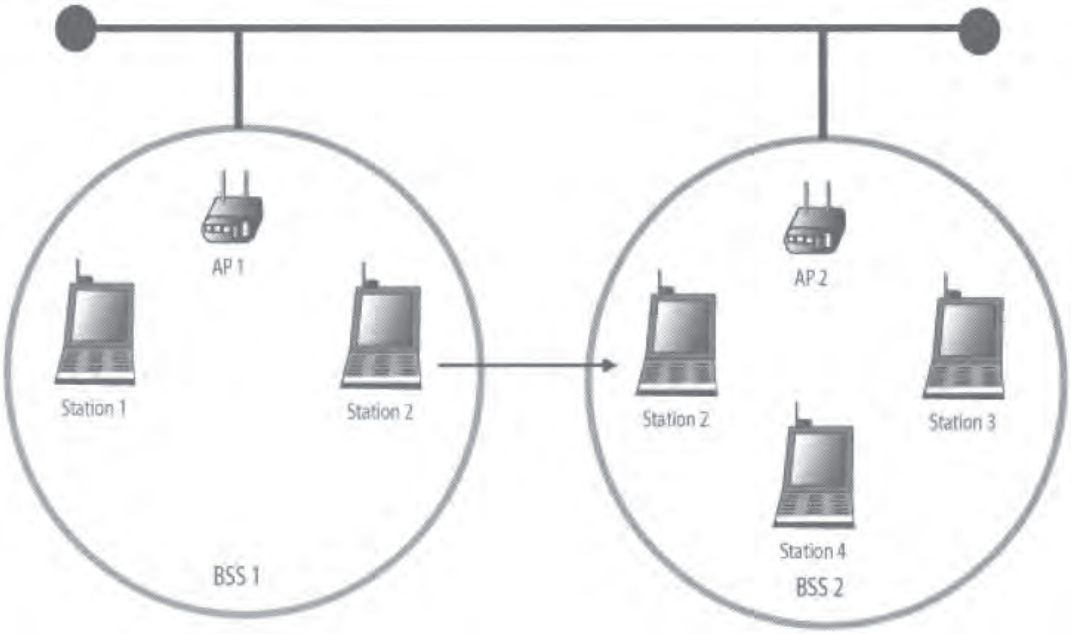

Figure 2–11: A station roaming within an ESS

Initially a protocol for AP-to-AP communication was not defined by IEEE 802.11. While leaving this protocol to vendors allows great flexibility in DS and AP functional design, APs from different vendors may not work together. In particular, the requirement that a mobile station has a single association at any given time is unlikely to be enforced. As IEEE 802.11 systems have grown in popularity, this limitation has become an impediment to WLAN market growth.

The IAPP affects not only APs. It also affects other devices in the network such as layer 2 networking devices, (e.g., bridges[8] and switches). From a business perspective, this gives opportunities for vendors to differentiate themselves in the market.

[8] Large local-area networks are often divided into segments. Bridges are devices used to connect these segments. Access points as members of an Ethernet network could be connected by bridges. Bridges are similar to repeaters. However, repeaters work at the physical layer, while bridges work at the MAC layer. The devices that perform a similar function at the network layer are called routers.

INTEL-1020
U.S. Patent 9,265,043

The first goal of IAPP is to ensure that a station is always associated with, at most, one AP, as it roams within the ESS. The goal of single association is achieved using a registration protocol developed by the IETF, called Service Location Protocol (SLP). The registration service can be hosted on an independent server or can be contained within an AP. The second goal of IAPP is to make the APs interoperable. The packets between APs are carried over using the user datagram protocol over IP (UDP/IP). Another objective of the IAPP is not to make IAPP a routing protocol. In other words, the IAPP should not directly deliver the data frames to the stations. Instead, the DS uses existing network functionality to deliver data frames. Therefore, the IAPP requires IEEE 802.11 stations to maintain a network layer address, e.g., IP address. When stations maintain a valid network layer address, with respect to forwarding traffic to stations, APs function much the same as IEEE 802.1D bridges. Additionally, the IAPP maps the address of APs (their BSSID) to DS network layer addresses, i.e., to their IP addresses.

The main operational program of the AP is called AP management entity (APME). The IAPP communicates with the local AP management entity. Several IAPP service primitives are defined that allow the AP management entity to cause the IAPP to perform some function or to communicate with other APs in the DS or a registration service. Several other service primitives indicate to the AP management entity that operations have taken place at other APs in the DS that can have an effect on information local to the AP. The AP management entity must be able to find and use a registration service. It must register as part of an ESS and look up the IP addresses of other APs in the ESS. The AP management entity uses the Service Location Protocol to accomplish this.

The IAPP provides services to an AP in which it resides through the IAPP Service Access Point (SAP). The SAP allows the management entity of the AP to invoke IAPP services and receive indications of service invocations at other APs in a single ESS. Recall that Figure 1–6 shows the four service types that exist at the SAP—requests, confirms, indications, and responses. Service requests and responses are issued to the IAPP by the entity at the next higher layer. The next higher layer is understood to be APME. Service confirms and indications are issued by the IAPP to the AP management entity.

Suppose a station is roaming in an ESS, as in Figure 2–11. Association with a new AP can be accomplished by two commands – association or reassociation.

INTEL-1020
U.S. Patent 9,265,043

If association is used, the address of the old AP is not specified and remains unknown by the IAPP. If reassociation is used, the station can inform the new AP of the address of the old AP, because the reassociation command contains the address of the AP the station was previously associated with. The IAPP works differently, depending on whether the roaming station chooses to use association or reassociation. Every time the station associates with a new AP, the APME generates an appropriate request to the IAPP. In turn, the IAPP attempts to enforce the single-station association requirement by sending an IAPP packet called ADD-notify to the DS. This packet notifies other APs and layer 2 devices (bridges) about the association.

This notification causes two changes. First, the forwarding tables will update. All future traffic for the newly associated station is forwarded to its new AP. Second, every AP that receives the ADD-notify packet checks whether the roaming station was not previously associated with it. If this is the case, the AP disassociates with this station and removes any context stored for the station.

A different request is issued when a station reassociates with an AP, since in this case the AP with which the station was previously associated is known. Again, forwarding tables are updated and the DS is notified of the reassociation. This time, however, the AP notifies the IAPP of the BSSID of the AP with which the station was previously associated. The IAPP translates the BSSID to the IP address of the AP. Using the IAPP protocol, a packet called MOVE-notify is sent over UDP and IP. This packet is sent from the AP directly to the old AP with which the reassociating mobile station was previously associated. This guarantees that the old AP will be notified of the new association.

All information stored at the old AP is forwarded to the new AP. The IAPP relies on a station making use of the Reassociation Request when roaming from one AP to another, instead of an Association Request. The reason is that with an Association Request, the IAPP does not specifically notify the old AP of the new association. This may result in the old AP maintaining context for the station that has roamed to a new AP for a longer time than is necessary. This maintenance not only wastes resources at the old AP, but, more importantly, it violates the single-AP association requirement of IEEE 802.11. The reason the old AP may not be notified of the new association is that the ADD-notify packet is sent, using the IAPP protocol over UDP and IP, on the local LAN segment. The packet is sent to

INTEL-1020
U.S. Patent 9,265,043

the subnet broadcast address, so that it will reach every device on the DSM local subnet. If the old AP is not on the local subnet, it will not receive this packet.

Note the security implications of roaming. When a station reassociates with a new AP, it must first authenticate with the new AP. If authentication with the new AP fails, the station is still authenticated and associated with the old AP. The old AP will not be informed of the unsuccessful authorization attempt. It is possible for authentication with the new AP to succeed, but reassociation to fail. In this case, the network will consider the station associated with the old AP.

While the IAPP is intended to transfer context among IEEE 802.11 access points, it can also be used to transfer context between access points supporting other standards, such as within the IEEE 802.15 and IEEE 802.16 families of standards. Furthermore, it can be used to support the transfer of authentication, authorization, and accounting (AAA) context between devices supporting IEEE 802.1X authentication. The IAPP is independent of the security scheme defined in IEEE 802.11. Authentication information is not carried between the APs. The message exchanges of the IAPP are not protected nor authenticated beyond the protection and authentication afforded by other protocol traffic in the ESS, unless those services are provided by the protocol carrying the IAPP messages. Usually, using IPSec to carry the IAPP messages would provide this protection.

IAPP, however, supports mobility among APs within a single ESS. Higher-layer handoff is required to support mobility among multiple ESSs.

MEDIUM-ACCESS MECHANISM AND REAL-TIME TRAFFIC OVER IEEE 802.11

When IEEE 802.11 started work on the MAC in 1990, it was envisioned that it would be used for computer data communications primarily; therefore, real-time data such as voice and audio was not given much consideration. Because of its CSMA nature, the MAC is optimal for computer data communications, but not for real-time data. Recently, however, data communications systems have been used increasingly to transport real-time data such as voice, audio, and video. All telecommunications service providers are looking at ways to generate additional revenue from their communication networks. Providing different levels of quality of service (QoS) is widely recognized as one of the most important ways to

INTEL-1020
U.S. Patent 9,265,043

achieve this. The nature of real-time traffic is different from data applications. The main QoS factors are:

- Bandwidth
- End-to-end delay or latency
- Latency jitter
- Signal quality and packet loss

QoS is also a business issue, and the business requirements are:

- Capability to offer service differentiation
- Service availability.

Real-time data has strict throughput, delay, and jitter requirements. For example, to maintain acceptable quality of a human conversation the latency for voice packets must be less than approximately 250 ms or 200 ms, and must remain constant. If a packet arrives with a delay exceeding this amount, it can be discarded, because it loses its value. Latency is usually addressed first by signal processing and protocol improvements. Jitter can be removed by buffering, though a large buffer may create latency problems. Networks that satisfy these requirements are said to provide QoS. Real-time data introduces different requirements towards the network.

Not all of these requirements are stricter. Real-time traffic relaxes certain requirements imposed by data traffic. In particular, real-world signals such as audio and video are somewhat tolerant of bit errors. Voice quality and video quality, as perceived by the human auditory and visual systems, are not severely affected by occasional bit errors. Therefore, real-time data tolerates much higher Bit Error Rate (BER) than asynchronous data. Data packets, of course, do not tolerate any bit errors.

Note that at present networks, including IEEE 802.11, do not take advantage of this property. In particular, if a packet has a few bit errors, it will be discarded and not passed on to the higher layers, even though the voice or video signal quality might be acceptable to a human recipient.

The QoS mechanism for WLANs is very different from the QoS mechanism in other networks in several respects. First, QoS is an objective, not a guarantee. It may be impossible, or may become impossible, for

INTEL-1020
U.S. Patent 9,265,043

the requested service quality to be provided. Second, due to the characteristics of the wireless medium, broadcast and multicast data may experience a lower quality of service, especially with regard to loss rate, compared to that of unicast data frames. Third, it becomes necessary for higher layer protocols to become "WLAN aware," at least to the extent of understanding that the QoS characteristics of a WLAN are subject to frequent, and sometimes substantial, dynamic changes. This seems to violate a fundamental objective of IEEE 802.11, but nevertheless is mandated by the wireless environment.

In this section, the different timing intervals that play a role are defined. Then the basic access mechanism for IEEE 802.11 is described. After this, the recently defined QoS facility is described.

For the medium-access protocol timing is very important. The time interval between frames is called interframe space (IFS). There are five timing intervals. From the shortest to the longest they are: short IFS (SIFS), point coordination function IFS (PIFS), distributed IFS (DIFS), arbitration IFS (AIFS), and extended IFS (EIFS).

SIFS and the slot time are fixed by the PHY. These timing intervals for the three main physical layers of 802.11 are given in Table 2-4. One reason for having SIFS is that at any given time, a device can only transmit or receive, but cannot do both. Furthermore, it takes some time for a device to switch from receive to transmit mode and back. This amount of time is called turnaround time. SIFS is equal to the sum of the turnaround time of the RF transceiver, the MAC processing delay, the PHY processing delay, and the RF delay.

PIFS is SIFS plus one slot time. The slot time is the sum of the time for the clear channel assessment, the turnaround time, the MAC processing delay and the air propagation time.

DIFS is SIFS plus two slot times. The DIFS is the time that separates the start of the contention window from the end of the previous transmission.

The EIFS is the sum of SIFS, DIFS, and the length of time it takes to transmit an ACK Control frame at the PHY's lowest mandatory data rate. The EIFS is long enough to protect the ACK from colliding with a transmission from a station that was unable to update its network allocation vector (NAV).

INTEL-1020
U.S. Patent 9,265,043

Figure 2–12 illustrates the relation between the SIFS, PIFS, and DIFS as they are measured on the medium. Arbitration interframe space (AIFS) is defined only for the QoS facility. The different interframe spaces are independent of the bit rate. It is convenient to think of these time periods as time gaps on the medium. They are fixed for each physical layer.

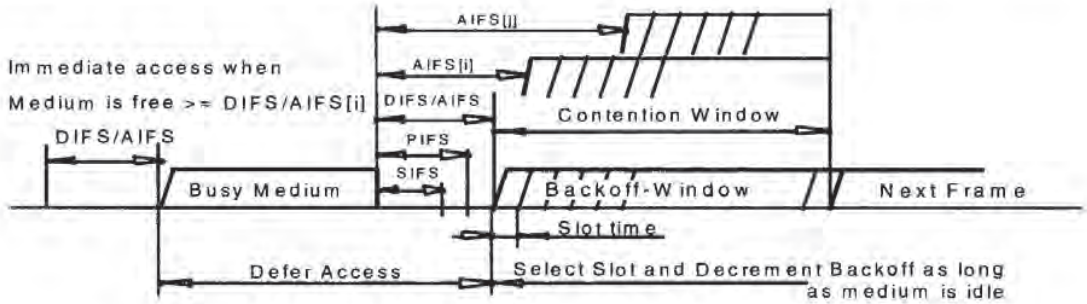

Figure 2–12: The basic access mechanism and interframe space relationships

Central to the medium-access mechanism is the concept of a coordination function (CF). The coordination function is a logical function that determines when a station operating within a Basic Service Set is permitted to transmit and receive data via the wireless medium. The interval of time when a particular station has the right to initiate transmissions onto the wireless medium is called transmission opportunity. The basic access mechanism in IEEE 802.11 is carrier sense multiple access with collision avoidance (CSMA/CA) with exponential backoff. The CSMA/CA mechanism is implemented by the DCF. In this type of mechanism, before transmitting data, a station checks whether the medium is busy, as illustrated in Figure 2–12. If the medium is busy, the station does not transmit. The medium should be available for an amount of time equal to the DIFS if the last frame detected on the medium was received correctly or to the EIFS, if the last frame detected on the medium was not received correctly. After this idle time, every station will generate a random backoff period for an additional deferral time before transmitting. This process minimizes collisions. This is the physical carrier sense mechanism.

The random backoff time is distributed according to a uniform distribution. The maximum extent of the uniform range is called contention window (CW). The distribution is uniform between 0 and CW. The time unit of the backoff timer is

INTEL-1020
U.S. Patent 9,265,043

called the slot time, which is equal to the maximum round-trip propagation delay. At each time slot, the channel is sensed to determine if there is activity on the channel. If the channel is idle for the duration of the slot, the timer is decremented by one time slot. If the channel is busy, the timer is not decremented. When the channel becomes idle again for a period greater than the DIFS, the backoff procedure continues from where it was interrupted. The process is repeated until the timer reaches zero. Only then will the station begin transmission. After a transmission, a station will wait for an acknowledgment (ACK) back to the transmitter, if a packet is received correctly. The acknowledgment is transmitted after the short IFS. This is necessary to ensure that an acknowledgment is transmitted before a new packet. No station will sense the channel before an acknowledgment is transmitted.

If no acknowledgment is received, the transmitting station will infer that an error has happened. Many circumstances may cause an error to occur—collision, other interference, etc. Error recovery is handled by retransmissions. The transmitting station increments a retry counter and attempts to retransmit. In this case, the contention window is doubled, a new backoff interval is selected, and the backoff countdown begins again. The backoff is exponential to avoid the zero-throughput state where the channel is jammed up with retransmissions. Every station maintains two retry counts: short retry count and long retry count. Their initial value is zero. The short retry count is incremented every time transmission of a short MAC frame (frame of length less than or equal to a certain threshold) fails. This short retry count is reset when a transmission of a short MAC frame succeeds. This long retry count is reset when a transmission of long MAC frame succeeds. The station long retry counter is reset to 0 whenever an acknowledgment frame is received in response to data transmission of length greater than a certain threshold, or whenever a frame with a group address is transmitted.

The contention window is increased every time an unsuccessful attempt to transmit causes either station retry counter to increment, until the contention window reaches a specified maximum. Once it reaches this maximum value the contention window remains at this value until it is reset. This improves the stability of the access protocol under high-load conditions. Retransmitted frames are identified with the Retry field set to 1. Retransmissions continue until the transmission is successful or until the relevant retry limit is reached, whichever

INTEL-1020
U.S. Patent 9,265,043

occurs first. When the retry counters reach their maximal values, retry attempts cease, and the data is discarded.

Figure 2–13 shows the backoff procedure when several stations have data to transmit. Initially the medium is occupied by station A, and stations B, C, and D defer. DIFS after station A completes its transmission, stations B, C, and D will enter the backoff procedure. The contention is won by station C, which chooses the shortest backoff interval. During the transmission of station C, stations B and D suspend their backoff counters. After station C completes its transmission, B and D again resume decrementing their counters. In the meantime station E also has data to transmit and will enter the contention. This time station D wins the contention. The process continues as shown in Figure 2–13.

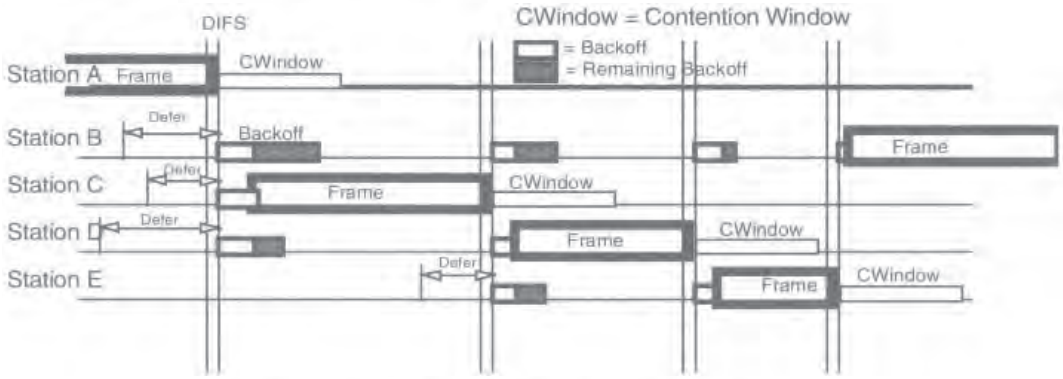

Figure 2–13: Backoff procedure

There is one situation when a station is not required to perform the random backoff before starting data transmission. A data packet arriving from the higher layer may be transmitted immediately, if the last post-backoff has been finished already (i.e., the queue was empty) and the channel has been idle for a minimum duration of DIFS. All the following packets must be transmitted after random backoff, until the transmission queue is empty again.

However, wireless networks can suffer from a hidden user phenomenon. One station may be able to communicate with two other stations, but these two stations may not be able to communicate between themselves. In this case, a station may sense the channel to be idle, even when there is a transmission. This increases the likelihood of a collision. The physical carrier sense of the CSMA/CA procedure is

INTEL-1020
U.S. Patent 9,265,043

inadequate to deal with the hidden node problem. However, there is also a virtual carrier sense. The virtual carrier sense mechanism deals with this problem by a reservation mechanism to announce the impending use of the channel. Every frame contains timing values, which are used to maintain special counters, called network allocation vectors (NAVs), which are present in all stations. NAV is implemented as a timer, which is continuously decremented regardless of whether the channel is busy or idle. The NAV is updated from the duration values of all received MAC headers, not just frames addressed to this particular station. Before sensing the medium with the physical carrier sense mechanism, a station would check its NAV. The station would attempt transmission only when the NAV has a value of zero, meaning that there is no transmission on the air. If either the physical or the virtual carrier sense mechanisms indicates that the medium is busy, stations select a random backoff interval.

To help deal with the hidden node problem there is an optional mechanism, called the RTS-CTS mechanism. The transmitting and receiving station can exchange messages called request-to-send (RTS) and clear-to-send (CTS). A station that has data to transmit sends a RTS packet. The intended recipient of the data sends a CTS packet. Both the RTS and CTS packets contain the source and destination addresses as well as the estimated time interval to complete the transmission and to return an acknowledgment.

A station that is not able to communicate with one of the two stations will receive only one packet—the RTS or the CTS. This is enough, however, for it to understand that the channel will be in a busy state and the duration of this state. The duration will be stored in the NAV. It is obvious that the RTS/CTS virtual carrier sense mechanism imposes overhead, which could be significant. The RTS/CTS mechanism is not used for short packets, for which the collision likelihood and cost in terms of retransmission time are small. To determine how short a packet is, there is a threshold called the RTS threshold. If the packet length is below the RTS threshold, RTS/CTS is avoided. In practice, the effectiveness of this reservation mechanism depends on the extent to which stations have identical operating ranges. Note that the RTS/CTS mechanism does not completely eliminate collisions. There may be collisions during the RTS or during the CTS packet, depending on whether the hidden nodes receive the RTS or the CTS packet. Collisions are more likely to occur during the transmission of the RTS signal. Note also that this reservation mechanism is unsuitable for broadcast or

INTEL-1020
U.S. Patent 9,265,043

multicast packets. If the RTS transmission fails, the short retry count and the station short retry count are incremented until the number of attempts reaches a certain limit.

Note that according to the DCF described previously, all stations are equal. The point coordination function (PCF) is a different coordination function, according to which stations and access points have different roles. While DCF is mandatory to implement, using the PCF is optional. The PCF introduces the concept of a superframe. Superframe is a frame that has two periods: contention-free period (CFP) and contention period (CP) (Figure 2–14). DCF is active in the contention period. PCF is active in the contention-free period. At the beginning of the superframe, if the channel is free, the PCF takes control over the channel. If the channel is busy (and therefore controlled by the DCF), the PCF defers. The PCF operates as follows. Stations can request that the AP registers them on a polling list. Then the AP regularly polls the stations one at a time for traffic. During the CFP, medium usage is entirely controlled by the AP, thereby eliminating the need for individual stations to contend for the medium. However, the AP must contend for the medium, and the start of the CFP can be delayed from its start time. Once the PCF gains access to the medium, the AP sends a beacon frame, begins to deliver queued data to the stations, and can poll stations by sending CF-Poll frames to those stations that have requested contention-free service. Stations that have data to transmit are allowed to send one frame in response to a CF-Poll. Stations that do not have data to transmit simply do not respond to the poll.

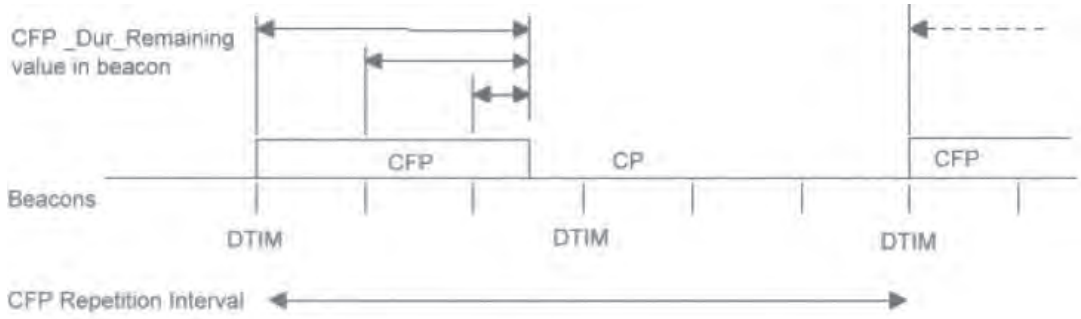

Figure 2–14: Beacons, contention, and contention-free periods

INTEL-1020
U.S. Patent 9,265,043

The beacon frame is a management frame that maintains the synchronization of the local timers in the stations and delivers protocol-related parameters. The point coordinator (PC), which is typically co-located with the AP, generates beacon frames at regular intervals. Thus every station knows when the next beacon frame will arrive. The intended period of the beacon frames is called targeted beacon transmission time (TBTT) and is announced in every beacon frame. Note that the beacon frame is required in pure DCF, even if there is only contending traffic. At the beacon period, the AP will schedule a beacon for transmission. If the medium is busy, the beacon transmission will be delayed, according to the CSMA mechanism. When the medium becomes available, the beacon will be transmitted before other data frames (Figure 2–15).

During the CFP, the maximum idle period is PIFS. If the AP does not receive data from a station being polled within PIFS of the polling frame, it can proceed and poll another station.

The PCF can transmit four types of frames: CF-Poll, CF-Ack-Poll, CF-Data-Poll, and CF-Data-Ack-Poll. To increase the aggregate throughput, the polling request is piggybacked on data and/or acknowledgment frames.

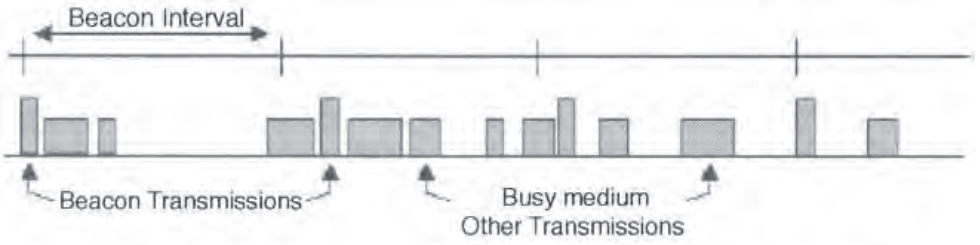

Figure 2–15: The beacon transmission is almost periodic

IEEE Std 802.11 does not address the question of order when there several stations to be polled, thus providing the opportunity for proprietary algorithms to be used. One approach could be to implement some type of prioritization, where some stations are polled before others. This prioritization can be dynamic. Another approach is to poll in a round-robin fashion, where for example the station with the lowest address is polled first, and so on. The PCF can be considered as a packet-switched, connection-oriented service.

INTEL-1020
U.S. Patent 9,265,043

To implement polling the AP must maintain a polling list. If the AP has completed polling all stations and traffic is light the CFP can be ended by the transmission of a CF-end frame. However, there is no mechanism to turn off retransmissions in the CFP.

The beacon frame notifies all of the other stations in the cell not to initiate transmissions for the length of the contention-free period. The length of the CFP can cover multiple beacon periods. Having silenced all other stations, the PCF station allows a given station to have a contention-free access. This allows stations to be even in power-save mode. In an ESS, a station in power-save mode initiates a frame exchange sequence by transmitting a PS-Poll frame to request data from an AP. In the event that neither an acknowledgment frame nor a Data frame is received from the AP in response to a PS-Poll frame, then the station can retry by transmitting another PS-Poll frame at its convenience.

If the AP sends a Data frame in response to a PS-Poll frame, but fails to receive the acknowledgment frame in response to this Data frame, the next PS-Poll frame from the same station may cause a retransmission of the last data. This duplicate data is filtered at the receiving station using the normal duplicate frame filtering mechanism. If the AP responds to a PS-Poll by transmitting an acknowledgment frame, then responsibility for the Data frame delivery error recovery shifts to the AP. The data is transferred in a subsequent frame exchange sequence, which is initiated by the AP. The AP will deliver one block of data to the station that transmitted the PS-Poll. If the power-save station that transmitted the PS-Poll returns to power-save mode after transmitting the acknowledgment frame in response to successful receipt of this block of data, but the AP fails to receive this acknowledgment, the AP will retry transmission of this block of data until the relevant retry limit is reached.

If a hidden station misses the previous beacon frames, then this hidden station will not have any knowledge about the TBTT and will not stop its operation based on DCF. If it has any data to transmit, the result will be collision during the CFP. The 802.11 standard tries to remedy this to some extent by requiring a station to set its NAV at the TBTT, irrespective of the reception of a beacon frame. However, if a station has not received any of the previous beacon frames, it cannot set its NAV at TBTT.

INTEL-1020
U.S. Patent 9,265,043

Beacons are transmitted even in an IBSS, in the absence of an AP. Because an AP is not present, there isn't an active point coordinator. However, every station at the TBTT selects a random delay and queues a beacon frame for transmission. When a beacon is received the queued frame is canceled. This scheme guarantees that every station in an IBSS will either receive a beacon or transmit one.

The PCF and DCF operate simultaneously and the PCF has priority over the DCF. The PIFS is shorter than the DIFS, so that the PCF will seize the medium before the DCF does.

The DCF and the PCF are the methods for medium access in the original IEEE 802.11 standard. The PCF was intended to carry periodic traffic such as speech, audio, and video signals. However, this scheme does not provide sufficient QoS to carry real-time traffic such as voice and video for several reasons.

Three main problems with the PCF led to the current activities within IEEE 802.11e to enhance the protocol: unpredictable beacon delays, unknown transmission duration of the polled stations, and lack of prioritization. The contention period may be of variable length and in turn, this may cause the contention-free period to start at different times. At TBTT, a point coordinator schedules the beacon as the next frame to be transmitted, and the beacon can be transmitted when the medium has been determined to be idle for at least PIFS. Depending on the wireless medium at this point in time—i.e., whether it is idle or busy around the TBTT—a delay of the beacon frame may occur. The time the beacon frame is delayed (i.e., the duration it is sent after the TBTT) postpones the transmission of time-bounded data that has to be delivered in CFP. This makes it impossible to maintain periodicity.

Another problem with the PCF is the unknown transmission time of polled stations. A station that has been polled by the PC is allowed to send a single frame that may be fragmented and of arbitrary length, up to the maximum of 2304 bytes (2312 bytes with encryption). Also, because different modulation and coding schemes are specified in the physical layers, the point coordinator does not control the duration of the transmission after polling. This destroys any attempt to provide QoS to other stations that are polled during the rest of the CFP.

Another disadvantage of the PCF is that there is no prioritization of data. If one station has several types of data to send, there is no distinction among them.

INTEL-1020
U.S. Patent 9,265,043

The original IEEE 802.11 of 1997 [B68] did not provide QoS when it was approved, and this was expected to be a significant disadvantage for it to succeed in the marketplace. Several influential companies were ready to mass-produce products based on different standards. It turned out that these concerns were unfounded. IEEE 802.11 became the most successful wireless standard, even before the QoS mechanism was in place. One reason the standard is so successful is that data is the primary application. Another reason is that it is possible to improve the original IEEE 802.11 and provide better QoS.

Enhancements in IEEE 802.11e

IEEE 802.11e will bring many enhancements—some general, some specific.

The first general enhancement is an option to allow stations to talk directly to other stations, bypassing the AP, even when there is one. The original IEEE 802.11 standard permits transmitting from one station to another only in ad-hoc mode. According to the original IEEE 802.11, in a BSS, when there is an AP, two stations cannot exchange information directly. If a station A wants to talk to another station B, station A first must send the data to the AP, and then the AP must send it to station B. This is not only inefficient in terms of channel usage, but it doubles the latency. The recent enhancement includes the ability for station-to-station transmission to bypass the AP and move instead directly between the relevant stations within the same BSS. This data transfer is set up using Direct Link Protocol (DLP). This protocol is necessary for several reasons. First, the protocol is used to exchange data rate and security information. DLP is also necessary because the intended recipient may be in power-save mode, in which case it can only be woken up by the AP. This protocol prohibits the stations going into power-save mode for the active duration of the Direct Stream. DLP does not apply in an IBSS, where frames are always sent directly from one station to another. The DLP handshake is illustrated in Figure 2–16.

INTEL-1020
U.S. Patent 9,265,043

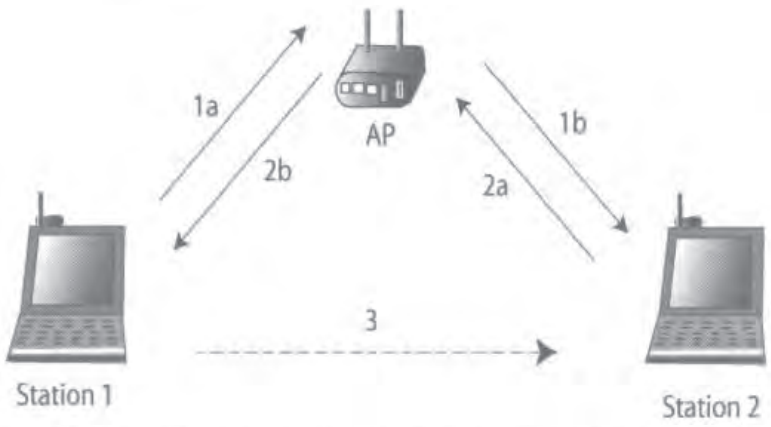

Figure 2–16: Steps involved in the Direct Link handshake.

First, a station 1 that has data to send will use the DLP and will send a DLP-request frame to the AP (arrow 1a in Figure 2–16). This request contains the rate set, and the capabilities of station 1, as well as the MAC addresses of stations 1 and 2. If station 2 is associated in the BSS, the AP will forward the DLP-request to the recipient, (1b). If station 2 accepts the request, it will send a DLP-response to the AP (2a), which contains information about its capabilities. The AP will forward the DLP-response to the first station (2b), after which the direct link becomes active and frames can be sent between the two stations. Note that a protection mechanism such as the RTS/CTS exchange is recommended during the Direct Link handshake.

The second general enhancement is the introduction of negotiable acknowledgments. The existing IEEE 802.11 standard relies on a positive acknowledgment mechanism for all frame exchanges. IEEE 802.11e introduces the notion of negotiable acknowledgments whereby a packet in a stream need not be acknowledged, or such acknowledgments can be aggregated, depending on the parameterization of a stream. Such negotiable acknowledgments lead to a more efficient utilization for the available channel and enable such applications a reliable multicast.

According to 802.11e, there are four acknowledgment types: normal (as in the legacy 802.11), no acknowledgment, no explicit acknowledgment, and block acknowledgment. Clearly, no acknowledgment is used whenever acknowledgment is not required. This could be appropriate for delay-sensitive

INTEL-1020
U.S. Patent 9,265,043

traffic. Because no acknowledgment is associated with lower reliability of data transfer, it is recommended to use RTS/CTS to increase the likelihood of successful data transfer. The option of no explicit acknowledgment allows acknowledgments to be piggybacked onto different frames.

Stations that have set up direct acknowledgment links with other stations are not required to go in sleep mode. These stations must remain in awake state.

The block acknowledgment mechanism allows a block of QoS frames to be transmitted, each separated by a SIFS period. This improves latency and jitter. There are two types of block acknowledgments: immediate and delayed. Immediate block acknowledgment is suitable for high-bandwidth, low-latency traffic, while the delayed block acknowledgment is suitable for applications that tolerate moderate latency.

The third general enhancement is the introduction and traffic prioritization and traffic parameterization. These are central to any QoS mechanism.

Traffic prioritization gives the ability to handle traffic differently according to the priority of that traffic. IEEE 802.11e is using a standard called 802.1D-1998 (specifically Annex H.2 of that standard), which adds three bits to the packet header and gives the LAN the capability to recognize eight levels of priority. The eight priority levels are designated by the integers: 7–0. The priority value can be indicated directly, or indirectly, in a traffic specification. By default, priority 7 is the highest priority and priority 1 is the lowest priority, with priority 0, which is used for best-effort traffic, ordered between priority 3 and priority 2. The resulting default ordering is {7, 6, 5, 4, 3, 0, 2, 1}. Every priority level may have different traffic specifications. A device can implement fewer than eight physical queues. In this case, it must provide a mapping from traffic categories and delivery priorities to the available queues. At least four physical queues are mandatory. These priority values can be specified directly according to Table 2–1, or indirectly according to traffic specification (TSPEC).

Another newly defined ability is the opportunity to handle traffic in accordance with a parameterized specification for each traffic type, allowing for more broadly differentiated traffic than would be available with simple prioritization. Again, the parameterization must be specified as part of the traffic and is not determined by the MAC.

INTEL-1020
U.S. Patent 9,265,043

Table 2–1: IEEE 802.1D priority levels

Priority	Traffic type
1 (lowest)	Background (BK) – aperiodic data
2	—
0	Best effort (BE) – aperiodic data
3	Excellent effort (EE) – isochronous data
4	Controlled load
5	Video (VI) – periodic
6	Voice (VO) – periodic
7 (highest)	Network control (NC)

The major novelty developed by IEEE 802.11e is the new coordination function, called hybrid-coordination function (HCF). It is a coordination function that combines aspects of the distributed coordination function and the point coordination function. The hybrid coordinator (HC) updates the NAV and appears to legacy stations as a point coordinator. However, it is a type of a point coordinator different from the one defined in the point-coordination function (PCF). The hybrid-coordination function has a contention-based and a contention-free-based channel access mechanism. It is important to realize that these are not two separate coordination functions; they are part of the hybrid-coordination function. The hybrid coordinator performs bandwidth management including the allocation of transmission opportunities to stations and the initiation of controlled contention intervals. The HCF is called hybrid because it is used during both the contention and contention-free periods.

In IEEE 802.11e, transmission opportunities are defined by a starting time and a maximum duration. These transmission opportunities are won in contention with other stations or granted during the CFP. The QoS CF-Poll from the HC can be sent after a PIFS idle period without any backoff.

The hybrid-coordination function uses a contention-based channel access method, called the enhanced distributed channel access (EDCA). Because it operates in the contention period, EDCF is similar to DCF. OoS support is realized with the introduction of the Traffic Categories (TCs). The prioritized channel access is realized with the QoS parameters per TC, which include AIFS[TC], CWmin[TC],

INTEL-1020
U.S. Patent 9,265,043

and CWmax[TC]. According to the EDCA, service differentiation is achieved through varying the amount of time a station senses the channel to be idle before backoff or transmission, varying the length of the contention window, and/or varying the duration a station may transmit after it acquires channel access.

The EDCA mechanism provides differentiated access to the wireless medium for the eight delivery priorities. A station desiring to initiate a transfer using the EDCA uses an equivalent to the DCF random backoff time mechanism, except that the station calculates and maintains a backoff timer and contention window for each queue—when there is data to be transmitted for that queue. Before each transmission (when the medium is busy), a station will defer until the medium is determined to be idle without interruption for a certain period of time. This interval is equal to the arbitration IFS for that queue (AIFS[i]) (rather than DIFS) when the last frame detected on the medium was received correctly. Note that AIFS[i] is variable, assigned either by a management entity, or by the AP and measured in time is equal to the integer AIFSN[i] multiplied by the slot time plus SIFS:

$$\text{AIFS[AC]} = \text{AIFSN[AC]} \times \text{slot time} + \text{SIFS} \qquad \text{Eq. 2–1}$$

The integer AIFSN must be greater than 2 for stations and greater than 1 for the AP. In this way, the AP has a higher priority for this channel access mechanism. The values of AIFSN for the different access categories is advertised in the beacons and probe response frames transmitted by the AP. An example is given in Table 2–2.

Table 2–2: Contention window and AIFSN parameters

AC	CWmin	CWmax	AIFSN
BK	CWmin	CWmax	7
BE	CWmin	CWmax	3
VI	(CWmin+1)/2 – 1	CWmin	2
VO	(CWmin+1)/4 – 1	CWmin	2

INTEL-1020
U.S. Patent 9,265,043

If the last frame was not received at this station with a correct MAC FCS value, the station defers transmission for EIFS – DIFS + AIFSN[AC] × slot time. The reason for using EIFS in the standard is the following. The virtual carrier sense mechanism indicates that the NAV at all stations is updated from the headers of all frames, not necessarily frames addressed to this station, nor frames from the same BSS. However, the header can only be interpreted from error-free frames, as ascertained by a correct value of the frame-check sequence (FCS). When activity on the wireless medium was detected, but a frame was not received correctly, the station knows that communication is under way. The station does not know the length of the period following the end of the frame. The EIFS is long enough to protect the acknowledgment from collision with a station that was not able to update its NAV. The backoff calculation uses the DCF method, with different contention window values for each queue. Similarly, the minimum and maximum contention window limits are not fixed per PHY, as with DCF, but are variable values, assigned to each traffic category either by a management entity or by the AP. In this way, each output queue contends for transmission opportunities.

At each transmission opportunity, a traffic scheduling entity selects a frame for transmission. If the backoff timers for more than one queue reach zero at the same slot, then the frame from the highest-priority queue is transmitted. These lower-priority queues will treat this as a collision and will set their contention window values as if they had experienced a transmit failure. The subtle difference is that retry bits are not set, as it would be done after an actual external collision on the wireless medium. The effect of the backoff procedure is that when multiple stations and/or queues of equal priority at one or more enhanced stations are deferring and go into random backoff, then the entity selecting the smallest backoff time using the random function will win the contention. In the case of queues of unequal priority at enhanced stations, queues with shorter AIFS[i] periods will win the contention. The higher-priority queue will lose the contention if its AIFS is longer. As stated previously, the Contention Window values, plus Short and Long Retry Counts are different for each queue *i:* CW[*i*], QSRC[*i*], and QLRC[*i*]. In the case of a collision the contention window increases in the same way as for the DCF. Provided that the contention window is less than the maximum, it is replaced by (CW[AC] + 1) × 2 – 1. Then out of this new, enlarged CW, another uniformly distributed backoff counter is drawn with a uniform distribution over the range [0,CW[AC]] inclusive, to reduce the probability of a

INTEL-1020
U.S. Patent 9,265,043

new collision. The CW never exceeds the parameter CWmax[TC], which is the maximum possible value for CW. The CW is reset to CWmin after every successful attempt to transmit or when the long and short retry counters reach a certain limit. Assigning shorter contention windows to data with higher priority ensures that, on average, higher-priority data will get through before lower-priority data. A continuation of a transmission opportunity obtained the EDCA mechanism can be granted SIFS after a successful transmission, provided the device has another frame to transmit with the same AC and the duration of this frame plus any acknowledgment, if required, will fit in the remaining TXOP duration.

Note that to ensure fairness and overall improved QoS to all AC, 802.11e provides a mechanism for dynamic priority assignment. Each station is required to maintain two variables for every access category: used time and admitted time. At the time of association the admitted time is set to zero. Later the station can request admitted time for a specific priority level. If the used time reaches or exceeds the admitted time the station will no longer transmit data at the given priority level. However, the same data can still be transmitted with a priority levels Background or Best effort.

A traffic specification (TSPEC) describes the traffic characteristics and the QoS requirements of a traffic stream (TS). The main purpose of the TSPEC is to reserve resources within the HC and modify the HC's scheduling behavior. It also allows other parameters to be specified that are associated with the traffic stream, such as acknowledgment policy. TSPEC parameters not specified by the application are determined based on assumptions about throughput, latency, and packet error rate requirements.

The second mechanism for access to the wireless medium, which is part of the HCF, is called HCF-controlled channel access (HCCA). It provides the capability for reservation of TXOPs with the HC. A station can make requests with the HC for TXOPs for its own transmissions, as well as transmissions from the HC to itself. The HC either accepts or rejects the request based on admission control policy. The admission control policy depends on factors such as the vendors' implementation of the scheduler, available channel capacity, etc. This mechanism is expected to be used for the transfer of periodic traffic such as voice and video. If the traffic stream is admitted, then the scheduler will service the station during a

INTEL-1020
U.S. Patent 9,265,043

service period (SP). This SP is a period of time during which one or more downlink frames are transmitted and one or more polled TXOPs are granted to the station. A station that is PS mode must wake-up before the start of the SP and may return to PS mode after the SP mode ends.

The scheduler is responsible for determining the mean data rate, packet size, and maximum service interval or delay bound. The schedule for an admitted stream is calculated in two steps. The first step is the calculation of the scheduled service period. In the second step, the TXOP duration for a given service period is calculated for the stream. In the first step, the scheduler finds the minimum of the maximum service periods for all admitted streams and then chooses a number lower than this minimum that is a submultiple of the beacon interval. This value is the service period for all stations with admitted streams. After the service period is calculated, the number of packets N_1 can be found from

$$N_1 = \frac{\text{mean data rate} \times SP}{L_1} \qquad \text{Eq. 2–2}$$

where L_1 is the chosen packet size. If the physical layer transmits at a data rate R, the transmission opportunity is the maximum to transmit N_1 packets of the chosen packet size or one packet of maximum size:

$$TXOP_i = \max\left(\frac{N_i \times L_i}{R} + O, \frac{L_{max}}{R} + O\right) \qquad \text{Eq. 2–3}$$

where L_{max} is the maximum packet size (2304 bytes). An example is shown in Figure 2–17. In this example the beacon interval is 100 ms and the maximum service period for the stream is 60 ms. As shown in the figure, the scheduler calculates a service period equal to 50 ms using the steps explained previously.

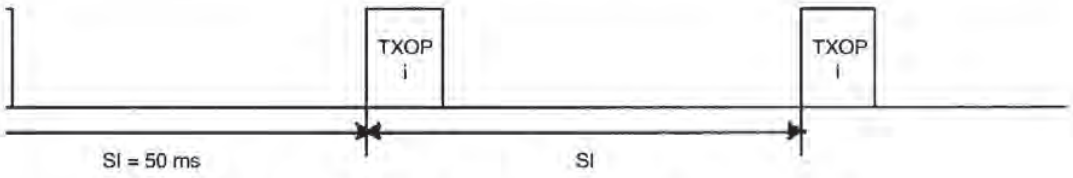

Figure 2–17: Schedule for traffic stream *i*

INTEL-1020
U.S. Patent 9,265,043

The same process is repeated continuously while the maximum service period for the admitted stream is larger than current service period. Stream $k + 1$ is admitted if the calculated $TXOP_{K+1}$ satisfies the inequality

$$\frac{TXOP_{K+1}}{SP} + \sum_{i=1}^{K} \frac{TXOP_{K+1}}{SP} \leq \frac{T - T_{CP}}{T} \qquad \text{Eq. 2–4}$$

where L_{max} is the beacon interval and T_{CP} is the contention period. Otherwise, the $k + 1$-th stream is rejected. In this way, all admitted streams have access to the channel.

The HCCA can use polling by issuing QoS (+) CF-Polls to stations during the contention-free period. However, because the hybrid coordinator can also grant polled transmission opportunities by sending QoS (+) CF-Poll frames during the contention period, it is not mandatory for the hybrid coordinator to use the contention-free period for QoS data transfers. Stations receiving a QoS (+) CF-Poll are required to respond within a short IFS period. If the polled station has no queued traffic to send, or if the data available to send are all too long to transmit within the specified transmission opportunity limit, the station will send a QoS Null frame. In the case of no queued traffic, this QoS Null has a QoS control field that reports a traffic category queue size of 0. In the case of insufficient transmission opportunity size, this QoS Null has a QoS control field that reports the requested transmission opportunity duration needed to send the selected data and traffic category identifier for the highest priority data that is ready for transmission.

It is clear that there are two ways a station can obtain a transmission opportunity. If the station receives a QoS (+)CF-Poll during the contention period or contention-free period, or if it wins an instance of EDCF contention (which will only occur during the contention period). In the case of a polled transmission opportunity, the entire transmission opportunity is protected by the NAV set by the duration of the frame that contained the QoS (+)CF-Poll function.

Under the hybrid-coordination function, the basic unit of allocation of the right to transmit onto the wireless channel is the transmission opportunity. Each transmission opportunity is defined by a particular starting time, relative to the end of a preceding frame, and a defined maximum length. The novelty brought by IEEE 802.11e is that TXOPs not only have a starting time, but they also have a maximum duration. The duration is between 32 and 8160 μs. TXOPs are allocated

INTEL-1020
U.S. Patent 9,265,043

by contention (EDCA-TXOP) or granted through HCCA (polled-TXOP). Every frame sent as a result of a polled transmission opportunity contains the transmission opportunity limit in its QoS control field. When polling is used in the original IEEE 802.11, only a single data frame can be transmitted in response. IEEE 802.11e allows, in response to a poll, the station to transmit as many data frames as will fit within the specified duration. These frame exchanges must be separated by a short IFS interval. Any station will not initiate transmission of a frame unless the transmission, and any acknowledgment or other immediate response expected from the peer MAC entity, can be completed before the end of the remaining duration of the transmission opportunity. However, data can be fragmented to fit in a transmission opportunity. Furthermore, no transmission opportunity nor transmission within it is allowed to extend across TBTT. Because this rule applies to transmission opportunities as well as transmissions within these transmission opportunities, it is the responsibility of the hybrid coordinator to ensure that the full duration of any granted transmission opportunity does not cross any of these boundaries. (See Figure 2–18.)

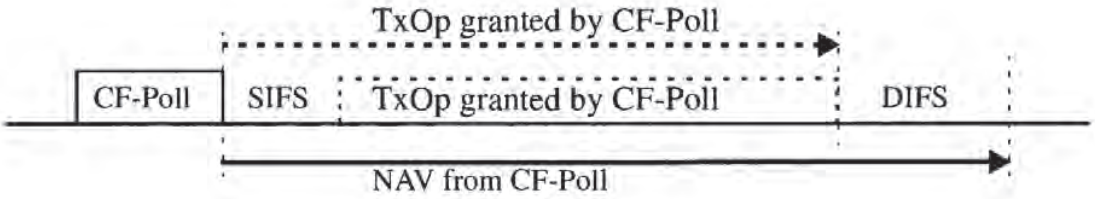

Figure 2–18: TXOP

There are certain rules about what type of data can be transmitted during a transmission opportunity. Management data and non-QoS data must be the only or the last frame exchange sequence in the transmission opportunity. Multiple control frames and QoS data frames can be sent in a single TXOP, with the nonfinal bit in all but the final frame set to 1 to indicate that more frames are coming. If an expected acknowledgment to a frame with nonfinal bit set to 1 is not received, the station awaiting the acknowledgment should transmit either its next frame or a QoS Null. All of these frames (management, control, QoS data, and non-QoS data) contain a duration field. The initial frame of a multiple frame exchange sequence contains a duration value, which is the remaining duration of the TXOP. The initial frame of the sole or final frame exchange sequence contains

INTEL-1020
U.S. Patent 9,265,043

a duration value that covers the actual remaining time needed to transmit the frame, plus one ACK time, plus one SIFS interval, plus one DIFS interval. This is important so that all stations can maintain their NAVs.

During the CP, each TXOP begins either when the medium is determined to be available under the EDCA rules, i.e., after AIFS plus backoff time, or when the station receives a special poll frame, the QoS CF-Poll, from the HC. The QoS CF-Poll from the HC can be sent after a PIFS idle period without any backoff. However, although the poll frame is a new frame as part of the upcoming 802.11e, the legacy stations also set their NAVs upon receiving this frame. Therefore, the HC can issue polled TXOPs in the CP using its prioritized medium access. During the CFP, the starting time and maximum duration of each TXOP is specified by the HC, again using the QoS CF-Poll frames. Stations will not attempt to get medium access on their own during the CFP, so only the HC can grant TXOPs by sending QoS CF-Poll frames. The CFP ends after the time announced in the beacon frame or by a CF-End frame from the HC.

The HC may allocate TXOPs to itself to initiate fame deliveries whenever it wants, but only after detecting the channel as being idle for PIFS. The duration values used in the QoS frame exchange reserve the medium for DIFS longer than the end of the sequence to permit continuation of a NAV-protected contention-free transfer by the concatenation of several contention-free blocks. This extra reservation of the wireless channel allows the hybrid coordinator to initiate a subsequent transmission opportunity with reduced risk of collision because stations cannot begin contending for the medium until DIFS after the last communication within the transmission opportunity.

The hybrid coordinator performs delivery of queued broadcast and multicast frames following DTIM (Delivery Traffic Indication Message) beacons in a contention-free period. The hybrid coordinator may use a longer contention-free period for QoS delivery and/or QoS polling by continuing with hybrid-coordination function frame exchange sequences, after broadcast/multicast delivery, for a duration not exceeding the maximum possible duration of the contention-free period. The HC may also operate as a PC, providing (non-QoS) CF-Polls to stations.

Because the hybrid coordinator is a type of point coordinator, the hybrid coordinator includes a contention-free Parameter Set element in the Beacon

INTEL-1020
U.S. Patent 9,265,043

frames it generates. This causes the BSS to appear to be a point-coordinated BSS. This causes all stations (other than the hybrid coordinator) to set their NAVs to the corresponding duration value at TBTT. This prevents most contention with the contention-free period by preventing nonpolled transmissions by all stations whether or not they are CF-Pollable.

In this way, all stations inherently obey the medium access rules of the hybrid-coordination function.

When a station updates its NAV setting using the duration value from a QoS CF-Poll containing the BSSID of this BSS, that station will also save the address of the holder of the transmission opportunity. If an RTS, management type, data type or QoS data type frame is received from a MAC address that matches this saved TXOP holder address, the station will send the appropriate response after SIFS, without regard for—and without resetting—its NAV. This saved TXOP holder address is cleared whenever the NAV is reset, the NAV counts down to 0, or when a QoS data type frame is received with an address that matches the saved TXOP holder address and has the non-final bit set to 0. The NAV is cleared when the station receives a CF-End frame or a QoS CF-Poll with a duration value equal to zero.

During the CP, the NAV can be reset in the following manner. When a QoS data frame is received from the TXOP holder, and the non-final bit in the QoS control field is set to 0, and the No ACK bit is set to 1, the NAV is reset. During the CP, when everything is the same, except that the No ACK bit in the QoS control field is set to 0, the NAV is reset.

During the CP, if an enhanced station receives a QoS data type frame that requires acknowledgment, when its NAV is set and the saved TXOP holder address is clear, the response after SIFS is an ACK frame, even if the frame being acknowledged was of a subtype including QoS CF-Poll. In this manner, the responding enhanced station indicates that it is unable to accept the TXOP conveyed by the QoS CF-Poll.

Another way to achieve improved NAV protection during either the CP or the CFP is to use the well-known RTS/CTS exchange. The sending of RTS during the CFP is used not to update the NAV, which is protected for the entire duration of the CFP. RTS still can be used to ensure that the addressed recipient station is

INTEL-1020
U.S. Patent 9,265,043

within range and awake, and to protect against hidden nodes. Stations can send an RTS as the first frame of any frame exchange sequence without regard for RTS frame threshold. The corresponding CTS will set the NAV in the vicinity of the receiver, and its duration field is set to the number of microseconds until the end of the TXOP.

The HC is a type of point coordinator, but differs from the point coordinator used in PCF in several significant ways. Most important is that the hybrid-coordination function frame exchange sequences may be used among stations during both the contention period and the contention-free period. Another significant difference is that QoS CF-Poll grants transmission opportunities with duration specified in the QoS (+) CF-Poll frame. As a result, stations can set their NAV appropriately, which is not possible if the duration is not specified. Stations also can transmit multiple frame exchange sequences within a given transmission opportunity. The hybrid coordinator not only delivers queued broadcast and multicast data, but can use a longer contention-free period for QoS delivery and/or QoS polling by continuing with hybrid-coordination function frame exchange sequences after broadcast/multicast delivery. Finally, the HC may also operate purely as a PC, providing (non-QoS) CF-Polls using all the applicable rules for the PCF.

It is interesting to consider the coexistence of the DCF, PCF, and HCF. The DCF and a point coordination function (either PCF or HCF) coexist in a manner that permits both to operate concurrently within the same network. When a point coordinator is operating in a basic service set, the point coordination function and the distributed coordination function access methods alternate, with a contention-free period followed by a contention period. When an HC is operating, there is a contention-free period and a contention period in each superframe, and stations treat the hybrid coordinator as if it were a point coordinator, using the distributed coordination function access method only during the contention period. The hybrid-coordination function access methods (polled and contention-based) operate concurrently, throughout the superframe. Concurrent operation allows the polled and contention-based access methods to alternate within intervals as short as the time to transmit a pair of frame exchange sequences. Note that it is possible to use both PCF and HCF in a single contention-free period. This, however, is extremely complex and does not seem advantageous.

INTEL-1020
U.S. Patent 9,265,043

The hybrid coordinator is always implemented in the access point. In an IBSS, there is no access point, no hybrid coordinator, and therefore no polled transmission opportunities. In an IBSS, QoS is supported only by the EDCA. In an IBSS priorities and differentiation of traffic classes are static parameters.

Since the QoS facility was standardized recently, there will be a lot of stations that do not support it. An interesting situation is when, in a basic service set, some of the stations support the QoS facility and some do not. Because the hybrid coordinator is a type of point coordinator, the hybrid coordinator includes a contention-free Parameter Set element in the Beacon frames it generates. This causes a BSS with enhanced stations to appear to be a point-coordinated BSS to legacy stations. This causes all stations (other than the hybrid coordinator) to set their NAVs to the corresponding duration value. This prevents most contention with the contention-free period. The HCF assumes that all non-QoS data has priority equivalent to best effort.

It should also be noted that real-time data does not require low-bit error rate like computer data. In particular if a voice or video packet has a few bit errors, then its quality is usually not affected significantly. However, at present this characteristic is not used by IEEE 802.11. The MAC will not pass to the higher layers any packets that contain bit errors. Research work in this area is likely to continue.

Some applications require timers at different stations to be synchronized. To support these applications the standard defines a MAC service. One way to accomplish synchronization across a BSS is by multicasting synchronization. In this method synchronization packets from the higher layers containing a time stamp and a sequence number are broadcasted. These packets can be recognized by their MAC headers. The time stamp in the synchronization packet would contain the higher layer clock value at the time when the previous synchronization packet was transmitted. The sequence number would include a value corresponding to the packet for which the time stamp is provided.

POWER-SAVING MECHANISM

The 802.11 MAC allows stations to enter a low-power mode. While in power-save (PS) mode, a device alternates between awake and doze states. The doze state is a

INTEL-1020
U.S. Patent 9,265,043

power-efficient state, and the device is in the awake state only for very short periods of time.

In a BSS, the station will first inform the AP when it is going to enter power-saving mode and the number of beacon periods that it will be in this mode. The AP is responsible then to buffer information for this station while it is in doze state. When the station returns to the awake state, these frames are delivered using the EDCA or HCCA mechanisms. The PS device will wake up just to receive a beacon, but it does not have to wake up for every beacon. The beacon contains information whether the AP has packets queued for it. If there are queued packets, the PS device will remain in the awake state and contend for the medium to send a poll frame to the AP. Upon receipt of a poll frame, the AP will contend for the medium to send these buffered packets. The contention-free period can also be used to deliver buffered frames. Service periods in response to a poll are called unscheduled service periods. The device must also awaken when multicast frames are to be delivered. These times are indicated in the beacon as the traffic indication map (TIM). The delivery TIM (DTIM) count is an integer value that is equal to the number of beacon frames that will occur between the delivery times of multicast frames. In this way power-saving stations do not have to wake-up for every beacon. A larger value of the DTIM count leads to greater power efficiency, but also to increased latency for multicast traffic.

To improve the QoS that is provided by the PS mechanism 802.11e introduced the notion of scheduled and unscheduled service periods. A scheduled service period is a service period that begins at a scheduled wake-up time. A scheduled SP starts at fixed intervals of time determined by the minimum service interval.

A station may be in power-save mode before the set up of DLP or block acknowledgment. Once DLP is set up with another station, the power save mode is suspended. In a BSS, block acknowledgment mode is suspended during the PS mode. In an IBSS, stations are required to suspend their power-save mode if they have set up block acknowledgment connections.

In an IBSS, a station that wants to enter power-saving mode must inform at least one other station that it is about to do so by setting appropriate power management bits within the header of transmitted frames. All stations in an IBSS before transmitting must estimate the power-saving state of the destination. If the destination is deemed to be in power-saving state, they should refrain from

INTEL-1020
U.S. Patent 9,265,043

transmitting and buffer the traffic. Then the power-saving device must wake up to receive every beacon. The device must also stay awake for a certain period of time after every beacon, called ad-hoc traffic indication message (ATIM) window. During this time period, other stations will make an announcement if they have data for the power-saving station. Multicast information will also be announced during this time period. If the power-saving station receives an ATIM frame, it must acknowledge the ATIM frame and remain awake to allow the other stations to deliver their data destined for it. Clearly, the power savings in an IBSS are less than the savings in a BSS, because the station is required to wake up at every beacon and remain awake for a certain period of time after the beacon.

IEEE 802.11 PHYSICAL LAYERS

The protocol data unit for all physical layers consists of a preamble and a header, followed by MAC data. The receiver uses the preamble for detection and synchronization. The header has information about the MAC data such as data rate and length. If the physical layer supports multiple data rates the preamble and the header are sent at the lowest data rate. In this way, the preamble and the header are sent in a way to ensure maximum range. The MAC frame that follows consists of the following basic components:

a) A header, which comprises frame control, duration, address, sequence control information, and, for QoS frames, QoS control information

b) A variable length frame body

c) A frame check sequence (FCS), which contains an IEEE 32-bit cyclic redundancy code (CRC).

This CRC protects all fields in the MAC header and frame body. Using the polynomial

$$G(x)= x^{32}+x^{26}+x^{23}+x^{22}+x^{16}+x^{12}+x^{11}+x^{10}+x^{8}+x^{7}+x^{5}+x^{4}+x^{2}+x+1$$

the 32-bit CRC is the ones complement of the sum modulo 2 of the following two quantities. The first quantity is the remainder of the division modulo 2 of $x^{k}(x^{31}+x^{30}+\ldots+x+1)$ by $G(x)$, where k is the total number of bits in the protected fields, the MAC header and frame body. The second quantity is the remainder of the division of the content of the protected fields treated as a

INTEL-1020
U.S. Patent 9,265,043

polynomial and multiplied by x^{32}, by $G(x)$. The CRC is an error-detection mechanism. At the receiver, the MAC passes to the higher layers only frames for which the received and calculated CRC values match.

IEEE 802.11 has at present six physical layers: three physical layers were defined initially, and three were developed later. The first three physical layers are direct-sequence spread-spectrum, frequency-hopping spread-spectrum, and infrared. They provide data rates of 1 and 2 Mb/s. The frequency-hopping physical layer uses two-level and four-level Gaussian frequency shift keying for transmission at 1 Mb/s and 2 Mb/s, respectively. In the 2.4 GHz band, the total number of 1-MHz-wide hop channels is 79, from 2402 MHz to 2480 MHz. These 79 channels are divided into three nonoverlapping sets. Frequency-hopping radios must hop at least at a rate of 2.5 hops per second with a minimum hop distance of 6 MHz. The minimum hop separation ensures that if one hop falls on a "bad" frequency (with a low signal-to-noise ratio), then the next hop will likely be on a "good" frequency. The frequency-hopping physical layer was used in some early products, but because the data rate is only 1 Mb/s or 2 Mb/s, the overall market was relatively small. The infrared physical layer is based on pulse-position modulation (PPM), achieves data rates of 1 Mb/s and 2 Mb/s, and requires line-of-sight. The frequency-hopping and infrared physical layers are not relevant at present because of their performance. The DSSS physical layer is considered next.

DSSS PHY

The DSSS physical layer operates in the 2.4 GHz band. The protocol data unit for this physical layer, like for all physical layers, consists of preamble, header, and MAC data. The preamble contains a 128-bit SYNC word, followed by a 16-bit start-of-frame delimiter. The preamble length in bits is 144 bits, and in time is 144 μs. Following the preamble, the header identifies what type of modulation is used for the MAC frame, the length of time required to transmit the MAC frame (measured in microseconds), and, as an error-detection mechanism, a 16-bit CRC. The 16-bit CRC is the ones complement of the remainder of the modulo-2 division of the contents of the header by the polynomial $x^{16} + x^{12} + x^5 + 1$. The 16-bit CRC is calculated before scrambling. If the 16 bits of the CRC match at the receiver, the PHY header has been correctly received. The preamble and the header are always sent at 1 Mb/s. The MAC data that follows can be transmitted at

INTEL-1020
U.S. Patent 9,265,043

1 Mb/s or 2 Mb/s. The modulation that the original standard supports is differential phase-shift keying.

Two types of differential phase shift keying are supported. The preamble and the header are transmitted using DBPSK. The MAC data is sent using either DBPSK, resulting in a data rate of 1 Mb/s, or using DQPSK resulting in a data rate of 2 Mb/s. The bit-to-symbol mapping for DBPSK is that 0 is encoded with no phase change, and 1 is encoded with a phase change of 180 degrees. For DQPSK the input bits are grouped in pairs of two. Then 00 is encoded with no phase change, 01 with a phase change of 90 degress rotating counterclockwise, 11 is encoded with a phase change of 180 degrees, and 10 of 270 degrees. Noncoherent demodulation can be used because a clock reference is not needed to recover the data. DBPSK is more robust compared with DQPSK. DBPSK achieves the same bit error rate at a lower signal-to-noise ratio compared with DQPSK.

The signal processing steps performed are scrambling, spreading, and modulation. Scrambling randomizes the data using the polynomial $1 + z^{-4} + z^{-7}$. The spreading method employs the 11-bit Barker sequence 1,–1,1,1,–1,1,1,1,–1,–1,–1. This Barker sequence is applied to a modulo-2 adder together with each of the information bits. The chipping rate is 11. The output of the adder is a signal, the data rate of which is 11 times higher than the information rate. This ensures a spreading factor and equivalently processing gain of 10.8 dB. The signal bandwidth becomes 22 MHz. Because the total bandwidth available is 83.5 MHz, with appropriate guard bands, there are three nonoverlapping 22-MHz-wide channels in the band. The first channel is centered at 2.412 GHz, the second is centered at 2.437 GHz, and the third is centered at 2.462 GHz. The maximum allowed output power in North America is 1W; however, many vendors have selected 100 mW as the default transmit power level. The reasons are that lower transmit output power does not drain batteries as fast, and 100 mW is the maximum output power per European regulations. Also, this decreases the range of operation of the access points, enabling vendors to sell more higher-margin access points.

IEEE 802.11b

IEEE 802.11b is discussed before IEEE 802.11a, because IEEE 802.11b is still based on spread-spectrum technology. IEEE 802.11b achieves data rates of

INTEL-1020
U.S. Patent 9,265,043

5.5 Mb/s and 11 Mb/s. Most of the products on the marketplace today are IEEE 802.11b. How does the IEEE 802.11b PHY work? Similarly to all physical layers a frame consists of a preamble, header, and MAC data. The preamble is used by the receiver to detect the signal and to synchronize with the transmitter. There are two preambles: long and short. The long preamble is identical to the IEEE 802.11 preamble described in the previous section. Because the long preamble provides backwards compatibility with IEEE 802.11, it is mandatory to implement. The short preamble is optional.

Two preambles are defined because the long is transmitted at 1 Mb/s, and therefore transmission of the preamble will take a significant amount of time. Equivalently, devices will be in a low-data-rate mode for a long time and will not take full advantage of the higher data rates that IEEE 802.11b provides. The short preamble helps to increase the actual throughput. This is an example of the trade-off between backwards compatibility and throughput, which is even more important for the other extensions of IEEE 802.11. If an IEEE 802.11b device uses the long preamble (which would be understood by legacy IEEE 802.11 stations) then the throughput is lower. If the long preamble is used, then the header is also according to IEEE 802.11. Legacy stations will be able to decode the header, and in particular will decode the information about the length of the MAC data that follows. From the length information, they can update their NAVs and can maintain silence during the transmission. If the long preamble is used, the MAC data that follows can be transmitted at 1, 2, 5.5, and 11 Mb/s. If the short preamble is used, the throughput would be higher, but legacy IEEE 802.11 equipment cannot decode the preamble. The short preamble uses a 56-bit SYNC field and is transmitted also at 1 Mb/s using DBPSK modulation and spreading using the 11-bit Barker code. If the short preamble is used, however, the header is transmitted at 2 Mb/s using DQPSK and Barker spreading. The MAC data that follows the header is transmitted at 2, 5.5, or 11 Mb/s. The actual data rate used is denoted in the header.

All transmitted bits are scrambled and descrambled using the polynomial $1 + z^{-4} + z^{-7}$, which is the same as for 802.11. The four modulation formats and data rates defined in IEEE 802.11b are: (1) the basic rate of 1 Mb/s modulated with DBPSK; (2) the extended rate of 2 Mb/s with DQPSK modulation; (3) and (4) two enhanced rates of 5.5 Mb/s and 11 Mb/s using CCK modulation. As an option alternative to CCK, PBCC can be used. The Barker code is the spreading

INTEL-1020
U.S. Patent 9,265,043

code for the basic and extended rates. The enhanced rates of 5.5 and 11 Mb/s do not use the Barker code. CCK modulation was chosen by the IEEE 802.11 Working Group, because it helps achieve backwards compatibility by maintaining the same 22 MHz-wide channels. CCK is a variation of M-ary orthogonal keying modulation. Spreading is achieved by using a spreading code with eight samples. Each of these eight chips is complex and obtained using QPSK modulation. The eight chips of the spreading code are

$$\left\{ e^{j(\varphi_1+\varphi_2+\varphi_3+\varphi_4)}, e^{j(\varphi_1+\varphi_3+\varphi_4)}, e^{j(\varphi_1+\varphi_2+\varphi_4)}, \right.$$
$$\left. -e^{j(\varphi_1+\varphi_4)}, e^{j(\varphi_1+\varphi_3+\varphi_4)}, e^{j(\varphi_1+\varphi_3)}, -e^{j(\varphi_1+\varphi_2)}, e^{j\varphi_1} \right\}$$

These eight values are obtained from a Walsh/Hadamard code, where φ_1 is added to all samples, φ_2 is added to every other sample, starting with the first, φ_3 is added to every other group of two of samples, and φ_4 is added to every other group of four samples. The fourth and seventh sample are multiplied by –1, or equivalently rotated by 180 degrees to optimize the correlation properties and minimize the DC offset. Walsh/Hadamard codes are also called complementary code and this is why the name CCK is used. The spreading code has eight samples, but the effective chipping rate is must be 11 to maintain the same channel bandwidth of 22 MHz. To achieve this the symbol rate is increased to 1.375 Msymbols/s.

For transmission at 5.5 Mb/s, four bits $d_0d_1d_2d_3$ are transmitted per symbol. The first two bits d_0d_1 encode φ_1 based on DQPSK—φ_1 is 0, 90, 180, or 270 degrees relative to the preceding symbol, if d_0d_1 are 00, 01, 11, or 10, correspondingly. In addition, starting with the second symbol, every other symbol is multiplied by –1. This multiplication by –1 can be represented by 180 degrees being added to φ_1 for the second, fourth, sixth, etc., symbols after the header. The other two bits d_2 and d_3 determine $\varphi_2 = d\pi_2 + \frac{\pi}{2}$ and $\varphi_4 = d_3\pi$, while $\varphi_3 = 0$.

For transmission at 11 Mb/s, eight bits $d_0d_1d_2d_3d_4d_5d_6d_7$ are transmitted per symbol. The first two bits d_0d_1 encode φ_1 exactly in the same way as for transmission at 5.5 Mb/s, with 180 degrees added to φ_1 for the second, fourth, sixth, etc., symbols after the header. The pairs d_3d_4, d_4d_5, and d_6d_7 encode φ_2,

INTEL-1020
U.S. Patent 9,265,043

φ_3, and φ_4, based on QPSK. These phases are 0, 90, 180, or 270 degrees counterclockwise, if the corresponding bit pair is 00, 01, 10, or 11, respectively.

IEEE 802.11b also defines an optional coding scheme called PBCC. It sometimes is referred to as high-performance mode, although it achieves the same data rates. It uses a 64-state rate one-half binary convolutional code and a cover sequence. BPSK is used for 5.5 Mb/s and QPSK for 11 Mb/s.

A PBCC modulator uses a 64-state rate-1/2 binary convolutional code (BCC) and a cover sequence, as shown in Figure 2–19.

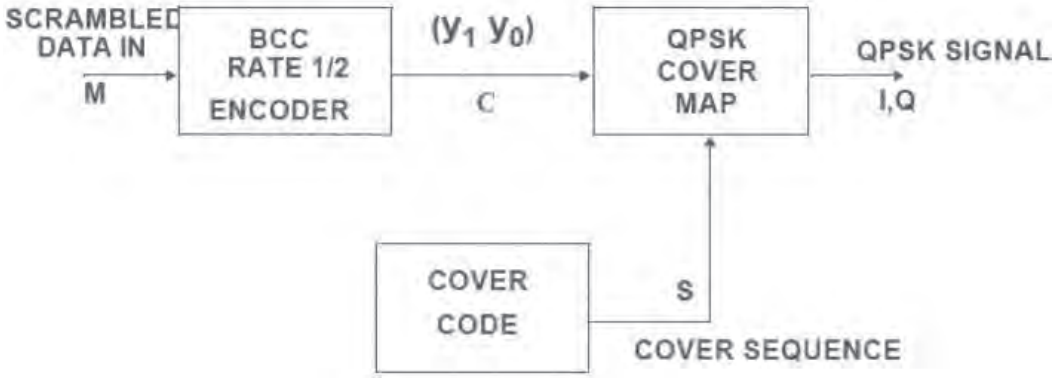

Figure 2–19: PBCC modulator

The binary convolutional encoder is shown in Figure 2–20. After the convolutional encoding, bit-to-symbol mapping is performed. For transmission at 5.5 Mb/s, the bit-to-symbol mapping is performed using BPSK. For transmission at 11 Mb/s, it is performed using QPSK.

The mapping from the output bits of the binary convolutional encoder to the BPSK and QPSK symbols is determined using a pseudorandom cover sequence. QPSK modulation using the pair of bits $y_1 y_0$ is done as shown in Figure 2–21 and BPSK modulation as shown in Figure 2–22.

INTEL-1020
U.S. Patent 9,265,043

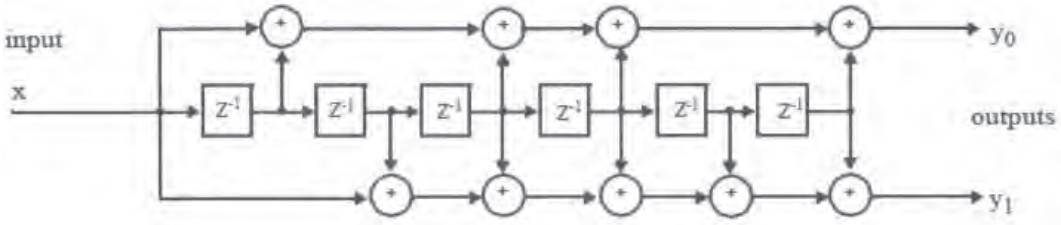

Figure 2–20: Binary convolutional encoder

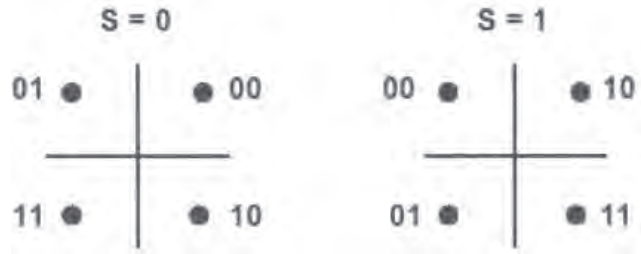

Figure 2–21: Cover code bit-to-symbol mapping in QPSK mode

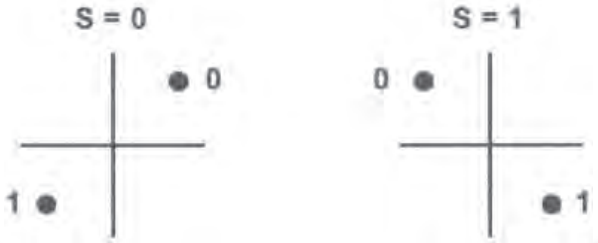

Figure 2–22: Cover code bit-to-symbol mapping in BPSK mode

The cover sequence has 256 bits, generated using the 16-bit seed sequence 0011001110001011. The first 16 bits of the cover sequence are identical to the seed sequence. To obtain the second, third, etc. groups of 16-bits in the cover sequence, the seed sequence is cyclically shifted on the left by 3, 6, etc. Every bit in the cover sequence determines how the bit-to-symbol mapping is performed for every bit in the frame. For frames that have more than 256 bits, the cover sequence is simply repeated.

INTEL-1020
U.S. Patent 9,265,043

IEEE 802.11b uses the same frequency channels as 802.11. The channels center frequencies are spaced 25 MHz apart to allow multiple WLANs to operate simultaneously in the same area without interfering. The North American channelization scheme is shown in Figure 2–23 and the European scheme is shown in Figure 2–24.

The range of operation of 802.11b is 150 feet for a floor divided into individual offices by concrete or sheet-rock, about 300 feet in semi-open indoor spaces such as offices partitioned into individual workspaces, and about 1000 feet in large open indoor areas.

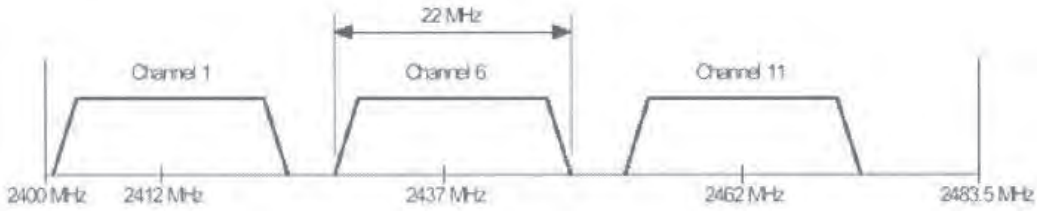

Figure 2–23: North American channelization for 802.11b

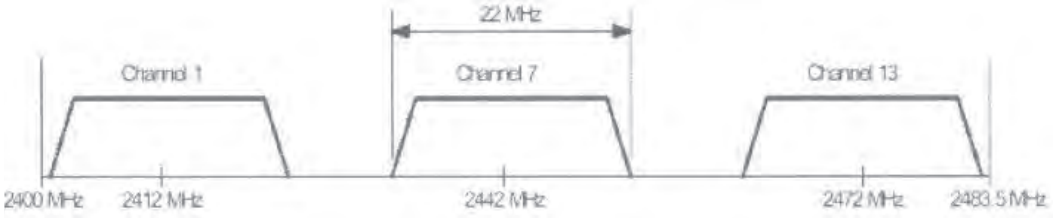

Figure 2–24: Nonoverlapping 802.11b channelization for Europe

IEEE 802.11a

As a result of the Telecommunications Act of 1996, in 1997 the FCC allocated bands in the 5 GHz region for unlicensed use [B41], [B42]. These bands are in total 300 MHz wide. As discussed in "Government regulations" on page 5, in 2003 the FCC made also available the band between 5.47 and 5.725 GHz.

This chunk of spectrum allows significantly higher speeds of communication than the 83.5-MHz-wide ISM band between 2.4 GHz and 2.4835 GHz. At the same

INTEL-1020
U.S. Patent 9,265,043

time, a technology called orthogonal frequency division multiplexing (OFDM) became very well developed, widely understood, and economically feasible for incorporation in mass-market wireless products—all the necessary characteristics for inclusion in a standard. In 1997, IEEE 802.11 seized the opportunity to create a powerful high-speed physical layer in the newly allocated UNII bands. This standard, IEEE 802.11a, was developed in a very short time—less than two years [B69]. At the present IEEE 802.11a products are available on the market.

Since the IEEE 802.11a PHY is using OFDM and not spread spectrum technology, it is worth comparing OFDM and spread spectrum. In high-data rate wireless communications, OFDM is superior for several reasons. The main reason is that high-data rates cannot be achieved with spread spectrum technology because the spectrum is limited. The benefits of spread spectrum are proportional to the spreading factor. However, in a limited band the spreading factor cannot be very high for a high data rate. OFDM is the technology of choice for high-data rate wireless communications. In OFDM the transmission bandwidth is divided into many narrow subchannels, which are transmitted in parallel. Ideally each subchannel is narrow enough so that the fading it experiences is flat. OFDM allows the subchannels to overlap, which is spectrally efficient. Since the subchannels overlap, how are they separated at the receiver? The subchannels can be separated at the receiver because they are orthogonal. OFDM combats ISI caused by multipath by using symbol periods several times longer than the delay spread of the channel impulse response.

Like the other physical layers, the OFDM PHY takes the MAC data and adds a preamble and a header. The preamble is used for detection, synchronization, automatic gain control (AGC) estimation, coarse and fine carrier frequency estimation, and channel estimation. The preamble consists of 10 short training symbols, each 0.8 μs long, and two long training symbols, each 3.2 μs. The long training symbols are preceded by a long guard interval of 1.6 μs. This guard interval is a cyclic prefix of a long training symbol. The total preamble is 16 μs long and is sent using BPSK and OFDM using convolutional coding rate of one-half. A header follows the preamble. The header contains information about the MAC data that follows. The header first has information about the data rate, the length of the transmission, and parity and tail bits. The signal field is transmitted using BPSK and OFDM using convolutional coding rate of one-half.

INTEL-1020
U.S. Patent 9,265,043

The following signal processing steps are performed: First, the data is scrambled for randomization. The scrambler is defined by the polynomial $x^7 + x^4 + 1$, (Figure 2–25) and is identical with the descrambler.

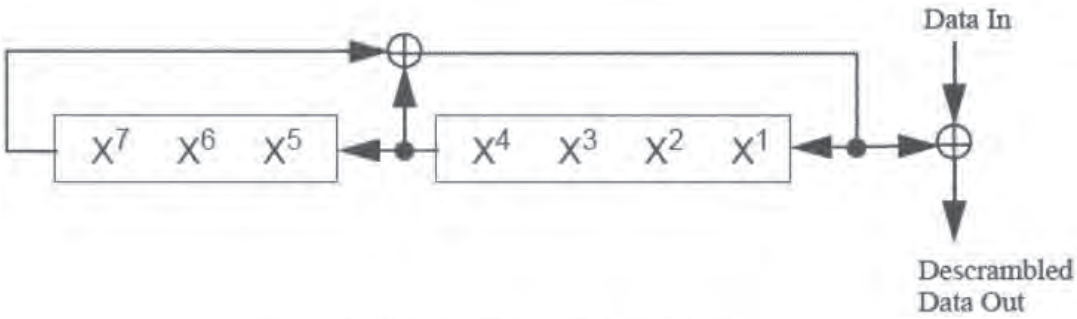

Figure 2–25: Scrambler circuit

The scrambler is initialized at the transmitter with a pseudorandom number. To provide the receiver with the capability to estimate the initial state of the scrambler, the seven least-significant bits of the header are initialized with zeros and then scrambled.

The bits after scrambling are encoded by a convolutional code of rate 1/2, 2/3, or 3/4. The rate-1/2 convolutional encoder is described by the polynomials 133 and 171 in octal notation (Figure 2–26). The higher data rates are obtained by puncturing omission of certain bits in the codeword. Decoding is performed by a hard-decision, or even better, soft-decision Viterbi decoder.

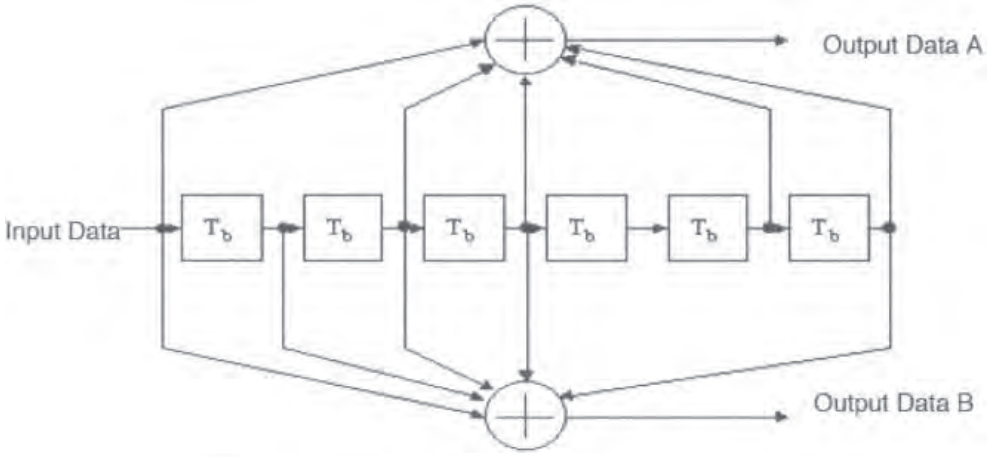

Figure 2–26: Rate one-half convolutional encoder

INTEL-1020
U.S. Patent 9,265,043

After convolutional encoding, the encoded bits are interleaved using a block interleaver with a block size corresponding to the number of bits in one OFDM symbol. The interleaver is defined by a two-step permutation. The first permutation ensures that adjacent coded bits are mapped onto nonadjacent subcarriers. The second ensures that adjacent coded bits are mapped alternately onto less and more significant bits of the constellation.

After interleaving, OFDM modulation is performed in the following way. From the interleaved bit sequence, complex symbols are obtained using BPSK, QPSK, 16 QAM, or 64 QAM modulation. The Gray-coded bit-to-symbol mappings are shown in Figure 2–27, Figure 2–28, and Figure 2–29. The value $(I + jQ)$ in these figures is multiplied by a normalization factor to ensure the same average power for all modulations. The normalization factor is 1 for BPSK, $1/\sqrt{2}$ for QPSK, $1/\sqrt{10}$ for 16-QAM, and $1/\sqrt{42}$ for 64-QAM. These complex symbols in groups of 48 become the subcarriers of one OFDM symbol. Four subcarriers in fixed locations are reserved as pilot tones, bringing the total number of nonzero subcarriers 52. In addition, 12 subcarriers are equal to zero, which effectively introduces a guard frequency band. The total number of subcarriers is 64, which allows a 64-point inverse fast Fourier transform (IFFT) to be used. In practice, the IFFT hardware is implemented using a radix-2 algorithm, or the computationally more efficient radix-4 algorithm.

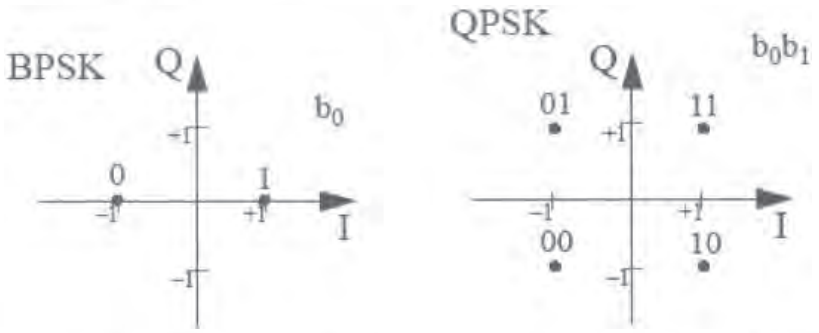

Figure 2–27: Bit-to-symbol mappings for BPSK and QPSK

INTEL-1020
U.S. Patent 9,265,043

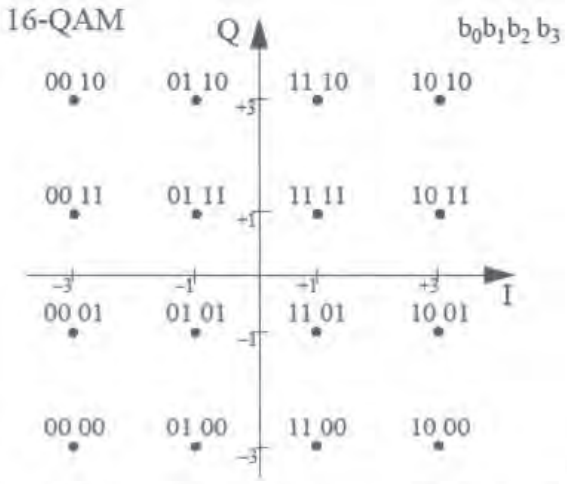

Figure 2–28: Bit-to-symbol mapping for 16-QAM

After the IFFT, a circular extension of the waveform is added as a cyclic prefix. Why is a cyclic prefix necessary? It is necessary to make linear convolution identical to circular convolution. The transmitted signal and the channel impulse response are linearly convolved. However, the product of the discrete Fourier transforms (DFT) of the transmitted signal and the channel impulse response corresponds to circular convolution. From digital signal processing, it is well known that to make circular convolution identical to linear convolution, a signal extension of appropriate length is necessary. The length of the cyclic prefix should be at least equal to the length of the channel impulse response. Without cyclic prefix the inter-symbol interference between adjacent OFDM symbols (e.g., the inter-block interference) is significant and will reduce the performance. The cyclic prefix resolves the inter-block interference problem. For 802.11a the cyclic prefix is 0.8 μs and therefore the inter-symbol interference will be reduced when the energy of the channel impulse response is contained within 0.8 μs.

If $r(t)$ is the complex baseband signal, the actual transmitted signal is

$$r_{RF}(t) = \mathrm{Re}\{r(t)e^{j2\pi f_c t}\} \qquad \text{Eq. 2–5}$$

where f_c is the carrier center frequency. The complex baseband signal, obtained using an IFFT is

INTEL-1020
U.S. Patent 9,265,043

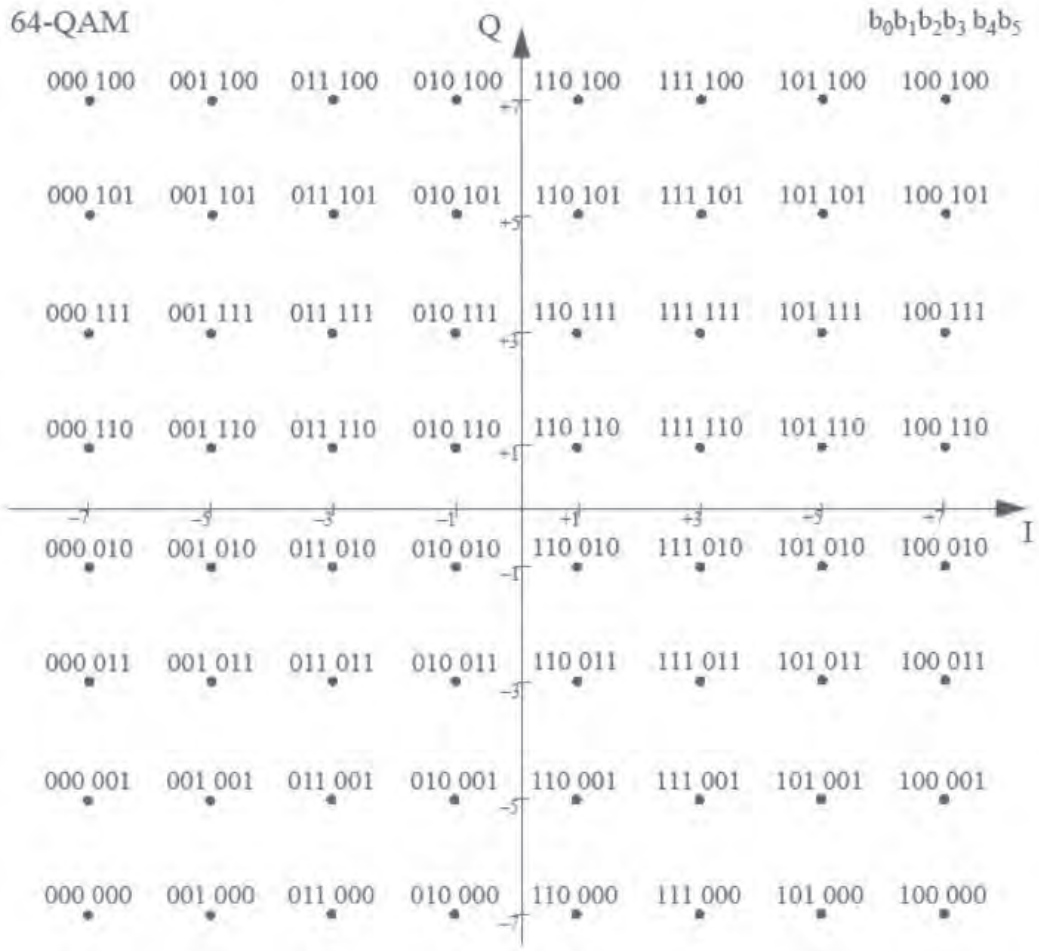

Figure 2–29: Bit-to-symbol mapping for 64-QAM

INTEL-1020
U.S. Patent 9,265,043

$$r(t) = w(t) \sum_{k=-26}^{k=26} C_k e^{j\frac{2\pi kt}{T_{FFT}}} \quad \text{Eq. 2–6}$$

where C_k are the complex symbols, obtained using BPSK or QAM modulation, and $T_{FFT} = 3.2\ \mu s$ is the FFT period, or the symbol period without the cyclic prefix. If the complex baseband signal is sampled with a sampling interval of T_{FFT} / N, or alternatively 3.2 μs/64 = 50*ns*, then it can be represented as

$$r(nT_{FFT} / N) = w(nT_{FFT} / N) \sum_{k=-26}^{k=26} C_k e^{j\frac{2\pi kn}{N}} \quad \text{Eq. 2–7}$$

or alternatively using discrete-time notation

$$r(n) = w(n) \sum_{k=-26}^{k=26} C_k e^{j\frac{2\pi kn}{N}} \quad \text{Eq. 2–8}$$

This makes OFDM modulation equivalent to an inverse DFT and allows FFT algorithms to be used. $w(t)$ is a window function. Two window functions are specified: rectangular window having values of 1 for every sample, and a window function, having value of 1 for every sample, except the first and the last sample, where the value is 0.5. The pilot tones are in carriers –21, –7, 7, and 21. Therefore, the contributions of the pilot subcarriers for the n-th OFDM symbol is obtained by an IFFT of $P_{-26,26}$, which is zero everywhere, except $P_{-26} = P_{-7} = P_7 = P_{21}$. To prevent the generation of spectral lines, the pilot carriers are BPSK modulated by pseudo-random binary sequence consisting of 1's and –1's. This pseudo-random binary sequence can be generated using the same scrambler in Figure 2–25, when the all-ones initial state is used, and all 1's are replaced with –1's and all 0's, are replaced with 1's. Every element from this pseudo-random sequence is used to obtain the four pilots of one OFDM symbol.

INTEL-1020
U.S. Patent 9,265,043

The short training symbol

$$r_{short}(t) = w(t) \sum_{k=-26}^{k=26} S_k e^{j\frac{2\pi kt}{T_{FFT}}} \quad \text{Eq. 2–9}$$

is obtained by an IFFT of the sequence $S_{-26,26}$, which is zero everywhere except $S_{-24} = S_{-16} = S_{-4} = S_{12} = S_{16} = S_{20} = S_{24} = \sqrt{13/6}(1+j)$ and

$S_{-20} = S_{-12} = S_{-8} = S_4 = S_8 = \sqrt{13/6}(-1-j)$.

One long training symbol

$$r_{long}(t) = w(t) \sum_{k=-26}^{k=26} L_k e^{j\frac{2\pi kt}{T_{FFT}}} \quad \text{Eq. 2–10}$$

is obtained using an IFFT of $L_{-26,26} = \{1, 1, -1, -1, 1, 1, -1, 1, -1, 1, 1, 1, 1, 1, 1, -1, -1, 1, 1, -1, 1, -1, 1, 1, 1, 1, 0, 1, -1, -1, 1, 1, -1, 1, -1, 1, -1, -1, -1, -1, -1, 1, 1, -1, -1, 1, -1, 1, -1, 1, 1, 1, 1\}$. To provide an opportunity for improved channel estimation accuracy, two long training symbols are used, preceded by a cyclic prefix. The preamble is the concatenation of the short and long training symbols.

A 64-point IFFT can be used. This requires that complex QAM symbols 1 to 26 are mapped to IFFT inputs 1 to 26, and complex QAM symbols –26 to –1 are mapped to IFFT inputs 38 to 63.

The data rates that 802.11a achieves are 6, 9, 12, 18, 24, 36, 48, and 54 Mb/s. One OFDM symbol is sent every 4 µs, of which 0.8 µs is the cyclic prefix. Since 250,000 symbols are sent every second, and one symbol uses 48 data carriers, a data rate of 6 Mb/s is achieved with BPSK modulation and a convolutional code of rate one-half ($24 \times 250{,}000 = 6$ Mb/s). Data rate of 54 Mb/s is achieved with 64 QAM and a convolutional code rate of ¾.

Note that the slot time for 802.11a is required to be 9 µs. The slot time is the sum of the Rx-to-Tx turnaround time, MAC processing delay, and the time to perform clear channel assessment (CCA). The Rx-to-Tx turnaround time should be under 2 µs, the MAC processing delay should also be under 2 µs, and the CCA time should be under 4 µs.

The bandwidth of one OFDM symbol is 20 MHz and since there are 64 carriers one carrier takes about 20 MHz/64, or 312 KHz. Since only 52 carriers are used, the power density can be calculated assuming that one symbol occupies 16 MHz.

INTEL-1020
U.S. Patent 9,265,043

Therefore, output power of 200 mW, or 23 dBm, corresponds to power density of 12.5 mW/MHz, which is equal to 11 dBm/MHz and –49 dBm/Hz. The total bandwidth in the UNII bands is 300 MHz, but there are guard bands. In each of the three UNII bands four non-overlapping 20 MHz channels can be used. Therefore the total number of different OFDM channels is 12. Immediately one sees that besides high speed OFDM has another advantage. Since there are 12 different OFDM channels compared with only three different 802.11b channels, the total capacity of 802.11a is much higher.[9] (See Figure 2–30 and Figure 2–31.)

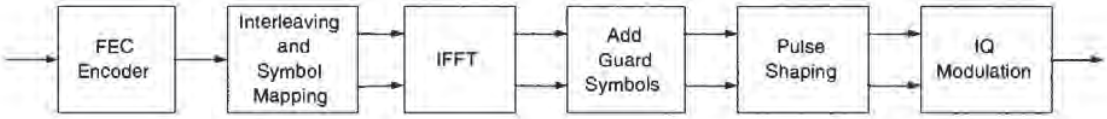

Figure 2–30: IEEE 802.11a transmitter according to the standard

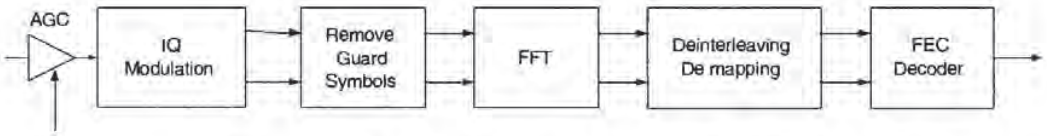

Figure 2–31: An implementation of a IEEE 802.11a receiver

Recently the IEEE 802.11 working group completed work on an amendment to IEEE 802.11a, called IEEE Std 802.11h. IEEE 802.11h defines how IEEE 802.11a devices implement dynamic frequency selection (DFS) and transmitter power control (TPC). IEEE 802.11h is necessary for two reasons. First, DFS and TPC are necessary is to satisfy European regulations, which require all unlicensed devices in the 5 GHz bands to implement DFS and TPC. Because the original IEEE 802.11a does not have DFS and TPC, IEEE 802.11a devices could not be sold in Europe. The second reason is that in the 5 GHz bands there are other wireless devices, such as satellite communication systems and radars, which are the primary users of the bands.

According to government regulations, these devices have the right-of-way, i.e., IEEE 802.11a must not cause interference to them. Anticipating widespread

[9] The capacity becomes even higher if the very recently allocated band of 5.47–5.725 GHz is also included.

INTEL-1020
U.S. Patent 9,265,043

adoption of WLAN, the radar community was concerned about interference from IEEE 802.11a. According to IEEE 802.11h, IEEE 802.11a devices must periodically test for the presence of radars. If radar is found in the 20-MHz-wide channel in which IEEE 802.11a is operating, the IEEE 802.11a devices must vacate the channel within a certain period of time. To accomplish this channel transfer, the IEEE 802.11a devices must find another channel and exchange messages to coordinate the channel transfer. If an AP is present, the AP decides which channel to transfer to. If an AP is not present, one of the stations assumes the role of DFS owner and it makes the decision as to which channel to transfer to. The channel switch announcement is made not according to the CSMA/CA rules of the DCF, but according to the PCF. In this way, IEEE 802.11h tries to satisfy the demands of the radar community by ensuring that the channel will be vacated as quickly as possible.

The second mechanism specified by IEEE 802.11h is TPC. TPC is a mechanism according to which the output power is adapted to regulatory constraints and channel conditions. Notice first that the maximum allowed output power levels in the lower, middle, and upper UNII bands are different. Furthermore, government regulations around the world are different. Therefore, devices can change their output power depending on the regulatory domain and frequency band of operation. TPC provides the capabilities of devices to use less than the maximum allowed output power in response to smaller path loss or larger value of the RSSI. The AP is responsible for informing stations about the maximum output power where the network is operating. The algorithm for determining the output power level is not specified in 802.11h. Note that DFS and TPC require feedback information about channel conditions. Therefore, IEEE 802.11h includes a message exchange according to which a device can be requested to measure channel conditions and report back. TPC and DFS together give the device agility and in general can be used to reduce interference and to promote coexistence with other wireless devices.

Another recent work within IEEE 802.11 is IEEE P802.11j, which will take into account Japanese regulations in the band 4.9–5 GHz and adapt IEEE 802.11a for operation in this frequency band.

INTEL-1020
U.S. Patent 9,265,043

IEEE 802.11g

As was discussed in the Introduction, there is always a desire to increase the data rate. The value of WLAN technology and, therefore, the user acceptance and market size depend substantially on the speed of connection. Almost as soon as IEEE 802.11b and IEEE 802.11a were approved, work began on another physical layer for the 2.4 GHz band. The requirement was at least to double the data rate of IEEE 802.11b while achieving backwards compatibility with IEEE 802.11b. Another motivation was the change in FCC regulations. While previous regulations required spread-spectrum technology in the 2.4 GHz band, the new regulations allow any digital modulation to be used.

There were two major competing technologies: an extension of PBCC and an OFDM technology, called DSSS-OFDM, which uses the preamble of IEEE 802.11b for backwards compatibility. These technologies were feverishly supported mainly by the companies that have relevant intellectual property: Texas Instruments™ (by its 1999 acquisition of Alantro Communications Inc., of Santa Rosa, CA) in the case of the PBCC and Intersil in the case of DSSS-OFDM. Neither of the technologies had qualified majority of 75% within the working group to win. As a result, the IEEE 802.11g standard is a compromise: it has one mandatory mechanism and two optional mechanisms.

The mandatory physical layer is in principal identical to 802.11a, but works in the 2.4 GHz band. 802.11g is the first standard for wireless data communications in the 2.4 GHz band which is not using spread-spectrum technology. Since 802.11a was standardized before 802.11g, and it was widely understood technology, the investment in 802.11a could be leveraged by using the same physical layer, and only changing the frequency band of operation. There are several small differences between the mandatory physical layer of 802.11g and 802.11a. The first difference is that the SIFS for 802.11a is 16 μs and the SIFS for 802.11g is 10 μs. 802.11a uses 16 μs to allow enough time for the time-consuming convolutional decoding. To give 802.11g equipment the same time to perform convolutional decoding, an 802.11g packet is followed by 6 μs of silence. The duration field of every packet includes this 6 μs signal extension. This is called a virtual extension of SIFS.

802.11g can achieve speeds of up to 54 Mb/s in the 2.4 GHz band. Recall that 802.11g is required to be backwards-compatible with 802.11b. This backwards

INTEL-1020
U.S. Patent 9,265,043

compatibility, however, is not achieved at the physical layer. A legacy 802.11b station cannot receive any part of an 802.11g transmission. Backwards compatibility is achieved at the MAC layer. In a network composed of 802.11b and 802.11g devices, 802.11g devices can use the RTS/CTS protocol to communicate. The RTS/CTS exchange is done using 802.11b so that other stations can update their NAVs and avoid transmitting during this period. Another protection mechanism that a station can use to reserve the wireless medium for certain duration is to send CTS to itself. Everything that is mandatory for 802.11b and 802.11a is mandatory for 802.11g as well, with the following main exception: when clear channel assessment (CCA) is performed to determine whether the medium is busy, it must detect OFDM and DSSS transmissions. There are two main approaches to implement CCA: energy thresholding and detection of the 802.11b or 802.11a preambles. The only difference with 802.11b is that support for the short preamble is now mandatory.

IEEE 802.11g supports three different preamble and header formats. The long preamble is the legacy of the original IEEE 802.11. The short preamble is legacy from IEEE 802.11b. The third preamble and header format is legacy from IEEE 802.11a.

The two main contenders for 802.11g were standardized as optional mechanisms.

The first optional mechanism is DSSS-OFDM. It uses the long and short preambles defined in 802.11 and 802.11b, preceding the OFDM part. In this way, it does not need RTS/CTS or CTS-to-self protection mechanisms, because legacy 802.11b stations will be able to update their NAVs.

The other optional technology, PBCC, achieves data rates of 22 and 33 Mb/s [B49]. The two modes are called PBCC-22 and PBCC-33. PBCC-22 uses a 256-state binary convolutional code with a rate of 2/3 and a cover sequence. The input bits are divided into pairs of adjacent bits. In each pair, the first bit is fed to the upper input of the convolutional encoder and the second bit is fed to the lower input of the convolutional encoder. The output of the convolutional code is mapped to an 8-PSK constellation. This yields a throughput of two information bits per PSK symbol. The cover sequence is the same that is used in the optional PBCC mode of 802.11b. PBCC-33 mode achieves a 33 Mb/s data rate by

INTEL-1020
U.S. Patent 9,265,043

increasing the clock frequency by 50 percent from 11 MHz to 16.5 MHz only for the data portion of the packet. (See Figure 2–32 and Figure 2–33.)

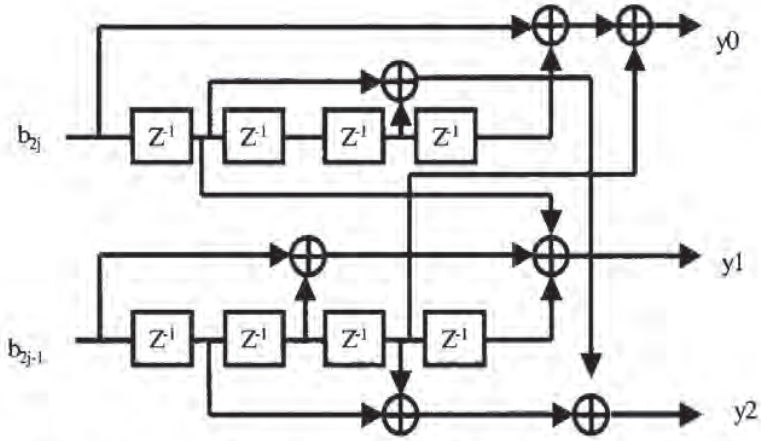

Figure 2–32: 256-state rate two-thirds binary convolutional encoder

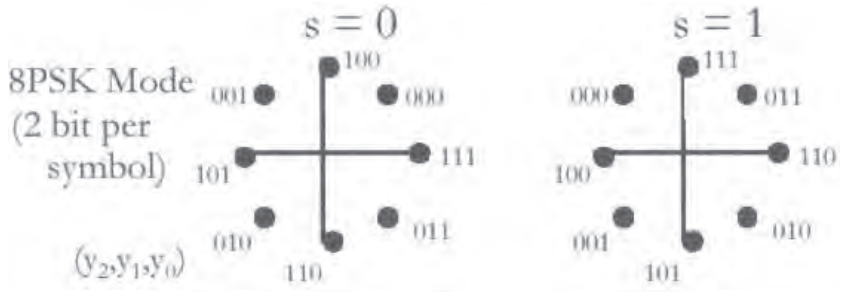

Figure 2–33: 8-PSK symbol mapping

Clearly 802.11g supports many data rates: the 6, 12, and 24 Mb/s data rates are mandatory for OFDM. In addition, the mandatory OFDM part of 802.11g has data rates of 18, 36, 48, and 54 Mb/s. The optional CCK-OFDM part of 802.11g provides the same data rates as 802.11a. The optional PBCC part of 802.11g provides data rates of 22 Mb/s and 33 Mb/s. The data rates that the 802.11 physical layers provide are summarized in Table 2–3.

Since every physical layer has multiple data rates, an important question is how to select the data rate? This is not specified in the standard and there are several

INTEL-1020
U.S. Patent 9,265,043

approaches. Products by Lucent select the transmission rate according to how many acknowledgements the device has successfully received at a given rate. Another approach is to monitor the channel during the RTS/CTS frame exchange. It is also possible to determine the data rate on the basis of the SNR during the immediately preceding packet exchange.

Table 2–3: 802.11 physical layers

Standard	Max. Data Rate (Mbs)	Band	Frequency Range (GHz)
IEEE 802.11	2	ISM	2.4–2.485
IEEE 802.11b	11	ISM	2.4–2.485
IEEE 802.11a	54	U-NII	Low 5.15–5.25 Mid 5.25–5.35 High 5.725–5.825
IEEE 802.11g	54	ISM	2.4–2.485

802.11g supports two slot times –20 μs, which should be used when there is legacy equipment in the network, and a shorter slot time of 9 μs, which provides higher throughput and could be used when there are only 802.11g devices in the network. In an IBSS only the long slot time should be used to facilitate ad-hoc networking. The time to perform clear channel assessment should be under 15 μs if the longer slot time is used, or under 4 μs if the shorter slot time is used. The timing parameters of the three physical layers (802.11b, 802.11a, and 802.11g) are summarized in Table 2-4. The maximum contention window for all physical layers is 1023 time slots.

INTEL-1020
U.S. Patent 9,265,043

Table 2–4: Timing parameters for 802.11b, 802.11a, and 802.11g

Parameter	802.11b	802.11a	802.11g
SIFS duration, μs	10	16	10 (with virtual extension of 6 μs)
Slot time, μs	20	9	20 or 9
Minimum contention window, in number of slots	31	15	15

WLAN INSTALLATION

If an AP can hear another AP or a distant station, due to the very nature of the CSMA/CA protocol, it will defer transmission, just as it would for a station within its area. Similarly, if more than one AP can hear a station, all of these APs will defer transmission, thus degrading the network performance. Therefore, there is a conflict.

The WLAN installation is a significant practical problem because network performance can be significantly affected by the installation. On one hand, there should be no gaps of coverage. On the other hand, the APs should be as far apart as possible. Clearly, WLAN installation is not a trivial task. The goal of WLAN installation is to achieve certain performance with minimum cost. The performance measures are most often data rate and coverage area. In addition, there may be areas where coverage is deliberately not desired. Coexistence with other wireless networks can also be a performance measure. The WLAN installation problem involves selection of antenna types, AP locations, and AP frequencies.

First, antennas should be considered. A wide variety are available for wireless LAN systems, the most common being the dipole antenna. Most access points on the market today have dipole antennas. The dipole antenna is omnidirectional, or it radiates equally in all directions around its axis or azimuth. It does not radiate along the axis of the antenna and therefore focuses its energy in a horizontal plane with a circular pattern outward. The radiation pattern is sometimes referred to as the donut pattern. The top image in Figure 2–34 shows an example of the

INTEL-1020
U.S. Patent 9,265,043

horizontal radiation pattern of a dipole antenna, and the lower image shows an example of the vertical radiation pattern.

If a dipole antenna is placed in the center of a single floor of a multistory building, most of its energy will be radiated along the length of that floor, with some small fraction sent to the floors above and below the access point. If the intended area is a large open area more or less symmetric in all directions, the omnidirectional antenna is certainly a good choice. In contrast to omnidirectional antennas, directional antennas concentrate their energy into a cone, also called a beam.

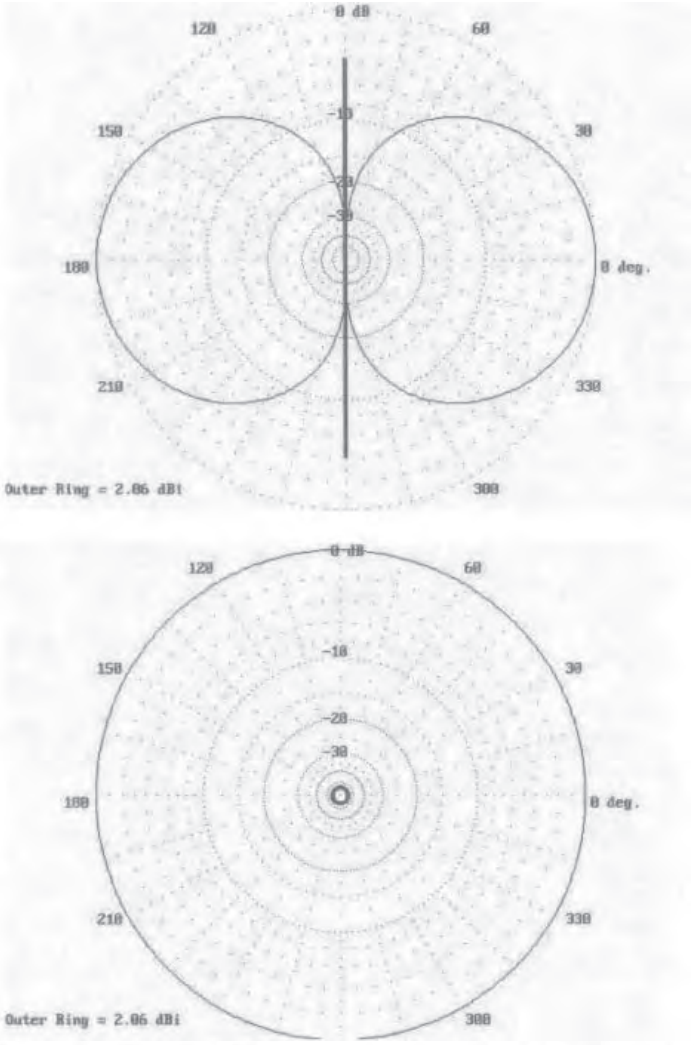

Figure 2–34: Horizontal and vertical pattern of a dipole antenna

INTEL-1020
U.S. Patent 9,265,043

Figure 2–35 shows the horizontal and vertical patterns of an antenna with horizontal 3-dB beamwidth of 90 degrees. A measure of the directivity is the antenna gain. While the total radiated energy is the same, directional antennas focus the energy in one direction, and therefore the gain in that direction is increased. For example, in Figure 2–35 the energy that would have been radiated to the back is redirected to the front. As a result, the antenna has a higher gain in the front and the range of the antenna in the front is greater. Note that an isotropic antenna is simply a theoretical model and it is not possible to construct one physically. An isotropic antenna has a gain equal to 0 dBi (dBi means dB in comparison with an isotropic antenna). A half-wave dipole has a gain of 2.14 dBi, or it concentrates 2.14 dB more energy in its direction of maximum radiation. In general, the gain of an antenna should take into account two other factors. The gain should be reduced by the ohmic losses in the antenna and by the mismatch factor.

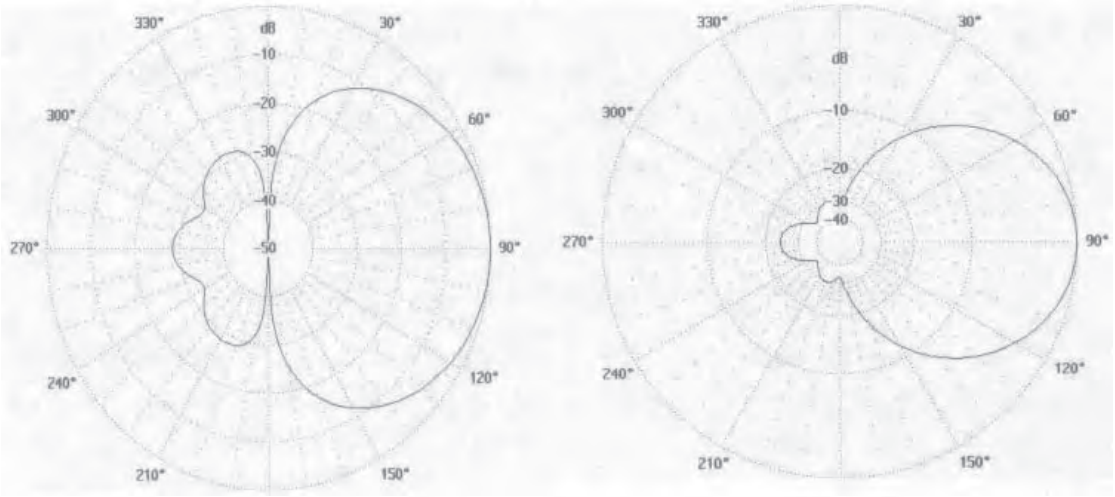

Figure 2–35: Horizontal and vertical pattern of a 90° directional antenna

Too much gain is not permitted, and regulations should always be observed. Wireless LAN access points have special connectors to keep from attaching nonpermissible antennas. Modifying the connector to accept higher gain antennas is not permitted. It is allowable to use higher gains than 6 dBi, provided the transmitter power is reduced by the same amount that the antenna exceeds 6 dBi.

INTEL-1020
U.S. Patent 9,265,043

This means that observing the gain and antenna pattern together is needed when choosing directional antennas. An antenna such as this could be placed in the corner of a work area, concentrating the energy into the work area and radiating very little outside the building. Directional antennas are normally available in patterns of 30, 45, 60, 80, 90, 120, and 180 degrees. In addition, some antennas have adjustable beamwidths, but are more expensive.

Another antenna issue is antenna diversity. Antenna diversity is simply taking the uncorrelated signals from two or more different antennas and combining them for added signal strength. Cellular and PCS systems have been using two receive antennas at the base station to reduce multipath fading for many years. It is now becoming common to use diversity in wireless LAN systems. There are several types of diversity: spatial, polarization, pattern (angle), temporal (time), and frequency diversity. Only spatial, polarization, and pattern diversity are practical for wireless LAN antenna systems. Currently, spatial diversity is the most widely used form of diversity. Some access point products have two antennas, such as Proxim's Harmony. One antenna is used as the primary transmitting and receiving port, while the other is periodically polled to see if it is receiving a stronger signal than the main antenna. It is known that signals received from diversity antennas spaced half a wavelength apart can be essentially uncorrelated, depending on the channel impulse response. In the 2.4 GHz range, half a wavelength is 62 mm, or 2.5 in., and in the 5.8 GHz range, half a wavelength is 26 mm, or about 1 in. Therefore, spatial diversity can be employed in consumer products.

Polarization diversity is the placement of antennas or antenna elements separated by an angle of 90 degrees. When a signal hits an object, such as a metal cabinet, a wall, etc., not only is there reflection, but there is also polarization change. Positioning two antennas with different polarizations can increase signal strength by up to 3 dB.

Another type of diversity that also makes sense is pattern or angle diversity. It involves using two or more different radiation patterns combined to form a more suitable pattern. Pattern diversity can be used in PC cards to make the radiation pattern for the laptop close to a hemisphere above the laptop. This radiation pattern would maximize robustness and coverage and can be achieved with a dipole antenna with a directional antenna (Figure 2–36).

INTEL-1020
U.S. Patent 9,265,043

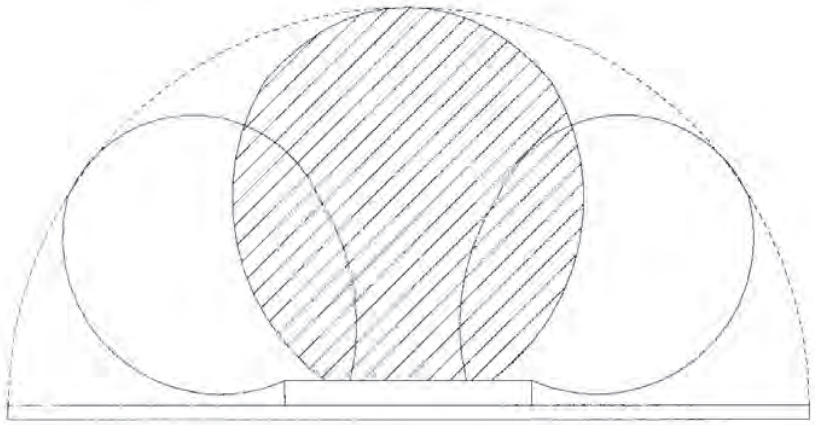

Figure 2–36: Pattern diversity on the laptop.

Note that the use of smart antennas makes the support of mobility more difficult, because even a small movement will have a significant effect. To support mobility with smart antennas, advanced adaptive or blind signal processing steps are needed.

The use of smart antennas [B104] with IEEE 802.11 WLANs has not been extensively considered. It is not impossible, but at present, the standard does not efficiently support smart antennas [B115]. While in the past, the cost of smart antennas has been high for consumer products, the price has been decreasing. In the future, perhaps they can be used by WLANs for increased performance. At present, some companies offer smart antenna solutions for 802.11. The main direct advantage of smart antennas is reduced transmit power. This makes frequency reuse possible, improving the capacity and battery life. One problem with smart antennas is that the position of the antenna to achieve the correct beam pattern is not known in advance. A possible solution is to send the RTS and CTS packets with omnidirectional pattern and obtain channel information during the RTS/CTS exchange [B112]. During the RTS/CTS exchange, stations determine the appropriate beamforming, which will be used during the data exchange. This still allows a great deal of flexibility and, therefore, is an opportunity for vendor differentiation. Smart antennas have been used together with PCF operation [B115].

INTEL-1020
U.S. Patent 9,265,043

To determine the AP and antenna placement, two approaches can be used. The first approach is to perform a site survey. The second approach is to use computer simulations. A site survey is done by performing direct RF measurements. Some prior planning is needed before the actual site survey is performed regarding usage overlap and number of users per AP. A 20 percent overlap in coverage area is recommended, because laptops on the edge may move in and out of the coverage area either by physical movement or the by fading of the access point's signal. The overlap in coverage creates a problem, however. The performance of the network will depend on the algorithm that the station uses to select an AP. An algorithm that selects the AP with the strongest signal will not always provide best overall results. Other criteria that could be used include capabilities and channel conditions. To facilitate AP selection and roaming, it is desireable for stations to receive information about channel conditions from the APs. This is not supported by the standard at present and is being investigated by 802.11k, a Task Group devoted to radio resource management. For example, the APs could include information about channel conditions and neighboring APs in their beacons.

In the site survey, floor plans and building material need to be considered. Concrete and steel attenuate the signal more than other building materials. RF interferers can be identified with spectrum analyzers. As a rule of thumb, APs should be kept at least 20 feet from microwave ovens or other access points.

An example of a floor plan is shown in Figure 2–37. Note that the main obstacles to RF transmission are the elevator shafts, marked with large Xs.

The floor plan shown in Figure 2–37 can be covered with some overlap, as shown in Figure 2–38.

The goal of installing a WLAN is to maximize coverage with as few access points as possible. It is difficult to use computer simulations. For example, floor plans may not be available in suitable electronic form, and it is difficult to take everything into the computer model. A site survey, together with directional and diversity antennas, will yield the best results in practice. Many offices have narrow and odd-shaped areas. Narrow beamwidth antennas can effectively cover long hallways, wider beamwidth antennas are more suitable in wide areas, and 90 degree antennas are a good choice for placement in corners. Furthermore, large spaces have large delay spreads, making diversity antennas effective. In the floor

INTEL-1020
U.S. Patent 9,265,043

plan in Figure 2–37, the use of directional antennas allows the area to be covered with four antennas instead of six (Figure 2–39) [B41].

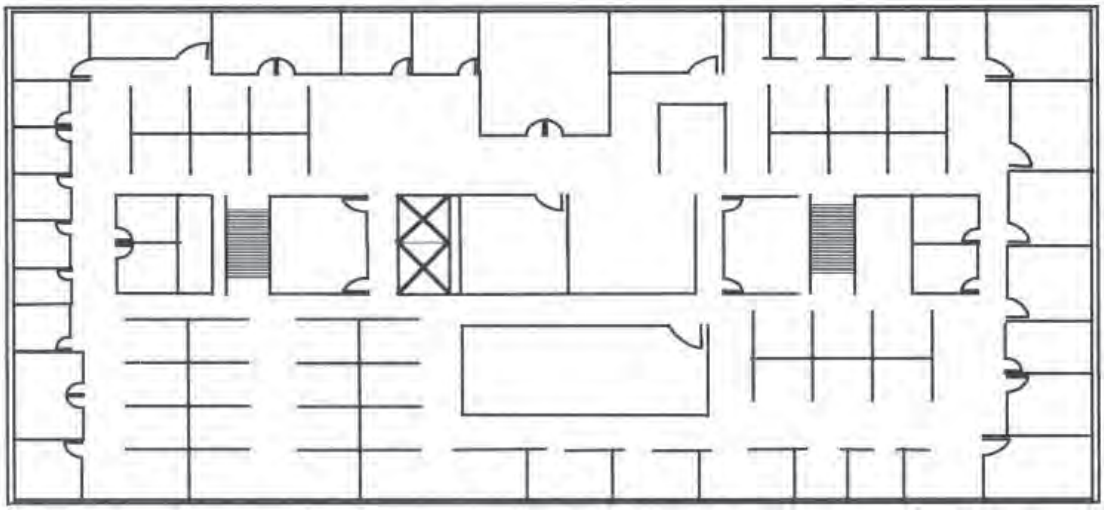

Figure 2–37: Typical floor plan

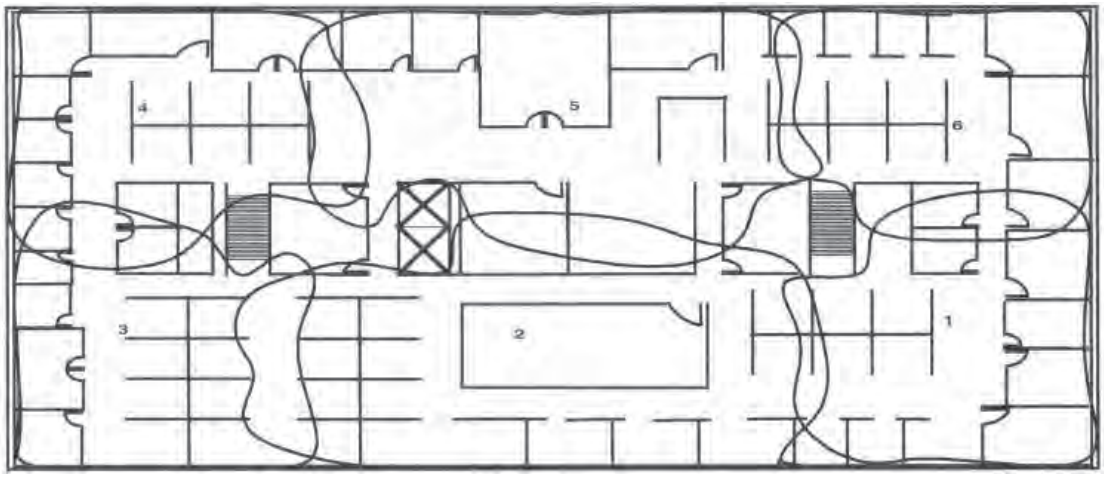

Figure 2–38: An example of an WLAN installation

INTEL-1020
U.S. Patent 9,265,043

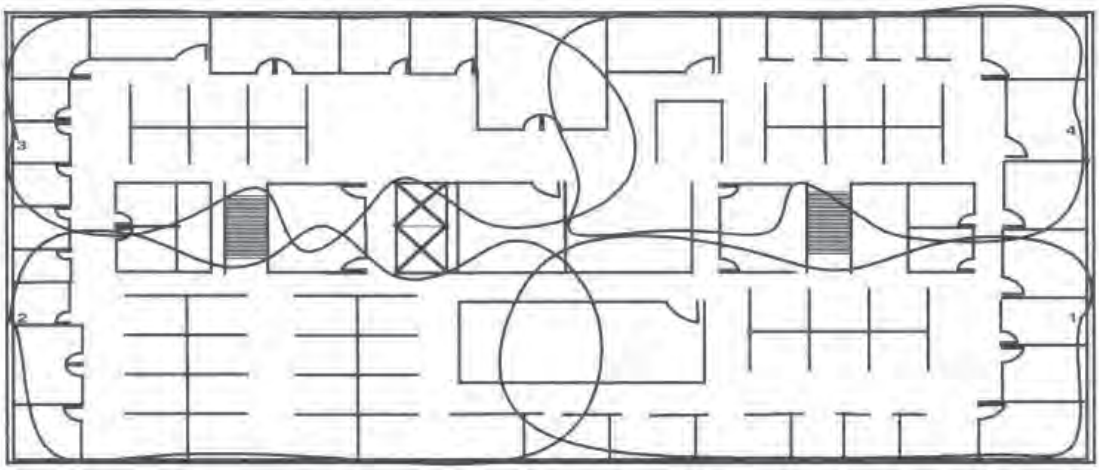

Figure 2–39: An installation of WLAN with directional antennas

After the APs have been located and their coverage areas measured, radio channels are assigned to the APs. A total of eleven DSSS radio channels are available in the 2.4 GHz band, and of these, three have minimal spectral overlap, as shown in Figure 2–23 and Figure 2–24. It is interesting to note that making these frequency assignments is essentially a map-coloring problem with three colors, and there are various algorithms that give optimal or near-optimal assignment of the three radio channels, given the AP locations and coverage areas.

Another factor to consider in the design of a WLAN is areas with high and low densities of users. Areas of high densities are convention centers, airports, etc. Obviously, high-density areas will need more APs. The number of necessary APs N can be determined in the following way:

$$N = \frac{Th_{\text{user}} \times N_{\text{users}} \times A}{\text{Data rate} \times \text{efficiency}} \qquad \text{Eq. 2–11}$$

where Th_{user} is the desired throughput per user, N_{users} is the number of users, A is the average activity rate per user, Data rate is the data rate at which users communicate with the APs, and efficiency is the throughput efficiency in percentage terms at this data rate. For example, for an 802.11a network, suppose that there are 100 users, and each user communicates with the APs at 24 Mb/s with efficiency equal to 0.6 (throughput of 60 percent), and each user is active

INTEL-1020
U.S. Patent 9,265,043

only 20 percent of the time. If the required throughput per user is 3 Mb/s, the number of APs necessary to provide the required throughput is

$$N = \frac{3 \times 100 \times 0.2}{24 \times 0.6} \qquad \text{Eq. 2–12}$$

Since the number of APs is an integer, in this example, 5 APs must be used to provide the desired performance.

It may be desirable to reduce the output power of some APs, because this will reduce their coverage areas. As a result, they will serve fewer users. To achieve this, some WLAN products allow one to set receiver threshold, thus controlling the size of the coverage area. Normally the receiver threshold will be set to a minimum to maximize the coverage area. In high-density areas, the threshold may be set to a higher value to reduce the coverage area.

Another factor to help select an appropriate AP is the QoS that can be expected. Therefore, as part of 802.11e APs will broadcast in the beacons information on the number of stations that are associated with them, channel utilization, and the remaining amount of bandwidth capacity available. Channel utilization is defined as percentage of time the AP has determined the medium to be busy. Therefore, roaming stations can select an AP that is likely to accept future admission control requests.

IEEE 802.11 Technology and Business Trends

The market success of 802.11 and widespread proliferation of devices has caused some additional technical problems. As was said in earlier, equipment must be certified by the regulatory body in the country of sale. In addition, most radio regulations do not permit transmissions by equipment that is not certified for those regulations. Wireless devices, however, do not understand in which country they are located. The 802.11d part of the standard tries to deal with this issue. Stations that implement IEEE 802.1d will maintain a list of countries in which they are certified to operate and the transmission parameters for these countries. These stations can obtain information about the regulatory domain in which they are located, without sending any transmissions. This regulatory domain information (such as the country codes defined in ISO/IEC 3166-1:1997) is in the beacon frame and stations will use only passive scanning initially. Active scanning is not

INTEL-1020
U.S. Patent 9,265,043

allowed at first. Only after obtaining the regulatory domain information, can the station determine if it is certified to operate in this regulatory domain. Only after this determination is done, can a station legally transmit and perform, for example, active scanning. 802.11d is not specific to any particular physical layer.

Yet another 802.11 project that was launched in Nov. 2002 is 802.11k for radio resource management. The goal is to provide interfaces for radio and network measurements. Radio resource management provides mechanisms that support intelligent admission of sessions, and distributions of traffic and power. The goal of radio resource management is to achieve optimized usage of radio resources and maximum system capacity. The current draft of P802.11k [B76] is still in an early stage of development. According to the draft, there are dedicated and concurrent measurements. Dedicated measurements require the interruption of the normal operation of a device. Concurrent measurements are performed during the course of normal operation. Measurement start times are random. 802.11k will also try to identify hidden nodes and produce hidden node reports.

Motivated by the need for higher data rates, 802.11 established Task Group n to develop a standard for an even higher data rate systems. One simple method to increase data rate is to use a higher-order modulation, for example 256 QAM, or 512 QAM. This, however is not efficient, because these modulation schemes require very high signal-to-noise ratios and the data rate will be increased only at very small distances, about 10 meters. Another technique to increase the data rate of 802.11a (and 802.11g) is to reduce the sampling interval from 50 *ns* to 25 *ns* to and preserve the same number of carriers—64. This will have the effect of reducing the symbol period by a factor of 2, from 4 μs to 2 μs. However, the time duration of the cyclic prefix is also reduced by 2 and becomes 400 *ns*. If the delay spread of the channel is longer than 400 *ns*, this approach will lead to increased ISI, if additional equalization is not used. Another simple approach is to double the number of subcarriers from 64 to 128. This also reduces the sampling interval to 25 *ns*, but preserves the symbol period and the duration of the cyclic prefix. These approaches perform a little better compared with the simple use of a higher-order modulation.

Another approach to increase the data rate is based on the following idea. IEEE 802.11a uses the same QAM constellation on all 48 data carriers. This encodes the same number of bits on all the carriers. A better approach is to use a different

INTEL-1020
U.S. Patent 9,265,043

QAM constellation for every carrier, i.e., to encode different number of bits per carrier [B25], [B100]. In a multicarrier system, the process of modulating different number of bits on each carrier based on the SNR of the carriers is often called bit loading. Bit loading is much better suited to the characteristics of a multipath channel. Bit loading also enables constellations sizes much higher than 64QAM to be used. One design requirement in a system that uses bit loading is that it requires a feedback mechanism to communicate the bit-allocation table from the receiver to the transmitter. This communication from the receiver to the transmitter must happen faster than substantial changes in the channel profile.

Further improvement in data rate can be obtained combining bit loading with coded modulation techniques (such as trellis coded modulation (TCM)). TCM maps the coded bits to points in the signal space such that the Euclidean distance between transmitted sequences is maximized. Typically the method of set partitioning is used to maximize the Euclidean distance. Since coded modulation schemes do not encode all information bits, coded modulation must be combined with bit loading in multipath channels to achieve the coding gain benefits. If the average number of bits per carrier remains the same, this technique will not lead to a higher data rate. Bit loading together with trellis coding allows for a higher average number of bits per carrier. Figure 2–40 [B25] illustrates the higher data rate that can be obtained with the new approach for the lower UNII band. The results on this figure are obtained assuming transmit power of –56.9 dBm/Hz (this is less than 40 mW), receiver noise figure of 10 dB, and an additional implementation loss of 5 dB. The thermal noise level is assumed to be –174 dBm/Hz. The gain of the receive and transmit antennas is assumed to be zero.

In Figure 2–40, the 802.11a system uses equal number of bits per carrier. The maximum data rate for the two systems is determined as the maximum data rate in which the packet error-rate is 10 percent. Since in the simulation the maximum output power is considerably less than the maximum per FCC regulations significant further improvement in range can be obtained. Yet another approach for high-data rate wireless systems is based on a technology called vector OFDM, or multiple-input multiple-output OFDM (MIMO OFDM) [B8], [B94], [B149], [B150]. This technology achieves considerably higher capacity, but requires computationally intensive signal processing operations. Finally, note that it is possible to combine several or all of the above-mentioned methods to achieve

INTEL-1020
U.S. Patent 9,265,043

even higher data rates. Some early research indicates that it is possible to achieve data rates in excess of 250 Mb/s [B43].

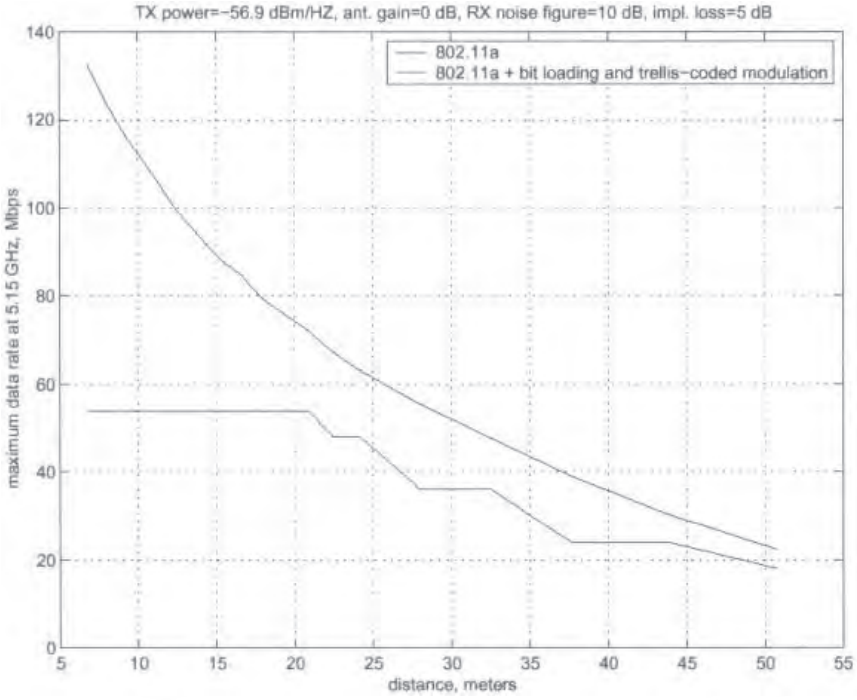

Figure 2–40: Comparison of the maximum data rates that can be achieved with the specified simulation parameters for a plain 802.11a system and 802.11a system with bit loading and trellis coding

The 802.11n Task Group realized from the beginning that throughput must be increased, and not just the maximum data rate. Throughput, however, depends primarily on the MAC protocol [B145]. Therefore, 802.11n will likely develop not only a physical layer, but also appropriate MAC modifications to ensure high throughput.

In July of 2003, within IEEE 802.11, a Study Group on Dedicated Short-Range Communications (DSRC) in the 5.9 GHz band was founded. As in the past, the direct motivation was the allocation in 1999 by the FCC of 75 MHz of bandwidth (5.85–5.925 GHz) right above the higher UNII band for communication between vehicles and roadside systems and between vehicles. Similar bands exist in Japan and Europe. In addition, further FCC rules were expected by the end of 2003.

INTEL-1020
U.S. Patent 9,265,043

Furthermore the U.S. Congress directed the Department of Transportation to develop a standard to ensure interoperability in the DSRC band. This resulted in the ASTM E2213 standard [B128], developed by the American Society for Testing and Materials (ASTM), and largely based on 802.11a. The standard is written as an amendment, specifically allowing 802.11a chipsets to be used in DSRC equipment. The MAC modifications introduced by ASTM E2213 can be satisfied using a simple firmware upgrade. The applications of this standard are numerous and include safety enhancements (intersection collision warning, etc.), roadside kiosks (updated map information, weather updates, road conditions), Internet access hot spots at service stations and parking lots, electronic toll payment, inventory tracking, etc. The 802.11 Study Group was established because the ASTM finds it more appropriate for the DSRC standard to be developed further and maintained by 802.11. As 802.11 evolves, it is easier to update DSRC together with 802.11.

The traditional view of wireless communication networks is that there are two network categories: voice-oriented and data-oriented. Each of these categories can be further subdivided into two subcategories depending on the intended coverage area: local area or wide area.

Wireless voice systems were developed long before wireless data systems. Wireless voice systems have been in use in the USA. since the 1930s when the first police vehicles were equipped with one-way and then two-way radios. Cellular telephony was technically developed at Bell Telephone Laboratories in the 1940s and commercially deployed on a wide scale in the 1980s. Simultaneously with the development of 802.11, cellular telephony has made enormous progress in its ability to handle not only voice, but also data. Third-generation cellular telephony will offer up to 2 Mb/s at a fixed location, although the real throughput will be much lower. In the meantime, before third-generation networks are widely deployed, the so-called 2.5 generation of technology will offer data rates up to 172 Kb/s, with actual speeds of about 42 Kb/s.

The main business trend in cellular telephony is the market saturation; slow growth in revenue from voice services and higher growth in revenue from data services. Clearly the fundamental long-term trend is the merger between voice-only and data-only networks. This merger of voice and data networks

INTEL-1020
U.S. Patent 9,265,043

requires seamless roaming between cellular and wireless local area networks [B101], [B117]. Already there are technologies and products on the market that enable single-subscriber identity on these two different networks. Future wireless networks will be a combination of several radio communication technologies, such as third generation radio access and wireless LAN [B12]. With top speeds of 54 Mb/s WLAN technology will continue to provide much superior bandwidth compared to any cellular technology.

At present commercially available WLAN devices are predominantly network adapters (NAs) and access points (APs). NAs are available as PC cards, physically separate from laptops. The market though is moving towards integrated devices. Many laptops sold at present have integrated WLAN interfaces with 802.11b, and some with 802.11a and 802.11g. In the next several years WLAN devices will be integrated in many consumer electronic products as well.

Today most large offices are equipped with appropriate wiring for conventional LANs. The first market for wireless LANs is large indoor areas and offices and buildings with wiring difficulties, such as historic buildings, etc., large indoor areas such as manufacturing floors, warehouses, convention and exhibition centers, libraries, etc. WLAN technology is increasingly the network of choice. In today's fast-paced world with frequent relocations of equipment and high number of temporary offices, portability significantly increases productivity. A major factor in the market growth of WLANs is that the market for laptops and other portable computing devices is growing faster than the desktop market. Obviously portable computing devices naturally call for portable LAN, i.e., WLAN. WLANs also will create new markets for equipment and services.

Overall, 802.11 not only improves business productivity, but is changing the dynamics of some business. For example, at conferences 802.11 allows instant polling, instant journalism, etc.

An interesting business question to consider is to what extent will WLAN compete and to what extent will it complement other wireless data communication technologies. There is no simple answer to this question—it depends on the business model. To some cellular service providers offering high-data rate service 802.11 may appear as a foe. Clearly combining 802.11 with directional antennas for extended coverage creates more significant competition to wide area wireless networks. Many companies have started providing high-speed data services using

INTEL-1020
U.S. Patent 9,265,043

WLANs in places such as coffee shops, fast food restaurants, airports, convention centers, etc. T-Mobile is the largest WLAN service provider at present. A number of start-up companies have business models involving providing WLAN service.

IEEE Std 802.11 is a mature and low-cost technology, and in addition operates in the unlicensed spectrum. At the same time mobile phone service providers are introducing 3G technology in the marketplace. The cost per customer of 3G networks is higher than that of WLAN because of higher equipment costs and operation in expensive spectrum. In addition, WLAN offer higher data rates compared with 3G networks. Table 2–5 shows a comparison between cellular telephony and 802.11.

Table 2–5: Comparison between cellular telephony and 802.11

Parameter	Cellular	802.11
Range	High	Low
Coverage	Ubiquitous	Localized to "hot spots"
Power consumption	Low	Medium
Spectrum	Licensed (very expensive)	Unlicensed
Security	Very secure	WEP is broken, but will be secure with IEEE Std 802.11i
Data rate	Low	High
Infrastructure cost	Very expensive	Very cheap

Wireless local area networks are used to provide data-only service. 802.11 is not used to provide voice service and this makes it more suitable as a complement to 3G systems. There are several market opportunities with regard to WLAN and 3G. First, since users value highly seamless mobility and service, some products enable seamless transition between WLAN and 3G networks. When WLAN coverage is available, stations can communicate at a higher data rate. In the absence of WLAN coverage, stations can communicate over 3G networks. The

INTEL-1020
U.S. Patent 9,265,043

two networks complement themselves: one with wide coverage, and the other with enhanced performance in isolated areas. A more interesting question is whether a complete outdoor cellular network can be designed, based on 802.11 [B22]. Such a network would have at the same time cost and performance advantages over current 2G, 2.5G, and 3G networks. The main problem of an 802.11-based cellular network is the limited transmission range. This problem was recently investigated by Clark, et al., in their article, "Outdoor IEEE 802.11 cellular networks: radio link performance" [B22]. The main conclusion is that yes, an outdoor cellular network based on WLAN is feasible, with a cell radius in the range of 0.7 to 3 km, assuming 802.11b and 1 Mb/s data rate. However, this requires additional signal processing, since the rms delay spread for outdoor urban areas is about 1 microsecond. To mitigate this large delay spread a RAKE receiver is necessary.

While the first big wave in the deployment of 802.11 was its use in offices, the next big wave is expected to be the "unwiring" of consumer electronics devices such as TV sets, CD and DVD players, etc. Since this will unleash the long-awaited demand for video over the Internet, this will likely have a significant and much-needed impact on the whole communications industry. The acceptance of 802.11a by the computer industry will likely push 802.11a as a link between consumer electronic and computer devices.

Not only is the number of Wi-Fi hot-spots growing, but there are places where these hot spots are adjacent and installed in a way to provide continuous coverage. Cities with the most installed access points are San Francisco, followed by New York, Seattle, Chicago, and San Jose. This does not include accidental hot spots leaking out of private homes and offices. In the years ahead 802.11 will become a truly universal standard, found everywhere in the world. Current market research estimates that 99 million people will be using 802.11 by the year of 2006.

INTEL-1020
U.S. Patent 9,265,043

Copyright © 2004 IEEE

Chapter 3 Standards for Wireless Personal Area Networking (WPAN)

INTRODUCTION

While the growth and success of WLAN is due to the market growth of laptops, the main market demand for wireless personal area networking (WPAN) technology comes from the fast growth of small and inexpensive devices many of which have Internet connectivity. Examples of these electronic devices are cellular phones, personal digital assistants, personal music players, digital cameras, etc. As was mentioned in Chapter 1, it is a basic law of communications technology that the value that these devices represent to consumers would be significantly enhanced if they were able to communicate among themselves and with other devices. Traditionally, proprietary special-purpose cables have been used to solve this short-range connectivity problem. However, most users find that cables not only limit mobility, but are outright frustrating. Therefore, it becomes quite desirable to develop connectivity technologies for interconnecting personal devices that do not require the use of cables. These connectivity technologies are known as wireless personal area networks (WPANs).

COMPARING WPAN AND WLAN

What are the similarities and differences between WPAN and WLAN? At a first glance, a WPAN may appear to resemble a WLAN, like IEEE 802.11. Both the WLAN and WPAN are short-range wireless data communications technologies. However, WLANs have been designed and are optimized for portable computing client devices such as notebook computers. WPAN devices are even more mobile. The two technologies differ in three ways: power levels and coverage, control of the media, and lifespan of the network.

Typical WLAN coverage distances are on the order of 100 m with transmit power of 100 mW, and about 500 m with transmit power of 1 W. Furthermore, power consumed by the device can be significant. Mobile devices typically operate on

INTEL-1020
U.S. Patent 9,265,043

batteries. Portable devices are normally used while in a fixed location and therefore usually will run from power supplied by wall sockets. WLANs are typically wireless extensions of wired LANs. The devices participating in a WLAN are mainly of one type (laptops), while the devices participating in a WPAN can be of several different types. WPAN devices cover areas of about 10 m with transmitted power of about 1 mW. WPAN technology includes additional measures to ensure low power consumption by the device, making battery operation possible for long enough periods.

Note the distinction between "mobile" and "portable" devices. Mobile devices typically operate on batteries. Portable devices are normally used while in a fixed location and therefore usually will run from power supplied by wall sockets.

Because WLANs cover relatively large areas with many uncoordinated devices, the mechanism for access to the medium must handle potential collisions. CSMA/CA, used in IEEE 802.11, is a very efficient protocol for WLANs. WPANs cover small areas where devices can be synchronized, significantly reducing collisions. As a result, the medium access protocol can be different from the one used for WLANs.

The final difference between WPAN and WLAN is the different lifespan of the networks. As bona fide members of a larger infrastructure, WLANs do not have inherent or implied lifespan. They have "existence" independent of their constituent devices. This is not true for WPANs. WPAN networks are ad-hoc, i.e., they are created in a spontaneous fashion, cover a small area, and exist for a small period of time. Connection establishment for ad-hoc networks must be very fast. Personal devices that participate in a WPAN are designed for their personal appeal and functionality. They are not designed to be members of an established networking infrastructure, although they may connect to it when necessary. In a WPAN, a device creates a connection that lasts only for as long as needed and has a finite lifespan. For example, a file transfer application may cause a connection to be created only long enough to accomplish its goal. When the application terminates, the connection between the two devices may be severed as well. The devices to which one's personal device is currently connected in a WPAN may bear no semblance to the device that it was previously connected to or it will connect to next. For example, a notebook computer may connect with a PDA at one moment, a digital camera at another moment, or a cellular phone at yet

INTEL-1020
U.S. Patent 9,265,043

another moment. At times, the notebook computer may be connected with any or all of these other devices.

The demands of the consumer market require WPAN technology that is inexpensive, consumes very little power, and efficiently supports QoS. Further, these devices will likely be used in many different environments such as offices, homes, convention centers, shopping malls, airports. Therefore, these devices must satisfy the marketing requirements dictated not only by the consumer market but the business market as well.

The range of operation is short, but there may not be a line of sight. The physical layer must work in a non-line-of-sight (NLOS) environment.

The IEEE 802.15 Working Group has taken a leading role in WPAN technology development. First, the Bluetooth protocol [B13] was standardized as IEEE Std 802.15.1-2002 [B78]. To achieve this, the IEEE 802 Working Groups received a nonexclusive license from the Bluetooth Special Interest Group. However, IEEE 802.15.1 not only standardized Bluetooth, but in the process of developing IEEE 802.15.1 on the basis of Bluetooth, IEEE 802.15.1 made significant contributions to the Bluetooth specification. Second, the 802.15 Working Group undertook serious work to develop a recommended practice for the coexistence of IEEE 802.11b and Bluetooth. The results of this effort became IEEE Std 802.15.2™-2003 [B79]. Third, IEEE 802.15 developed two other WPAN technologies—a high-rate WPAN (IEEE Std 802.15.3™-2003) [B80] and a low-rate WPAN (IEEE Std 802.15.4™-2003) [B81]. This chapter is devoted to these technologies.

The organization of the chapter follows the standardization work within IEEE 802.15. First, IEEE 802.15.1 is presented. The coexistence mechanism is discussed next, followed by a discussion on the high-rate WPAN. And finally, the low-rate WPAN is presented.

IEEE 802.15.1

Overview

The first WPAN standard is Bluetooth, developed by an industry consortium called Bluetooth Special Interest Group (SIG). The original intent was that Bluetooth would be like a cable replacement with a range of operation of about 10 m, very little power consumption, and very inexpensive. Thus Bluetooth is

INTEL-1020
U.S. Patent 9,265,043

optimized for power-conscious, battery-operated, small size, lightweight personal devices.

Led initially by Ericsson®, Nokia®, IBM®, Intel®, Toshiba® and later joined by 3Com, Motorola®, Lucent, and Microsoft®, the Bluetooth SIG has about 2000 member companies, which demonstrates the enormous interest in the industry [B53].

A Bluetooth WPAN supports both synchronous communication channels for toll-quality voice communication and asynchronous communications channels for data communications. Unlike IEEE 802, the Bluetooth standard covers all layers in the OSI model. However, an IEEE 802-style MAC and PHY are not used. The precise mapping between the elements of the Bluetooth protocol and IEEE 802-style MAC and PHY is cumbersome and debatable. It was attempted but ultimately abandoned by the IEEE 802.15 Working Group.

Figure 3–1 shows the protocol stacks in the OSI 7-layer model and the Bluetooth wireless technology and their relation as it pertains to this standard. As shown in Figure 3–1, the logical link control (LLC) and MAC sublayers together encompass the functions intended for the data link layer of the OSI model.

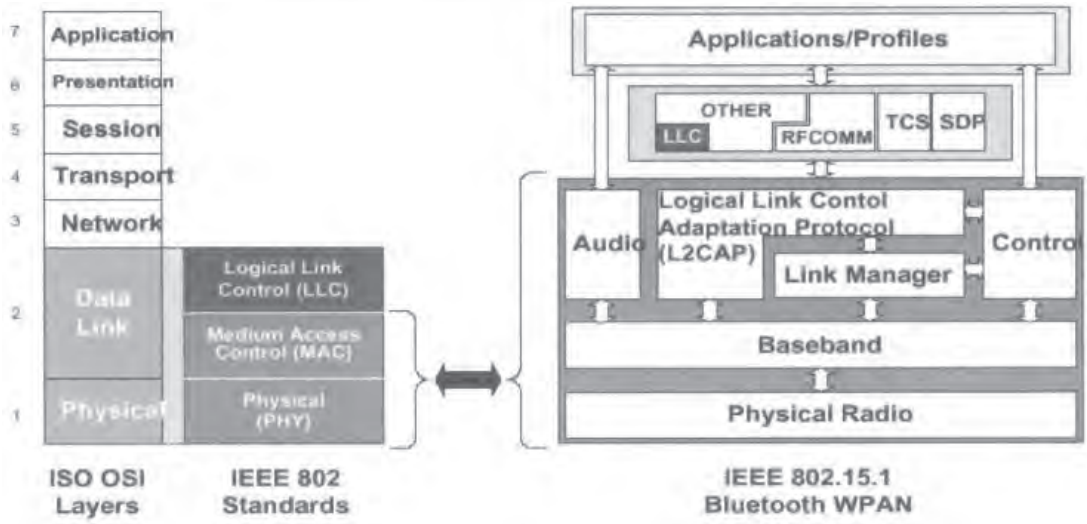

Figure 3–1: The mapping between the Bluetooth protocol and the IEEE 802 architecture

INTEL-1020
U.S. Patent 9,265,043

IEEE 802.15.1 is derived from the Bluetooth version 1.1 foundation core [B13]. To make Bluetooth an IEEE 802 standard, the Bluetooth SIG granted the IEEE a nonexclusive license to copy and adapt the Bluetooth specification.

To ensure the commercial success of Bluetooth and enable the creation of interoperable, and interactive applications, the Bluetooth specification in addition to the communication protocols includes interoperable applications. These applications support various usage scenarios addressed in the specification. IEEE 802.15.1 covers only a subset of the communications protocols in the Bluetooth specification related to PHY and MAC protocols as identified in Figure 3–1. It includes both Bluetooth-specific protocols such as the link-manager protocol (LMP) and the logical link control and adaptation protocol (L2CAP) and non-Bluetooth-specific protocols, which are grouped in the "Other" box in Figure 3–1.

The RFCOMM layer is a serial port emulation layer for enabling legacy applications over Bluetooth links. The TCS is a telephony control and signaling layer for advanced telephony applications. The SDP is a service discovery protocol layer allowing Bluetooth devices to inquire other devices for the services that they can provide. These layers are not part of IEEE 802.15; they are specified only in the Bluetooth specification.

The L2CAP protocol supports higher-level protocol multiplexing, and the requirement for this multiplexing segmentation and reassembly. Due to the small size of the packets, large upper-layer packets need to be segmented before transmission over the air.

The Bluetooth specification refers to the control protocol, (the rightmost block in Figure 3–1), as the host-control interface (HCI). It has two functions. First, as shown in the figure, it provides physical interface to higher layers. Because IEEE 802 standards do not specify implementation, this type of interface is outside the scope of IEEE 802 protocols. Second, it provides access to hardware status and control registers. This function is implemented by the management sublayer in IEEE 802 standards.

IEEE 802.15.1 Physical layer

The Bluetooth WPAN operates in the unlicensed 2.4 GHz, the Industrial, Scientific, Medical (ISM) band. The center frequency of these channels for the

INTEL-1020
U.S. Patent 9,265,043

United States, Europe, and most of the world is $f_c = 2402 + k$ MHz, where $k = 0,\ldots,78$. France has different regulations in the 2.4 GHz band, and the center frequency is $f_c = 2454 + k$ MHz, where k takes values between 0 and 22.

According to the output power, the equipment is divided into three power classes. Class 1 equipment has a maximum transmit power of +20 dBm (100 mW) and minimum output power of 0 dBm (1 mW). These power level requirements are, like in all wireless 802 standards, at the antenna connector of the equipment. The output power of class 1 devices can be adjusted in steps ranging between 2 dB and 8 dB between the maximum and minimum power levels. This adjustment of power levels in a monotonic sequence is called power control and is required for class 1 equipment. Power control achieves not only power efficiency, but minimizes potential interference to other wireless systems.

Note that the transmitter will not know how to adjust its output power level without feedback from the receiver. Therefore, power control can be used only if the receiver can provide feedback; i.e., the receiver can perform measurements (such as received signal strength indicator [RSSI] levels) and there is a protocol to exchange information with the transmitter about the power levels. Clearly, power control cannot be used before a connection is set up. At connection set up the devices exchange information about their capabilities. If the receiver is not capable of providing appropriate feedback, then the transmitting device must switch from class 1 to class 2 or class 3.

Class 2 equipment has a maximum transmit output power of 2.5 mW (4 dBm) and a minimum output power of 0.25 mW (–6 dBm). Class 3 equipment has a maximum output power of 1 mW (0 dBm) and no minimum. Power control capability for class 2 and class 3 devices can be used, but is optional.

A fast frequency hopping (1600 hops/sec) transceiver is used to combat interference and fading in this band. A slotted channel is used, which has a slot-duration of 1/1600=625 μs. Each packet is transmitted on a different frequency in the hopping sequence. Channel spacing is 1 MHz. High-power efficiency and low-transceiver cost cannot be achieved with modulation schemes like quadrature amplitude modulation (QAM). Bluetooth uses a constant-envelope modulation called Gaussian Frequency Shift Keying (GFSK). A binary 1 is represented by a positive frequency deviation, and a binary 0 is represented by a negative frequency deviation. The transmitted data has a symbol rate of 1 Msymbol/s.

INTEL-1020
U.S. Patent 9,265,043

Baseband

Network architecture

The basic network is called a piconet. A frequency-hopping channel based on the address of the master defines each piconet. All devices participating in communications in a given piconet are synchronized to the frequency-hopping channel for the piconet using the clock of the master of the piconet. Slaves communicate only with their master in a point-to-point fashion, under the control of the master. The master's transmissions may either be point-to-point or point-to-multipoint.

A piconet is a WPAN formed by a Bluetooth device serving as a master and one or more Bluetooth devices serving as slaves. Up to seven slaves can be active in the piconet. In addition, up to 255 slaves can remain locked to the master in a so-called parked state. These parked slaves cannot be active on the channel, but remain synchronized to the master.

Note that the names "master" and "slave" only refer to the protocol—all Bluetooth units are identical. Although certain applications may require that some devices act always as masters or slaves, this situation is outside the scope of the 802.15.1 standard. A slave device at one communications session could be a master in another and vice versa.

A scatternet is a collection of piconets, i.e., a scatternet is a collection of piconets that overlap. A Bluetooth device may participate in several piconets at the same time. Each piconet can only have a single master. Slaves can participate in different piconets on a time-division multiplex basis. In addition, a master in one piconet can be a slave in another piconet. The piconets are not frequency-synchronized. Because each piconet has a different master, the piconets hop independently, each with their own channel-hopping sequence and phase as determined by the respective master. In addition, the packets carried on the channels are preceded by different channel access codes as determined by the master device addresses. (The manufacturer assigns each Bluetooth device a unique 48-bit Bluetooth device address.) As more piconets are added, the probability of collisions increases—a graceful degradation of performance results, as is common in frequency-hopping spread-spectrum systems.

If multiple piconets cover the same area, a unit can participate in two or more overlaying piconets by applying time multiplexing. To participate on the proper channel, it should use the associated master device address and proper clock offset

INTEL-1020
U.S. Patent 9,265,043

to obtain the correct phase. A Bluetooth unit can act as a slave in several piconets, but only as a master in a single piconet. Because two piconets with the same master are synchronized and use the same hopping sequence, they are one and the same piconet.

The scatternet topology poses the question of efficient routing, which is not answered in the 802.15.1 standard. One suitable routing protocol has been developed in Prabhu and Chockalingam's paper, "A routing protocol and energy efficient techniques in Bluetooth scatternets" [B120]. The protocol in [B120] uses the available battery power in the devices as a cost metric. This routing protocol exploits the master-slave switch and the power control capability to achieve near optimal performance with respect to battery life. The master may drain its batter faster than the slaves; therefore, the master-slave switch can be performed (depending on the remaining battery power) and can increase network lifetime.

From the architecture of IEEE 802.15.1, it is clear that it provides a point-to-point connection (master-to-slave or slave-to-master) or a point-to-multipoint connection (broadcast messages from the master to all the slaves). The topology of a Bluetooth network is shown in Figure 3–2. By definition, the master is represented by the Bluetooth unit that initiates the connection to one or more slave units.

Every Bluetooth unit has an internal system clock that determines the timing and hopping of the transceiver. The Bluetooth clock is derived from a free, always running native clock. These native clocks obviously are not synchronized. Synchronization is very important for the Bluetooth standard, because the medium access mechanism follows TDMA principles. To achieve synchronization with other units, offsets are used that are added to the native clock. The timing and the frequency hopping on the channel of a piconet is determined by the Bluetooth clock of the master. After the piconet is established, the master clock is communicated to the slaves. Each slave adds an offset to its native clock to be synchronized to the master clock. Because the clocks are free-running, the offsets have to be updated regularly. The clock determines critical periods and triggers the events in the Bluetooth receiver.

INTEL-1020
U.S. Patent 9,265,043

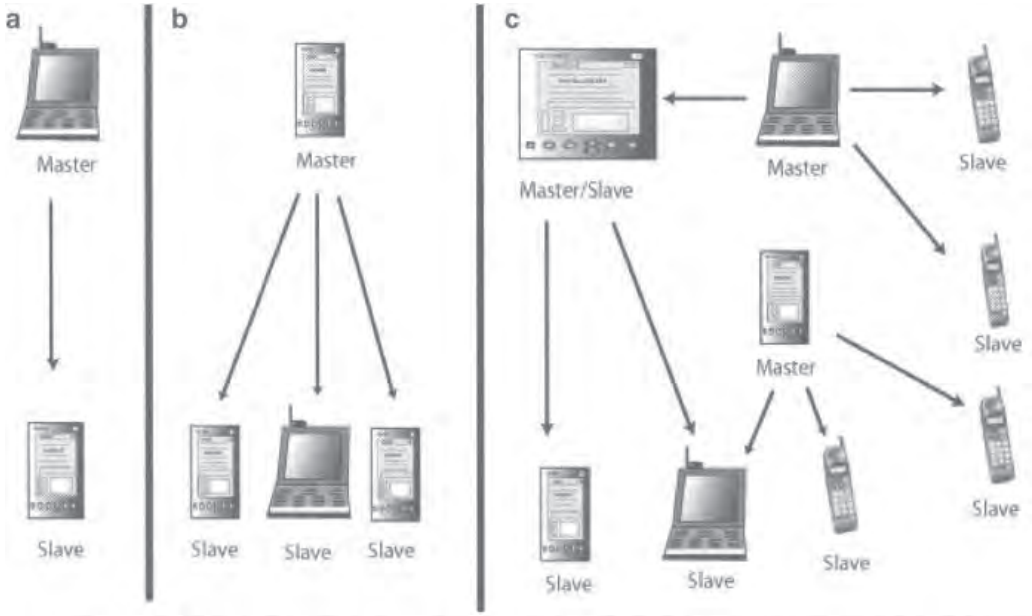

Figure 3–2: Topology of a Bluetooth network: (a) point-to-point, (b) point-to-multipoint, and (c) scatternet

Channel access and QoS

The channel is divided into time slots, each 625 μs long. The time slots are numbered according to the Bluetooth clock of the piconet master. The slot numbering ranges from 0 to $2^{27} - 1$ and is periodic. The channel is represented by a pseudorandom hopping sequence, which hops through the 79 or 23 RF channels. The hopping sequence is unique for the piconet and is determined by the Bluetooth device address of the master. The Bluetooth clock of the master determines the phase in the hopping sequence.

On the channel, information is exchanged through packets. A packet nominally covers a single slot, but it can be extended up to either three or five slots, as shown in Figure 3–3. The number of slots must be an odd number. Each packet is transmitted on a different hop frequency. The RF frequency is fixed for the duration of the packet. The nominal hop rate is 1600 hops/s.

INTEL-1020
U.S. Patent 9,265,043

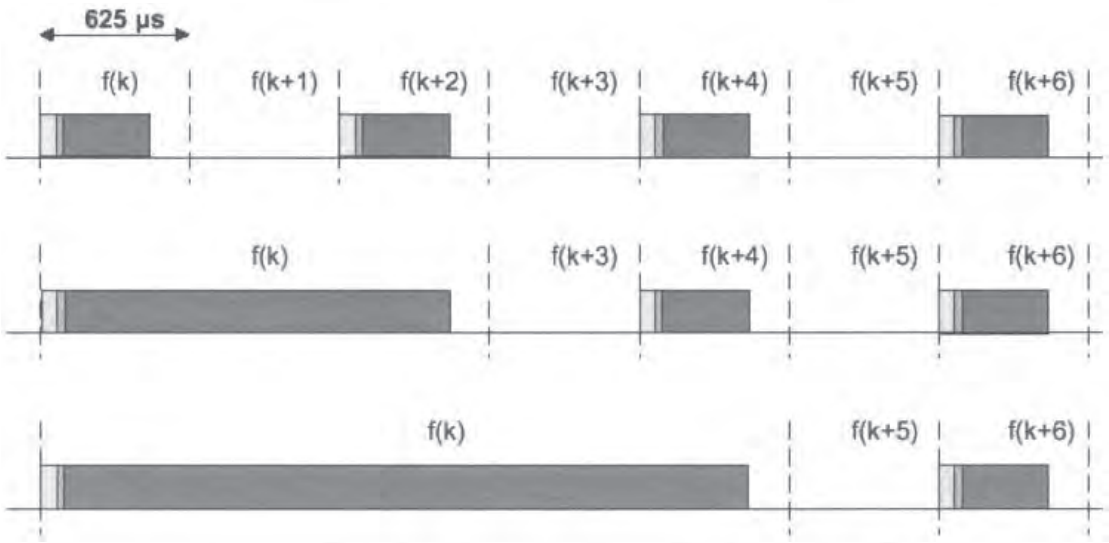

Figure 3–3: One-, three-, and five-slot packets

Each packet consists of three entities: the access code, the header, and the payload. The access code and header are of fixed size: 72 bits and 54 bits, respectively. The payload can range from zero to a maximum of 2745 bits, depending on the packet type. Packets may consist of the (shortened) access code only, of the access code-header, or of the access code-header-payload. The header has information about the packet type and to whom it is addressed. The access code is used for detection, synchronization, DC offset compensation, and piconet and device identification. The access code identifies all packets exchanged on the channel of the piconet—all packets sent in the same piconet are preceded by the same channel access code. The access code is also used by the receiver to perform robust detection, timing synchronization, and offset compensation. For example, detection and synchronization can be performed by sliding correlation of the received signal against the access code. If sliding correlation is computationally complex, other methods can also be used. The access code is also used in paging and inquiry procedures. In this case, the access code itself is used as a message, and neither a header nor a payload is present.

There are three error correction schemes defined for Bluetooth: 1/3 rate forward error correction (FEC) and 2/3 rate FEC, as well as an automatic repeat request (ARQ) scheme for the data. The purpose of the FEC scheme on the data payload is to reduce the number of retransmissions. However, in a reasonably error-free

INTEL-1020
U.S. Patent 9,265,043

environment, FEC gives unnecessary overhead that reduces the throughput. Therefore, while packet headers are always protected by a 1/3 rate FEC, the payload is not always protected. The 1/3 FEC is a simple three-times repetition code. The repetition code is implemented by repeating the bit three times. The other FEC scheme is a (15,10) shortened Hamming code. This code can correct all single errors and detect all double errors in each codeword. The third FEC approach is ARQ.

Before the user information is sent over the air interface, several bit manipulations are performed in the transmitter to increase reliability and security. To the packet header, an HEC is added, both the header and the payload bits are scrambled, and FEC coding is applied (Figure 3–4 and Figure 3–5). Both the header and the payload are scrambled with a data-whitening word in order to randomize the data from highly redundant patterns and to reduce DC bias in the packet. Scrambling is performed by a bitwise XOR with the output of a linear feedback shift register (LFSR), defined by the polynomial $D^7 + D^4 + 1$. In the receiver, the inverse processes are carried out. The received data is descrambled using the same whitening word.

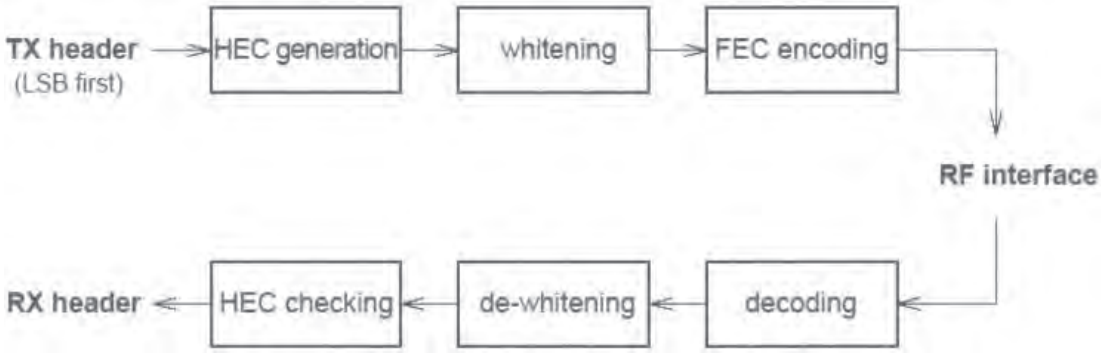

Figure 3–4: Header bit processing

Figure 3–4 shows the processes carried out for the packet header both at the transmit and the receive side. All header bit processes are mandatory. For the payload, similar processes are performed. The packet type determines which processes are carried out. Figure 3–5 shows the processes that may be carried out on the payload. In addition to the processes defined for the packet header, encryption can be applied on the payload. Only whitening and de-whitening are

INTEL-1020
U.S. Patent 9,265,043

mandatory for every payload; all other processes are optional and depend on the packet type and the mode enabled. In Figure 3–5, the optional processes are indicated by dashed blocks. After scrambling at the transmitter, FEC encoding is performed. Descrambling is performed at the receiver after FEC decoding.

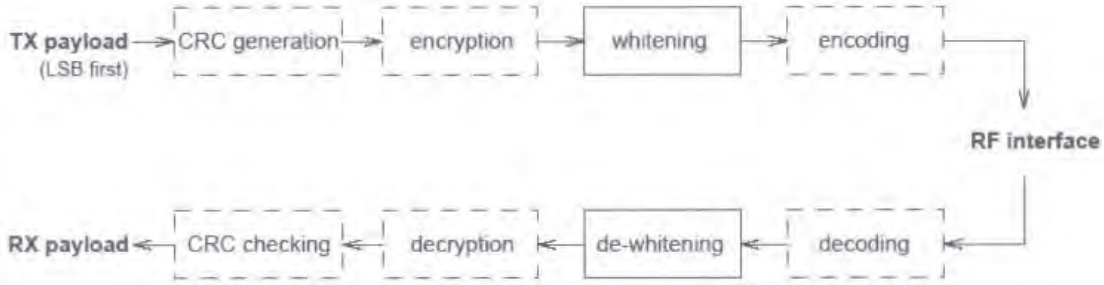

Figure 3–5: Payload bit processing

The piconet is synchronized by the system clock of the master. The master never adjusts its system clock during the existence of the piconet. The slaves adapt their native clocks with a timing offset in order to match the master clock. By comparing the exact RX timing of the received packet with the estimated RX timing, this timing offset is updated each time a packet is received from the master.

Several features are included into Bluetooth to ensure a low-power operation. These features include appropriate signal processing when handling the packets, general architecture, and protocol.

It should be noted that it is much easier for IEEE 802.15.1 to support QoS than it is for IEEE 802.11. The fundamental reason is that IEEE 802.15.1 does not use a CSMA MAC. To emulate full duplex transmission, a Time-Division Duplexing (TDD) scheme is used. The master and slaves transmit alternatively. This controlling mechanism is similar to the point coordination function for WLANs. Recall that the time period for WLANs in which a point coordinator is active is called the contention-free period. In IEEE 802.15.1, because a distributed coordinator is absent, all the time can be considered "contention free". The network is time-multiplexed. The master polls its collection of IEEE 802.15.1 WPAN slaves for transmissions, thus regulating the bandwidth assigned to them. The master can enforce any QoS policy that it desires.

INTEL-1020
U.S. Patent 9,265,043

The jitter can be controlled by the packet size. One-slot packets lead to less jitter. In addition, employing a frequency-hopping scheme with the small-sized slots provides noise resilience to interference from other wireless technologies operating in the same band. The master assigns a three-bit active member address to each slave active in its piconet. The all-zero active member address is reserved for broadcast messages. A slave only accepts packets with a matching active member address and broadcast packets.

The master always starts the transmission in the even-numbered slots, whereas the slaves start their transmission in the odd-numbered slots. For greater efficiency, IEEE 802.15.1 uses a combination of circuit and packet switching. Two link types are defined in a piconet—Synchronous Connection-Oriented (SCO) link and Asynchronous Connection-Less (ACL) link.

The connection between two Bluetooth devices is initially an ACL type. The ACL link is a point-to-multipoint link between the master and all the slaves participating on the piconet. In the slots not reserved for the SCO link(s), the master can establish an ACL link on a per-slot basis to any slave, including the slaves already participating in an SCO link. The ACL link provides a packet-switched connection between the master and all active slaves participating in the piconet. If a slave receives an ACL packet from the master in an even-indexed slot, it is allowed to return an ACL packet in the following odd-indexed slot. ACL packets not addressed to a specific slave are considered as broadcast packets and are read by every slave. If the master sends a broadcast message in a even-numbered slot, no slave is allowed to return a packet in the following odd-numbered slot.

Although the ACL link does not have any latency guarantees, a time interval is defined equal to the maximum time between subsequent transmissions from the master to a particular slave on an ACL link. This time interval is called poll interval, T_{poll}, and is used to provide dynamic bandwidth allocation and latency control. To change the bandwidth allocated to a particular ACL connection, the poll interval can be changed. Either the master or the slave can request a new poll interval. This will allow the master and slave to dynamically negotiate the quality of service as needed.

After a connection has been made and the ACL link is established, the master or the slave can request to establish an SCO link by sending a request with

INTEL-1020
U.S. Patent 9,265,043

parameters that specify the timing, packet type, and coding that will be used on the SCO link. The SCO link reserves slots separated by the SCO interval, T_{SCO}. The first slot reserved for the SCO link is defined by T_{SCO} and the SCO offset, D_{SCO}. After that, the SCO slots follows periodically with the SCO interval. The other device may not accept the suggested SCO parameters and can suggest another possible set of SCO parameters. In this way, the master and the slaves agree on a set of SCO parameters. At any time, the master or the slave can request a change in the SCO parameters or remove the SCO link. The SCO link is a point-to-point symmetric link between a master and a single slave. The master maintains the SCO link by using reserved slots at regular intervals. The SCO link can be considered as a circuit-switched connection because it consists of reserved slots between the master and the slave. Because SCO packets carry time-critical data, they are never retransmitted. The master will send SCO packets at regular intervals to the slave in the reserved master-to-slave slots. The SCO slave will respond with an SCO packet in the following slave-to-master slot. Normally, only the slave that is addressed by the master in a even-numbered slot can return a packet in the next slave-to-master slot. An exception is made for the SCO link to support periodic traffic efficiently. If an SCO link already exists, a slave is still allowed to return an SCO packet in the reserved SCO slot even if the slave is unable to decode its address in the preceding time slot.

For each of the SCO and ACL links, several different packet types are defined. These packet types differ according to the amount of error correction and the number of slots they occupy. There are five packet types common to the ACL and SCO links: ID, NULL, POLL, FHS, and DM1. The ID, NULL, POLL, and FHS packets are link-control packets. The ID packet consists only of the access code and is used in inquiry and paging. The DM1 packet is used for control information on either link and can carry data.

The FHS packet is used to maintain synchronization between the master and all slaves. At connection setup and during a master-slave switch the FHS packet is transferred from the master to the slave. This packet will establish the timing and frequency synchronization. The slaves adjust their RX/TX timing according to the reception of the FHS packet (and not according to the reception of the page message). The FHS packet contains the master's real-time Bluetooth clock, the master's 48-bit Bluetooth device address, and some additional information like

INTEL-1020
U.S. Patent 9,265,043

the class of device. The FHS packet contains all information to construct the channel access code.

The SCO packets do not include CRC. There are three pure SCO packets. The HV1 packet is protected by a rate 1/3 FEC. The packet length is 240 bits, and because of the FEC, this packet carries 80 information bits. The HV2 packet is protected by a rate 2/3 FEC and has the same overall length of 240 bits. To achieve higher throughput, the HV3 packet is not protected by FEC and carries 240 information bits. If HV1, HV2, or HV3 packets are used for a 64 Kb/s speech signal, one packet will carry 1.25 ms, 2.5 ms, or 3.75 ms of speech. Taking into account the slot period, one packet must be sent every second, fourth, or sixth slot, or T_{SCO} is equal to 2, 4, or 6, correspondingly for HV1, HV2, and HV3. For all of these packets, the hopping rate is 1600 hops/second. Therefore, for a 64 Kb/s speech signal, one HV1, HV2, or HV3 packet is sent correspondingly 800, 400, or 266.6 times each second.

The DV packet is a data-voice SCO packet. This packet carries an asynchronous data field in addition to a synchronous (voice) field. The voice field of the DV packet is 80 bits long and is not protected by FEC. The data field is up to 150 bits long and is protected by a rate 2/3 FEC. The DV packet can support 64 Kb/s for voice, in addition to 57.6 Kb/s for data. In each DV packet, the voice information is new, but the data information might be retransmitted if the previous transmission failed. The parameters of the SCO packets are summarized in Table 3–1.

Table 3–1: SCO packet type parameters

Packet type	HV1	HV2	HV3	DV
Payload header (bytes)	None	None	None	1 data part only
Payload length (bytes)	10	20	30	20
FEC code	1/3	2/3	None	2/3 data part only

There are seven ACL packet types, including the DM1 packet. Six of these packets have 16-bit CRC and can be retransmitted if not acknowledged. These six packet types are DM*x* and DH*x*, where *x* is equal to the slot duration of the packet and can be 1, 3, or 5. DM stands for data-medium rate and DH stands for

INTEL-1020
U.S. Patent 9,265,043

data-high rate. DM1 has 18 information bytes and a 16-bit CRC protected together by a rate 2/3 FEC. The DH packets are not protected by FEC. The seventh ACL packet type, AUX1, is just like DM1, but has no CRC.

The parameters of the ACL packets are summarized in Table 3–2, and the data rates that these packets achieve are summarized in Table 3–3. Note that for the DM1 and DH1 packets, the data rate for symmetric and nonsymmetric operation is the same. The reason is that these packets are one-slot packets. The use of three-slot and five-slot packets reduces the transmission opportunities (Figure 3–3). The data rates for nonsymmetric operation for DM3, DH3, DM5, and DH5 are obtained by assuming that in one direction three-slot and five-slot packets are used, while in the reverse direction one-slot packets are used. The data rates for symmetric operation are obtained by assuming that multislot packets are used in both directions.

Table 3–2: ACL packet type parameters

Packet type	DH1	DM1	DH3	DM3	DH5	DM5
Slot time	1	1	3	3	5	5
Payload header (bytes)	1	1	2	2	2	2
Payload length (bytes)	0–27	0–17	0–183	0–121	0–339	0–224
FEC code	None	2/3	None	2/3	None	2/3

Table 3–3: Data rate for ACL packets

ACL Packet type	Symmetric	Nonsymmetric
DM1	108.8	108.8/108.8
DH1	172.8	172.8/172.8
DM3	258.1	387.2/54.4
DH3	390.4	575.6/86.4
DM5	286.7	478/36.3
DH5	433.9	721/57.6
AUX1	185.6	185.6/185.6

INTEL-1020
U.S. Patent 9,265,043

Note that the data throughput for a given packet type depends on the quality of the RF channel. As an option, quality measurements in the receiver of one device can be used to dynamically control the packet type transmitted from the remote device for optimization of the data throughput.

According to the automatic repeat request scheme, DM, DH and the data field of DV packets are transmitted and retransmitted until acknowledgment of a successful reception is returned or a timeout is exceeded. The cyclic redundancy check (CRC) is used to determine whether the payload is correct or not. The ARQ scheme works only on the payload in the packet and only the payload that has a CRC. The packet header and the SCO payload are not protected by the ARQ scheme. Broadcast packets are checked on errors using the CRC, but no ARQ scheme is applied because the individual slaves never acknowledge broadcast packets. Because broadcast packets are not acknowledged, each broadcast packet is normally repeated for a fixed number of times.

Retransmissions can be stopped if transmitting the next packet is more important than retransmitting the previous packet. It is clear that the ARQ scheme can cause variable delay in the traffic flow. Also note that although a Bluetooth unit may be in a position where it cannot receive new information, it can still continue to transmit information—the flow control is separate for each direction.

Bluetooth handles voice particularly well. Bluetooth supports three different voice coding formats: μ-law log PCM, A-law log PCM, and CVSD, all at 64 kb/s. The public switched telephone network (PSTN) in the United States, Canada, and Japan uses μ-law log PCM. The PSTN in the rest of world uses A-law log PCM. Because Bluetooth devices do not know which country they are used in, they can perform conversion between A-law and μ-law log PCM. Voice coding can be turned off to achieve a linear PCM data link at 64 kb/s. The 64 kb/s log PCM with either A-law or μ-law companding is done according to ITU-T recommendation G.711. In addition, a sophisticated delta modulator, an innovative type of CVSD, is supported. The modulation scheme follows the waveform where the output bits indicate whether the prediction value is smaller or larger then the input waveform. To reduce slope overload effects, syllabic companding is applied—the step size is adapted according to the average signal slope. The input to the CVSD encoder is linear PCM.

INTEL-1020
U.S. Patent 9,265,043

Because there are many voice-compression algorithms, and many of them provide higher compression ratios than CVSD, it is easy to wonder why CVSD was chosen. CVSD was chosen because bit errors are likely, and the quality of the voice depends on the robustness of the voice-coding scheme to bit errors. Compression algorithms achieving high-compression ratios are normally very sensitive to bit errors in the compressed bit stream. In some cases, occasional bit errors can lead to disastrous results. On the other hand, CVSD is rather insensitive to random bit errors. These random bit errors are experienced in voice signals as white background noise. However, when a packet is rejected because either the channel access code or the HEC test was unsuccessful, a whole speech segment will be lost. Additional signal processing can be performed to account for the missing speech parts. This is why the voice payload in the HV2 packet is protected by a 2/3 rate FEC. For errors that are detected but cannot be corrected, the receiver should try to minimize the audible effects at the application layer. For instance, from the 15-bit FEC segment with uncorrected errors, the 10-bit information part as found before the FEC decoder could be used. The HV1 packet is protected by a three-bit repetition FEC. For this code, the decoding scheme will always assume zero or one-bit errors. Thus, no detectable but uncorrectable error events for HV1 packets exist.

Data or link control information is exchanged between the master and the SCO slave using the DV or DM1 packets. Any ACL type of packet can be used to send data or link control information to any other ACL slave.

Bluetooth devices have two major states: standby and connection. By default, every Bluetooth unit is in the stand-by state. In this state, the Bluetooth unit is in a low-power mode. The device may leave the standby state to scan for page or inquiry messages, or to perform page or inquiry itself.

In order to establish new connections, the procedures inquiry and paging are used. The inquiry procedure enables a unit to discover which units are in range, and what their device addresses and clocks are. The inquiry procedure is used in applications where the other device addresses are unknown. During an inquiry substate, the discovering unit collects the Bluetooth device addresses and clocks of all units that respond to the inquiry message. It can then, if desired, make a connection to any one of them by means of the page procedure. The receiver in the page scan or inquiry scan substate correlates against the inquiry or the device access codes with a matching

INTEL-1020
U.S. Patent 9,265,043

correlator. The device access code is used during page, page scan, and page response substates. There is one general inquiry access code (GIAC) for general inquiry operations, and there are 63 dedicated inquiry access codes (DIACs) for dedicated inquiry operations. Various software applications may take advantage of these different inquiry access codes to discover not all devices, but only devices with certain capabilities. For example, one can inquire about printers, facsimile machines, etc. In the page substate, the paging device transmits the device access code (ID packet) corresponding to the slave to be connected rapidly on a large number of different hop frequencies. Because the ID packet is a very short packet, the hop rate is increased from 1600 hops/s to 3200 hops/s. In a single TX slot interval, the paging master transmits on two different hop frequencies. Correspondingly, in a single RX slot interval, the paging receiver listens on two different hop frequencies. This is also necessary, because there could be a large offset between the clocks of the devices. After a successful page attempt, the paging device will enter the connection state as a master, and the paged unit will enter the connection state as a slave.

In the connection state, the connection has been established and packets can be sent back and forth. During the connection state, the Bluetooth units can be in several modes of operation: active mode, sniff mode, hold mode, and park mode.

In the active mode, the Bluetooth unit actively participates on the channel. The hold mode is typically entered when there is no need to send data for a relatively long time. The transceiver can then be turned off in order to save power. But the hold mode can also be used if a device wants to discover or be discovered by other Bluetooth devices or wants to join other piconets. Each device decides what it will actually do during the hold time. The messages exchanged before entering the hold mode contain a parameter, hold instant, which specifies the instant in the future at which the hold will become effective. Before entering the hold mode, the master and slave agree on the time duration the slave will be in the hold mode. The master or the slave can request to enter hold mode. Only the ACL link can be in the hold mode, in which ACL packets are neither transmitted nor accepted. To maintain QoS, SCO links will still be supported. With the hold mode, capacity can be made free to do other things like scanning, paging, inquiring, or attending another piconet. The unit in hold mode may not do any of these things and just be in a low-power sleep mode. During the hold mode, the slave unit keeps its active member address.

INTEL-1020
U.S. Patent 9,265,043

The park and sniff modes are similar to the hold mode. When a slave does not need to participate on the piconet channel, but still wants to remain synchronized to the channel, it can enter the park mode, which is a low-power mode with very little activity in the slave. Devices that are in park mode give up their active member address but are assigned a unique parked member address. This parked member address can be used by the master to unpark that slave. To support parked slaves, the master establishes a beacon channel when one or more slaves are parked. The master communicates with parked slaves using this beacon channel. The beacon instants follow periodically at the predetermined beacon interval. Parked slaves sleep most of the time. However, at the beacon instants they wake up to resynchronize to the channel by receiving a packet from the master. Any packet exchanged on the channel can be used for synchronization. The parked slave does not have to wake up at every beacon instant. At the beacon instants, the master can unpark a parked slave, change the park mode parameters, transmit broadcast information to all parked slaves, or let the parked slaves request access to the channel. The master or the slave can request park mode.

Two parameters control the listening activity in the slave. The sniff attempt parameter determines for how many slots the slave shall listen, beginning at the sniff slot, even if it does not receive a packet with its own active member address. The sniff timeout parameter determines for how many additional slots the slave shall listen if it continues to receive only packets with its own active member address.

Active slaves listen in every master-to-slave slot for packets. If an active slave is not addressed, it may sleep until the next new master transmission period. Sniff mode is used to reduce the duty cycle of the slave. With the sniff mode, the time slots where the master can start transmission to a specific slave are reduced; that is, the master can only start transmission in specified time slots. These so-called sniff slots are spaced regularly with a certain interval, called a sniff interval, T_{sniff}. Before entering a sniff mode, the devices negotiate duration of the sniff interval and a sniff offset, D_{sniff}. The offset determines the time of the first sniff slot; after that the sniff slots follows periodically with the sniff interval. Both slaves and master can request to enter sniff mode. The sniff mode is the least power efficient, followed by the hold mode, and then by the parked mode.

In a scatternet, a device can be simultaneously a master in one piconet and a slave in several other piconets. A master or slave can become a slave in another piconet

INTEL-1020
U.S. Patent 9,265,043

by being paged by the master of this other piconet. On the other hand, a unit participating in one piconet can page the master or slave of another piconet. Time multiplexing is used to switch between piconets. In case of ACL links only, a unit can request to enter the hold or park mode in the current piconet. During this time, the unit may join another piconet by just changing the channel parameters. Units in the sniff mode may have sufficient time to visit another piconet in between the sniff slots. If SCO links are established, other piconets can only be visited in the nonreserved slots in between. This is only possible if there is a single SCO link using HV3 packets. In the four slots in between, the unit can visit one other piconet.

Because the multiple piconets are not synchronized, guard time shall be left to account for misalignment. This means that only two slots can effectively be used to visit another piconet in between the HV3 packets. Because the clocks of two masters of different piconets are not synchronized, a slave unit participating in two piconets has to take care of two offsets to match time with both master clocks. Because the two master clocks drift independently, regular updates of the offsets are required in order for the slave unit to synchronize to both masters.

Sometimes a master-slave switch is desirable:

- When a unit paging the master of an existing piconet wants to join this piconet because, by definition, the paging unit initially is master of a "small" piconet only involving the pager (master) and the paged (slave) unit.
- When a slave in an existing piconet wants to set up a new piconet, involving itself as master and the current piconet master as slave. This case implies a double role of the original piconet master; it becomes a slave in the new piconet while still maintaining the original piconet as master.
- When a slave wants to fully take over an existing piconet, i.e., the switch also involves transfer of other slaves of the existing piconet to the new piconet, a much more complicated example.

INTEL-1020
U.S. Patent 9,265,043

The master-slave switch requires careful support at the application layer (such as how to handle security and transfer all kind of slave information from old to new master). This procedure is not defined in the 802.15.1 standard, which leaves it to vendors to implement. This is one weakness in the standard, especially from the standpoint of security.

Clearly, this third instance can be achieved by allowing the new master to setup a completely new piconet through the conventional paging scheme. However, that would require individual paging of the old slaves and thus take an unnecessarily long time. Instead, letting the new master use the old master's timing knowledge is more efficient. As a consequence of the master-slave switch, the slaves in the piconet have to be transferred to the new piconet, changing their timing and their hopping scheme.

How does the master-slave switch happen? For the master and slave involved in the role switch, it reverses their TX and RX timing—a TDD switch. Moreover, because the piconet parameters are derived from the device address and clock of the master, a master-slave switch inherently involves a redefinition of the piconet as well—a piconet switch. Assume unit A wants to become master; unit B was the former master. Then there are basically two alternative master-slave switch scenarios—either the slave takes the initiative or the master takes the initiative.

First, both slave A and master B do the TDD switch but keep the former hopping scheme (still using the device address and clock of unit B); there is no piconet switch yet. Unit A now becomes the master, unit B, the slave. The FHS packet is sent by master A using the "old" piconet parameters. The active member address in the FHS packet header is the previous address used by unit A. From this moment, Slave B starts to use the active member address formerly used by unit A in its slave role. After the FHS acknowledgment, which consists of the ID packet and is sent by the slave on the old hopping sequence, both master A and slave B turn to the new channel parameters of the new piconet as indicated by the FHS packet.

Bluetooth security

WPANs are ad hoc, and security is an important requirement. To provide security, measures must be taken both at the application layer and the data-link layer.

INTEL-1020
U.S. Patent 9,265,043

The goal of the security mechanism at the data-link layer is to provide efficient support for different applications. To be appropriate for ad-hoc networks—i.e., for a peer environment—the authentication and encryption routines in every unit must be implemented in an identical fashion. Four different entities are used for maintaining security at the data-link layer: a public address that is unique for each user; two secret keys; a a random number, which is different for each new transaction.

The public address is the 48-bit Bluetooth device address, which is unique for each Bluetooth unit. The Bluetooth addresses are publicly known and can be obtained, for example, from an inquiry by a Bluetooth unit.

The two secret keys are derived during initialization and are never further disclosed. One is used for authentication and called authentication key, the other is used for encryption and called encryption key.

The fourth entity, the random number, can be derived from a random or pseudorandom process in every Bluetooth unit. It is not a static parameter and changes frequently. Random numbers are used for many purposes within the security functions, e.g., for the challenge-response scheme, for generating authentication, and encryption keys. Within Bluetooth, the requirements placed on the random numbers used are that they be nonrepeating within the lifetime of the authentication key and randomly generated.

First consider first the authentication key and then the encryption key. The length of the authentication key is always 128 bits. The authentication key is also referred to as the link key. This link key is shared between two or more parties and is the base for all security transactions between these parties. The current link key is used for all authentications and all generation of encryption keys in the on-going session. Because piconets are ad-hoc networks, every unit participates in a piconet for a finite time interval. This time interval is called a session. Thus, the session terminates when the unit disconnects from the piconet. The link keys are either semipermanent or temporary. A semipermanent link key is stored in nonvolatile memory and may be reused after the current session is terminated. The lifetime of a temporary link key is limited by the lifetime of the current

INTEL-1020
U.S. Patent 9,265,043

session; it cannot be reused in a later session. Actually, there are four types of authentication (or link) keys:

- The unit key K_A
- The combination key K_{AB}
- The temporary key K_{master}
- The initialization key K_{init}

For a Bluetooth unit, the combination key K_{AB} and the unit key K_A are functionally indistinguishable; the difference is in the way they are generated. The unit key, K_A is generated in a single unit, which can be denoted as unit A. The combination key K_{AB} is derived from information in, for example, two units A and B and is therefore always dependent on two units. The combination key is derived for each new combination of two Bluetooth units. Whether a unit key or a combination key is used depends on the application. If a Bluetooth unit has little memory to store keys or has to communicate with a large group of users, it can use its own unit key. In that case, it has to store only a single key. Preferably, applications that require a higher security level will use the combination keys. These applications will require more memory because a combination key for each link to a different Bluetooth unit has to be stored.

The initialization key, K_{init}, is used as link key during the initialization process when no link key has been defined or when a link key has been lost.

Typically, in a point-to-multipoint configuration where the same information is to be distributed securely to several recipients, a common key is useful. To achieve this, a special link key (denoted master key) can temporarily replace the current link keys. The master key, K_{master}, is a link key only used during the current session.

The link (or alternatively, authentication) keys must be generated and distributed among the Bluetooth units in order to be used in the authentication procedure. Because the link keys must be secret, they cannot be obtained through an inquiry routine in the same way as the Bluetooth addresses. The exchange of the keys takes place during an initialization phase. All initialization procedures consist of the following five parts: (1) generation of an initialization key; (2) generation of link key; (3) link key exchange; (4) authentication; (5) generating of encryption key in each unit (optional).

INTEL-1020
U.S. Patent 9,265,043

The first step is the generation of an initialization key. This initialization key can be derived, for example, from a random number, the Bluetooth device address, and user-specified PIN code. The initialization key is used temporarily. After the units have performed the link key exchange, the initialization key is discarded. A fraudulent Bluetooth unit may try to test a large number of PINs by claiming another Bluetooth device address each time. It is the application's responsibility to take countermeasures against this threat. This initialization is carried out separately for each two units that want to implement authentication and encryption.

The second and third steps are the generation and exchange of link keys. At initialization, the two applications running at the two devices must determine which key will be used as the link key—K_A, K_B, or K_{AB}. (Which key will be selected is not specified in 802.15.1.) Typically, this will be the unit key of the device with restricted memory capabilities, because this unit only has to remember its own unit key. If this unit key is chosen, then this unit key is transferred to the other party and stored as link key for that particular party. If it is desired to use the combination key, this key is first generated during the initialization procedure. The combination key is the combination of the two random numbers generated in units A and B. First, each unit generates a random number and then the two units exchange the two random numbers that they have generated. When the random numbers and have been mutually exchanged, each unit recalculates the other unit's contribution to the combination key. This is possible because each unit knows the Bluetooth device address of the other unit. The result is stored in unit A as the link key for the connection with device B, and in unit B as the link key for the connection with device A. For any new connection established between units A and B, they will use the agreed upon link key, instead of once again going through this procedure. However, if no link key is available, the devices will perform the initialization procedure again.

The unit keys K_A and K_B are generated when the Bluetooth unit is in operation for the first time; i.e., not during each initialization. Once created, these unit keys are stored in nonvolatile memory and (almost) never changed. For example, if one of these unit keys is being used as the link key for a particular pair of devices, and this key is changed after initialization, the other device will possess a wrong link key.

INTEL-1020
U.S. Patent 9,265,043

After the link key exchange, the fourth step—mutual authentication—can be initiated. The authentication used in Bluetooth uses a challenge-response scheme. In this algorithm, a claimant's knowledge of a secret key is checked through a two-step protocol using symmetric secret keys. Authentication succeeds if the claimant and the verifier share the same secret key. In the challenge-response scheme, the verifier challenges the claimant to authenticate a random input (the challenge), with an authentication code and return the result to the verifier. The claimant calculates a response (which is a function of the challenge), the claimant's Bluetooth device address, and a secret key. The response is sent back to the verifier, which checks if the response was correct. Successful calculation of the authentication procedure requires that two devices share a secret key. Both the master and the slave can be verifiers. Mutual authentication is achieved by first performing the authentication procedure in one direction and then immediately by performing the authentication procedure in the opposite direction. Certain applications only require a one-way authentication. However, in some peer-to-peer communications, one might prefer a mutual authentication in which each unit is subsequently the challenger (verifier) in two authentication procedures. In general, the application indicates who has to be authenticated by whom. As a side effect of a successful authentication procedure, an auxiliary parameter (the Authenticated Ciphering Offset [ACO]) will be computed. This ACO is used in the generation of the encryption key. If an authentication is successful, the value of ACO should be retained. When the authentication attempt fails, a certain waiting interval must pass before the verifier can initiate a new authentication attempt to the same claimant or before it will respond to an authentication attempt initiated by a unit claiming the same identity as the suspicious unit. For each subsequent authentication failure with the same Bluetooth address, the waiting interval is increased exponentially. The waiting interval is limited to a certain maximum. The maximum waiting interval depends on the implementation. The waiting time shall exponentially decrease to a minimum when no new failed attempts are being made during a certain time period. This procedure prevents an intruder from repeating the authentication procedure with a large number of different keys. To make the system somewhat less vulnerable to denial-of-service attacks, the Bluetooth units should keep a list of individual waiting intervals for each unit with which it has established contact.

The fifth step is optional and is performed only if encryption is turned on.

INTEL-1020
U.S. Patent 9,265,043

After the initialization procedure, the units can proceed to communicate, or the link can be disconnected. If encryption is implemented, the E_0 algorithm is used with the proper encryption key derived from the current link key.

The length of the encryption key can vary between 8 and 128 bits. The size of the encryption key is configurable for two important reasons—one political and one technical. The first reason is the government regulations regarding the export of cryptographic algorithms. The technical reason to make the encryption key of variable length is to provide an easy upgrade path for future hardware and software implementations.

According to U.S. government regulations, the export of technology with strong encryption outside the United States and Canada is not allowed. Increasing the effective key size is the simplest way to combat the increased computing power of attackers. Furthermore, government regulations do not allow users to be able to set the encryption key size. Thus, although the encryption key length is variable, it should be a factory preset entity. Therefore, the Bluetooth baseband processing does not accept an encryption key given from higher software layers.

Encryption can be turned on and off many times, and each time encryption is activated, a new encryption key is derived from the current link key. Thus the lifetime of the encryption key does not necessarily correspond to the lifetime of the authentication key. The authentication key is more static than the encryption key. Only the particular application running on the Bluetooth device decides when, or if, to change it. The purpose of separating the authentication key and encryption key is to facilitate the use of a shorter encryption key without weakening the strength of the authentication procedure.Note that there are no governmental restrictions on the strength of authentication algorithms.

The encryption key, K_C, is derived from the current authentication key, a 96-bit Ciphering Offset number (COF), and a 128-bit random number. This COF is either derived from the Bluetooth device address of the master or is set to the value of ACO as computed during the authentication procedure. The encryption key is automatically changed each time the unit enters the encryption mode.

It is quite possible for the master to use separate encryption keys for each slave in a point-to-multipoint configuration with encryption activated. However, if the application requires more than one slave to listen to the same payload, each slave

INTEL-1020
U.S. Patent 9,265,043

must be addressed individually. This clearly reduces throughput. Moreover, a Bluetooth unit cannot be expected to be capable of switching between two or more encryption keys in real time. Thus, the master cannot use different encryption keys for broadcast messages and individually addressed traffic. Instead of addressing each slave individually, even for broadcast messages, the master may tell several slave units to use a common link key. This key is called K_{master}. Because the encryption key is derived from the link key, the encryption key will be common as well. However, this common link key is only of temporary interest because many applications use broadcast messages only part of the time.

To create a broadcast group, the master will issue a command to the slaves to replace their respective current link key by the new (temporary) master key. Before encryption can be activated, the master also has to generate and distribute a common random number to all participating slaves. Using this random number and the newly derived master key, each slave can generate independently the new common encryption key. Then the master can broadcast encrypted information. Note that because the encryption key is common, all the slaves in the broadcast group must support the length of this encryption key. Those slaves that cannot must be excluded from the broadcast group. The broadcast group must be a trusted domain, because each slave in possession of the master key can eavesdrop on all encrypted traffic, not only on the traffic intended for itself. If there is no more broadcast traffic or if there is a change in the security policy, the master can tell all participants to fall back to their old link keys simultaneously.

In certain circumstances, modifying the link keys is desirable. A link key based on a unit key can be changed, but not very easily. The unit key is created once during the first use. Changing the unit key is a less desirable alternative, because several units may share the same unit key as link key. One example is a printer whose unit key is distributed among all users using the printer's unit key as a link key. Changing the unit key will require reinitialization of all units trying to connect. In certain cases, this might be desirable; for example, to deny access to previously allowed units. Another possibility is starting up an entirely new initialization procedure. In that case, user interaction may be necessary because a PIN is required in the authentication and encryption procedures.

The encryption procedure is discussed next. User information can be protected by encryption of the packet payload; the access code and the packet header are never

INTEL-1020
U.S. Patent 9,265,043

encrypted. The encryption of the payloads is carried out with a stream cipher called E_0 that is resynchronized for every payload. The overall principle is shown in Figure 3–6.

The stream cipher shown in Figure 3–6, consists of three steps. The first step is the generation of the payload key. The payload key is constructed from the master Bluetooth address, 26 bits of the master real-time clock and the encryption key as input. The encryption key, K_C, is derived from the current link key, COF, and a random number. The master transmits the random number before entering encryption mode. This random number is publicly known because it is transmitted as plain text over the air. The algorithm is reinitialized at the start of every new packet. By using the real-time clock, at least one bit is changed for every new packet. The payload key generator is very simple—it merely combines the input bits in an appropriate order.

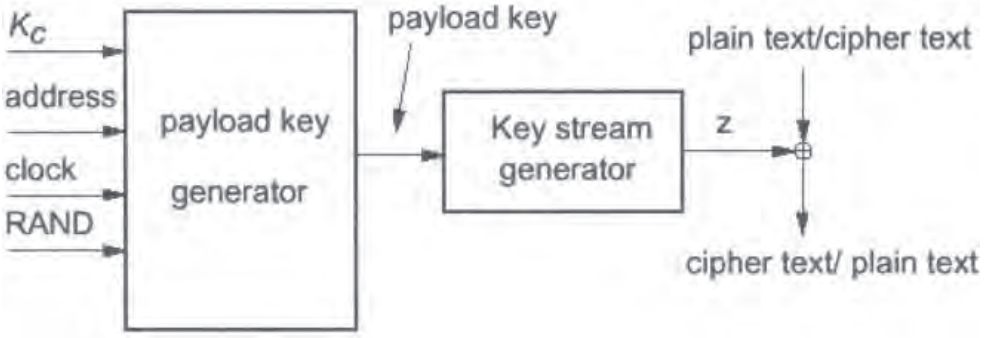

Figure 3–6: A block diagram of the E_0 stream cipher

The main part of the cipher algorithm is the second step. The second step generates the key stream bits, and the third step performs the encryption or decryption. The actual encryption is performed by adding the ciphering bits bit-wise modulo-2 to the data bits. The payload is ciphered after the CRC bits are appended, but before the FEC encoding. Each packet payload is ciphered separately. The cipher is symmetrical; decryption is performed in exactly the same way using the same key as used for encryption.

Because the encryption key is of variable length, each Bluetooth device has a parameter defining the maximal allowed key length. Every application must define a number indicating the smallest acceptable key size for that particular

INTEL-1020
U.S. Patent 9,265,043

application. Before generating the encryption key, the involved units must negotiate to decide what key size to actually use. This negotiation can happen by the master sending a suggested value and the slave deciding whether to accept it or not. If the suggested key length is not acceptable to the slave, it will in turn suggest a different key length. This procedure is repeated until a key length agreement is reached, or, under the control of the application one unit aborts the negotiation. If the agreed key length is too short for the application, the units cannot exchange encrypted information.

The master and every slave in the piconet must agree first whether to use encryption or not. If encryption is used, they must agree whether to use encryption only on the point-to-point packets or on both the unicast and multicast packets. If a slave has a semi-permanent link key (i.e., a combination key or a unit key), it can receive encrypted data only on slots individually addressed to it. In particular, it will assume that broadcast messages are not encrypted.

LMP and L2CAP

The protocol used for link set-up, security, and control is called link manager protocol (LMP). The place of the LMP in the protocol stack is illustrated in Figure 3–7. LMP sits on top of the baseband protocol. LMP messages are transferred in the payload and are distinguished by a reserved value in the payload header. The LMP messages are exchanged between LMP entities and are not propagated to higher layers. Link manager messages have higher priority than user data. LMP messages are not acknowledged. LMP messages are used when devices enter hold, park, or sniff modes. The LMP also performs power control. If the RSSI value differs too much from the preferred value of a Bluetooth device, it can request an increase or a decrease of the other device's TX power. If a device does not support power control requests, this is indicated in the supported features list. Upon receipt of this message, the output power is increased or decreased one step, where the step size is defined in the radio specification. At the master side, the TX power is completely independent for different slaves; a request from one slave can only affect the master's TX power for that same slave.

INTEL-1020
U.S. Patent 9,265,043

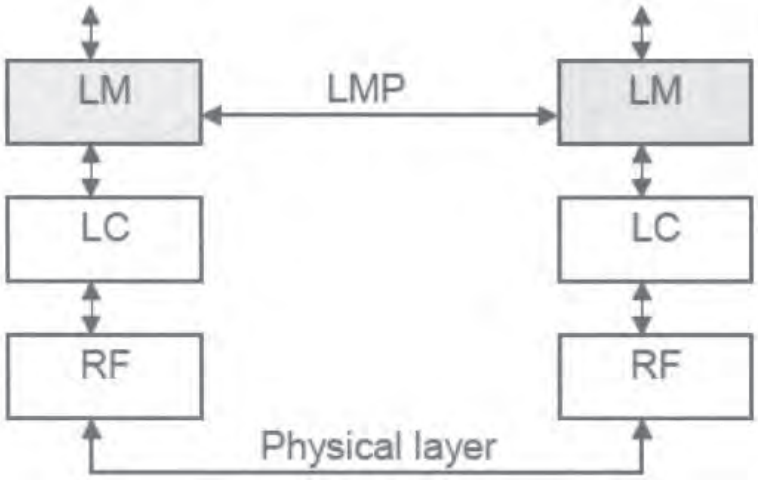

Figure 3–7: The LMP in Bluetooth protocol stack

The link controller (LC) cannot guarantee that it will deliver a message to the remote device within a certain period of time. Therefore, there can be no guarantees that LMP messages will be delivered within a certain period of time.

Each Bluetooth link has a timer that is used for link supervision. This timer is used to detect link loss caused by devices moving out of range, a device's power-down, or other similar failure cases.

The Link Manager plays a role in setting up ACL and SCO links, etc. The LMP plays a role in the paging mechanism. In addition to the mandatory paging scheme, Bluetooth defines optional paging schemes. LMP provides a means to negotiate the paging scheme, which is to be used the next time a unit is paged.

Another protocol of the Bluetooth specification is the logical link control and adaptation protocol (L2CAP). This protocol supports higher-level protocol multiplexing, packet segmentation and reassembly, and the conveying of quality of service information. Figure 3–8 illustrates how L2CAP fits into the Bluetooth Protocol Stack. L2CAP lies above the Baseband Protocol and interfaces with other communication protocols such as the Bluetooth Service Discovery Protocol (SDP), RFCOMM, and Telephony Control (TCS). Voice-quality channels for audio and telephony applications are usually run over Baseband SCO links. Also,

INTEL-1020
U.S. Patent 9,265,043

packetized audio data, such as IP Telephony, may be sent using communication protocols running over L2CAP.

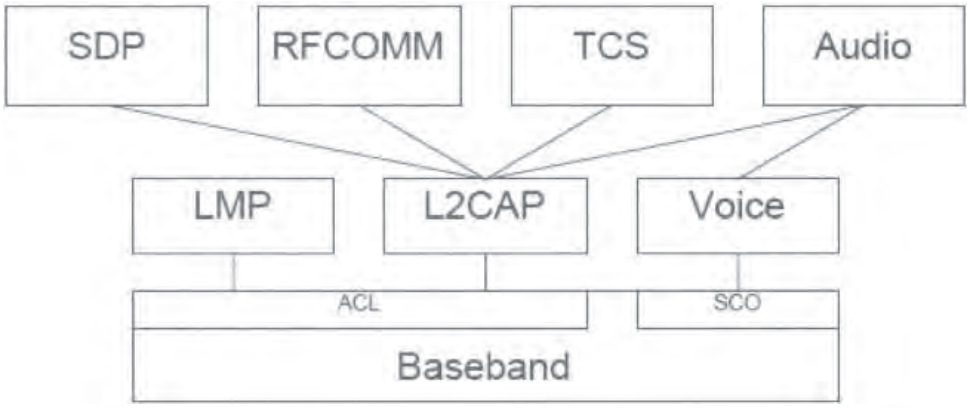

Figure 3–8: L2CAP in the Bluetooth protocol stack

COEXISTENCE AMONG WIRELESS STANDARDS

Overview

The major advantage of wireless communication devices is that they do not require wires. One disadvantage is that, well, because they do not require wires different wireless devices can operate in the same frequency band. Multiple-access methods have been thoroughly investigated, but until recently no one studied the operation of several different wireless systems in the same frequency band. Multiple-access methods consider how multiple devices can use the spectrum according to the same protocol. They do not (and cannot) consider how different wireless systems can coexist in the same frequency band. Every wireless system is designed assuming that it has the entire spectrum in which it operates available to itself. What will happen if there are two or more different wireless systems in the same frequency and deployed in proximity to each other?

Devices that transmit at a relatively higher power and/or use protocols more robust to interference will get their data through. The other devices will not get their data through, or at least their throughput will be reduced. This situation can be called the "big stick" policy. Devices that have a bigger stick will work. It is obvious that as the number of wireless devices based on different standards grows,

INTEL-1020
U.S. Patent 9,265,043

the lack of coexistence among the devices can discourage consumers to use more wireless devices.

This issue seems important in home networking applications, where the number of wireless devices is expected to skyrocket in the next several years. Therefore, measures need to be taken to ensure that different wireless systems will be able to function, even when they use the same frequency band. In other words, replace the "big stick" policy with the "good citizen" policy. The IEEE 802—more specifically, the IEEE 802.15 Working Group—is the first standards-making body to consider this problem.

Before the recommended practice for coexistence is discussed, the meanings of coexistence, interoperability, interworking, and backwards compatibility must be defined. *Coexistence* is the ability of one system to perform a task in a given shared environment in which there are other systems that may or may not be using the same set of rules. *Interoperability* is the ability of two systems to perform a given task using a single set of rules. The ability of two systems to perform a task when each system implements a different set of rules is called interworking. If a new system can interwork with an old system, the new system is backwards compatible with the old system.

Many parameters affect the interference between two different wireless systems, such as separation among antennas, amount of data traffic, power levels, and modulation types. These are physical layer parameters. However, it is important to note that interference is not only a physical layer problem. In most of the previous research, the role of the MAC layer in interference and coexistence problems was not considered. Not only the MAC protocol can play a very important part in the interference scenario and affect the overall system performance, but the reverse is also true—interference can significantly impact the performance of the MAC layer and higher layers.

The IEEE 802.15.2 task group was formed to develop a recommended practice for coexistence between Bluetooth/IEEE 802.15.1 and IEEE 802.11. When IEEE 802.15.2 was formed, it was expected that WPAN devices based on Bluetooth/IEEE 802.15.1 and WLAN devices based on IEEE 802.11b will have the largest market share of devices operating in the 2.4 GHz ISM band. The recommended practice considers the IEEE 802.11b direct-sequence spread-spectrum standard at data rates of 1, 2, 5.5, and 11 Mb/s. Both IEEE 802.11 and

INTEL-1020
U.S. Patent 9,265,043

IEEE 802.15 are continuing to work on additional standards. Future revisions of the recommended practice may consider the coexistence issue among those newer WLAN and WPAN standards.

For IEEE 802.11, the data rate typically will be adjusted by the application to be the maximum data rate that results in a packet error rate of 10 percent. If the number of bits per packet is N, the packet error rate can be determined from the bit error rate (BER) as

$$\text{PER} = 1 - (1 - \text{BER})^N \qquad \text{Eq. 3–1}$$

At larger distances as the packet error rate increases the data rate will decrease. IEEE 802.11b is occupying a 22 MHz-wide band within the 2.4 GHz ISM band. IEEE 802.15.1 hops on 79 frequencies within the entire band. There would be a collision if IEEE 802.15.1 hops on a frequency occupied by IEEE 802.11b. Out of the 79 frequencies, 22 are subject to interference. The decrease of the data rate by the application can lead to instability. For example, due to interference from IEEE 802.15.1, it is possible for the application to lower the data rate of IEEE 802.11. At a lower data rate, however, packets become longer and can experience interference for longer periods of time, leading to increase in interference, and so on.

In its turn, IEEE 802.11b can cause interference to IEEE 802.15.1, affecting 22 out of the 79 channels and typically causing throughput decrease.

Note that transmitter power control and rate scaling is a simple coexistence mechanism. Power control can be effective if IEEE 802.11b and Bluetooth/IEEE 802.15.1 devices are designed to limit their transmit powers near the threshold to obtain the required performance. To implement power control, a transmit power control mechanism must be implemented. All IEEE 802.11b devices currently support multiple transmit rates, i.e., 1, 2, 5.5 and 11Mb/s. As a result, all IEEE 802.11b devices currently implement a rate shifting/control algorithm. Systems monitor SNR, SINR, PER, etc., to determine which rate should be used. The maximum rate is always desired. The rate is shifted down when packets cannot be successfully decoded at current rate. The rate control algorithm of IEEE 802.11b devices can be extended to incorporate the highest mandatory rate at lower transmit powers. The rate shift algorithm would shift to the highest possible rate with lower transmit power whenever possible.

INTEL-1020
U.S. Patent 9,265,043

The recommended practice, developed by the IEEE 802.15.2 Task Group [B79], has developed more sophisticated coexistence approaches. It starts with a detailed computer model of the mutual interference between IEEE 802.11 WLAN and IEEE 802.15.1 WPAN. The parameters of this model are the PHY and MAC models, network topology, user density, propagation model, and data traffic model. The coexistence model quantifies the effect of mutual interference of WLAN and WPAN on each other. This computer model is based on detailed simulation models of the RF channel, and the PHY and MAC of the standards. The recommended practice is considering only the DCF channel access mechanism for IEEE 802.11.

The remainder of this subsection will provide an overview of the coexistence mechanisms, followed by a discussion on the collaborative coexistence mechanisms, which, in turn is followed by a discussion on noncollaborative approaches. Because channel classification is one important part of the coexistence mechanism, it is considered in "Channel classification" on page 174. The analytic and computer MAC and PHY models are discussed in "PHY and MAC models " on page 176.

After deliberation, the IEEE 802.15.2 Task Group developed two types of coexistence mechanisms—collaborative and noncollaborative. In collaborative mechanisms, WPAN and WLAN exchange information over a separate link between one another to minimize mutual interference. The type of link and the protocol for communication over the link are not specified in the recommended practice. This is an opportunity for additional proprietary firmware and software for products, which support both IEEE 802.11b and Bluetooth. Collaborative coexistence mechanisms can be implemented when, for example, one device is a member of both the IEEE 802.11b and IEEE 802.15.1 networks. Noncollaborative coexistence mechanisms do not require exchange of information between the two networks. A total of eight coexistence mechanisms are specified in the recommendation—three collaborative and five noncollaborative.

The first collaborative method is simple. If the two networks have a separate link, then their transmissions can be scheduled in different time segments. This technique is called alternating wireless medium access.

The second collaborative method is packet traffic arbitration. In this method, the IEEE 802.11b and IEEE 802.15.1 networks can decide what type of packet to

INTEL-1020
U.S. Patent 9,265,043

send to minimize the interference between them. The first and the second collaborative mechanisms require joint action by both networks.

The third collaborative mechanism is to perform deterministic interference suppression by IEEE 802.11b. In this method, IEEE 802.11b uses a programmable notch filter to filter out the narrowband IEEE 802.15.1 interference. Clearly, no action is required by the IEEE 802.15.1 devices.

The first noncollaborative method is the IEEE 802.11b device to perform adaptive interference suppression. In this case, the notch filter must be adaptive to filter out the narrowband IEEE 802.15.1 interference. In this method no action is required on the part of the IEEE 802.15.1 devices.

There are four noncollaborative methods, in which no action on the side of IEEE 802.11b is required. In these four methods, only the IEEE 802.15.1 network takes measures to avoid or minimize interference. These methods are adaptive packet selection, packet scheduling on the ACL link, packet selection on the SCO link, and adaptive frequency hopping.

Collaborative mechanisms are recommended, when possible. If many devices are close together and the potential for interference is high, the alternate wireless medium access mechanism is recommended. If the potential for interference is low, the packet scheduling, combined with deterministic interference suppression is recommended.

Finally, it must be noted that two or more coexistence mechanisms can be used simultaneously. Again, this is not specified in the recommended practice and is an area for developing intellectual property. Some vendors regard the coexistence performance as an area of product differentiation in the marketplace.

Collaborative methods

The first collaborative method is alternating wireless medium access. Recall that an IEEE 802.11b access point sends out a beacon at approximately periodic time intervals. According to this coexistence method, a portion of the IEEE 802.11b beacon period is used for WLAN traffic, and a portion is used for IEEE 802.15.1 traffic. Information about the length of the WLAN and WPAN periods is included in the beacon. Because the two networks transmit in nonoverlapping time periods, there is no interference between them. This method clearly requires

INTEL-1020
U.S. Patent 9,265,043

synchronization. To achieve synchronization between the networks, an IEEE 802.11 station must send a synchronization signal to the IEEE 802.15.1 master. This is simple to accomplish if the master of the IEEE 802.15.1 piconet and an IEEE 802.11b station are in the same device. If ACL packets are used, it can be guaranteed that WPAN devices do not transmit during the WLAN portion of the beacon period, because a slave is not allowed to send an ACL packet if it is not addressed the previous time slot. However, a slave is allowed to send an SCO packet, even if it is not addressed the previous slot. As a result, there can be no guarantee that an SCO packet will not be sent during the time allocated for WLAN. Therefore, this coexistence method is not suitable for situations with SCO connections.

The second collaborative method is packet traffic arbitration (PTA). In this method, the IEEE 802.11b station and the IEEE 802.15.1 unit are collocated. Every packet about to be transmitted is submitted to a block called the packet traffic arbitrator for approval or rejection. If the transmission will not result in collision between the networks, the packet traffic arbitrator grants approval. If the requests from the two networks are simultaneous or almost simultaneous and there would be a collision, prioritization is used. One of the requests will be approved, and the other will be delayed. Priority could be assigned in a deterministic or random fashion. The deterministic approach is to assign priority levels in the following order from highest to lowest: IEEE 802.11b acknowledgment packets, IEEE 802.15.1 SCO packets, IEEE 802.11b data packets, and IEEE 802.15.1 ACL packets. Furthermore particular implementations can use different fairness criteria such as PER or BER to grant or approve requests. Both SCO and ACL traffic are supported, and SCO traffic is given higher priority than the ACL traffic in scheduling. The PTA method is somewhat similar to alternating wireless medium access, in the sense that only one of the networks transmits at any given time. However, medium access is divided dynamically.

Deterministic interference suppression is the third collaborative mechanism. It is based on the following idea. Because the bandwidth of IEEE 802.15.1 is only 1 MHz, it can be considered as a narrowband interferer to IEEE 802.11b, which has bandwidth of 22 MHz. A notch filter can suppress the IEEE 802.15.1 signal. This notch filter is placed at the IEEE 802.11b receiver. Because the IEEE 802.15.1 signal is hopping within the entire band, the notch filter must be

INTEL-1020
U.S. Patent 9,265,043

programmable and must hop synchronously with IEEE 802.15.1. Therefore, the IEEE 802.11b receiver must have an integrated IEEE 802.15.1 unit to provide knowledge of the hopping sequence and timing. Deterministic interference suppression works only at the physical layer, and prevents interference from IEEE 802.15.1 to IEEE 802.11b, but not vice versa.

Noncollaborative methods

If an IEEE 802.11b device is not colocated with an IEEE 802.15.1 device, it is not possible to use collaborative coexistence methods. While deterministic interference suppression is not possible, adaptive interference suppression is possible. It is again entirely based on signal processing to suppress IEEE 802.15.1 interference to IEEE 802.11b. Because the IEEE 802.11b receiver does not know the hopping sequence and the associated timing, this method requires an adaptive filter.

The adaptive packet selection and scheduling, another noncollaborative mechanism, is possible because IEEE 802.15.1 systems use various packet types with different packet lengths and degree of error protection. The interference between IEEE 802.15.1 and IEEE 802.11b depends on the type of packet used. For instance, if the channel is dominated by interference from IEEE 802.11b network, packet loss will be mainly due to collisions between IEEE 802.15.1 and IEEE 802.11 systems, instead of bit errors resulting from noise. Note that FEC methods by design provide better protection against errors resulting from noise sources. However, FEC is not effective against interference. Therefore, packet types that do not include FEC will be shorter and could provide better throughput.

The key idea is to adapt the transmission according to channel conditions. Note the difference with PTA method. Because packet traffic arbitration is a collaborative mechanism, it does not have to perform channel estimation. Adaptive packet selection is noncollaborative and relies on channel estimation. Channel estimation can be done in a variety of ways. For example the RSSI, BER, and PER all provide information about the channel. It is also possible to combine these measures. The recommended practice does not specify how channel estimation can be done.

This coexistence mechanism can be considered as an optimization problem. Given the channel conditions, select first the packet type, and then the time of

INTEL-1020
U.S. Patent 9,265,043

packet transmission so that interference is minimized and the resulting throughput is within a certain range. One simple solution to this optimization problem is to adjust the time of packet transmission so that the IEEE 802.15.1 devices transmit during hops outside the WLAN frequencies and refrain from transmitting while in-band. This method does not require any action on the part of the IEEE 802.11b devices.

The first algorithmic problem is adaptive packet selection. Bluetooth specifies a variety of packet types with different combinations of payload length, slots occupied, FEC codes, and ARQ options. Choosing the packet type that will achieve maximum network performance can be tricky. First, consider an SCO link. Bluetooth provides four types of packets that can be sent over an SCO link: HV1, HV2, HV3, and DV packet. The different packets differ mostly in the FEC code that is used. The ACL link supports six packet types. They differ in the coding rate and the number of occupied slots. Again, the different SCO and ACL packets allow trade-offs to be made mainly between throughput and generated interference. Within this trade-off will be the solution to the optimization problem.

Consider first the case of a range-limited SCO connection. A receiver can decide that a connection is range-limited by low values of the RSSI. In a range-limited connection, random bit errors, not interference, are the dominant problem. In this case, choosing a packet type with stronger error-correcting capabilities will increase the performance of the SCO link. Therefore, for range-limited applications, HV1 packet is preferred over HV2 packet, and HV2 packet is preferred over HV3 packet.

In the case of interference-limited SCO connection, the dominant source of bit errors is the interference produced by the IEEE 802.11b network. A receiver can decide that a connection is interference-limited by high values of the RSSI and high BER, for example. In this case, lowering the coding rate to increase the protection will cause the Bluetooth device to generate more packets, because of the additional parity bits that must be sent. In turn, this creates more interference to the IEEE 802.11b network and ultimately instability, significantly decreasing the total network throughput. This unstable situation will end when the hopping frequency is out of the IEEE 802.11b band. Therefore, in interference-limited scenarios (as in Bluetooth and IEEE 802.11b coexistence scenarios), HV3 packets will lead to better results over HV2 packets, and HV2 packets over HV1 packets.

INTEL-1020
U.S. Patent 9,265,043

For the ACL packets, the situation is similar. When the Bluetooth network performance is range-limited, ACL packets with FEC protections (which include DM1, DM3, and DM5) will obtain better results. On the other hand, when the system is interference-limited, bandwidth efficient packets such as DH1, DH3, or DH5 will produce better results.

The third noncollaborative technique is ACL packet scheduling. This approach consists of two components: (1) frequency classification and (2) master delay policy. In the frequency classification phase, the Bluetooth devices measure channel conditions on the different frequencies and classify the frequencies as "good" or "bad." Again, the actual method for channel classification is not specified, but RSSI, BER, and PER are good criteria. The results are entered into a frequency classification table for every receiver. These tables should be kept at the master of the piconet. The goal of the so-called master delay policy is to avoid a packet transmission in a "bad" receiving channel. Because the master controls the network this policy must be implemented only at the master. The basic idea is to wait for a "good" frequency in the frequency-hopping pattern. Furthermore, because following each master's transmission, there is a slave transmission, the master checks both the slave's receiving frequency and its own receiving frequency before choosing to transmit a packet in a given frequency hop.

This scheduling policy is effective in reducing packet loss and delay, especially for multislot Bluetooth packets. In addition, packets are not transmitted when the channel is bad, which saves power. Also, channels occupied by other devices are avoided, which eliminates interference to other systems. Therefore, scheduling is a "good-citizen" policy. However, it must be noted that scheduling works only for ACL packets, because SCO packets must be sent at fixed intervals.

To further increase the robustness of the ACL link, the adaptive frequency hopping mechanism can be used in conjunction with a packet-scheduling algorithm.

Another noncollaborative mechanism is adaptive frequency hopping. It is a method according to which the hopping sequence is dynamically modified according to the frequency classification table, and devices hop over only the good frequencies. Therefore, channel classification and channel classification exchange between the master and all the active slaves on the piconet is a prerequisite for adaptive frequency hopping.

INTEL-1020
U.S. Patent 9,265,043

Adaptive frequency hopping (AFH) (Figure 3–9) consists of the three distinct components:

- The legacy hop kernel
- The partition sequence generator
- The frequency remapping function

The first component of the AFH mechanism is the legacy hop kernel, which generates the hopping sequence as defined in IEEE 802.15.1. The second component of the AFH mechanism is the partition sequence generator. Partition mapping is a method to select one channel from the partition. Each channel in the partition should have the same probability of selection on average. The mapping is pseudorandom. The pseudorandom mapping device maps the selected channel from the original hopping sequence into one channel of the partition specified by the partition sequence. The partition sequence generator makes sure that the new hopping sequence simultaneously appears to be random and is within the set of the good channels.

The output of the partition sequence generator is then used as an input to the final component of the AFH mechanism, the frequency remapping function, which generates an adapted hopping sequence with the appropriate structure. The frequency remapping function remaps (if necessary) the hopping frequency produced by the legacy kernel on to the set of the good channels S_G or the set of bad channels to be kept S_{BK}. Note that when the input to the frequency remapping function is a constant signal of one, it produces an adapted hopping sequence that only hops over the good channels. The remapping of the hopping frequencies is fairly straightforward. If the legacy hopping frequency is already in the set that is specified by the partition sequence, then the output of the frequency remapping function is equal to the legacy hopping frequency. However, if the legacy hopping frequency is not in the required set, then the legacy hopping frequency is remapped to a different frequency.

AFH would work well as long as the number of the good channels is above a certain minimum. Note that in the United States and other countries, regulatory bodies have set a limit on the minimal number of channels in the hopping sequence. System robustness also requires a minimal number of hopping channels. Depending on the number of available good channels, some bad

INTEL-1020
U.S. Patent 9,265,043

channels may have to be placed in the hopping sequence. If the number of good channels is below the minimum, first the SCO packets are allocated to the good channels, and what is left is distributed among the ACL packets. A level of QoS is thus guaranteed in interference environment.

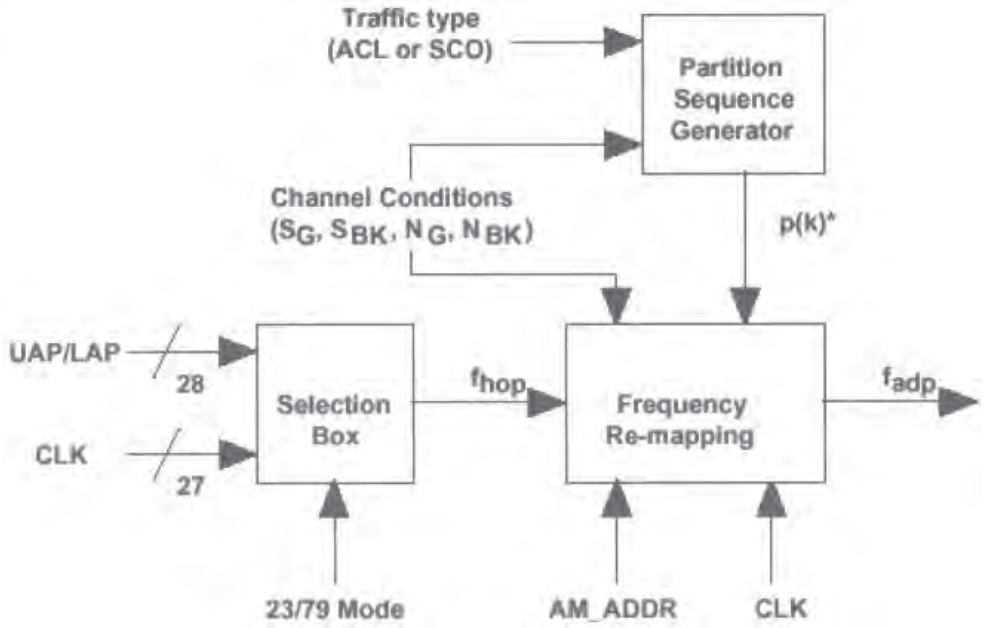

Figure 3–9: Block diagram of the adaptive frequency-hopping mechanism

Channel classification

Channel classification is required in both of the two noncollaborative mechanisms. Adaptive packet selection and scheduling adapts the packet types and transmission timing to the channel condition of the current hopping channel. Adaptive frequency hopping generates the new hopping sequence based on the result of channel classification. The goal of channel classification is to determine the quality of each channel. As discussed earlier, channels experiencing a high degree of interference are classified as "bad" channels, and channels with low interference are classified as "good" channels. The channel classification mechanism is not specified in the standard and is left to the technology and equipment vendors to implement. There are three sources of errors, each one leading to a packet error: access code error, header error, and payload error. In the

INTEL-1020
U.S. Patent 9,265,043

case of access code error, the packet will not be detected. Because it does not know what was transmitted, an actual receiver cannot measure exactly the packet error rate. However, a receiver can compute the ratio between the number of packets with a CRC error (error in the payload) to the number of correctly received packets.

High values for the RSSI and the PER indicate that the errors are due to interference. A low value of the RSSI might indicate that the distance between the devices is significant, but it would not be a reason for classifying a channel as "bad." Every slave and the master perform channel classification. Suppose that S_{ij} is the assessment by slave i of channel j, $i = 1,\ldots N_s$, where N_s is the number of slaves, and $j = 1,\ldots N_c$, N_c being the number of channels (79 or 23, depending on the regulatory domain). The master's assessment of channel j is M_j. S_{ij} and M_j take only binary values, 1, if channel j is assumed to be good, and 0, if it is assumed to be bad. The master uses the classification lists from the slaves and its own classification list to produce a final classification list. The quality of the j-th channel is assumed to be:

$$Q_j = \frac{M + \sum_{i=1}^{N_s} S_{i,j}}{N_s + 1} \qquad \text{Eq. 3–2}$$

$$1 \le N_s \le 7$$
$$0 \le j \le N_c$$

where the quality of one channel is a linear combination of the assessments of the master and the slave,

$$W_{ij} = \alpha M_j + (1 - \alpha) S_{ij} \qquad \text{Eq. 3–3}$$

To ultimately classify channel j as good or bad, a threshold is applied to the value of Q_j. The threshold is not defined in the standard. If Q_j is below a threshold, the channel j is assumed to be bad. The master then compiles the final list of "good" and "bad" channels to be distributed to every supporting device in the piconet. The master should make the final decision for the channel classification of the piconet.

INTEL-1020
U.S. Patent 9,265,043

In the network there could be legacy devices that are not capable of hopping adaptively. Recall that the hopping sequence is unchanged for all devices at the good channels. This requires the master to transmit broadcast packets only on the good channels. Thus, full support for legacy devices is maintained.

To reduce the time that classification will take, it is possible to reduce the number of measurements at each channel. The procedure is to group channels into blocks and classify the blocks instead of the channels. This will, however, compromise the accuracy of the measurements at each channel.

One-slot packets (such as DM1 or DH1 packets) may also shorten channel classification time. Classification should be performed constantly, both offline and online. Classification when a piconet is not established or the piconet is on hold and no traffic is being exchanged is called offline classification. This classification can involve background RSSI measurements.

PHY and MAC models

The recommended practice relies on analytical models for the physical layer transmitters of IEEE 802.11b and IEEE 802.15.1. Two cases are considered for the simulation results. The first case is the performance of IEEE 802.11b when Bluetooth is considered interference. The second case is the performance of Bluetooth when IEEE 802.11 is considered interference. The model is supplied with device positions and transmission parameters. The path loss model that is used is described by the free-space propagation described by Equation 1–6 with a coefficient 2 up to a distance of 8 m, and coefficient of 3.3 at distances over 8 m. To calculate the BER it can be assumed that an IEEE 802.15.1 device transmits two packets (Figure 3–10). An IEEE 802.11b device transmits a single frame, consisting of a header part and a payload (Figure 3–10). Notice the overlap among the packets. In general, there are six periods of stationarity. A new period of stationarity starts at the end of the IEEE 802.11b header, because the payload is transmitted using a different modulation technique than the header. By definition, during the period of stationarity, the transmit power and modulation type do not change, and the position of the devices is constant. This enables the BER to be calculated.

INTEL-1020
U.S. Patent 9,265,043

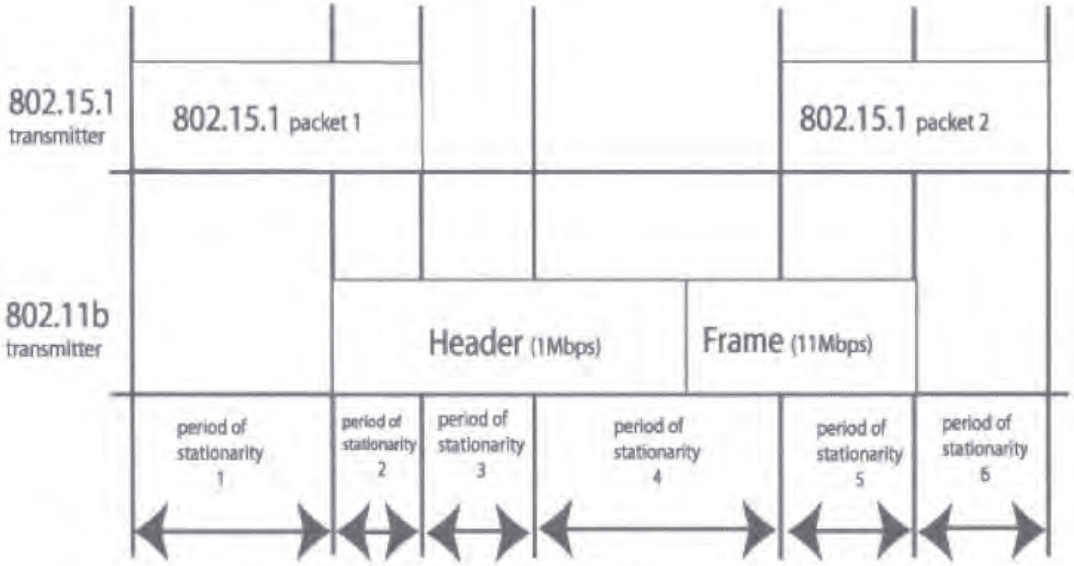

Figure 3–10: An example of interference between IEEE 802.15.1 and IEEE 802.11b with six different periods of stationarity

Figure 3–11 shows a diagram of the BER calculation. The figure shows that the received power of the intended transmission is the EIRP minus the path loss. The interference power at the receiver is the EIRP of the interfering transmission minus the path loss minus an additional factor (called spectrum factor) to account for the combined effect of receiver and transmitter masks. The total interference power is the sum of all interference powers. The SIR is the ratio of the received signal power to the total received interference power. Based on the SIR, first the the Symbol Error Rate is calculated for each modulation type. Then the BER is calculated from the Symbol Error Rate. Note that this model does not consider receiver noise. One simple way to account for receiver noise is simply to decrease the received signal power by a constant, equal to the expected value of the receiver noise. The recommended practice provides analytical formulae for evaluating the BER for the different modulation types in IEEE 802.11 and IEEE 802.15.1.

In addition to an analytical model, the recommended practice includes detailed Monte Carlo simulations. While the analytical model uses transmitter power and distance as input parameters, the simulation model uses signal-to-noise ratio and the signal-to-interference ratio. In both cases, the output is the BER.

INTEL-1020
U.S. Patent 9,265,043

To simulate the MAC of IEEE 802.15.1 and IEEE 802.11, the OPNET modeler, a network technology development environment, is used. Each of the IEEE 802.11 and Bluetooth MAC protocols is implemented as a state machine. Events trigger transitions from one state to another. The MAC models were interfaced to the PHY models to simulate a complete system. At the end of each packet transmission, a list is generated consisting of all interfering packets, the collision duration, the timing offset, the frequency, the power, and the topology used. This list is then passed to the PHY module along with a stream of bits representing the packet being transmitted.

For Bluetooth, two types of applications—voice and Internet traffic—are considered. Packet loss measures the number of packets discarded at the MAC layer due to errors in the bit stream. This measure is calculated after performing error correction. The residual number of errors in the Bluetooth voice packets measures the number of errors that remain in the packet payload after error correction is performed.

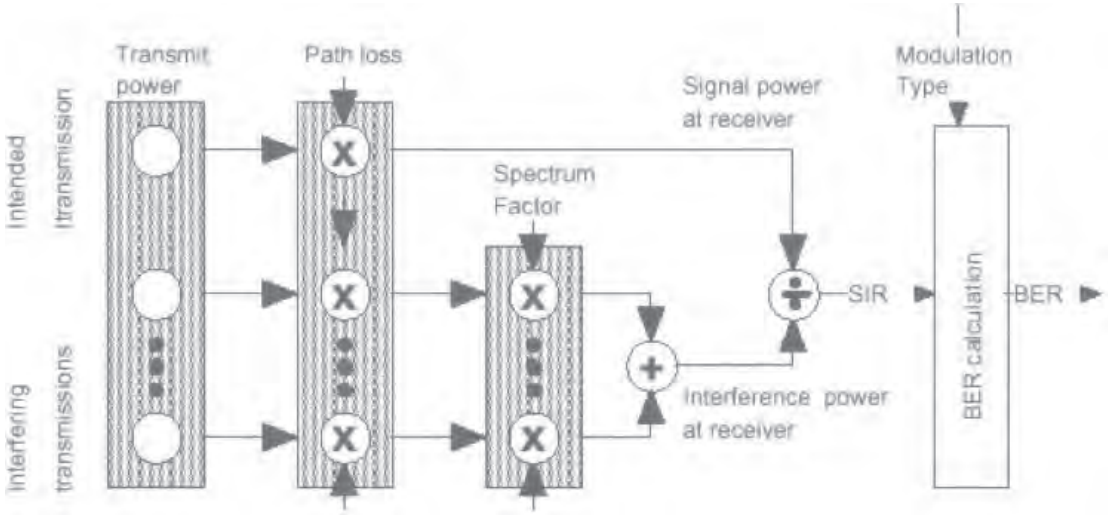

Figure 3–11: BER calculation process

INTEL-1020
U.S. Patent 9,265,043

HIGH-RATE WPAN

Overview

As discussed in this book's introduction, a fundamental law of communications is the constant desire to increase the data rate. In November of 1999, before IEEE 802.15.1 was completed, several companies (including Kodak™, Motorola, Cisco, and Aware) initiated the work on a higher-speed WPAN technology, which became IEEE 802.15.3. Future portable consumer electronics devices, such as digital cameras, provided the main motivation. These devices must support multimedia traffic that requires high bit rates. The high-rate WPAN technology must support:

a) ad-hoc connections
b) QoS
c) low cost
d) high speed—up to 55 Mb/s

Because multimedia is an especially important application, IEEE 802.15.3 decided that it would not use the MAC of IEEE 802.15.1, i.e., IEEE 802.15.3 will not only be a physical layer. In this sense, IEEE 802.15 is different from IEEE 802.11, because IEEE 802.11 has only one MAC. The fact that there is a market for a high-rate WPAN shows that Bluetooth is not sufficient. Bluetooth is not fast enough for all applications. Studies have shown that consumers want to transfer large files in less time than possible with Bluetooth. An area where Bluetooth is especially inadequate is video. The second answer comes from the fundamental principles of communications—higher data rate is always desirable, due to a combination of consumer demand and competition among vendors in the marketplace. If Bluetooth is not fast enough, yet another question is why not 802.11b, 802.11a, or 802.11g?. While 802.11 enjoys considerable advantage in the marketplace, as always, when there is a market need, there are several competing technologies trying to satisfy this market need.

Figure 3–12 shows a reference model of IEEE 802.15.3. In addition to a physical layer and a medium-access control layer, as shown in Figure 3–12, both the MAC and PHY layers have management entities, called correspondingly the MAC layer management entity and PHY layer management entity (MLME and PLME, respectively). These management entities interact with the Station Management

INTEL-1020
U.S. Patent 9,265,043