GLOBAL EDITION

Wireless Communication Networks and Systems

Cory Beard • William Stallings

ALWAYS LEARNING

PEARSON

Wireless Communication Networks and Systems

Global Edition

Cory Beard

University of Missouri-Kansas City

William Stallings

Global Edition Contributions by
Mohit Tahiliani
National Institute of Technology Karnataka

PEARSON

Boston • Columbus • Hoboken • Indianapolis • New York • San Francisco
Amsterdam • Cape Town • Dubai • London • Madrid • Milan • Munich • Paris • Montreal
Toronto • Delhi • Mexico City • São Paulo • Sydney • Hong Kong • Seoul • Singapore • Taipei • Tokyo

Vice President and Editorial Director, ECS: *Marcia J. Horton*
Executive Editor: *Tracy Johnson (Dunkelberger)*
Editorial Assistant: *Kelsey Loanes*
Assistant Acquisitions Editor, Global Edition: *Murchana Borthakur*
Associate Project Editor, Global Edition: *Binita Roy*
Program Manager: *Carole Snyder*
Director of Product Management: *Erin Gregg*
Team Lead Product Management: *Scott Disanno*
Project Manager: *Robert Engelhardt*
Media Team Lead: *Steve Wright*
R&P Manager: *Rachel Youdelman*
R&P Senior Project Manager: *Timothy Nicholls*
Procurement Manager: *Mary Fischer*
Senior Specialist, Program Planning and Support: *Maura Zaldivar-Garcia*
Inventory Manager: *Bruce Boundy*
Senior Manufacturing Controller, Production, Global Edition: *Trudy Kimber*
VP of Marketing: *Christy Lesko*
Director of Field Marketing: *Demetrius Hall*
Product Marketing Manager: *Bram van Kempen*
Marketing Assistant: *Jon Bryant*
Cover Designer: *Lumina Datamatics*
Cover Image: © *artemisphoto / Shutterstock*
Full-Service Project Management: *Mahalatchoumy Saravanan, Jouve India*

Pearson Education Limited
Edinburgh Gate
Harlow
Essex CM20 2JE
England

and Associated Companies throughout the world

Visit us on the World Wide Web at:
www.pearsonglobaleditions.com

Authorized adaptation from the United States edition, entitled Wireless Communication Networks and Systems, 5th edition, ISBN 978-0-13-359417-1, by Cory Beard and William Stallings, published by Pearson Education © 2016.

ISBN 10: 1-292-10871-1
ISBN 13: 978-1-292-10871-1

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

10 9 8 7 6 5 4 3 2 1

Typeset in 10/12 Times Ten LT Std by Jouve India.

Printed and bound by Courier Westford in the United States of America.

For my loving wife, Tricia

—WS

For Michelle, Ryan, and Jonathan,
gifts from God to me

—CB

This page intentionally left blank

CONTENTS

PREFACE

OBJECTIVES

Wireless technology has become the most exciting area in telecommunications and networking. The rapid growth of mobile telephone use, various satellite services, the wireless Internet, and now wireless smartphones, tablets, 4G cellular, apps, and the Internet of Things are generating tremendous changes in telecommunications and networking. It is not an understatement to say that wireless technology has revolutionized the ways that people work, how they interact with each other, and even how social structures are formed and transformed. This book provides a unified overview of the broad field of wireless communications. It comprehensively covers all types of wireless communications from satellite and cellular to local and personal area networks. Along with the content, the book provides over 150 animations, online updates to technologies after the book was published, and social networking tools to connect students with each other and instructors with each other.

The organization of the book reflects an attempt to break this massive subject into comprehensible parts and to build, piece by piece, a survey of the state of the art. The title conveys a focus on all aspects of wireless systems—wireless communication techniques, protocols and medium access control to form wireless networks, then the deployment and system management to coordinate the entire set of devices (base stations, routers, smartphones, sensors) that compose successful wireless systems. The best example of an entire wireless system is 4G Long Term Evolution (LTE).

For those new to the study of wireless communications, the book provides comprehension of the basic principles and topics of fundamental importance concerning the technology and architecture of this field. Then it provides a detailed discussion of leading-edge topics, including Gigabit Wi-Fi, the Internet of Things, ZigBee, and 4G LTE-Advanced.

The following basic themes serve to unify the discussion:

- **Technology and architecture:** There is a small collection of ingredients that serves to characterize and differentiate wireless communication and networking, including frequency band, signal encoding technique, error correction technique, and network architecture.
- **Network type:** This book covers the important types of wireless networks, including wireless LANs, wireless personal area networks, cellular, satellite, and fixed wireless access.
- **Design approaches:** The book examines alternative principles and approaches to meeting specific communication requirements. These considerations provide the reader with comprehension of the key principles that will guide wireless design for years to come.
- **Standards:** The book provides a comprehensive guide to understanding specific wireless standards, such as those promulgated by ITU, IEEE 802, and 3GPP, as well as standards developed by other organizations. This emphasis reflects the importance of such standards in defining the available products and future research directions in this field.
- **Applications:** A number of key operating systems and applications (commonly called "apps") have captivated the attention of consumers of wireless devices. This book examines the platforms and application development processes to provide apps that make wireless devices easily accessible to users.

The book includes an extensive online glossary, a list of frequently used acronyms, and a bibliography. Each chapter includes problems and suggestions for further reading. Each chapter also includes, for review, a list of key words and a number of review questions.

INTENDED AUDIENCES

This book is designed to be useful to a wide audience of readers and students interested in wireless communication networks and systems. Its development concentrated on providing flexibility for the following.

- **Variety of disciplines:** The book provides background material and depth so those from several disciplines can benefit.
  - Those with **computer science** and **information technology** backgrounds are provided with accessible and sufficient background on signals and systems. In addition to learning about all of the wireless systems, they can especially study complete systems like the Evolved Packet System that supports LTE and mobile device operating systems and programming.
  - Those from **electrical engineering**, **computer engineering**, and **electrical engineering technology** (and even other areas of engineering) are given what they need to know about networking and protocols. Then this book provides material sufficient for a senior undergraduate communications course with no prerequisite of another communication course. It provides substantial depth in Chapters 6 through 10 on wireless propagation, modulation techniques, OFDM, CDMA, and error control coding. The technologies in the later chapters of the book can then be used as examples of these techniques. This book not only provides fundamentals but also understanding of how they are used in current and future wireless technologies.
- **Ranges of experience:** Those who are novices with wireless communications, or even communication technologies themselves, are led through the knowledge they need to become proficient. And those with existing knowledge learn about the latest advances in wireless networking.
- **Levels of depth:** This book offers options for the level of depth used to cover different topics. Most notably Chapter 5, entitled Overview of Wireless Communications, provides tutorial-level coverage of the important wireless concepts needed to understand the rest of the book. For those needing more detailed understanding, however, Chapters 6 through 10 cover the same topics in more depth for fuller understanding. This again makes the book accessible to those with a variety of interests, level of prior knowledge, and expertise.

PLAN OF THE TEXT

The objective of this book is to provide a comprehensive technical survey of wireless communications fundamentals, wireless networks, and wireless applications. The book is organized into four parts as illustrated in Figure P.1.

Part One, Technical Background: Provides background material on the process of data and packet communications, as well as protocol layers, TCP/IP, and data networks.

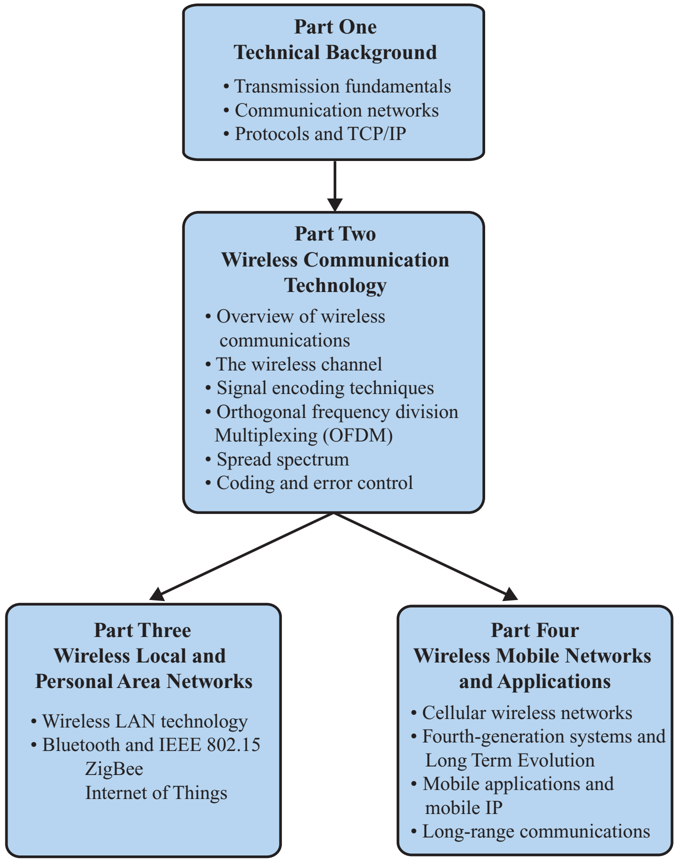


Figure P.1 Wireless Topics

Part Two, Wireless Communication Technology: Covers all of the relevant information about the process of sending a wireless signal and combating the effects of the wireless channel. The material can be covered briefly with Chapter 5, Overview of Wireless Communications, or through five chapters on the wireless channel (antennas and propagation), signal encoding, OFDM, spread spectrum, and error control coding.

Part Three, Wireless Local and Personal Area Networks: Provides details on IEEE 802.11, IEEE 802.15, Bluetooth, the Internet of Things, and ZigBee.

Part Four, Wireless Mobile Networks and Applications: Provides material on mobile cellular systems principles, LTE, smartphones, and mobile applications. It also covers long-range communications using satellite, fixed wireless, and WiMAX.

The book includes a number of pedagogic features, including the use of over 150 animations and numerous figures and tables to clarify the discussions. More details are given below. Each chapter also includes a list of key words, review questions, homework problems, and suggestions for further reading. The book also includes an extensive online glossary, a list of frequently used acronyms, and a reference list.

ORDER OF COVERAGE

With a comprehensive work such as this, careful planning is required to cover those parts of the text most relevant to the students and the course at hand. The book provides some flexibility. For example, the material in the book need not be studied sequentially. As a matter of fact, it has been the experience of the authors that students and instructors are more engaged if they are able to dive into the technologies themselves as soon as possible. One of the authors in his courses has routinely studied IEEE 802.11 (Chapter 11) before concentrating on the full details of wireless communications. Some physical layer details may need to be skipped at first (e.g., temporarily skipping Sections 11.5 and 11.6), but students are more engaged and able to perform projects if they've studied the actual technologies earlier.

The following are suggestions concerning paths through the book:

- Chapter 5, Overview of Wireless Communications, can be substituted for Chapters 6 through 10. Conversely, Chapter 5 should be omitted if using Chapters 6 through 10.
- Part Three can be covered before Part Two, omitting some physical layer details to be revisited later. Part Two should precede Part Four, however.
- Chapters 2 through 4 can be left as outside reading assignments. Especially by using animations provided with the book, some students can be successful studying these topics on their own.
- Within Part Three, the chapters are more or less independent and can be studied in either order depending on level of interest.
- The chapters in Part Four can also be studied in any order, except Chapters 13 and 14 on cellular systems and LTE should be studied as a unit.
- Computer science and information technology courses could focus more on Wi-Fi, IEEE 802.15, and mobile applications in Chapters 11, 12, and 15, then proceed with projects on MAC protocols and mobile device programming.
- Electrical engineering and engineering technology students can focus on Chapters 6 through 10 and proceed with projects related to the modulation and error control coding schemes used for IEEE 802.11 and LTE.

ANIMATIONS

Animations provide a powerful tool for understanding the complex mechanisms discussed in this book, including forward error correction, signal encoding, and protocols. Over 150 Web-based animations are used to illustrate many of the data communications and protocol concepts in this book.

The animations progressively introduce parts of diagrams and help to illustrate data flow, connection setup and maintenance procedures, error handling, encapsulation, and the ways technologies perform in different scenarios. For example, see Figure P.2 and its animation. This is actually Figure 13.7. From the print version, the animations can be accessed through the QR code next to the figure or through the book's Premium Web site discussed below. Walking step-by-step through the animation can be accomplished with a click or tap on the animation. This figure shows possible choices of handoff decisions at different locations between two base stations. The original figure might be difficult for the reader to first understand, but the animations give good enhanced understanding by showing the figure piece-by-piece with extra explanation. These animations provide significant help to the reader to understand the purpose behind each part of the figure.

INSTRUCTOR SUPPORT MATERIALS

The major goal of this text is to make it as effective a teaching tool for this exciting and fast-moving subject as possible. This goal is reflected both in the structure of the book and in the supporting material. The text is accompanied by the following supplementary material to aid the instructor:

- **Solutions manual:** Solutions to all end-of-chapter Review Questions and Problems.
- **Supplemental problems:** More problems beyond those offered in the text.
- **Projects manual:** Suggested project assignments for all of the project categories listed in the next section.
- **PowerPoint slides:** A set of slides covering all chapters, suitable for use in lecturing.
- **PDF files:** Reproductions of all figures and tables from the book.
- **Wireless courses:** Links to home pages for courses based on this book. These pages may be useful to other instructors in providing ideas about how to structure their course.
- **Social networking:** Links to social networking sites that have been established for instructors using the book, such as on Facebook and LinkedIn, where instructors can interact.

All of these support materials are available at the **Instructor Resource Center (IRC)** for this textbook, which can be reached through the publisher's Web site

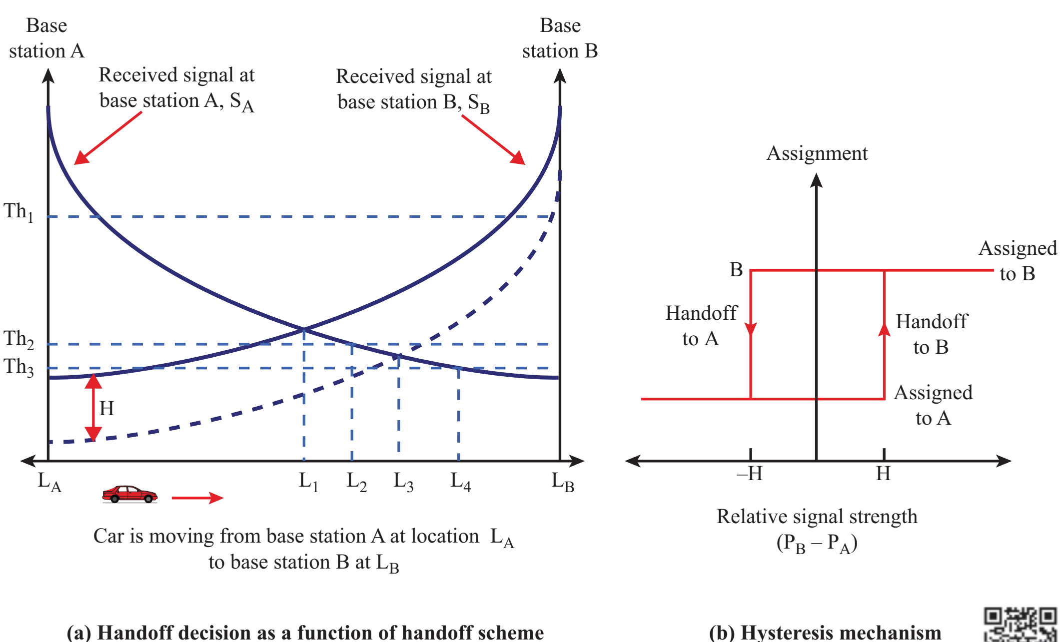


Figure P.2 Handoff Between Two Cells

www.pearsonglobaleditions.com/Stallings. To gain access to the IRC, please contact your local Pearson sales representative.

The **Companion Web site**, at www.corybeardwireless.com, includes technology updates, Web resources, etc. This is discussed in more detail below in the section about student resources.

PROJECTS AND OTHER STUDENT EXERCISES

For many instructors, an important component of a wireless networking course is a project or set of projects by which students get hands-on experience to reinforce concepts from the text. This book provides an unparalleled degree of support for including a projects component in the course. The IRC not only provides guidance on how to assign and structure the projects but also includes a set of User's Manuals for various project types plus specific assignments, all written especially for this book. Instructors can assign work in the following areas:

- **Practical exercises:** Using network commands, students gain experience in network connectivity.
- **Wireshark projects:** Wireshark is a protocol analyzer that enables students to study the behavior of protocols. A video tutorial is provided to get students started, in addition to a set of Wireshark assignments.
- **Simulation projects:** Students can use different suggested simulation packages to analyze network behavior. The IRC includes a number of student assignments.
- **Performance modeling projects:** Multiple performance modeling techniques are introduced. The IRC includes a number of student assignments.
- **Research projects:** The IRC includes a list of suggested research projects that would involve Web and literature searches.
- **Interactive assignments:** Twelve interactive assignments have been designed to allow students to give a specific set of steps to invoke, or devise a sequence of steps to achieve a desired result. The IRC includes a set of assignments, plus suggested solutions, so that instructors can assess students' work.

This diverse set of projects and other student exercises enables the instructor to use the book as one component in a rich and varied learning experience and to tailor a course plan to meet the specific needs of the instructor and students.

RESOURCES FOR STUDENTS

A substantial amount of original supporting material for students has been made available online, at two Web locations. The **Companion Web site**, at www.corybeardwireless.com, includes the following.

- **Social networking tools**: Students using the book can interact with each other to share questions and insights and develop relationships. Throughout the lifetime of the book, various social networking tools may become prevalent; new social networking sites will be developed and then links and information about them will be made available here.

- **Useful Web sites:** There are links to other relevant Web sites which provide extensive help in studying these topics. Links to these are provided.
- **Errata sheet:** An errata list for this book will be maintained and updated as needed. Please e-mail any errors that you spot from the link at corybeardwireless.com.
- **Documents:** These include a number of documents that expand on the treatment in the book. Topics include standards organizations and the TCP/IP checksum.
- **Wireless courses:** There are links to home pages for courses based on this book. These pages may be useful to other instructors in providing ideas about how to structure their course.

Purchasing this textbook new also grants the reader twelve months of access to the **Premium Content site**, which includes the following:

- **Animations:** Those using the print version of the book can access the animations by going to this Web site. The QR codes next to the book figures give more direct access to these animations. The ebook version provides direct access to these animations by clicking or tapping on a linked figure.
- **Glossary:** List of key terms and definitions.
- **Appendices:** Three appendices to the book are available on traffic analysis, Fourier analysis, and data link control protocols.
- **Technology updates:** As new standards are approved and released, new chapter sections will be developed. They will be released here before a new edition of the text is published. The book will therefore not become outdated in the same way that is common with technology texts.

To access the Premium Website, click on the *Premium Website* link at www.pearsonglobaleditions.com/Stallings and enter the student access code found on the card in the front of the book.

William Stallings also maintains the Computer Science Student Resource Site, at computersciencestudent.com. The purpose of this site is to provide documents, information, and useful links for computer science students and professionals. Links are organized into four categories:

- **Math:** Includes a basic math refresher, a queuing analysis primer, a number system primer, and links to numerous math sites
- **How-to:** Advice and guidance for solving homework problems, writing technical reports, and preparing technical presentations
- **Research resources:** Links to important collections of papers, technical reports, and bibliographies
- **Miscellaneous:** A variety of useful documents and links

ACKNOWLEDGMENTS

This book has benefited from review by a number of people, who gave generously of their time and expertise. The following professors and instructors provided reviews: Alex Wijes-Iinha (Towson University), Dr. Ezzat Kirmani (St. Cloud State University), Dr. Feng Li (Indiana University-Purdue University Indianapolis), Dr. Guillermo A. Francia III (Jacksonville State University), Dr. Kamesh Namuduri (University of North Texas), Dr. Melody Moh (San Jose State University), Dr. Wuxu Peng (Texas State University), Frank E. Green (University of Maryland, Baltimore County), Gustavo Vejarano (Loyola Marymount University), Ilker Demirkol (Rochester Institute of Tech), Prashant Krishnamurthy (University of Pittsburgh), and Russell C. Pepe (New Jersey Institute of Technology).

Several students at the University of Missouri-Kansas City provided valuable contributions in the development of the figures and animations. Bhargava Thondapu and Siva Sai Karthik Kesanakurthi provided great creativity and dedication to the animations. Pedro Tonhozi de Oliveira, Rahul Arun Paropkari, and Naveen Narasimhaiah also devoted themselves to the project and provided great help.

Kristopher Micinski contributed most of the material on mobile applications in Chapter 15.

Finally, we thank the many people responsible for the publication of the book, all of whom did their usual excellent job. This includes the staff at Pearson, particularly our editor Tracy Johnson, program manager Carole Snyder, and production manager Bob Engelhardt. We also thank Mahalatchoumy Saravanan and the production staff at Jouve India for an excellent and rapid job. Thanks also to the marketing and sales staffs at Pearson, without whose efforts this book would not be in front of you.

Pearson wishes to thank and acknowledge the following people for reviewing the Global Edition:

Moumita Mitra Manna, Bangabasi College, Kolkata
Dr. Chitra Dhawale, P. R. Patil Group of Educational Institutes, Amravati
Nikhil Marriwala, Kurukshetra University

About the Authors

Dr. William Stallings has authored 17 textbooks, and counting revised editions, over 40 books on computer security, computer networking, and computer architecture. In over 30 years in the field, he has been a technical contributor, a technical manager, and an executive with several high-technology firms. Currently he is an independent consultant whose clients have included computer and networking manufacturers and customers, software development firms, and leading-edge government research institutions. He has 13 times received the award for the best Computer Science textbook of the year from the Text and Academic Authors Association.

He created and maintains the Computer Science Student Resource Site at ComputerScienceStudent.com. This site provides documents and links on a variety of subjects of general interest to computer science students (and professionals). He is a member of the editorial board of Cryptologia, a scholarly journal devoted to all aspects of cryptology.

Dr. Stallings holds a PhD from MIT in computer science and a BS from Notre Dame in electrical engineering.

Dr. Cory Beard is an Associate Professor of Computer Science and Electrical Engineering at the University of Missouri-Kansas City (UMKC). His research areas involve the prioritization of communications for emergency purposes. This work has involved 3G/4G cellular networks for public safety groups, MAC layer performance evaluation, call preemption and queuing, and Internet traffic prioritization. His work has included a National Science Foundation CAREER Award.

He has received multiple departmental teaching awards and has chaired degree program committees for many years. He maintains a site for book-related social networking and supplemental materials at corybeardwireless.com.

This page intentionally left blank

CHAPTER 1

INTRODUCTION

1.1 **Wireless Comes of Age**

1.2 **The Global Cellular Network**

1.3 **The Mobile Device Revolution**

1.4 **Future Trends**

1.5 **The Trouble with Wireless**

LEARNING OBJECTIVES

After studying this chapter, you should be able to:

- Describe how wireless communications have developed.
- Explain the purposes of various generations of cellular technology.
- Describe the ways mobile devices have revolutionized and will continue to revolutionize society.
- Identify and describe future trends.

This book is a survey of the technology of wireless communication networks and systems. Many factors, including increased competition, the introduction of digital technology, mobile device user interface design, video content, and social networking have led to unprecedented growth in the wireless market. In this chapter, we discuss some of the key factors driving this wireless networking revolution.

1.1 WIRELESS COMES OF AGE

Guglielmo Marconi invented the wireless telegraph in 1896.[1] In 1901, he sent telegraphic signals across the Atlantic Ocean from Cornwall to St. John's Newfoundland, a distance of about 3200 km. His invention allowed two parties to communicate by sending each other alphanumeric characters encoded in an analog signal. Over the last century, advances in wireless technologies have led to the radio, the television, communications satellites, mobile telephone, and mobile data. All types of information can now be sent to almost every corner of the world. Recently, a great deal of attention has been focused on wireless networking, cellular technology, mobile applications, and the Internet of Things.

Communications satellites were first launched in the 1960s; today satellites carry about one-third of the voice traffic and all of the television signals between countries. Wireless networking allows businesses to develop WANs, MANs, and LANs without a cable plant. The IEEE 802.11 standard for wireless LANs (also known as Wi-Fi) has become pervasive. Industry consortiums have also provided seamless short-range wireless networking technologies such as ZigBee, Bluetooth, and Radio Frequency Identification tags (RFIDs).

The cellular or mobile telephone started with the objective of being the modern equivalent of Marconi's wireless telegraph, offering two-party, two-way communication. Early generation wireless phones offered voice and limited data services through bulky devices that gradually became more portable. Current third and

[1]The actual invention of radio communications more properly should be attributed to Nikola Tesla, who gave a public demonstration in 1893. Marconi's patents were overturned in favor of Tesla in 1943 [ENGE00].

fourth generation devices are for voice, texting, social networking, mobile applications, mobile Web interaction, and video streaming. These devices also include cameras and a myriad of sensors to support the device applications. The areas of coverage for newer technologies are continually being expanded and focused on key user populations.

The impact of wireless communications has been and will continue to be profound. Very few inventions have been able to "shrink" the world in such a manner, nor have they been able to change the way people communicate as significantly as the way wireless technology has enabled new forms of social networking. The standards that define how wireless communications devices interact are quickly converging, providing a global wireless network that delivers a wide variety of services.

Figure 1.1 highlights some of the key milestones in the development of wireless communications.[2] Wireless technologies have gradually migrated to higher frequencies. As will be seen in later chapters, higher frequencies enable the support of greater data rates and throughput but require higher power, are more affected by obstructions, and have shorter effective range.

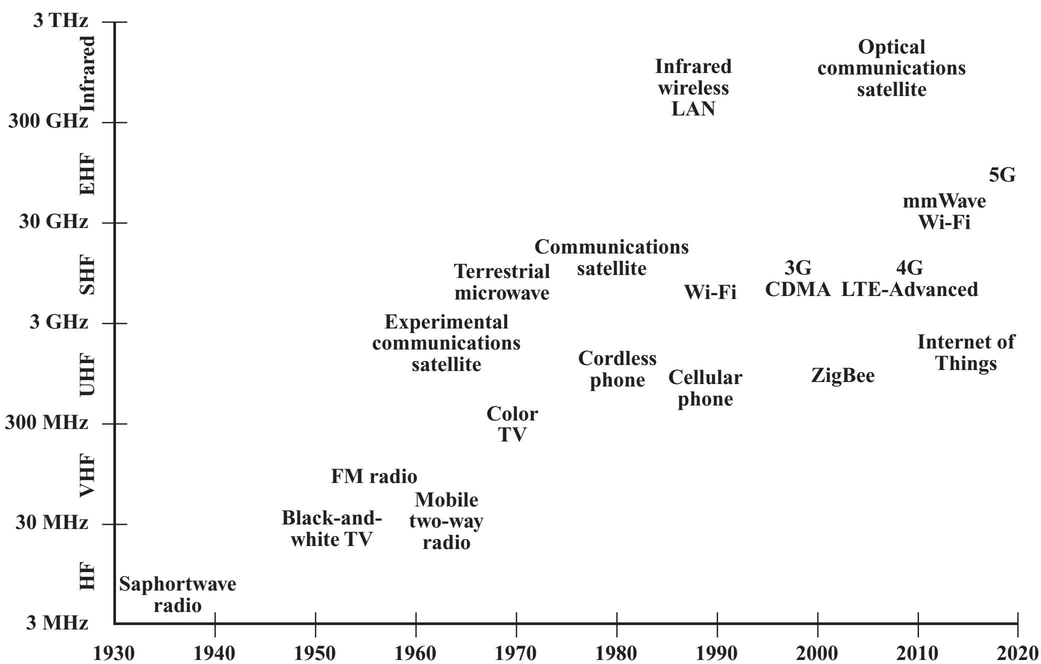


Figure 1.1 Some Milestones in Wireless Communications

[2]Note the use of a log scale for the *y*-axis. A basic review of log scales is in the math refresher document at the Computer Science Student Resource Site at computersciencestudent.com.

1.2 THE GLOBAL CELLULAR NETWORK

The cellular revolution is apparent in the growth of the mobile phone market alone. In 1990, the number of users was approximately 11 million [ECON99]. Today, according to 4G Americas, that number is over seven billion. There are a number of reasons for the dominance of mobile devices. Mobile devices are convenient; they move with people. In addition, by their nature, they are location aware. Mobile cellular devices communicate with regional base stations that are at fixed locations. In many geographic areas, mobile telephones are the only economical way to provide phone service to the population. Operators can erect base stations quickly and inexpensively when compared with digging up ground to lay cables in harsh terrain.

Today there is no single cellular network. Devices support several technologies and generally work only within the confines of a single operator's network. To move beyond this model, work is being done to define and implement standards.

The dominant first-generation wireless network in North America was the Advanced Mobile Phone System (AMPS). The key second generation wireless systems are the Global System for Mobile Communications (GSM), Personal Communications Service (PCS) IS-136, and PCS IS-95. The PCS standard IS-136 uses time division multiple access (TDMA); GSM uses a combination of TDMA and frequency division multiple access (FDMA), and IS-95 uses code division multiple access (CDMA). 2G systems primarily provide voice services, but also provide some moderate rate data services.

The two major third-generation systems are CDMA2000 and Universal Mobile Telephone Service (UMTS). Both use CDMA and are meant to provide packet data services. CDMA2000 released 1xRTT (1 times Radio Transmission Technology) and then 1xEV-DO (1 times Evolution-Data Only) through Release 0, Revision A, and Revision B. The competing UMTS uses Wideband CDMA. It is developed by the Third Generation Partnership Project (3GPP); its first release was labeled Release 99 in 1999, but subsequent releases were labeled Releases 4 onward.

The move to fourth generation mainly involved competition between IEEE 802.16 WiMAX, described in Chapter 15, and Long Term Evolution (LTE), described in Chapter 14. Both use a different approach than CDMA for high spectral efficiency in a wireless channel called orthogonal frequency division multiplexing (OFDM). The requirements for 4G came from directives by the International Telecommunication Union (ITU), which said that 4G networks should provide all-IP services at peak data rates of up to approximately 100 Mbps for high-mobility mobile access and up to approximately 1 Gbps for low-mobility access. LTE, also developed by 3GPP, ended up the predominant technology for 4G, and 3GPP Release 8 was its first release. Although LTE Release 8 does not meet the ITU requirements (even though marketers have called it "4G LTE"), the later Release 10 achieves the goals and is called LTE-Advanced. There are a wide number of Release 8 deployments so far but much fewer Release 10 upgrades.

1.3 THE MOBILE DEVICE REVOLUTION

Technical innovations have contributed to the success of what were originally just mobile phones. The prevalence of the latest devices, with multi-megabit Internet access, mobile apps, high megapixel digital cameras, access to multiple types of wireless networks (e.g., Wi-Fi, Bluetooth, 3G, and 4G), and several on-board sensors, all add to this momentous achievement. Devices have become increasingly powerful while staying easy to carry. Battery life has increased (even though device energy usage has also expanded), and digital technology has improved reception and allowed better use of a finite spectrum. As with many types of digital equipment, the costs associated with mobile devices have been decreasing.

The first rush to wireless was for voice. Now, the attention is on data; some wireless devices are only rarely used for voice. A big part of this market is the wireless Internet. Wireless users use the Internet differently than fixed users, but in many ways no less effectively. Wireless smartphones have limited displays and input capabilities compared with larger devices such as laptops or PCs, but mobile apps give quick access to intended information without using Web sites. Because wireless devices are location aware, information can be tailored to the geographic location of the user. Information finds users, instead of users searching for information. Tablet devices provide a happy medium between the larger screens and better input capabilities of PCs and the portability of smartphones.

Examples of wireless technologies that are used for long distance are cellular 3G and 4G, Wi-Fi IEEE 802.11 for local areas, and Bluetooth for short distance connections between devices. These wireless technologies should provide sufficient data rates for the intended uses, ease of connectivity, stable connections, and other necessary quality of service performance for services such as voice and video. There are still improvements needed to meet these requirements in ways that are truly invisible to end users.

For many people, wireless devices have become a key part of how they interact with the world around them. Currently, this involves interaction with other people through voice, text, and other forms of social media. They also interact with various forms of multimedia content for business, social involvement, and entertainment. In the near future, many envision advanced ways for people to interact with objects and machines around them (e.g., the appliances in a home) and even for the devices themselves to perform a more active role in the world.

1.4 FUTURE TRENDS

As 4G LTE-Advanced and higher speed Wi-Fi systems are now being deployed, many see great future untapped potential to be realized. Great potential exists for Machine to Machine (MTM) communications, also called the Internet of Things (IoT). The basic idea is that devices can interact with each other in areas such as healthcare, disaster recovery, energy savings, security and surveillance, environmental awareness, education, inventory and product management, manufacturing, and many others. Today's current smart devices could interact with myriads of objects equipped with wireless networking capabilities. This could start with information dissemination to enable data mining and decision support, but could also involve capabilities for automated remote adaptation and control. For example, a residential home could have sensors to monitor temperature, humidity, and airflow to assess human comfort levels. These sensors could also collaborate with home appliances, heating and air conditioning systems, lighting systems, electric vehicle charging stations, and utility companies to provide homeowners with advice or even automated control to optimize energy consumption. This would adjust when homeowners are at home conducting certain activities or away from home. Eventually these wirelessly equipped objects could interact in their own forms of social networking to discover, trust, and collaborate.

Future wireless networks will have to significantly improve to enable these capabilities. Some envision a 100-fold increase in the number of communication devices. And the type of communication would involve many short messages, not the type of communication supported easily by the current generations of the technologies studied in this book. If these communications were to involve control applications between devices, the real-time delay requirements would be much more stringent than that required in human interaction.

Also, the demands for capacity will greatly increase. The growth in the number of subscribers and per-user throughput gives a prediction of a 1000-fold increase in data traffic by 2020. This has caused the development of the following technologies for what may be considered 5G (although the definition of 5G has not been formalized). Some of these will be better understood after studying the topics in this book, but we provide them here to set the stage for learning expectations.

- **Network densification** will use many small transmitters inside buildings (called femtocells) and outdoors (called picocells or relays) to reuse the same carrier frequencies repeatedly.
- **Device-centric architectures** will provide connections that focus on what a device needs for interference reduction, throughput, and overall service quality.
- **Massive multiple-input multiple-output (MIMO)** will use 10 or more than 100 antennas (both on single devices and spread across different locations) to focus antenna beams toward intended devices even as the devices move.
- **Millimeter wave (mmWave)** frequencies in the 30 GHz to 300 GHz bands have much available bandwidth. Even though they require more transmit power and have higher attenuation due to obstructions and atmosphere, massive MIMO can be used to overcome those limitations.
- **Native support for mobile to mobile (MTM) communication** will accommodate low data rates, a massive number of devices, sustained minimum rates, and very low delays.

Throughout this book, the reader will see the methods by which technologies such as Wi-Fi have expanded and improved. We will review the foundational technologies and see the ways in which new directions such as OFDM and LTE-Advanced have created dramatic improvements. This provides excellent preparation so that researchers and practitioners will be ready to participate in these future areas.

1.5 THE TROUBLE WITH WIRELESS

Wireless is convenient and often less expensive to deploy than fixed services, but wireless is not perfect. There are limitations, political and technical difficulties, that may ultimately hamper wireless technologies from reaching their full potential. Two issues are the wireless channel and spectrum limitations.

The delivery of a wireless signal does not always require a free line-of-sight path, depending on the frequency. Signals can also be received through transmission through objects, reflections off objects, scattering of signals, and diffraction around the edges of objects. Unfortunately, reflections can cause multiple copies of the signal to arrive at the receiver at different times with different attenuations. This creates the problem of **multipath fading** when the signals add together and can cause the signal to be significantly degraded. Wireless signals also suffer from noise, interference from other users, and Doppler shifting caused by movement of devices.

A series of approaches are used to combat these problems of wireless transmission. All are discussed in this book.

- **Modulation** sends digital data in a signal format that sends as many bits as possible for the current wireless channel.
- **Error control coding**, also known as channel coding, adds extra bits to a signal so that errors can be detected and corrected.
- **Adaptive modulation and coding** dynamically adjusts the modulation and coding to measurements of the current channel conditions.
- **Equalization** counteracts the multipath effects of the channel.
- **Multiple-input multiple-output** systems use multiple antennas to point signals strongly in certain directions, send simultaneous signals in multiple directions, or send parallel streams of data.
- **Direct sequence spread spectrum** expands the signal over a wide bandwidth so that problems in parts of the bandwidth are overcome because of the wide bandwidth.
- **Orthogonal frequency division multiplexing** breaks a signal into many lower rate bit streams where each is less susceptible to multipath problems.

Spectrum regulations also affect the capabilities of wireless communications. Governmental regulatory agencies allocate spectrum to various types of uses, and wireless communications companies frequently spend large amounts of money to acquire spectrum. These agencies also give rules related to power and spectrum sharing approaches. All of this limits the bandwidth available to wireless communications. Higher frequencies have more available bandwidth but are harder to use effectively due to obstructions. They also inherently require more transmission power. Transition from today's 1 GHz to 5 GHz bands to millimeter wave (mmWave) bands in the 30 GHz to 300 GHz range is of increasing interest since they have more bandwidth available.

Part One

Technical Background

CHAPTER 2

TRANSMISSION FUNDAMENTALS

LEARNING OBJECTIVES

After studying this chapter, you should be able to:

- Distinguish between digital and analog information sources.
- Explain the various ways in which audio, data, image, and video can be represented by electromagnetic signals.
- Discuss the characteristics of analog and digital waveforms.
- Explain the roles of frequencies and frequency components in a signal.
- Identify the factors that affect channel capacity.
- Compare and contrast various forms of wireless transmission.

The purpose of this chapter is to make this book self-contained for the reader with little or no background in data communications. For the reader with greater interest, references for further study are supplied at the end of the chapter.

2.1 SIGNALS FOR CONVEYING INFORMATION

In this book, we are concerned with electromagnetic signals used as a means to transmit information. An electromagnetic signal is a function of time, but it can also be expressed as a function of frequency; that is, the signal consists of components of different frequencies. It turns out that the **frequency domain** view of a signal is far more important to an understanding of data transmission than a **time domain** view. Both views are introduced here.

Time Domain Concepts

Viewed as a function of time, an electromagnetic signal can be either analog or digital. An **analog signal** is one in which the signal intensity varies in a smooth fashion over time. In other words, there are no breaks or discontinuities in the signal. A **digital signal** is one in which the signal intensity maintains a constant level for some period of time and then changes to another constant level.[1] Figure 2.1 shows examples of both kinds of signals. The analog signal might represent speech, and the digital signal might represent binary 1s and 0s.

The simplest sort of signal is a **periodic signal**, in which the same signal pattern repeats over time. Figure 2.2 shows an example of a periodic analog signal (sine wave) and a periodic digital signal (square wave). Mathematically, a signal $s(t)$ is defined to be periodic if and only if

$$s(t + T) = s(t) \qquad -\infty < t < +\infty$$

where the constant T is the period of the signal (T is the smallest value that satisfies the equation). Otherwise, a signal is **aperiodic**.

[1]This is an idealized definition. In fact, the transition from one voltage level to another will not be instantaneous, but there will be a small transition period. Nevertheless, an actual digital signal approximates closely the ideal model of constant voltage levels with instantaneous transitions.

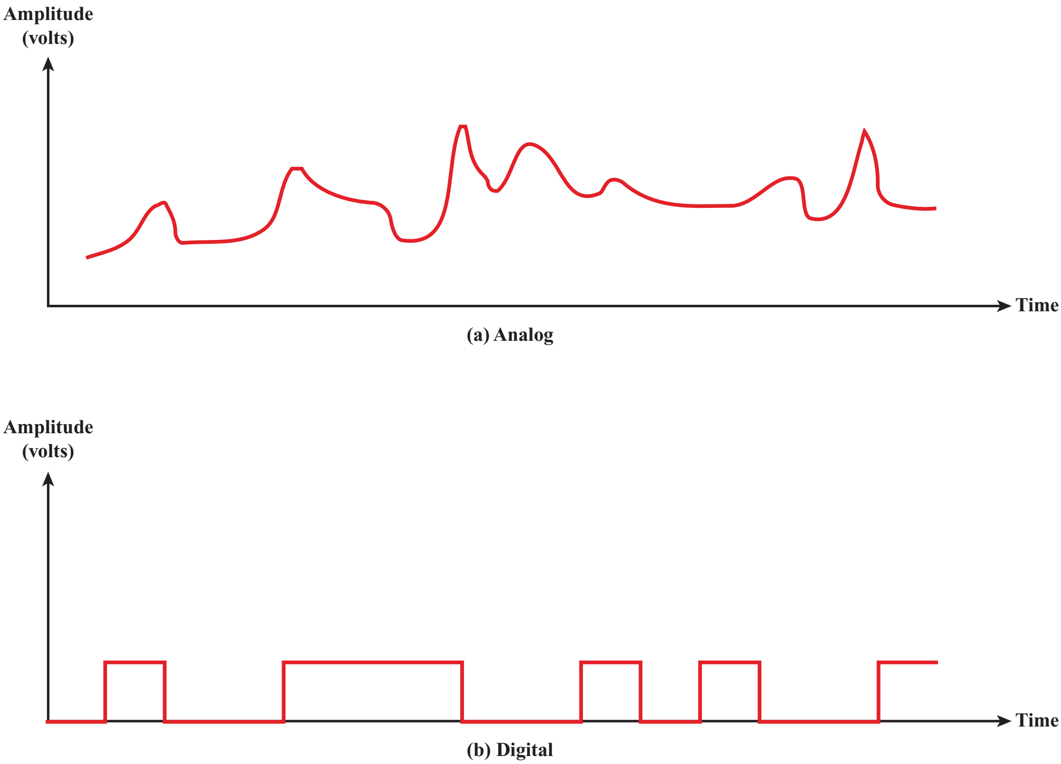


Figure 2.1 Analog and Digital Waveforms

The sine wave is the fundamental analog signal. A general sine wave can be represented by three parameters: peak amplitude (A), frequency (f), and phase (ϕ). The **peak amplitude** is the maximum value or strength of the signal over time; typically, this value is measured in volts. The **frequency** is the rate [in cycles per second, or Hertz (Hz)] at which the signal repeats. An equivalent parameter is the **period** (T) of a signal, which is the amount of time it takes for one repetition; therefore, $T = 1/f$. **Phase** is a measure of the relative position in time within a single period of a signal, as illustrated later.

The general sine wave can be written as

$$s(t) = A \sin(2\pi ft + \phi) \tag{2.1}$$

A function with the form of Equation (2.1) is known as a **sinusoid**. Figure 2.3 shows the effect of varying each of the three parameters. In part (a) of the figure, the frequency is 1 Hz; thus the period is $T = 1$ second. Part (b) has the same frequency and phase but a peak amplitude of 0.5. In part (c) we have $f = 2$, which is equivalent to $T = 0.5$. Finally, part (d) shows the effect of a phase shift of $\pi/4$ radians, which is 45 degrees (2π radians $= 360° = 1$ period).

In Figure 2.3 the horizontal axis is time; the graphs display the value of a signal at a given point in space as a function of time. These same graphs, with a change of

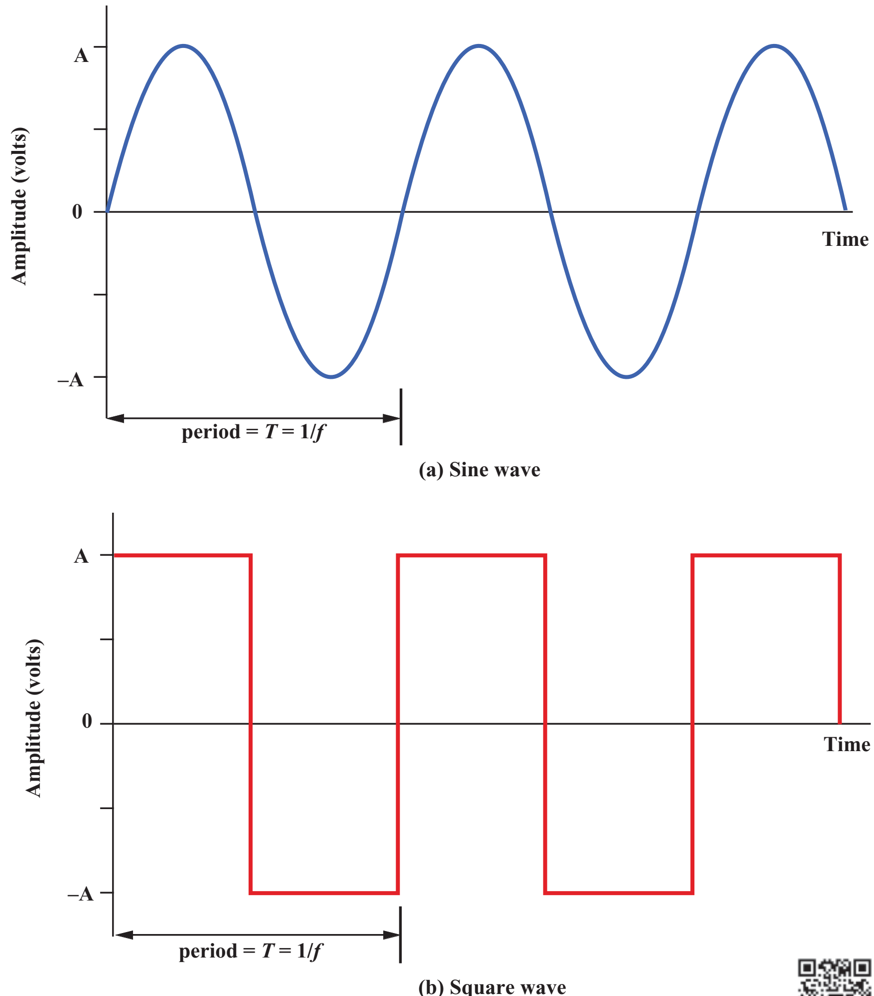
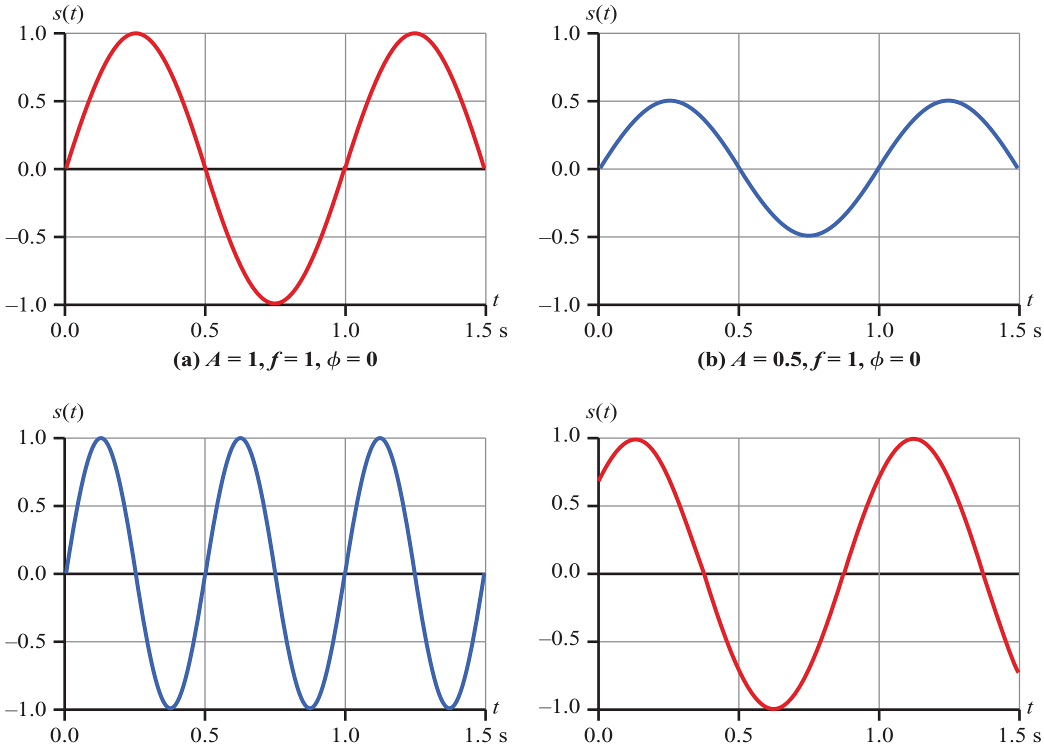


Figure 2.2 Examples of Periodic Signals

scale, can apply with horizontal axes in space. In that case, the graphs display the value of a signal at a given point in time as a function of distance. For example, for a sinusoidal transmission (e.g., an electromagnetic radio wave some distance from a radio antenna or sound some distance from loudspeaker) at a particular instant of time, the intensity of the signal varies in a sinusoidal way as a function of distance from the source.

There is a simple relationship between the two sine waves, one in time and one in space. The **wavelength** (λ) of a signal is the distance occupied by a single cycle, or, put another way, the distance between two points of corresponding phase of two consecutive cycles. Assume that the signal is traveling with a velocity v. Then the wavelength is related to the period as follows: $\lambda = vT$. Equivalently, $\lambda f = v$. Of particular relevance to this discussion is the case where $v = c$, the speed of light in free space, which is approximately 3×10^8 m/s.



(a) $A = 1, f = 1, \phi = 0$ **(b)** $A = 0.5, f = 1, \phi = 0$

(c) $A = 1, f = 2, \phi = 0$ **(d)** $A = 1, f = 1, \phi = \pi/4$

Figure 2.3 $s(t) = A \sin(2\pi ft + \phi)$

Frequency Domain Concepts

In practice, an electromagnetic signal will be made up of many frequencies. For example, the signal

$$s(t) = (4/\pi) \times (\sin(2\pi ft) + (1/3)\sin(2\pi(3f)t))$$

is shown in Figure 2.4c. The components of this signal are just sine waves of frequencies f and $3f$; parts (a) and (b) of the figure show these individual components. There are two interesting points that can be made about this figure:

- The second frequency is an integer multiple of the first frequency. When all of the frequency components of a signal are integer multiples of one frequency, the latter frequency is referred to as the **fundamental frequency**. The other components are called **harmonics**.
- The period of the total signal is equal to the period of the fundamental frequency. The period of the component $\sin(2\pi ft)$ is $T = 1/f$, and the period of $s(t)$ is also T, as can be seen from Figure 2.4c.

It can be shown, using a discipline known as Fourier analysis, that any signal is made up of components at various frequencies, in which each component is a sinusoid. By adding together enough sinusoidal signals, each with the appropriate

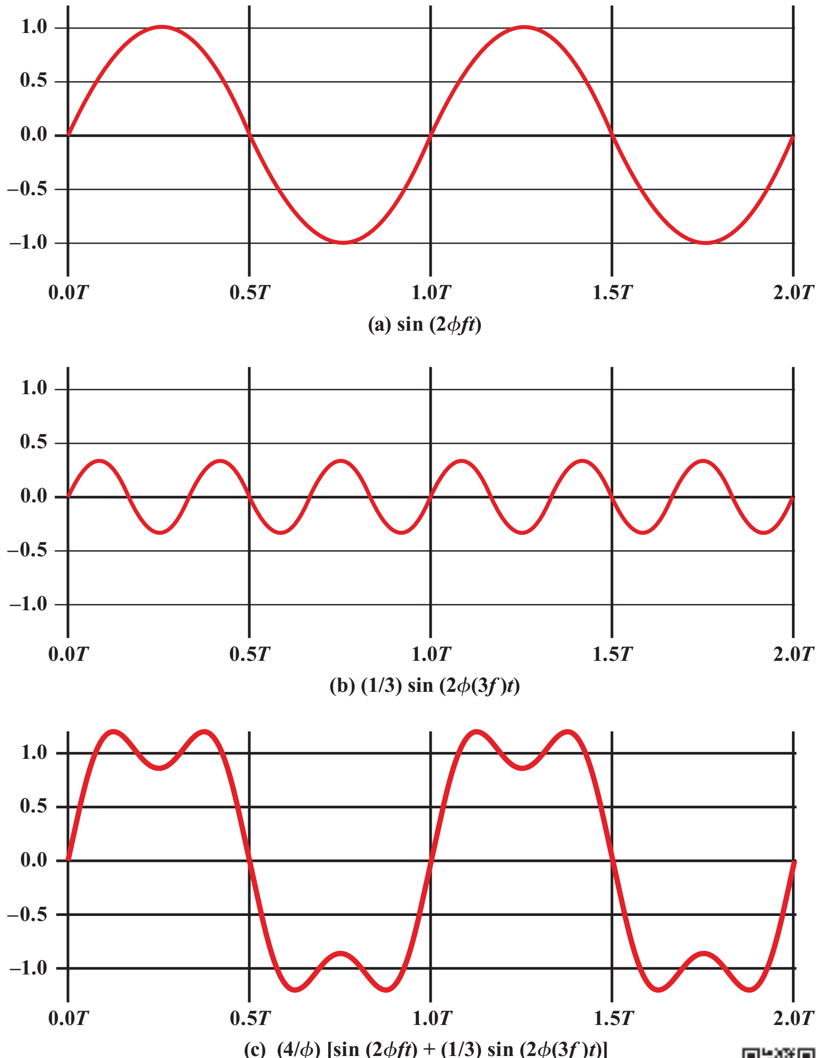


Figure 2.4 Addition of Frequency Components ($T = 1/f$)

amplitude, frequency, and phase, any electromagnetic signal can be constructed. Put another way, any electromagnetic signal can be shown to consist of a collection of periodic analog signals (sine waves) at different amplitudes, frequencies, and phases. The importance of being able to look at a signal from the frequency perspective (frequency domain) rather than a time perspective (time domain) should become clear as the discussion proceeds. For the interested reader, the subject of Fourier analysis is introduced in online Appendix B.

The **spectrum** of a signal is the range of frequencies that it contains. For the signal of Figure 2.4c, the spectrum extends from f to $3f$. The **absolute bandwidth** of a signal is the width of the spectrum. In the case of Figure 2.4c, the bandwidth is $3f - f = 2f$. Many signals have an infinite bandwidth, but with most of the energy contained in a relatively narrow band of frequencies. This band is referred to as the **effective bandwidth**, or just **bandwidth**.

Relationship between Data Rate and Bandwidth

There is a direct relationship between the information-carrying capacity of a signal and its bandwidth: The greater the bandwidth, the higher the information-carrying capacity. As a very simple example, consider the square wave of Figure 2.2b. Suppose that we let a positive pulse represent binary 0 and a negative pulse represent binary 1. Then the waveform represents the binary stream 0101. . . . The duration of each pulse is $1/(2f)$; thus the data rate is $2f$ bits per second (bps). What are the frequency components of this signal? To answer this question, consider again Figure 2.4. By adding together sine waves at frequencies f and $3f$, we get a waveform that begins to resemble the square wave. Let us continue this process by adding a sine wave of frequency $5f$, as shown in Figure 2.5a, and then adding a sine wave of frequency $7f$, as shown in Figure 2.5b. As we add additional odd multiples of f, suitably scaled, the resulting waveform approaches that of a square wave more and more closely.

Indeed, it can be shown that the frequency components of the square wave with amplitudes A and $-A$ can be expressed as follows:

$$s(t) = A \times \frac{4}{\pi} \times \sum_{k \text{ odd}, k=1}^{\infty} \frac{\sin(2\pi kft)}{k}$$

This waveform has an infinite number of frequency components and hence an infinite bandwidth. However, the peak amplitude of the kth frequency component, kf, is only $1/k$, so most of the energy in this waveform is in the first few frequency components. What happens if we limit the bandwidth to just the first three frequency components? We have already seen the answer, in Figure 2.5a. As we can see, the shape of the resulting waveform is reasonably close to that of the original square wave.

We can use Figures 2.4 and 2.5 to illustrate the relationship between data rate and bandwidth. Suppose that we are using a digital transmission system that is capable of transmitting signals with a bandwidth of 4 MHz. Let us attempt to transmit a sequence of alternating 0s and 1s as the square wave of Figure 2.5c. What data rate can be achieved? We look at three cases.

Case I. Let us approximate our square wave with the waveform of Figure 2.5a. Although this waveform is a “distorted” square wave, it is sufficiently close to the square wave that a receiver should be able to discriminate between a binary 0 and a binary 1. If we let $f = 10^6$ cycles/second $= 1$ MHz, then the bandwidth of the signal

$$s(t) = \frac{4}{\pi} \times \left[\sin((2\pi \times 10^6)t) + \frac{1}{3}\sin((2\pi \times 3 \times 10^6)t) + \frac{1}{5}\sin((2\pi \times 5 \times 10^6)t)\right]$$

is $(5 \times 10^6) - 10^6 = 4$ MHz. Note that for $f = 1$ MHz, the period of the fundamental frequency is $T = 1/10^6 = 10^{-6} = 1\ \mu s$. If we treat this waveform as a bit string of 1s and 0s, one bit occurs every $0.5\ \mu s$, for a data rate of $2 \times 10^6 = 2$ Mbps. Thus, for a bandwidth of 4 MHz, a data rate of 2 Mbps is achieved.

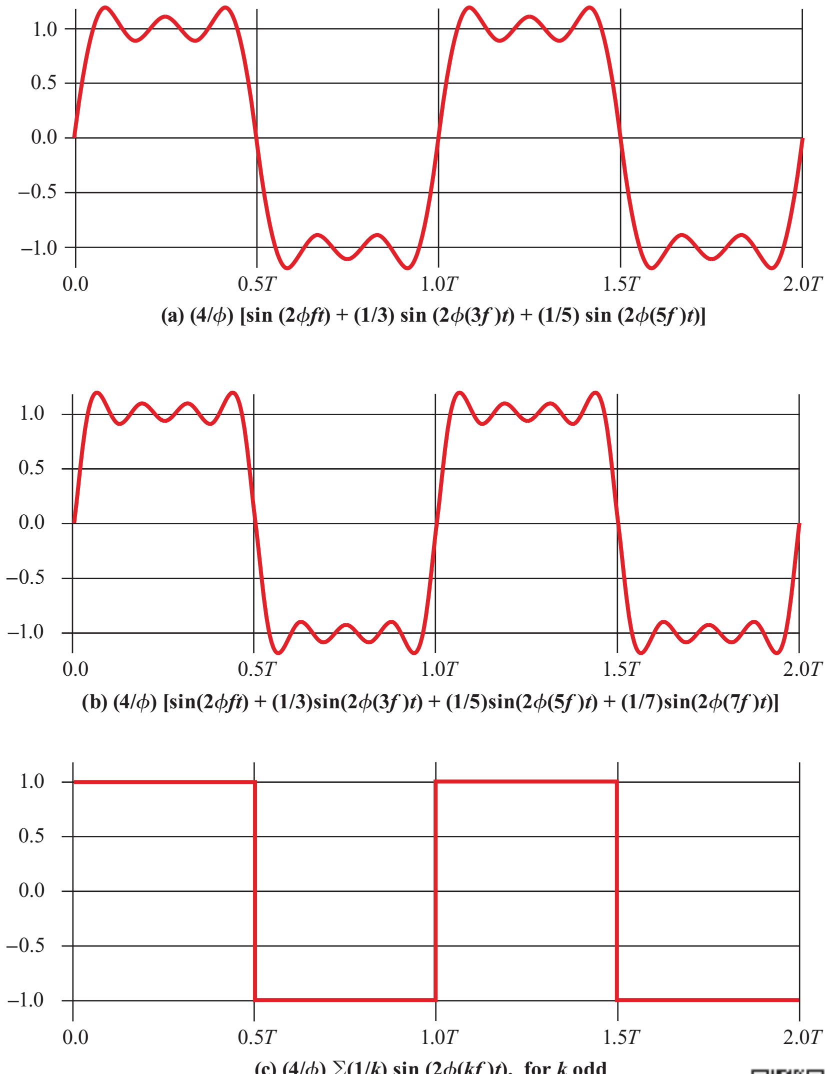


Figure 2.5 Frequency Components of Square Wave ($T = 1/f$)

Case II. Now suppose that we have a bandwidth of 8 MHz. Let us look again at Figure 2.5a, but now with $f = 2$ MHz. Using the same line of reasoning as before, the bandwidth of the signal is $(5 \times 2 \times 10^6) - (2 \times 10^6) = 8$ MHz. But in this case $T = 1/f = 0.5\,\mu$s. As a result, one bit occurs every 0.25 μs for a data rate of 4 Mbps. Thus, other things being equal, by doubling the bandwidth, we double the potential data rate.

Case III. Now suppose that the waveform of Figure 2.4c is considered adequate for approximating a square wave. That is, the difference between a positive and negative pulse in Figure 2.4c is sufficiently distinct that the waveform can be used successfully to represent a sequence of 1s and 0s. Assume as in Case II that

$f = 2$ MHz and $T = 1/f = 0.5\ \mu s$, so that one bit occurs every 0.25 μs for a data rate of 4 Mbps. Using the waveform of Figure 2.4c, the bandwidth of the signal is $(3 \times 2 \times 10^6) - (2 \times 10^6) = 4$ MHz. Thus, a given bandwidth can support various data rates depending on the ability of the receiver to discern the difference between 0 and 1 in the presence of noise and other impairments.

To summarize,

- **Case I:** Bandwidth = 4 MHz; data rate = 2 Mbps
- **Case II:** Bandwidth = 8 MHz; data rate = 4 Mbps
- **Case III:** Bandwidth = 4 MHz; data rate = 4 Mbps

We can draw the following conclusions from the preceding discussion. In general, any digital waveform using rectangular pulses will have infinite bandwidth. If we attempt to transmit this waveform as a signal over any medium, the transmission system will limit the bandwidth that can be transmitted. Furthermore, for any given medium, the greater the bandwidth transmitted, the greater the cost. Thus, on the one hand, economic and practical reasons dictate that digital information be approximated by a signal of limited bandwidth. On the other hand, limiting the bandwidth creates distortions, which makes the task of interpreting the received signal more difficult. The more limited the bandwidth, the greater the distortion and the greater the potential for error by the receiver.

2.2 ANALOG AND DIGITAL DATA TRANSMISSION

The terms *analog* and *digital* correspond, roughly, to *continuous* and *discrete*, respectively. These two terms are used frequently in data communications in at least three contexts: data, signals, and transmission.

Briefly, we define **data** as entities that convey meaning, or information. **Signals** are electric or electromagnetic representations of data. **Transmission** is the communication of data by the propagation and processing of signals. In what follows, we try to make these abstract concepts clear by discussing the terms *analog* and *digital* as applied to data, signals, and transmission.

Analog and Digital Data

The concepts of analog and digital data are simple enough. **Analog data** take on continuous values in some interval. For example, voice and video are continuously varying patterns of intensity. Most data collected by sensors, such as temperature and pressure, are continuous valued. **Digital data** take on discrete values; examples are text and integers.

The most familiar example of analog data is **audio**, which, in the form of acoustic sound waves, can be perceived directly by human beings. Figure 2.6 shows the acoustic spectrum for human speech and for music. Frequency components of typical speech may be found between approximately 100 Hz and 7 kHz. Although much of the energy in speech is concentrated at the lower frequencies, tests have shown that frequencies below 600 or 700 Hz add very little to the intelligibility of speech

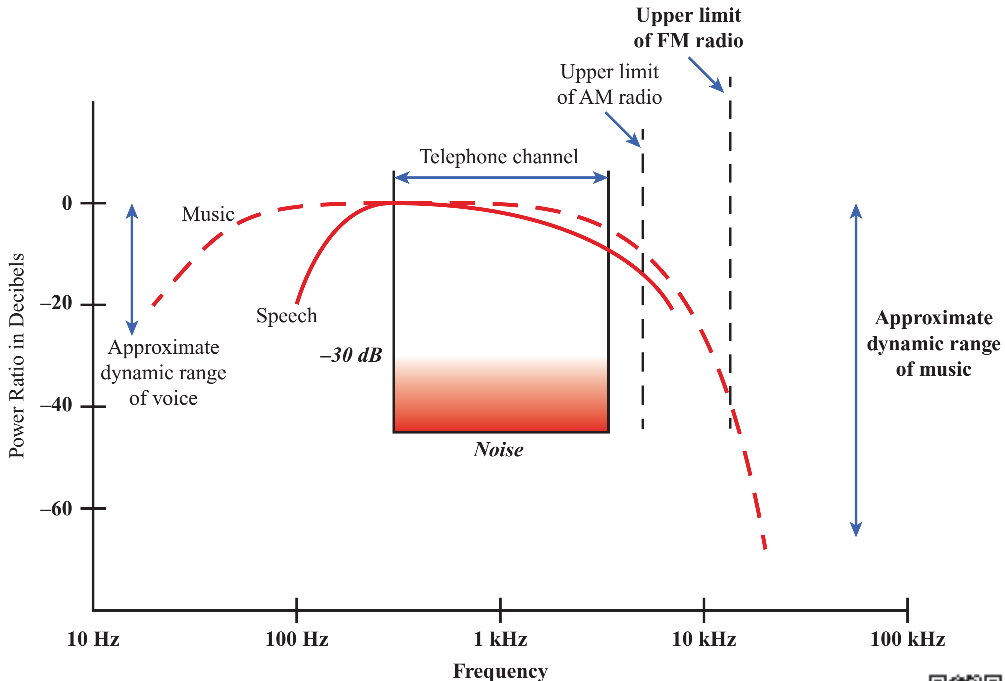


Figure 2.6 Acoustic Spectrum of Speech and Music

to the human ear. Typical speech has a dynamic range of about 25 dB[2]; that is, the power produced by the loudest shout may be as much as 300 times greater than that of the least whisper.

Analog and Digital Signaling

In a communications system, data are propagated from one point to another by means of electromagnetic signals. An **analog signal** is a continuously varying electromagnetic wave that may be propagated over a variety of media, depending on frequency; examples are copper wire media, such as twisted pair and coaxial cable; fiber optic cable; and atmosphere or space propagation (wireless). A **digital signal** is a sequence of voltage pulses that may be transmitted over a copper wire medium; for example, a constant positive voltage level may represent binary 0 and a constant negative voltage level may represent binary 1.

The principal advantages of digital signaling are that it is generally cheaper than analog signaling and is less susceptible to noise interference. The principal disadvantage is that digital signals suffer more from attenuation than do analog signals. Figure 2.7 shows a sequence of voltage pulses, generated by a source using two voltage levels, and the received voltage some distance down a conducting medium. Because of the attenuation, or reduction, of signal strength at higher frequencies, the pulses become rounded and smaller. It should be clear that this attenuation can lead rather quickly to the loss of the information contained in the propagated signal.

[2]The concept of decibels is explained in Appendix 2A.

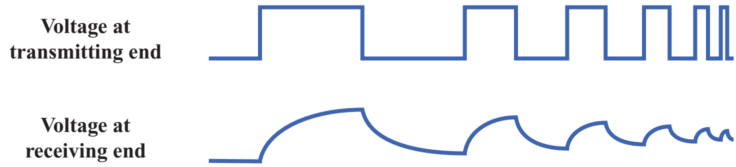


Figure 2.7 Attenuation of Digital Signals

Both analog and digital data can be represented, and hence propagated, by either analog or digital signals. This is illustrated in Figure 2.8. Generally, analog data are a function of time and occupy a limited frequency spectrum. Such data

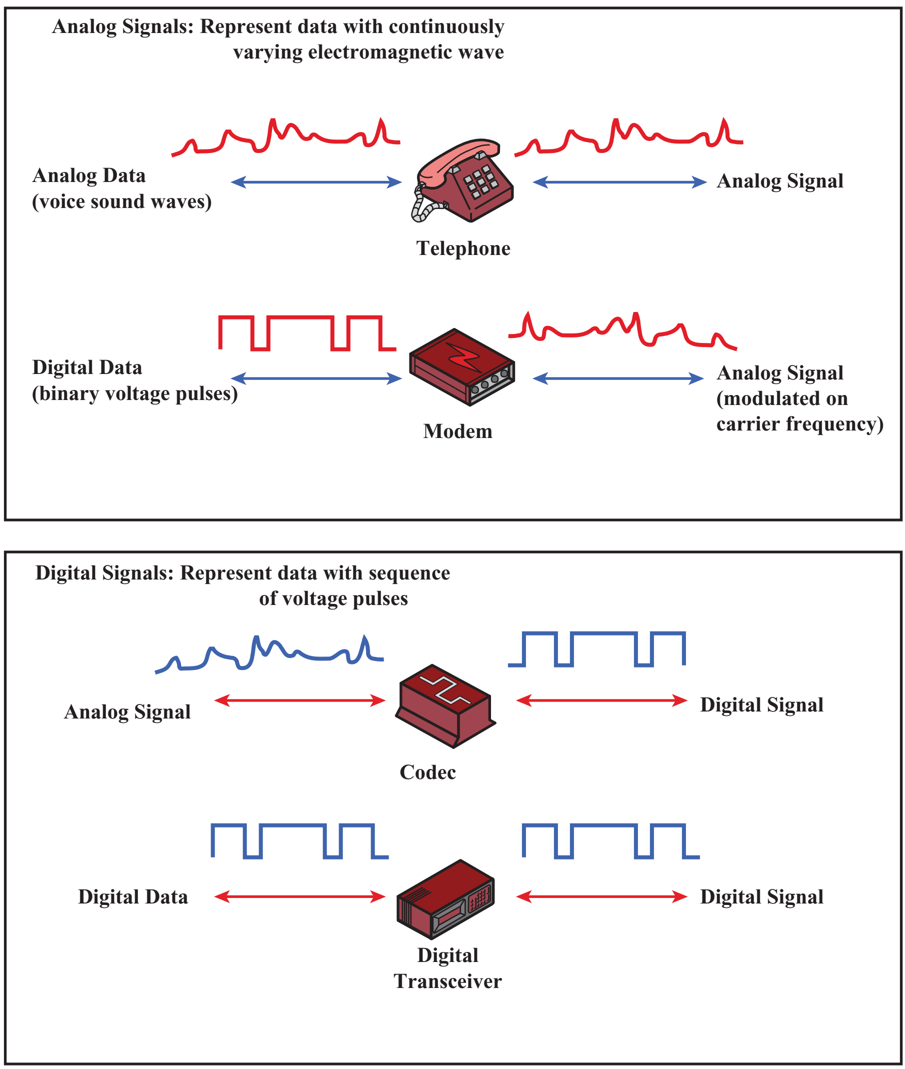


Figure 2.8 Analog and Digital Signaling of Analog and Digital Data

can be directly represented by an electromagnetic signal occupying the same spectrum. The best example of this is voice data. As sound waves, voice data have frequency components in the range 20 Hz to 20 kHz. As was mentioned and shown in Figure 2.6, most of the speech energy is in a much narrower range, with the typical speech range of between 100 Hz and 7 kHz. The standard spectrum of voice telephone signals is even narrower, at 300 to 3400 Hz, and this is quite adequate to propagate speech intelligibly and clearly. The telephone instrument does just that. For all sound input in the range of 300 to 3400 Hz, an electromagnetic signal with the same frequency–amplitude pattern is produced. The process is performed in reverse to convert the electromagnetic energy back into sound.

Digital data can also be represented by analog signals by use of a modem (modulator–demodulator). The modem converts a series of binary (two-valued) voltage pulses into an analog signal by modulating a carrier frequency. The resulting signal occupies a certain spectrum of frequency centered about the carrier and may be propagated across a medium suitable for that carrier. At the other end of the line, a modem demodulates the signal to recover the original data.

In an operation very similar to that performed by a modem, analog data can be represented by digital signals. The device that performs this function for voice data is a codec (coder–decoder). In essence, the codec takes an analog signal that directly represents the voice data and approximates that signal by a bit stream. At the other end of the line, a codec uses the bit stream to reconstruct the analog data. This topic is explored subsequently.

Finally, digital data can be represented directly, in binary form, by two voltage levels. To improve propagation characteristics, however, the binary data are often encoded into a more complex form of digital signal, as explained subsequently.

Each of the four combinations (Table 2.1a) just described is in widespread use. The reasons for choosing a particular combination for any given communications task vary. We list here some representative reasons:

- **Digital data, digital signal:** In general, the equipment for encoding digital data into a digital signal is less complex and less expensive than digital-to-analog equipment.
- **Analog data, digital signal:** Conversion of analog data to digital form permits the use of modern digital transmission and switching equipment for analog data.
- **Digital data, analog signal:** Some transmission media, such as optical fiber and satellite, will only propagate analog signals.
- **Analog data, analog signal:** Analog data are easily converted to an analog signal.

Analog and Digital Transmission

Both analog and digital signals may be transmitted on suitable transmission media. The way these signals are treated is a function of the transmission system. Table 2.1b summarizes the methods of data transmission. **Analog transmission** is a means of transmitting analog signals without regard to their content; the signals may represent analog data (e.g., voice) or digital data (e.g., data that pass through a modem). In either case, the analog signal will suffer attenuation that limits the length of the

Table 2.1 Analog and Digital Transmission

(a) Data and Signals

	Analog Signal	Digital Signal
Analog Data	Two alternatives: (1) signal occupies the same spectrum as the analog data; (2) analog data are encoded to occupy a different portion of spectrum.	Analog data are encoded using a codec to produce a digital bit stream.
Digital Data	Digital data are encoded using a modem to produce analog signal.	Two alternatives: (1) signal consists of two voltage levels to represent the two binary values; (2) digital data are encoded to produce a digital signal with desired properties.

(b) Treatment of Signals

	Analog Transmission	Digital Transmission
Analog Signal	Is propagated through amplifiers; same treatment whether signal is used to represent analog data or digital data.	Assumes that the analog signal represents digital data. Signal is propagated through repeaters; at each repeater, digital data are recovered from inbound signal and used to generate a new analog outbound signal.
Digital Signal	Not used	Digital signal represents a stream of 1s and 0s, which may represent digital data or may be an encoding of analog data. Signal is propagated through repeaters; at each repeater, stream of 1s and 0s is recovered from inbound signal and used to generate a new digital outbound signal.

transmission link. To achieve longer distances, the analog transmission system includes amplifiers that boost the energy in the signal. Unfortunately, the amplifier also boosts the noise components. With amplifiers cascaded to achieve long distance, the signal becomes more and more distorted. For analog data, such as voice, quite a bit of distortion can be tolerated and the data remain intelligible. However, for digital data transmitted as analog signals, cascaded amplifiers will introduce errors.

Digital transmission, in contrast, is concerned with the content of the signal. We have mentioned that a digital signal can be propagated only a limited distance before attenuation endangers the integrity of the data. To achieve greater distances, repeaters are used. A repeater receives the digital signal, recovers the pattern of ones and zeros, and retransmits a new signal. Thus, the attenuation is overcome.

The same technique may be used with an analog signal if the signal carries digital data. At appropriately spaced points, the transmission system has retransmission devices rather than amplifiers. The retransmission device recovers the digital data from the analog signal and generates a new, clean analog signal. Thus, noise is not cumulative.

2.3 CHANNEL CAPACITY

A variety of impairments can distort or corrupt a signal. A common impairment is **noise**, which is any unwanted signal that combines with and hence distorts the signal intended for transmission and reception. Noise and other impairments are discussed in Chapter 6. For the purposes of this section, we simply need to know that noise is something that degrades signal quality. For digital data, the question that then arises is to what extent these impairments limit the data rate that can be achieved. The maximum rate at which data can be transmitted over a given communication path, or channel, under given conditions is referred to as the **channel capacity**.

There are four concepts here that we are trying to relate to one another:

- **Data rate:** This is the rate, in bits per second (bps), at which data can be communicated.
- **Bandwidth:** This is the bandwidth of the transmitted signal as constrained by the transmitter and the nature of the transmission medium, expressed in cycles per second, or Hertz.
- **Noise:** For this discussion, we are concerned with the average level of noise over the communications path.
- **Error rate:** This is the rate at which errors occur, where an error is the reception of a 1 when a 0 was transmitted or the reception of a 0 when a 1 was transmitted.

The problem we are addressing is this: Communications facilities are expensive and, in general, the greater the bandwidth of a facility, the greater the cost. Furthermore, all transmission channels of any practical interest are of limited bandwidth. The limitations arise from the physical properties of the transmission medium or from deliberate limitations at the transmitter on the bandwidth to prevent interference from other sources. Accordingly, we would like to make as efficient use as possible of a given bandwidth. For digital data, this means that we would like to get as high a data rate as possible at a particular limit of error rate for a given bandwidth. The main constraint on achieving this efficiency is noise.

Nyquist Bandwidth

To begin, let us consider the case of a channel that is noise free. In this environment, the limitation on data rate is simply the bandwidth of the signal. A formulation of this limitation, due to Nyquist, states that if the rate of signal transmission is $2B$, then a signal with frequencies no greater than B is sufficient to carry the signal rate. The converse is also true: Given a bandwidth of B, the highest signal rate that can be carried is $2B$. This limitation is due to the effect of intersymbol interference, such as is produced by delay distortion.[3] The result is useful in the development of digital-to-analog encoding schemes.

Note that in the preceding paragraph, we referred to signal rate. If the signals to be transmitted are binary (take on only two values), then the data rate that can be

[3]Delay distortion of a signal occurs when the propagation delay for the transmission medium is not constant over the frequency range of the signal.

supported by B Hz is $2B$ bps. As an example, consider a voice channel being used, via modem, to transmit digital data. Assume a bandwidth of 3100 Hz. Then the capacity, C, of the channel is $2B = 6200$ bps. However, as we shall see in Chapter 7, signals with more than two levels can be used; that is, each signal element can represent more than one bit. For example, if four possible voltage levels are used as signals, then each signal element can represent two bits. With multilevel signaling, the Nyquist formulation becomes

$$C = 2B \log_2 M$$

where M is the number of discrete signal elements or voltage levels. Thus, for $M = 8$, a value used with some modems, a bandwidth of $B = 3100$ Hz yields a capacity $C = 18{,}600$ bps.

So, for a given bandwidth, the data rate can be increased by increasing the number of different signal elements. However, this places an increased burden on the receiver: Instead of distinguishing one of two possible signal elements during each signal time, it must distinguish one of M possible signals. Noise and other impairments on the transmission line will limit the practical value of M.

Shannon Capacity Formula

Nyquist's formula indicates that, all other things being equal, doubling the bandwidth doubles the data rate. Now consider the relationship among data rate, noise, and error rate. The presence of noise can corrupt one or more bits. If the data rate is increased, then the bits become "shorter" in time, so that more bits are affected by a given pattern of noise. Thus, at a given noise level, the higher the data rate, the higher the error rate.

Figure 2.9 is an example of the effect of noise on a digital signal. Here the noise consists of a relatively modest level of background noise plus occasional larger spikes of noise. The digital data can be recovered from the signal by sampling the received waveform once per bit time. As can be seen, the noise is occasionally sufficient to change a 1 to a 0 or a 0 to a 1.

All of these concepts can be tied together neatly in a formula developed by the mathematician Claude Shannon. As we have just illustrated, the higher the data rate, the more damage that unwanted noise can do. For a given level of noise, we would expect that a greater signal strength would improve the ability to receive data correctly in the presence of noise. The key parameter involved in this reasoning is the signal-to-noise ratio (SNR, or S/N),[4] which is the ratio of the power in a signal to the power contained in the noise that is present at a particular point in the transmission. Typically, this ratio is measured at a receiver, because it is at this point that an attempt is made to process the signal and eliminate the unwanted noise. For convenience, this ratio is often reported in **decibels (dB)**:

$$\mathrm{SNR_{dB}} = 10 \log_{10} \frac{\text{signal power}}{\text{noise power}}$$

[4]Some of the literature uses SNR; others use S/N. Also, in some cases the dimensionless quantity is referred to as SNR or S/N and the quantity in decibels is referred to as $\mathrm{SNR_{db}}$ or $(\mathrm{S/N})_{\mathrm{db}}$. Others use just SNR or S/N to mean the dB quantity. This text uses SNR and $\mathrm{SNR_{db}}$.

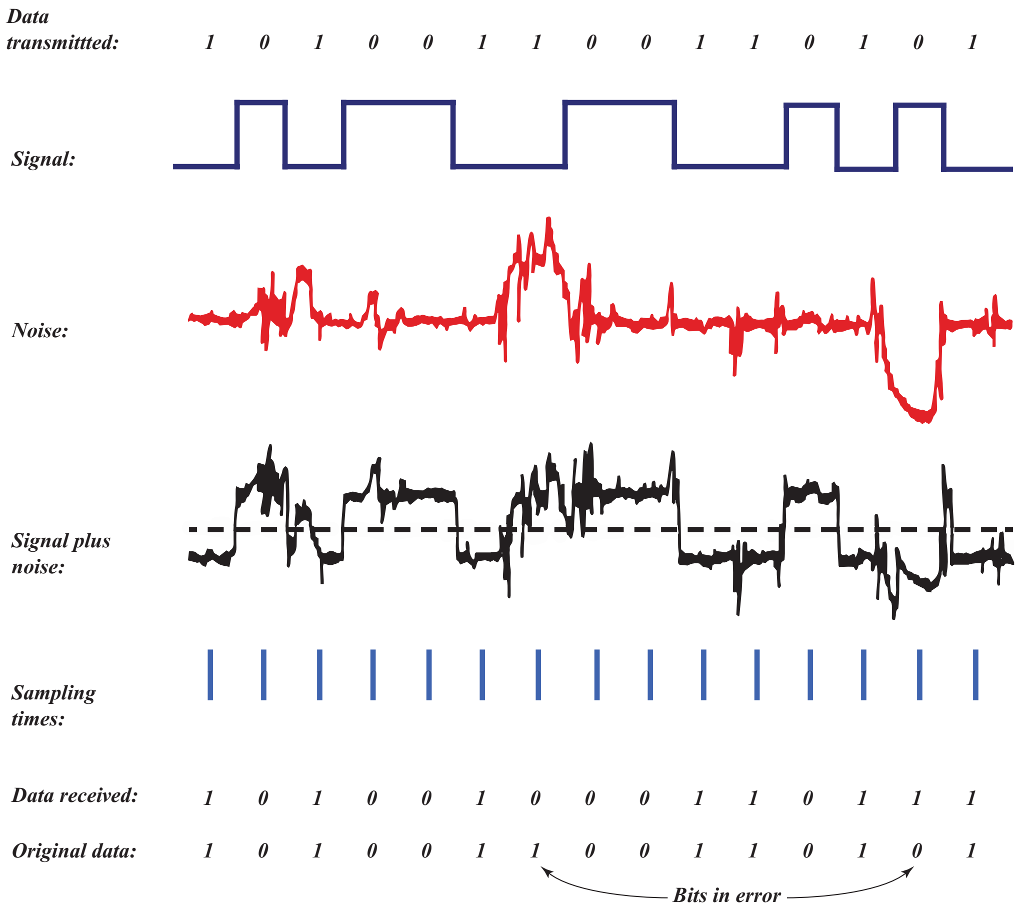


Figure 2.9 Effect of Noise on a Digital Signal

This expresses the amount, in decibels, that the intended signal exceeds the noise level. A high SNR will mean a high-quality signal.

The signal-to-noise ratio is important in the transmission of digital data because it sets the upper bound on the achievable data rate. Shannon's result is that the maximum channel capacity, in bits per second, obeys the equation

$$C = B \log_2(1 + \text{SNR})$$

where C is the capacity of the channel in bits per second and B is the bandwidth of the channel in Hertz. The Shannon formula represents the theoretical maximum that can be achieved. In practice, however, only much lower rates are achieved. One reason for this is that the formula assumes white noise (thermal noise). Impulse noise is not accounted for, nor are attenuation distortion or delay distortion. Various types of noise and distortion are discussed in Chapter 6.

The capacity indicated in the preceding equation is referred to as the error-free capacity. Shannon proved that if the actual information rate on a channel is less

Example 2.1 Let us consider an example that relates the Nyquist and Shannon formulations. Suppose that the spectrum of a channel is between 3 MHz and 4 MHz and $\text{SNR}_{\text{dB}} = 24\text{ dB}$. Then

$$B = 4\text{ MHz} - 3\text{ MHz} = 1\text{ MHz}$$
$$\text{SNR}_{\text{dB}} = 24\text{ dB} = 10\log_{10}(\text{SNR})$$
$$\text{SNR} = 251$$

Using Shannon's formula,

$$C = 10^6 \times \log_2(1 + 251) \approx 10^6 \times 8 = 8\text{ Mbps}$$

This is a theoretical limit and, as we have said, is unlikely to be reached. But assume we can achieve the limit. Based on Nyquist's formula, how many signaling levels are required? We have

$$C = 2B \log_2 M$$
$$8 \times 10^6 = 2 \times (10^6) \times \log_2 M$$
$$4 = \log_2 M$$
$$M = 16$$

than the error-free capacity, then it is theoretically possible to use a suitable signal code to achieve error-free transmission through the channel. Shannon's theorem unfortunately does not suggest a means for finding such codes, but it does provide a yardstick by which the performance of practical communication schemes may be measured.

Several other observations concerning the preceding equation may be instructive. For a given level of noise, it would appear that the data rate could be increased by increasing either signal strength or bandwidth. However, as the signal strength increases, so do the effects of nonlinearities in the system, leading to an increase in intermodulation noise. Note also that, because noise is assumed to be white, the wider the bandwidth, the more noise is admitted to the system. Thus, as B increases, SNR decreases.

2.4 TRANSMISSION MEDIA

In a data transmission system, the **transmission medium** is the physical path between transmitter and receiver. Transmission media can be classified as guided or unguided. In both cases, communication is in the form of electromagnetic waves. With **guided media**, the waves are guided along a solid medium, such as copper twisted pair, copper coaxial cable, or optical fiber. The atmosphere and outer space are examples of **unguided media**, which provide a means of transmitting electromagnetic signals but do not guide them; this form of transmission is usually referred to as **wireless transmission**.

The characteristics and quality of a data transmission are determined both by the characteristics of the medium and the characteristics of the signal. In the case of guided media, the medium itself is usually more important in determining the limitations of transmission. For unguided media, the bandwidth of the signal produced by the transmitting antenna is usually more important than the medium in determining transmission characteristics. One key property of signals transmitted by antenna is directionality. In general, signals at lower frequencies are omnidirectional; that is, the signal propagates in all directions from the antenna. At higher frequencies, it is possible to focus the signal into a directional beam.

Figure 2.10 depicts the electromagnetic spectrum and indicates the frequencies at which various guided media and unguided transmission techniques operate. In the remainder of this section, we provide a brief overview of unguided, or wireless, media.

For unguided media, transmission and reception are achieved by means of an antenna. For transmission, the antenna radiates electromagnetic energy into the medium (usually air), and for reception, the antenna picks up electromagnetic waves from the surrounding medium. There are basically two types of configurations for wireless transmission: directional and omnidirectional. For the directional configuration, the transmitting antenna puts out a focused electromagnetic beam; the transmitting and receiving antennas must therefore be carefully aligned. In the omnidirectional case, the transmitted signal spreads out in all directions and can be received by many antennas.

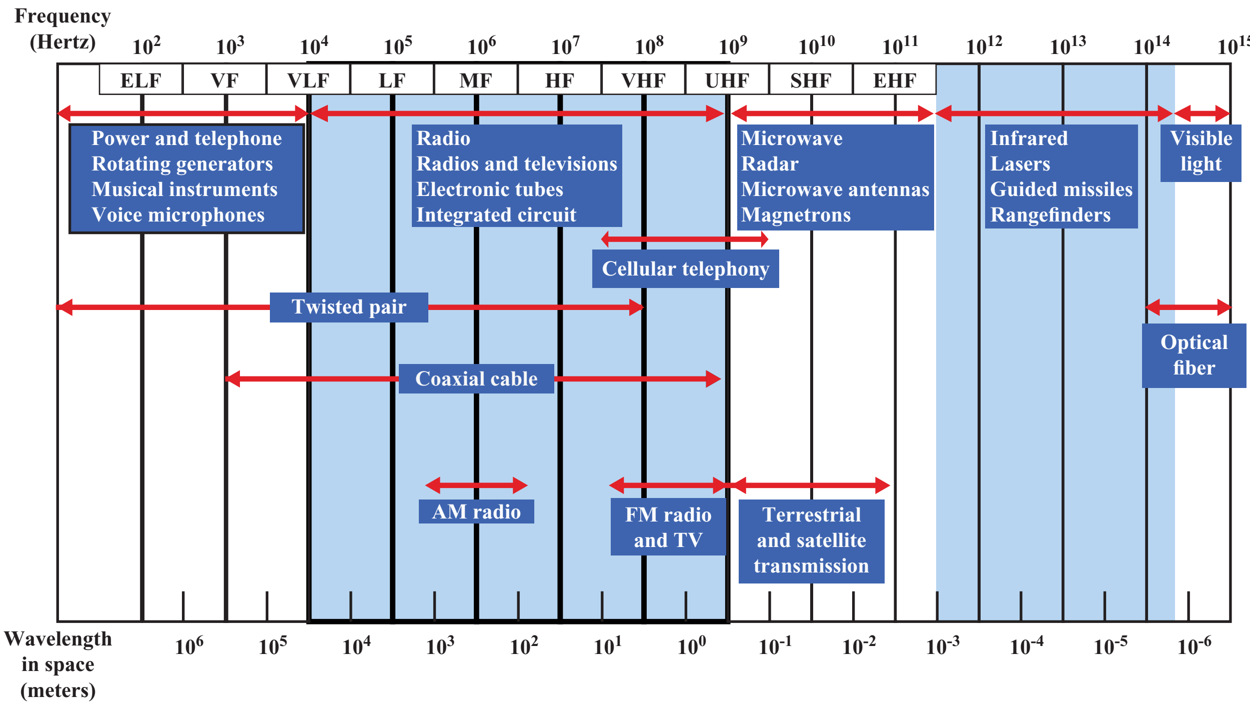
Three general ranges of frequencies are of interest in our discussion of wireless transmission. Frequencies in the range of about 1 GHz (gigahertz $= 10^9$Hz) to 100 GHz are referred to as **microwave frequencies**. At these frequencies, highly directional beams are possible, and microwave is quite suitable for point-to-point transmission. Microwave is also used for satellite communications. Frequencies in the range 30 MHz to 1 GHz are suitable for omnidirectional applications. We refer to this range as the radio range.

Another important frequency range, for local applications, is the infrared portion of the spectrum. This covers, roughly, from 3×10^{11} to 2×10^{14}Hz. Infrared is useful in local point-to-point and multipoint applications within confined areas, such as a single room.

Terrestrial Microwave

Physical Description The most common type of microwave antenna is the parabolic "dish." A typical size is about 3 m in diameter. The antenna is fixed rigidly and focuses a narrow beam to achieve line-of-sight transmission to the receiving antenna. Microwave antennas are usually located at substantial heights above ground level to extend the range between antennas and to be able to transmit over intervening obstacles. To achieve long-distance transmission, a series of microwave relay towers is used, and point-to-point microwave links are strung together over the desired distance.

Applications A primary use for **terrestrial microwave** systems is in long-haul telecommunications service, as an alternative to coaxial cable or optical fiber. The microwave facility requires far fewer amplifiers or repeaters than coaxial cable over



ELF = Extremely low frequency
VF = Voice frequency
VLF = Very low frequency
LF = Low frequency
MF = Medium frequency
HF = High frequency
VHF = Very high frequency
UHF = Ultrahigh frequency
SHF = Superhigh frequency
EHF = Extremely high frequency

Figure 2.10 Electromagnetic Spectrum for Telecommunications the same distance but requires line-of-sight transmission. Microwave is commonly used for both voice and television transmission.

Another increasingly common use of microwave is for short point-to-point links between buildings. This can be used for closed-circuit TV or as a data link between local area networks. Short-haul microwave can also be used for the so-called bypass application. A business can establish a microwave link to a long-distance telecommunications facility in the same city, bypassing the local telephone company.

Two other important uses of microwave are examined in some detail in Part Three: cellular systems and fixed wireless access.

Transmission Characteristics Microwave transmission covers a substantial portion of the electromagnetic spectrum. Common frequencies used for transmission are in the range 2 to 40 GHz. The higher the frequency used, the higher the potential bandwidth and therefore the higher the potential data rate. Table 2.2 indicates bandwidth and data rate for some typical systems.

As with any transmission system, a main source of loss is attenuation. For microwave (and radio frequencies), the loss can be expressed as

$$L = 10 \log \left(\frac{4\pi d}{\lambda} \right)^2 \text{dB} \tag{2.2}$$

where d is the distance and λ is the wavelength, in the same units. Thus, loss varies as the square of the distance. In contrast, for twisted pair and coaxial cable, loss varies exponentially with distance (linear in decibels). Thus repeaters or amplifiers may be placed farther apart for microwave systems—10 to 100 km is typical. Attenuation is increased with rainfall. The effects of rainfall become especially noticeable above 10 GHz. As frequency increases, λ decreases and loss increases. Another source of impairment is interference. With the growing popularity of microwave, transmission areas overlap and interference is always a danger. Thus the assignment of frequency bands is strictly regulated.

The most common bands for long-haul telecommunications are the 4- to 6-GHz bands. With increasing congestion at these frequencies, the 11-GHz band is now coming into use. The 12-GHz band is used as a component of cable TV systems. Microwave links are used to provide TV signals to local CATV installations; the signals are then distributed to individual subscribers via coaxial cable. Higher-frequency microwave is being used for short point-to-point links between buildings; typically, the 22-GHz band is used. The higher microwave frequencies are less useful for longer

Table 2.2 Typical Digital Microwave Performance

Band (GHz)	Bandwidth (MHz)	Data Rate (Mbps)
2	7	12
6	30	90
11	40	135
18	220	274

distances because of increased attenuation but are quite adequate for shorter distances. In addition, at the higher frequencies, the antennas are smaller and cheaper.

Satellite Microwave

Physical Description A communication satellite is, in effect, a microwave relay station. It is used to link two or more ground-based microwave transmitter/receivers, known as earth stations, or ground stations. The satellite receives transmissions on one frequency band (uplink), amplifies or repeats the signal, and transmits it on another frequency (downlink). A single orbiting satellite will operate on a number of frequency bands, called *transponder channels*, or simply *transponders*.

Applications The communication satellite is a technological revolution as important as fiber optics. Among the most important applications for satellites are

- Television distribution
- Long-distance telephone transmission
- Private business networks

Because of their broadcast nature, satellites are well suited to television distribution and are being used extensively in the United States and throughout the world for this purpose. In its traditional use, a network provides programming from a central location. Programs are transmitted to the satellite and then broadcast down to a number of stations, which then distribute the programs to individual viewers. One network, the Public Broadcasting Service (PBS), distributes its television programming almost exclusively by the use of satellite channels. Other commercial networks also make substantial use of satellite, and cable television systems are receiving an ever-increasing proportion of their programming from satellites. The most recent application of satellite technology to television distribution is direct broadcast satellite (DBS), in which satellite video signals are transmitted directly to the home user. The dropping cost and size of receiving antennas have made DBS economically feasible, and DBS is now commonplace. DBS is discussed in Chapter 16.

Satellite transmission is also used for point-to-point trunks between telephone exchange offices in public telephone networks. It is the optimum medium for high-usage international trunks and is competitive with terrestrial systems for many long-distance intranational links.

Finally, there are a number of business data applications for satellite. The satellite provider can divide the total capacity into a number of channels and lease these channels to individual business users. A user equipped with antennas at a number of sites can use a satellite channel for a private network. Traditionally, such applications have been quite expensive and limited to larger organizations with high-volume requirements.

Transmission Characteristics The optimum frequency range for satellite transmission is in the range 1 to 10 GHz. Below 1 GHz, there is significant noise from natural sources, including galactic, solar, and atmospheric noise, and human-made interference from various electronic devices. Above 10 GHz, the signal is severely attenuated by atmospheric absorption and precipitation.

Most satellites providing point-to-point service today use a frequency bandwidth in the range 5.925 to 6.425 GHz for transmission from earth to satellite (uplink) and a bandwidth in the range 3.7 to 4.2 GHz for transmission from satellite to earth (downlink). This combination is referred to as the 4/6-GHz band. Note that the uplink and downlink frequencies differ. For continuous operation without interference, a satellite cannot transmit and receive on the same frequency. Thus signals received from a ground station on one frequency must be transmitted back on another.

The 4/6-GHz band is within the optimum zone of 1 to 10 GHz but has become saturated. Other frequencies in that range are unavailable because of sources of interference operating at those frequencies, usually terrestrial microwave. Therefore, the 12/14-GHz band has been developed (uplink: 14 to 14.5 GHz; downlink: 11.7 to 12.2 GHz). At this frequency band, attenuation problems must be overcome. However, smaller and cheaper earth-station receivers can be used. It is anticipated that this band will also saturate, and use is projected for the 20/30-GHz band (uplink: 27.5 to 30.0 GHz; downlink: 17.7 to 20.2 GHz). This band experiences even greater attenuation problems but will allow greater bandwidth (2500 MHz versus 500 MHz) and even smaller and cheaper receivers.

Several properties of satellite communication should be noted. First, because of the long distances involved, there is a propagation delay of about a quarter second from transmission from one earth station to reception by another earth station. This delay is noticeable in ordinary telephone conversations. It also introduces problems in the areas of error control and flow control, which we discuss in later chapters. Second, **satellite microwave** is inherently a broadcast facility. Many stations can transmit to the satellite, and a transmission from a satellite can be received by many stations.

Broadcast Radio

Physical Description The principal difference between broadcast radio and microwave is that the former is omnidirectional and the latter is directional. Thus broadcast radio does not require dish-shaped antennas, and the antennas need not be rigidly mounted to a precise alignment.

Applications ***Radio*** is a general term used to encompass frequencies in the range of 3 kHz to 300 GHz. We are using the informal term ***broadcast radio*** to cover the VHF and part of the UHF band: 30 MHz to 1 GHz. This range covers FM radio and VHF television. This range is also used for a number of data networking applications.

Transmission Characteristics The range 30 MHz to 1 GHz is an effective one for broadcast communications. Unlike the case for lower-frequency electromagnetic waves, the ionosphere is transparent to radio waves above 30 MHz. Thus transmission is limited to the line of sight, and distant transmitters will not interfere with each other due to reflection from the atmosphere. Unlike the higher frequencies of the microwave region, broadcast radio waves are less sensitive to attenuation from rainfall.

As with microwave, the amount of attenuation due to distance for radio obeys Equation (2.2), namely $10 \log\left(\frac{4\pi d}{\lambda}\right)^2$ dB. Because of the longer wavelength, radio waves suffer relatively less attenuation.

A prime source of impairment for broadcast radio waves is multipath interference. Reflection from land, water, and natural or human-made objects can create multiple paths between antennas. This effect is frequently evident when TV reception displays multiple images as an airplane passes by.

Infrared

Infrared communications is achieved using transmitters/receivers (transceivers) that modulate noncoherent infrared light. Transceivers must be within the line of sight of each other either directly or via reflection from a light-colored surface such as the ceiling of a room.

One important difference between infrared and microwave transmission is that the former does not penetrate walls. Thus the security and interference problems encountered in microwave systems are not present. Furthermore, there is no frequency allocation issue with infrared, because no licensing is required.

2.5 MULTIPLEXING

In both local and wide area communications, it is almost always the case that the capacity of the transmission medium exceeds the capacity required for the transmission of a single signal. To make efficient use of the transmission system, it is desirable to carry multiple signals on a single medium. This is referred to as ***multiplexing***.

Figure 2.11 depicts the multiplexing function in its simplest form. There are n inputs to a multiplexer. The multiplexer is connected by a single data link to a demultiplexer. The link is able to carry n separate channels of data. The multiplexer combines (multiplexes) data from the n input lines and transmits over a higher-capacity data link. The demultiplexer accepts the multiplexed data stream, separates (demultiplexes) the data according to channel, and delivers them to the appropriate output lines.

The widespread use of multiplexing in data communications can be explained by the following:

1. The higher the data rate, the more cost-effective the transmission facility. That is, for a given application and over a given distance, the cost per kbps declines with an increase in the data rate of the transmission facility. Similarly, the cost

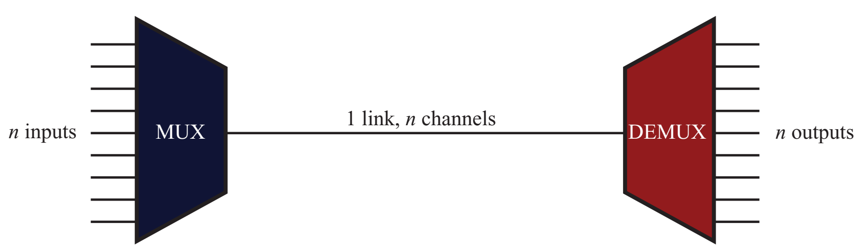


Figure 2.11 Multiplexing

of transmission and receiving equipment, per kbps, declines with increasing data rate.

2. Most individual data communicating devices require relatively modest data rate support.

The preceding statements were phrased in terms of data communicating devices. Similar statements apply to voice communications. That is, the greater the capacity of a transmission facility, in terms of voice channels, the less the cost per individual voice channel, and the capacity required for a single voice channel is modest.

Two techniques for multiplexing in telecommunications networks are in common use: **frequency division multiplexing (FDM)** and **time division multiplexing (TDM)**.

FDM takes advantage of the fact that the useful bandwidth of the medium exceeds the required bandwidth of a given signal. A number of signals can be carried simultaneously if each signal is modulated onto a different carrier frequency and the carrier frequencies are sufficiently separated so that the bandwidths of the signals do not overlap. Figure 2.12a depicts a simple case. Six signal sources are fed into a multiplexer that modulates each signal onto a different frequency ($f_1, \ldots, f_6$). Each signal requires a certain bandwidth centered on its carrier frequency, referred to as a **channel**. To prevent interference, the channels are separated by **guard bands**, which are unused portions of the spectrum (not shown in the figure).

An example is the multiplexing of voice signals. We mentioned that the useful spectrum for voice is 300 to 3400 Hz. Thus, a bandwidth of 4 kHz is adequate to carry the voice signal and provide a guard band. For both North America (AT&T standard) and internationally (International Telecommunication Union Telecommunication Standardization Sector [ITU-T] standard), a standard voice multiplexing scheme is twelve 4-kHz voice channels from 60 to 108 kHz. For higher-capacity links, both AT&T and ITU-T define larger groupings of 4-kHz channels.

TDM takes advantage of the fact that the achievable bit rate (sometimes, unfortunately, called bandwidth) of the medium exceeds the required data rate of a digital signal. Multiple digital signals can be carried on a single transmission path by interleaving portions of each signal in time. The interleaving can be at the bit level or in blocks of bytes or larger quantities. For example, the multiplexer in Figure 2.12b has six inputs that might each be, say, 9.6 kbps. A single line with a capacity of 57.6 kbps could accommodate all six sources. Analogously to FDM, the sequence of time slots dedicated to a particular source is called a channel. One cycle of time slots (one per source) is called a frame.

The TDM scheme depicted in Figure 2.12b is also known as **synchronous TDM**, referring to the fact that time slots are preassigned and fixed. Hence the timing of transmission from the various sources is synchronized. In contrast, asynchronous TDM allows time on the medium to be allocated dynamically. Unless otherwise noted, the term *TDM* will be used to mean synchronous TDM.

A generic depiction of a synchronous TDM system is provided in Figure 2.13. A number of signals [$m_i(t), i = 1, n$] are to be multiplexed onto the same transmission medium. The signals carry digital data and are generally digital signals. The incoming data from each source are briefly buffered. Each buffer is typically one bit or one character in length. The buffers are scanned sequentially to form a composite

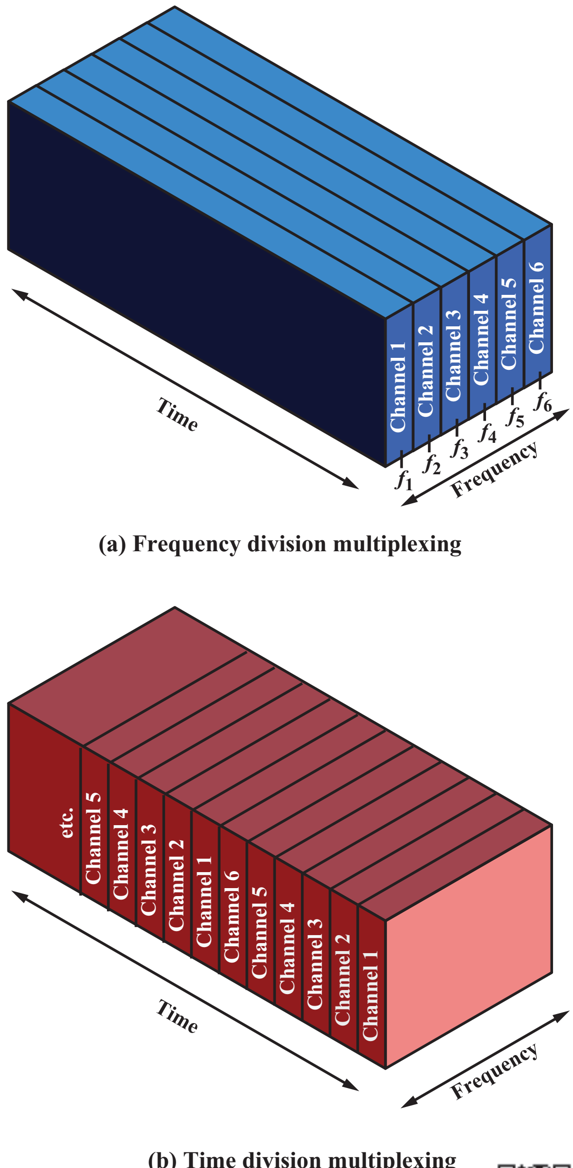


Figure 2.12 FDM and TDM

digital data stream $m_c(t)$. The scan operation is sufficiently rapid so that each buffer is emptied before more data can arrive. Thus, the data rate of $m_c(t)$ must at least equal the sum of the data rates of the $m_i(t)$. The digital signal $m_c(t)$ may be transmitted directly, or passed through a modem so that an analog signal is transmitted. In either case, transmission is typically synchronous.

The transmitted data may have a format something like Figure 2.13b. The data are organized into frames. Each frame contains a cycle of time slots. In each frame,

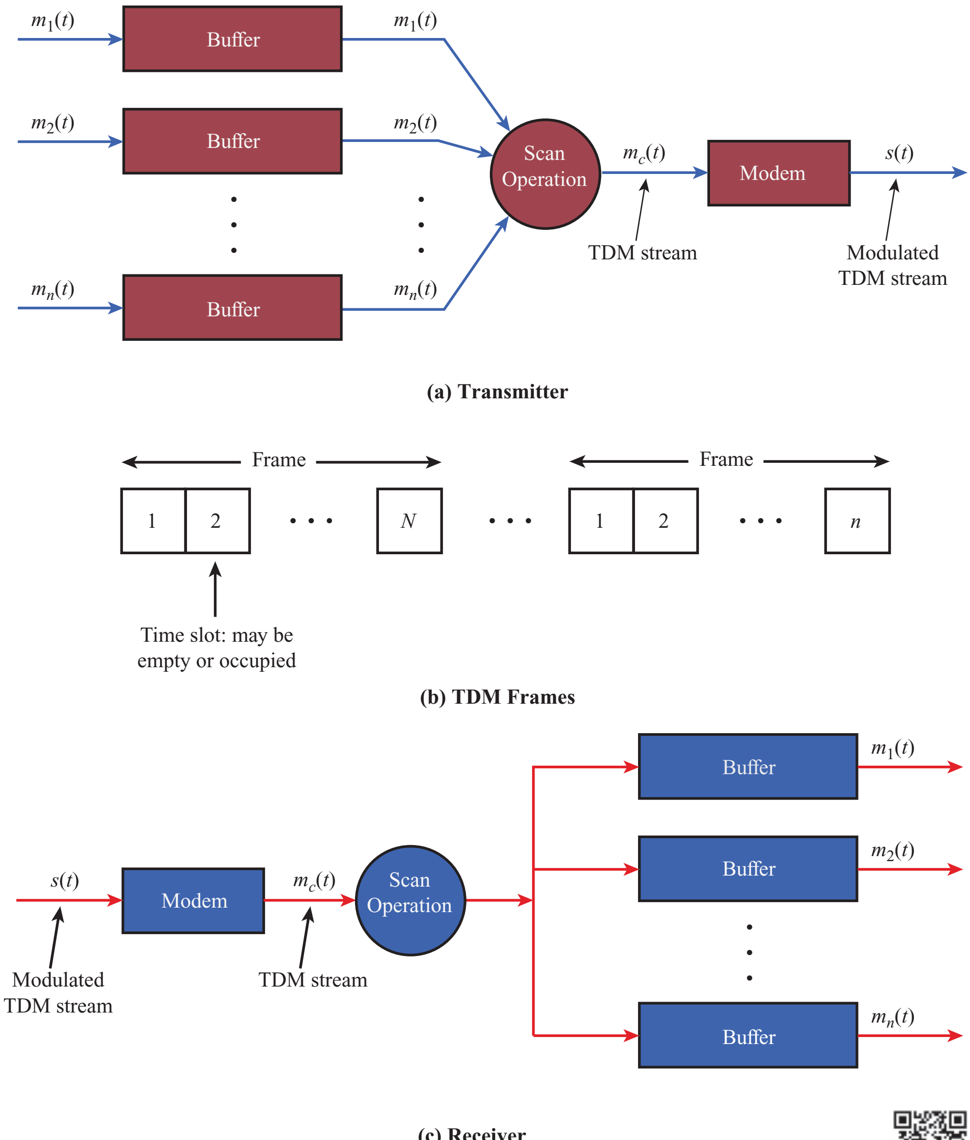


Figure 2.13 Synchronous TDM System

one or more slots are dedicated to each data source. The sequence of slots dedicated to one source, from frame to frame, is called a channel. The slot length equals the transmitter buffer length, typically a bit or a byte (character).

The byte-interleaving technique is used with asynchronous and synchronous sources. Each time slot contains one character of data. Typically, the start and stop bits of each character are eliminated before transmission and reinserted by the receiver, thus improving efficiency. The bit-interleaving technique is used with synchronous sources and may also be used with asynchronous sources. Each time slot contains just one bit.

At the receiver, the interleaved data are demultiplexed and routed to the appropriate destination buffer. For each input source $m_i(t)$, there is an identical output source that will receive the input data at the same rate at which it was generated.

Synchronous TDM is called synchronous not because synchronous transmission is used but because the time slots are preassigned to sources and fixed. The time slots for each source are transmitted whether or not the source has data to send. This is, of course, also the case with FDM. In both cases, capacity is wasted to achieve simplicity of implementation. Even when fixed assignment is used, however, it is possible for a synchronous TDM device to handle sources of different data rates. For example, the slowest input device could be assigned one slot per cycle, while faster devices are assigned multiple slots per cycle.

One example of TDM is the standard scheme used for transmitting PCM voice data, known in AT&T parlance as T1 carrier. Data are taken from each source, one sample (7 bits) at a time. An eighth bit is added for signaling and supervisory functions. For T1, 24 sources are multiplexed, so there are $8 \times 24 = 192$ bits of data and control signals per frame. One final bit is added for establishing and maintaining synchronization. Thus a frame consists of 193 bits and contains one 7-bit sample per source. Since sources must be sampled 8000 times per second, the required data rate is $8000 \times 193 = 1.544$ Mbps. As with voice FDM, higher data rates are defined for larger groupings.

TDM is not limited to digital signals. Analog signals can also be interleaved in time. Also, with analog signals, a combination of TDM and FDM is possible. A transmission system can be frequency divided into a number of channels, each of which is further divided via TDM.

2.6 RECOMMENDED READING

[STAL14] covers all of the topics in this chapter in greater detail. [FREE05] is also a readable and rigorous treatment of the topics of this chapter. A thorough treatment of both analog and digital communication is provided in [COUC13].

COUC13 Couch, L. *Digital and Analog Communication Systems*. Upper Saddle River, NJ: Pearson, 2013.

FREE05 Freeman, R. *Fundamentals of Telecommunications*. New York: Wiley, 2005.

STAL14 Stallings, W. *Data and Computer Communications, Tenth Edition*. Upper Saddle River, NJ: Pearson, 2014.

2.7 KEY TERMS, REVIEW QUESTIONS, AND PROBLEMS

Key Terms

analog data	aperiodic	channel capacity
analog signal	bandwidth	decibel (dB)
analog transmission	broadcast radio	digital data

(continued)

digital signal	microwave	synchronous TDM
digital transmission	multiplexing	terrestrial microwave
frequency	noise	time division multiplexing (TDM)
frequency division multiplexing (FDM)	peak amplitude	time domain
frequency domain	period	transmission media
fundamental frequency	periodic	unguided media
guided media	phase	wavelength
harmonics	radio	wireless
infrared	satellite microwave	
	spectrum	

Review Questions

2.1 Differentiate between an analog and a digital electromagnetic signal.
2.2 What are three important characteristics of a periodic signal?
2.3 How many radians are there in a complete circle of 360 degrees?
2.4 What is the relationship between the frequency and period of a periodic signal?
2.5 What is the relationship between a signal's spectrum and its bandwidth?
2.6 What is fundamental frequency?
2.7 Define the absolute bandwidth of a signal.
2.8 What key factors affect channel capacity?
2.9 Differentiate between infrared and microwave transmission.
2.10 What are some major advantages and disadvantages of microwave transmission?
2.11 What is direct broadcast satellite?
2.12 Why must a satellite have distinct uplink and downlink frequencies?
2.13 Indicate some significant differences between broadcast radio and microwave.
2.14 Why is multiplexing so cost-effective?
2.15 How is interference avoided by using frequency division multiplexing?
2.16 How is efficiency improved by the byte-interleaving technique in TDM systems?

Problems

2.1 A signal has a period of 2 ms. What is its fundamental frequency?

2.2 Express the following in the simplest form you can:
a. $\sin(2\pi ft - \pi) + \sin(2\pi ft + \pi)$
b. $\sin 2\pi ft + \sin(2\pi ft - \pi)$

2.3 Sound may be modeled as sinusoidal functions. Compare the wavelength and relative frequency of musical notes. Use 330 m/s as the speed of sound and the following frequencies for the musical scale.

Note	C	D	E	F	G	A	B	C
Frequency	264	297	330	352	396	440	495	528

2.4 If the solid curve in Figure 2.14 represents $\sin(2\pi t)$, what does the dotted curve represent? That is, the dotted curve can be written in the form $A \sin(2\pi ft + \phi)$; what are A, f, and ϕ?

2.5 Decompose the signal $(1 + 0.1 \cos 5t)\cos 100t$ into a linear combination of sinusoidal function, and find the amplitude, frequency, and phase of each component. *Hint*: Use the identity for $\cos a \cos b$.

2.6 Find the period of the function $f(t) = (10 \cos t)^2$.

2.7 Consider two periodic functions $f_1(t)$ and $f_2(t)$, with periods T_1 and T_2, respectively. Is it always the case that the function $f(t) = f_1(t) + f_2(t)$ is periodic? If so, demonstrate this fact. If not, under what conditions is $f(t)$ periodic?

2.8 Figure 2.5 shows the effect of eliminating higher-harmonic components of a square wave and retaining only a few lower harmonic components. What would the signal look like in the opposite case; that is, retaining all higher harmonics and eliminating a few lower harmonics?

2.9 Using Shannon's formula, find the channel capacity of a teleprinter channel with a 450-Hz bandwidth and a signal-to-noise ratio of 5 dB.

2.10 A signal element in a digital system encodes an 8-bit word.
  a. If the digital system is required to operate at 4800 bps, what is the minimum required bandwidth of the channel?
  b. Repeat part (a) for a system that is required to operate at 19200 bps.

2.11 Study the works of Shannon and Nyquist on channel capacity. Each places an upper limit on the bit rate of a channel based on two different approaches. How are the two related?

2.12 Given the narrow (usable) audio bandwidth of a telephone transmission facility, a nominal SNR of 56 dB (400,000), and a distortion level of <0.2%,
  a. What is the theoretical maximum channel capacity (kbps) of traditional telephone lines?
  b. What is the actual maximum channel capacity?

2.13 Given a channel with an intended capacity of 14 Mbps, the signal-to-noise ratio is 127. What should the bandwidth of the channel be to achieve this capacity?

2.14 Show that doubling the transmission frequency or doubling the distance between transmitting antenna and receiving antenna attenuates the power received by 6 dB.

2.15 Fill in the missing elements in the following table of approximate power ratios for various dB levels.

Decibels	1	2	3	4	5	6	7	8	9	10
Losses			0.5							0.1
Gains			2							10

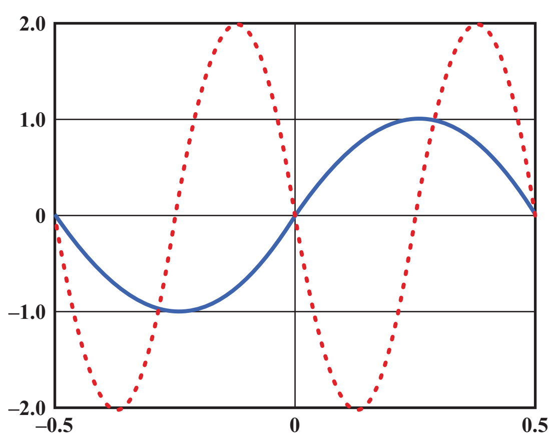


Figure 2.14 Figure for Problem 2.4

2.16 If an amplifier has a voltage ratio of 15.8, what is its voltage gain?

2.17 An amplifier has an output of 50 mW. What is its output in dBm?

APPENDIX 2A DECIBELS AND SIGNAL STRENGTH

An important parameter in any transmission system is the signal strength. As a signal propagates along a transmission medium, there will be a loss, or *attenuation*, of signal strength. To compensate, amplifiers may be inserted at various points to impart a gain in signal strength.

It is customary to express gains, losses, and relative levels in decibels because

- Signal strength often falls off exponentially, so loss is easily expressed in terms of the decibel, which is a logarithmic unit.
- The net gain or loss in a cascaded transmission path can be calculated with simple addition and subtraction.

The decibel is a measure of the ratio between two signal levels. The decibel gain is given by

$$G_{dB} = 10 \log_{10} \frac{P_{out}}{P_{in}}$$

where

G_{dB} = gain, in decibels
P_{in} = input power level
P_{out} = output power level
$\log_{10}$ = logarithm to the base 10 (from now on, we will simply use log to mean $\log_{10}$)

Table 2.3 shows the relationship between decibel values and powers of 10.

There is some inconsistency in the literature over the use of the terms *gain* and *loss*. If the value of G_{dB} is positive, this represents an actual gain in power. For example, a gain of 3 dB means that the power has doubled. If the value of G_{dB} is negative, this represents an actual loss in power. For example a gain of -3 dB means that the power has halved, and this is a loss of power. Normally, this is expressed by saying there is a loss of 3 dB. However, some of the literature would say that this is a loss of -3 dB. It makes more sense to say that a negative gain corresponds to a positive loss. Therefore, we define a decibel loss as

$$L_{dB} = -10 \log_{10} \frac{P_{out}}{P_{in}} = 10 \log_{10} \frac{P_{in}}{P_{out}} \tag{2.3}$$

Table 2.3 Decibel Values

Power Ratio	dB	Power Ratio	dB
10^1	10	10^{-1}	–10
10^2	20	10^{-2}	–20
10^3	30	10^{-3}	–30
10^4	40	10^{-4}	–40
10^5	50	10^{-5}	–50
10^6	60	10^{-6}	–60

Example 2.2 If a signal with a power level of 10 mW is inserted onto a transmission line and the measured power some distance away is 5 mW, the loss can be expressed as $L_{dB} = 10 \log(10/5) = 10(0.3) = 3\text{ dB}$.

Note that the decibel is a measure of relative, not absolute, difference. A loss from 1000 mW to 500 mW is also a loss of 3 dB.

The decibel is also used to measure the difference in voltage, taking into account that power is proportional to the square of the voltage:

$$P = \frac{V^2}{R}$$

where

P = power dissipated across resistance R
V = voltage across resistance R

Thus

$$L_{dB} = 10 \log \frac{P_{in}}{P_{out}} = 10 \log \frac{V_{in}^2/R}{V_{out}^2/R} = 20 \log \frac{V_{in}}{V_{out}}$$

Example 2.3 Decibels are useful in determining the gain or loss over a series of transmission elements. Consider a series in which the input is at a power level of 4 mW, the first element is a transmission line with a 12-dB loss (−12 dB gain), the second element is an amplifier with a 35-dB gain, and the third element is a transmission line with a 10-dB loss. The net gain is $(-12 + 35 - 10) = 13$ dB. To calculate the output power P_{out},

$$G_{db} = 13 = 10 \log(P_{out}/4\text{ mW})$$
$$P_{out} = 4 \times 10^{1.3}\text{ mW} = 79.8\text{ mW}$$

Decibel values refer to relative magnitudes or changes in magnitude, not to an absolute level. It is convenient to be able to refer to an absolute level of power or voltage in decibels so that gains and losses with reference to an initial signal level may be calculated easily. The **dBW (decibel-Watt)** is used extensively in microwave applications. The value of 1 W is selected as a reference and defined to be 0 dBW. The absolute decibel level of power in dBW is defined as

$$\text{Power}_{dBW} = 10 \log \frac{\text{Power}_W}{1\text{ W}}$$

Example 2.4 A power of 1000 W is 30 dBW, and a power of 1 mW is −30 dBW.

Another common unit is the **dBm (decibel-milliWatt)**, which uses 1 mW as the reference. Thus 0 dBm = 1 mW. The formula is

$$\text{Power}_{dBm} = 10 \log \frac{\text{Power}_{mW}}{1\text{ mW}}$$

Note the following relationships:

$$+30\text{ dBm} = 0\text{ dBW}$$
$$0\text{ dBm} = -30\text{ dBW}$$

CHAPTER 3

Communication Networks

LEARNING OBJECTIVES

After studying this chapter, you should be able to:

- Explain the roles and scope of wide, local, and metropolitan area networks.
- Define circuit switching and describe the key elements of circuit-switching networks.
- Define packet switching and describe the key elements of packet-switching technology.
- Discuss the relative merits of circuit switching and packet switching and analyze the circumstances for which each is most appropriate.

This chapter provides an overview of various approaches to communication networking. The chapter begins with a survey of different type of networks based on geographic extent. Then circuit-switching and packet-switching networks are examined.

3.1 LANS, MANS, AND WANS

Local area networks (LANs), metropolitan area networks (MANs), and **wide area networks (WANs)** are all examples of communications networks. Figure 3.1 illustrates these categories, plus some special cases. By way of contrast, the typical range of parameters for a multiple-processor computer is also depicted.

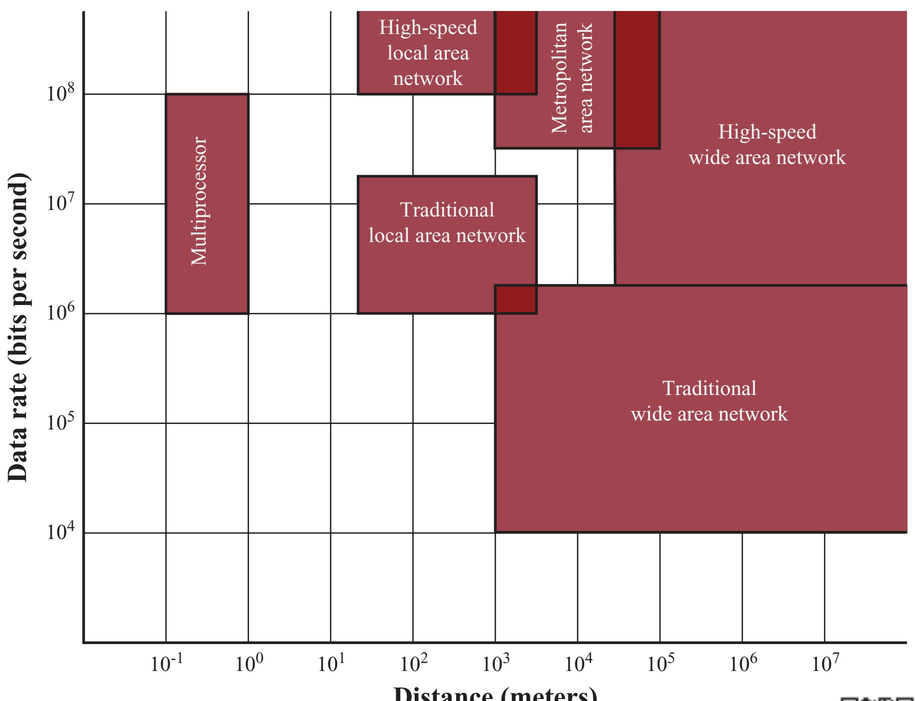


Figure 3.1 Comparison of Multiprocessor Systems, LANs, MANs, and WANs

Wide Area Networks

WANs cover a large geographical area, may require the crossing of public right-of-ways, and may rely at least in part on circuits provided by a common carrier. Typically, a WAN consists of a number of interconnected switching nodes. A transmission from any one device is routed through these internal nodes to the specified destination device.

Traditionally, WANs have been implemented using one of two technologies: circuit switching and packet switching. Subsequently, frame relay and asynchronous transfer mode (ATM) assumed major roles. While ATM and, to some extent, frame relay are still widely used, their use is gradually being supplanted by services based on gigabit Ethernet and Internet Protocol technologies.

Local Area Networks

As with WANs, a LAN is a communications network that interconnects a variety of devices and provides a means for information exchange among those devices. There are several key distinctions between LANs and WANs:

1. The scope of the LAN is small, typically a single building or a cluster of buildings. This difference in geographic scope leads to different technical solutions.
2. It is usually the case that the LAN is owned by the same organization that owns the attached devices. For WANs, this is less often the case, or at least a significant fraction of the network assets are not owned. This has two implications. First, care must be taken in the choice of LAN, since there may be a substantial capital investment (compared with charges for WANs) for both purchase and maintenance. Second, the network management responsibility for a LAN falls solely on the user.
3. The internal data rates of LANs are typically much greater than those of WANs.

A simple example of a LAN that highlights some of its characteristics is shown in Figure 3.2. All of the devices are attached to a shared transmission medium. A transmission from any one device can be received by all other devices attached to the same network.

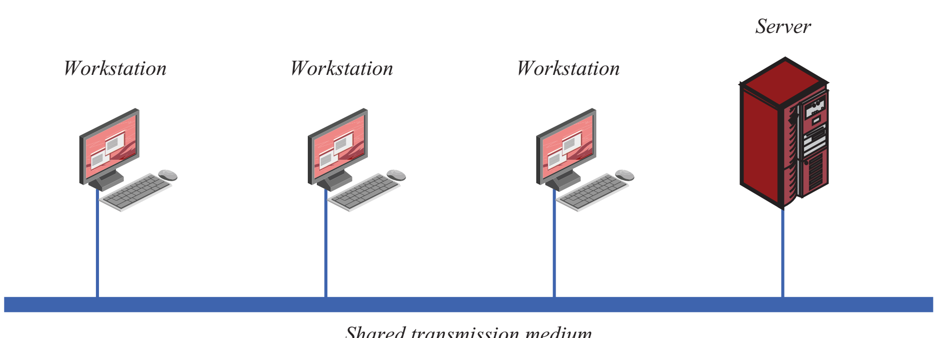


Figure 3.2 A Simple Local Area Network