WIRELESS COMMUNICATIONS

For the Students of B.E./B.Tech. of all Technical Universities of India

Er. RISHABH ANAND

S. CHAND

WIRELESS COMMUNICATIONS

For the students of B.E. / B.Tech. of all Technical Universities of India

Er. Rishabh Anand
B.Tech.

S. CHAND & COMPANY LTD.
(AN ISO 9001 : 2008 COMPANY)
RAM NAGAR, NEW DELHI - 110 055

S. CHAND & COMPANY LTD.

(An ISO 9001 : 2008 Company)

Head Office: 7361, RAM NAGAR, NEW DELHI - 110 055
Phone: 23672080-81-82, 9899107446, 9911310888
Fax: 91-11-23677446
Shop at: **schandgroup.com;** e-mail: **info@schandgroup.com**

Branches :

AHMEDABAD	:	1st Floor, Heritage, Near Gujarat Vidhyapeeth, Ashram Road, **Ahmedabad** - 380 014, Ph: 27541965, 27542369, ahmedabad@schandgroup.com
BENGALURU	:	No. 6, Ahuja Chambers, 1st Cross, Kumara Krupa Road, **Bengaluru** - 560 001, Ph: 22268048, 22354008, bangalore@schandgroup.com
BHOPAL	:	Bajaj Tower, Plot No. 243, Lala Lajpat Rai Colony, Raisen Road, **Bhopal** - 462 011, Ph: 4274723. bhopal@schandgroup.com
CHANDIGARH	:	S.C.O. 2419-20, First Floor, Sector - 22-C (Near Aroma Hotel), **Chandigarh** -160 022, Ph: 2725443, 2725446, chandigarh@schandgroup.com
CHENNAI	:	152, Anna Salai, **Chennai** - 600 002, Ph: 28460026, 28460027, chennai@schandgroup.com
COIMBATORE	:	No. 5, 30 Feet Road, Krishnasamy Nagar, Ramanathapuram, **Coimbatore** -641045, Ph: 0422-2323620 coimbatore@schandgroup.com **(Marketing Office)**
CUTTACK	:	1st Floor, Bhartia Tower, Badambadi, **Cuttack** - 753 009, Ph: 2332580; 2332581, cuttack@schandgroup.com
DEHRADUN	:	1st Floor, 20, New Road, Near Dwarka Store, **Dehradun** - 248 001, Ph: 2711101, 2710861, dehradun@schandgroup.com
GUWAHATI	:	Pan Bazar, **Guwahati** - 781 001, Ph: 2738811, 2735640 guwahati@schandgroup.com
HYDERABAD	:	Padma Plaza, H.No. 3-4-630, Opp. Ratna College, Narayanaguda, **Hyderabad** - 500 029, Ph: 24651135, 24744815, hyderabad@schandgroup.com
JAIPUR	:	1st Floor, Nand Plaza, Hawa Sadak, Ajmer Road, **Jaipur** - 302 006, Ph: 2219175, 2219176, jaipur@schandgroup.com
JALANDHAR	:	Mai Hiran Gate, **Jalandhar** - 144 008, Ph: 2401630, 5000630, jalandhar@schandgroup.com
JAMMU	:	67/B, B-Block, Gandhi Nagar, **Jammu** - 180 004, (M) 09878651464 **(Marketing Office)**
KOCHI	:	Kachapilly Square, Mullassery Canal Road, Ernakulam, **Kochi** - 682 011, Ph: 2378207, cochin@schandgroup.com
KOLKATA	:	285/J, Bipin Bihari Ganguli Street, **Kolkata** - 700 012, Ph: 22367459, 22373914, kolkata@schandgroup.com
LUCKNOW	:	Mahabeer Market, 25 Gwynne Road, Aminabad, **Lucknow** - 226 018, Ph: 2626801, 2284815, lucknow@schandgroup.com
MUMBAI	:	Blackie House, 103/5, Walchand Hirachand Marg, Opp. G.P.O., **Mumbai** - 400 001, Ph: 22690881, 22610885, mumbai@schandgroup.com
NAGPUR	:	Karnal Bag, Model Mill Chowk, Umrer Road, **Nagpur** - 440 032, Ph: 2723901, 2777666 nagpur@schandgroup.com
PATNA	:	104, Citicentre Ashok, Govind Mitra Road, **Patna** - 800 004, Ph: 2300489, 2302100, patna@schandgroup.com
PUNE	:	291/1, Ganesh Gayatri Complex, 1st Floor, Somwarpeth, Near Jain Mandir, **Pune** - 411 011, Ph: 64017298, pune@schandgroup.com **(Marketing Office)**
RAIPUR	:	Kailash Residency, Plot No. 4B, Bottle House Road, Shankar Nagar, **Raipur -** 492 007, Ph: 09981200834, raipur@schandgroup.com **(Marketing Office)**
RANCHI	:	Flat No. 104, Sri Draupadi Smriti Apartments, East of Jaipal Singh Stadium, Neel Ratan Street, Upper Bazar, **Ranchi** - 834 001, Ph: 2208761, ranchi@schandgroup.com **(Marketing Office)**
SILIGURI	:	122, Raja Ram Mohan Roy Road, East Vivekanandapally, P.O., **Siliguri-**734001, Dist., Jalpaiguri, (W.B.) Ph. 0353-2520750 **(Marketing Office)**
VISAKHAPATNAM	:	Plot No. 7, 1st Floor, Allipuram Extension, Opp. Radhakrishna Towers, Seethammadhara North Extn., **Visakhapatnam** - 530 013, (M) 09347580841, visakhapatnam@schandgroup.com **(Marketing Office)**

© 2012, Er. Rishabh Anand

All rights reserved. *No part of this publication may be reproduced or copied in any material form (including photo copying or storing it in any medium in form of graphics, electronic or mechanical means and whether or not transient or incidental to some other use of this publication) without written permission of the copyright owner. Any breach of this will entail legal action and prosecution without further notice.*

Jurisdiction : *All disputes with respect to this publication shall be subject to the jurisdiction of the Courts, tribunals and forums of New Delhi, India only.*

First Edition 2012

ISBN : 81-219-4055–9 **Code :** 10 540

PRINTED IN INDIA

By Rajendra Ravindra Printers Pvt. Ltd., 7361, Ram Nagar, New Delhi -110 055
and published by S. Chand & Company Ltd., 7361, Ram Nagar, New Delhi -110 055.

Dedicated to
My Mother whose love and affection made me capable of writing this book
and
Shri Guru Maharaj Ji of Shri Anandpur.

PREFACE

Today, the subject of ***Wireless Communications*** embraces an ever increasing body of knowledge. It spans a broad range of functions, from basic kinds of signaling and circuitry used, to kind of cables and wireless broadcast techniques used to transfer data from sender to receiver. The fundamental concepts and techniques covered in the book are based on those used in concepts and techniques covered in the book are based on those used in existing techniques in various parts of the world. We present a large number of examples that pertain to be most popular in Asia, USA, Japan and Canada. The main objective was to provide the students and practicing engineers with a clear and logical presentation of the basic concepts and principles of wireless and cellular systems.

The importance of Wireless Communications is well known in various engineering fields. The book is structured to cover the key aspects of the subject Wireless Communications.

The extent coverage in this book is meant to give learner enough material to help him understand how wireless communication works.

The book uses plain, lucid language to explain fundamentals of this subject. The book provides logical method of explaining various complicated concepts and stepwise methods to explain the important topics. Each chapter is well supported with necessary illustrations, practical examples and solved problems. All care has been taken to make students comfortable in understanding the basic concepts of the subject.

This book is for everyone who wants to explore the concepts in wireless communication and discover the techniques behinds the communication.

The book not only covers the entire scope of the subject but explains the philosophy of the subject. This makes the understanding of this subject but explains the philosophy of the subject. This makes the understanding of this subject more clear and makes it more interesting. The book will be very useful not only to the students but also to the subject teachers.

The author hopes that the book will fulfill the pressing need of interested readers and welcome any suggestions towards the improvement of the book.

Er. Rishabh Anand

Disclaimer : While the author of this book have made every effort to avoid any mistake or omission and have used their skill, expertise and knowledge to the best of their capacity to provide accurate and updated information. The author and S. Chand does not give any representation or warranty with respect to the accuracy or completeness of the contents of this publication and are selling this publication on the condition and understanding that they shall not be made liable in any manner whatsoever. S.Chand and the author expressly disclaim all and any liability/responsibility to any person, whether a purchaser or reader of this publication or not, in respect of anything and everything forming part of the contents of this publication. S. Chand shall not be responsible for any errors, omissions or damages arising out of the use of the information contained in this publication. Further, the appearance of the personal name, location, place and incidence, if any; in the illustrations used herein is purely coincidental and work of imagination. Thus the same should in no manner be termed as defamatory to any individual.

ACKNOWLEDGEMENT

We all marvel at the beautiful rose, the rose, with all its beauty and grandeur. But seldom we pause and thank the kind of gardener who patiently manured and watered it. The book I have written is the fruit of such unseen hands.

I wish to express my profound thanks to the people around me who helped to make this book a reality.

I am especially grateful to my teachers and my colleagues specially Mrs. Aparna Mahajan, Mr. Sanjeev Kumar, Mrs. Rashmi Mahajan, Wg.cdr. P. Patnayak, Gp. Capt. H.S. Dua, Wg.cdr. Onkar Singh, Mrs. Dimple Saproo, Mrs. Naresh Kumari, Mrs. Chitra Kaul, Director M.M. Jha, Mrs. Savita Khazanchi, Mrs. Meenakshi and other faculty members who influenced my academic grown during my study.

There was, in and always will be the love of my students whom I taught, for what is written in this book has been the fruit of experience that I taught them. Finally, thanks to my parents and to rest of my family for their patience and support during the long hours of writing this book and above all there is the one Almighty whose humble children we are. It is his blessings we cherish and pray for. It is the blessing I wish for you.

I deeply express my heartful thanks to the S. Chand & Co. Ltd., for publishing this book in such a beautiful get-up and well in time.

Er. Rishabh Anand

CONTENTS

1. Introduction to wireless commnuncation system **3–15**

1.1 Introduction 3
1.2 Evolution of mobile radio communications 3
1.3 Mobile radio systems around the world 7
1.4 Examples of wireless communication systems 8
1.5 Comparison of mobile communication system 13

2. Modern wireless communication system **16–34**

2.1 Introduction 16
2.2 Second generation (2G) cellular networks 17
2.3 Third generation (3G) wireless networks 25
2.4 Fourth generation (4G) wireless networks 30
2.5 Comparison with 3G 31
2.6 Future trends in wireless communication 32
2.7 Requirements in wireless communications 32

3. Mobile Radio Propagation: Large-Scale Path Loss, Small-Scale Fading an Multipath **35–81**

3.1 Introduction to radio wave propagation 35
3.2 Free space propagation model 35
3.3 The three basic propagation mechanisms 38
3.4 Reflection 39
3.5 Diffraction 42
3.6 Scattering 45
3.7 Practical link budget design using path loss models 46
3.8 Outdoor propagation models 47
3.9 Indoor propagation models 51
3.10 Small scale multipath propagation 52
3.11 Small scale multipath measurement 54
3.12 Parameters of mobile multipath channels 55
3.13 Type of small scale fading 58
3.14 Two-ray rayleigh fading model 61
3.15 Multipath shape factor 61

3.16 Fading rate variance relationships 62
3.17 Comparison to omnidirectional propagation 63
3.18 Second-order statistics using shape factors 64
3.19 Revisiting classical channel models with shape Factors 66

4. Spread Spectrum Modulation Techniques 82–102
4.1 Introduction 82
4.2 Advantage of spread spectrum techniques 82
4.3 Classification of spread spectrum system 83
4.4 Spread spectrum modulation techniques 83
4.5 Pseudo-noise (pn) sequences 83
4.6 Direct sequence spread spectrum (ds–ss) 85
4.7 Frequency hopped spread spectrum (fh–ss) 87
4.8 Performance of direct sequence spread spectrum 89
4.9 Performance of frequency hopping spread spectrum 91
4.10 Modulation performance in fading and multipath channels 93

5. Equalization and Diversity Techniques 103–126
5.1 Introduction 103
5.2 Fundamentals of equalization 104
5.3 Equalizers in a communications receiver 106
5.4 Survey of equalization techniques 106
5.5 Linear equalizers 108
5.6 Non-linear equalization 110
5.7 Diversity technique 113
5.8 Rake receiver 120

6. Speech Coding and Quantization Techniques 127–142
6.1 Introduction 127
6.2 Characteristics of speech signals 128
6.3 Quantization techniques 130
6.4 Vocoders 132
6.5 Linear predictive coders (LPC) 134

7. Multiple Access Techniques for Wireless Communication 143–166
7.1 Introduction 143
7.2 Introduction to multiple access 144
7.3 Frequency division multiple access (FDMA) 145
7.4 Time division multiple access (TDMA) 147

7.5 Space division multiple access (SDMA) 149
7.6 Comparison among different multiple access Technology 150
7.7 Aloha 151
7.8 Code division multiple access (CDMA) 153
7.9 Capacity of cellular system 154

8. The Cellular Concept-System Design Fundamentals 167–198

8.1 Introduction 167
8.2 The cellular concept 167
8.3 Frequency reuse 168
8.4 Channel assignment strategies 170
8.5 Handoff strategies 171
8.6 Interference and system capacity 174
8.7 Power control for reducing interference 177
8.8 Improving coverage and capacity in cellular sytems 177
8.9 Comparison among cell splitting cell sectoring micro cell zone concept increasing number of repeaters 183

9. Wireless Networking 199–224

9.1 Introduction to wireless networks 199
9.2 Difference between wireless and fixed telephone networks 201
9.3 The public switched telephone network (PSTN) 201
9.4 Limitations in wireless networking 203
9.5 Merging wireless networks and the pstn 204
9.6 Evolution and development of wireless networks 204
9.7 Traffic routing in wireless networks 209
9.8 Common channel signaling (CCS) 212
9.9 Integrated services digital network (ISDN) 213
9.10 Broadband isdn and atm 214
9.11 Signaling system number 7 (SS7) 215
9.12 Personal communication services/networks (PCS/PCNS) 219
9.13 Packet vs. Circuit switching for PCN 220
9.14 Protocol for network access 220
9.15 Network databases 221

10. Wireless Systems and Standards 225–250

10.1 Introduction 225
10.2 Global system for mobile (GSM) 225

10.3 Cdma digital cellular standard (LS-95) 237
10.4 IMT-2000 - a third generation communication 240
10.5 UMTS-universal mobile telecommunication system 241
10.6 PACS- personal access communication system 243
10.7 Amps and etacs 244
10.8 United states digital cellular (IS-54 and IS-136) 246
10.9 Wireless local loop (WLL) 246
10.10 Bluetooth 247

11. Satellite Communication 251–269

11.1 Introduction 251
11.2 Element of satellite communication 252
11.3 Reason for the downlink frequency to be lower than the uplink frequency 253
11.4 Satellite frequency allocation 253
11.5 Satellite orbit mechanism 256
11.6 Orbital parameters 259
11.7 Satellite subsystem 261
11.8 Multiple access techniques 263
11.9 Satellite applications 264

12. Modulation Techniques for Mobile Radio 270–306

12.1 Introduction 270
12.2 Frequency modulation vs. Amplitude modulation 271
12.3 Amplitude modulation 272
12.4 Angle modulation 276
12.5 Digital modulation–an overview 280
12.6 Linear modulation techniques 282
12.7 Differential phase shift keying (DPSK) 285
12.8 Quadrature phase shift keying (QPSK) 286
12.9 Qpsk transmission and detection techniques 288
12.10 $\pi/4$ QPSK 289
12.11 $\pi/4$ QPSK transmission techniques 290
12.12 $\pi/4$ QPSK detection techniques 291
12.13 Constant envelope modulation 294
12.14 Combined linear and constant envelope modulation techniques 298

13. Architecture and Applications of Wireless Networks **307–377**

13.1 Wireless usb 307
13.2 Wireless network design analysis 308
13.3 GPRS (general packet radio service) 310
13.4 Wireless application protocol (WAP) 315
13.5 Bluetooth 319
13.6 Cellular digital packet data (CDPD) 320
13.8 Connectionless protocol 328
13.9 Connection-oriented protocol 329
13.10 Internet protocol 330
13.11 IP address 333
13.12 Osi model 339
13.13 Transmission control protocol 345
13.14 User datagram protocol (UDP) 357
13.15 Wireless location service architectures 362

Appendices **379–403**

Model Question Papers **405–410**

You see, wire telegraph is a kind of a very, very long cat. You pull his tail in New York and his head is meowing in Los Angeles. Do you understand this? And radio operates exactly the same way: you send signals here, they receive them there. The only difference is that there is no cat.

Albert Einstein
(1879-1955)

1 Chapter

INTRODUCTION TO WIRELESS COMMUNICATION SYSTEMS

1.1. *Introduction.* **1.2.** *Evolution of Mobile Radio Communications.* **1.3.** *Mobile Radio Systems Around the World.* **1.4.** *Examples of Wireless Communication Systems.* **1.5.** *Comparison of Mobile Communication System.*

1.1 INTRODUCTION

The ability to communicate with people on the move has evolved remarkably since Guglielmo Marconi first demonstrated radio's ability to provide continuous contact with ships sailing the English channel. That was in 1897, and since then new wireless communications methods and services have been enthusiastically adopted by people throughout the world. Particularly during the past ten years, the mobile radio communications industry has grown by orders of magnitude, fueled by digital and RF circuit fabrication improvements, new large-scale circuit integration, and other miniaturization technologies which make portable radio equipment smaller, cheaper, and more reliable. Digital switching techniques have facilitated the large scale deployment of affordable, easy-to-use radio communication networks. These trends will continue at an even greater pace during the next decade.

1.2 EVOLUTION OF MOBILE RADIO COMMUNICATIONS

The understanding of radio waves is fundamental to wireless communications, but simply knowing that electromagnetic waves exist is a relatively recent historical event. In the short period since that time, there have been numerous milestones in the development of radio communications. Some of these milestones are the following:

1833	Carl Friedrich Gauss and Wilhelm Eduard Weber deploy first regularly used **electromechanical telegraph.**
1843	In the year 1843, Michael Faraday studied about the wireless communication and discovered that the space is capable of conducting electricity.
1844	Data communication began with Samuel Morse, who in 1844 invented and pioneered the telegraph, which used **Morse code** (consisting of interweaving dots and dashes) as its method for delivering communication over vast distances.

1845	The existence of the electro-magnetic fields had been discovered by the British physicist Michael Faraday in 1845. Further work was carried out by the British physicist James Maxwell and the German physicist Heinrich Hertz.
1867	James predicts existence of electromagnetic (EM) waves by the use of more complex experiments. He gave a number of laws in the field of electromagnetics known as **maxwell's equation.** He believed he was guided by the great Michael Faraday who had died in 1867.
1876	The **telephone** was invented by Alexander Graham Bell in 1876.
1887	Heinrich Hertz proves existence of EM waves, first spark transmitter generates a spark in a receiver several metres away.
1893	Jagadish Chandra Bose created the world's first wireless communication link by using millimeter waves for the remote control of a gun.
1896	Guglielmo Marconi developed the first Wireless telegraph system.
1897	**"The Birth of Radio"** - Marconi awarded patent for wireless telegraph. First **"Marconi station"** established on Needles island to communicate with English coast.
1898	Marconi awarded English patent No. *7777* for tuned communication. Wireless telegraphic connection between English and France was established.
1901	Marconi and his team received a record-breaking wireless communication in January, 1901 that covered a distance of 186 miles over Atlantic oceans.
1902	First **bidirectional** communication across Atlantic occurred.
1903	In 1903, the first transatlantic wireless communication originating in the United States was successfully transmitted, a message from US president Theodore Roosevelt to king Edward VII of the United Kingdom.
1909	Marconi awarded with **Nobel Prize** in physics for his work in wireless communication.
1914	First **Voice over radio transmission.**
1920	**Mobile receivers** installed in police cars in Detroit.
1923	The first **two way radio** was developed in Australia in the year 1923. The Victorian police were the first ones to use this type of a radio for wireless communication in cars.
1927	First **Commercial Radio Telephone** service operated between Britain and U.S.
1928	Motorola was founded in Chicago in 1928 and is the world leader in providing mission critical wireless communication solutions to public safety agencies.
1934	The **Federal Communication Commission (FCC)** an independent governmental agency, has the responsibility of regulating both wired and wireless communication in the United States created in 1934.
1935	Armstrong demonstrated the **Frequency Modulation (FM).**
1940	Majority of police systems converted to FM.
1942	On August 11, 1942 a patent (US patent 2292387) was granted to Hedy Lamarr for her co-invention of **spread-spectrum** broadcast communications technologies which form the basis of modern wireless communications.

1943	Tesla was somewhat vindicated in 1943 nine months after his death. The US supreme court held that some of Marconi's most important patents on wireless communication were invalid. Reversing the old decision, they awarded priority to Tesla.
1946	First car-based mobile telephone setup in St. Louis, using **'Push-to-talk'** technology. First **Interconnection of mobile users** to public switched telephone network (PSTN).
1948	Claude Shannon publishes two bench mark papers on **Information Theory,** containing the basis for **data compression (source encoding and error detection)** and **correction (channel encoding).**
1949	FCC recognizes mobile radio as new class of service.
1950	TD-2, the first **Terrestrial Microwave Telecommunication** system, installed to support 2400 telephone circuits.
1950's	Late in the decade, the first **Paging Access Control Equipment (PACE)** paging systems established.
1956	The world's first fully automated system for mobile teleophony is launched in Stokholm on 25th April. MTA starts with eight subscribers.
1960's	Number of mobile users are more than 1.4 million. Early in the decade, the improved Mobile Telephone System (IMTS) developed with simultaneous transmit and receive, more channels and greater power.
1962	The first **communication satellite,** Telstar launched into orbit.
1963	The first commercial communication system that used digital transmission via wire was developed.
1964	The **International Telecommunication Satellite Consortium** (INTELSAT) established and in 1965 launches the Early bird geostationary satellite.
1968	Defense Advanced Research Projects Agency - US (DARPA) selected BBN to develop the Advanced Research Projects Agency Network (ARPANET) the father of the modern Internet.
1969	Nordic Telecom Administrators meet for a conference on common Nordic Standard for Mobile Telephony. The conference accepts the proposal and appoints a working group called **Nordic Mobile Telephone Group (NMT).**
1970's	Packet switching emerges as an efficient means of data communications, with the X.25 standard emerging late in the decade.
1971	MTD a manually operated mobile telephone system is introduced.
1975	At a Nordic telecom conference, NMT is approved as a technical standard.
1976	Bell mobile phone has 543 pay customers using 12 channels in the New York city, waiting list is 3700 people, service is poor due to blocking.
1979	NTT/Japan deployed first **cellular communication system.**
1981	On 1st September, Saudi Arabia becomes the first in the world to launch a commercial NMT system. Ericsson and SPA are the suppliers.
1982	Denmark brings NMT-450 in January, Finland launches NMT in March. The first **GSM** meeting is held in Stockholm in December with 31 people from 11 countries participating.

1983	American **Motorola** releases the first **hand-held cellular mobile telephone, Dyna TAC.** The Advanced Mobile Phone System (AMPS) invented by Bell Labs, first installed in the US with geographic regions divided into 'cells' (*i.e.,* cellular telephone).
1984	The Finnish company **Nokia** launches its first portable car-telephone, the **Nokia Talkman.**
1985	**QUALCOMM Inc.** is founded in July, 1985 and today the company is a major competitor in the fastest-growing, digital communications technology for 3G wireless communication services. It has several CDMA patents.
1989	American Motorola new **Micro TAC** is the world's smallest and lightest mobile telephone so far.
1990	Groupe Special Mobile group decides on the final specifications for the new mobile system.
1991	Officially the GSM starts in July. But most countries delay their start until 1992 due to a shortage of telephones.
	Ericsson delivers the first GSM system to Mannesmann in Germany.
1992	The working group for the third generation mobile telephony UMTS is setup in September.
	The Briton, Neil Papworth is thought to have sent the first SMS on December 3rd from his PC to Richard Jarvis at the British GSM operator Vodafone.
	Ericsson introduced the first pocket sized GSM mobile telephone.
1993	IS-95 code division multiple access (CDMA) spread-spectrum digital cellular system deployed in US.
1994	GSM system deployed in US relabeled **"Global System for Mobile".**
1995	FCC auctions off frequencies in personal communication system (PCS) band at 1.8 GHz for mobile telephony.
1997	Number of cellular telephone user in U.S is more than 50 million and **Comviq** launches the first prepaid refill cards in Sweden in March.
1998	The analogue mobile network reaches its peak with 91 million subscribers in the world.
1999	The Japanese NTT DOCOMO launches its I-mode *i.e.,* internet services on the mobile phone. The technology becomes a success.
2000	All around the world operators are competing for 3G telephony licences.
2001	In April, **Ericsson and Vodafone** make the first 3G call in England. Japanese Sony and Ericsson decide to form a joint company for developing and manufacturing mobiles.
2002	By the end of the year there are 1.17 billion mobile subscribers in the world.
2003	The 3G operator launches around the world.
2005	Mobile Wimax challenges 3G as a standard for both telephony and data traffic.
2006	In August, Samsung shows off a mobile phone connection with 4G technology which will give a data speed of between 10 and 100 Mbps.
	The number of mobile subscribers in the world are 2.5 billion.

2007	The forecast predicts three billion mobile users in the world (*i.e.,* 45% of the world's inhabitants).

1.3. MOBILE RADIO SYSTEMS AROUND THE WORLD

Many mobile radio standards have been developed for wireless systems throughout the world, and more standards are likely to issue. Table 1.1 through Table 1.3 show the more common paging, cordless, cellular, and personal communications standards used in North America, Europe and Japan.

Table 1.1: Major Mobile Radio Standards in North America

Standard	*Type*	*Year of Introduction*	*Multiple Access*	*Frequency Band*	*Modulation*	*Channel Bandwidth*
AMPS	Cellular	1983	FDMA	824-894 MHz	FM	30 kHz
NAMPS	Cellular	1992	FDMA	824-894 MHz	FM	10 kHz
USDC	Cellular	1991	TDMA	824-894 MHz	π/4-DQPSK	30 kHz
CDPD	Cellular	1993	FH/Packet	824-894 MHz 824-894 MHz	GMSK.	30 kHz
IS-95	Cellular/ PCS	1993	CDMA	1.8-2.0 GHz	QPSK/BPSK	1.25 MHz.
GSC	Paging	1970s	Simplex	Several	FSK	12.5 kHz
POCSAG	Paging	1970s	Simplex	Several	FSK	12.5 kHz
FLEX DCS-1900	Paging	1993	Simplex	Several	4-FSK	15 kHz
(GSM)	PCS	1994	TDMA	1.85-1.99 GHz	GMSK	200 kHz
PACS	Cordless/ PCS	1994	TDMA/ FDMA	1.85-1.99 GHz	π/4-DQPSK	300 kHz
MIRS	SMR/PCS	1994	TDMA	Several	16-QAM	25 kHz
iDen	SMR/PCS	1995	TDMA	Several	16-QAM	25 kHz

Table 1.2: Major Mobile Radio Standards in Europe

Standard	*Type*	*Year of Introduction*	*Multiple Access*	*Frequency Band*	*Modulation*	*Channel Bandwidth*
ETACS	Cellular	1985	FDMA	900 MHz	FM	25 kHz
NMT-450	Cellular	1981	FDMA	450-470 MHz	FM	25 kHz
NMT-900	Cellular	1986	FDMA	890-960 MHz	FM	12.5 kHz
GSM	Cellular/PCS	1990	TDMA	890-960 MHz	GMSK	200 kHz
C-450	Cellular	1985	FDMA	450-465 MHz	FM	20 kHz/10 kHz
ERMES	Paging	1993	FDMA	Several	4-FSK	25 kHz
CT2	Cordless	1989	FDMA	864-868 MHz	GFSK	100 kHz

Standard	*Type*	*Year of Introduction*	*Multiple Access*	*Frequency Band*	*Modulation*	*Channel Bandwidth*
DECT	Cordless	1993	TDMA	1880-1900 MHz	GFSK	1.728 MHz
DCS-1800	Cordless/ PCS	1993	TDMA	1710-1880 MHz	GMSK	200 kHz

Table 1.3: Major Mobile Radio Standards in Japan

Standard	*Type*	*Year of Introduction*	*Multiple Access*	*Frequency Band*	*Modulation*	*Channel Bandwidth*
JTACS	Cellular	1988	FDMA	860-925 MHz	FM	25 kHz
PDC	Cellular	1993	TDMA	810-1501 MHz	Jt/4-DQPSK	25 kHz
NTT	Cellular	1979	FDMA	400/800 MHz	FM	25 kHz
NTACS	Cellular	1993	FDMA	843-925 MHz	FM	12.5 kHz
NTT	Paging	1979	FDMA	280 MHz	FSK	12.5 kHz
NEC	Paging	1979	FDMA	Several	FSK	10 kHz
PHS	Cord-less	1993	TDMA	1895-1907 MHz	Tt/4-DQPSK	300 kHz

The CT2 and Digital European Cordless Telephone (DECT) standards to exhibit in Europe are the two most popular cordless telephone standards throughout Europe and Asia. The CT2 system makes use of microcells which cover small distances, usually less than 100 m, using base stations with antennas mounted on street lights or on sides of buildings. The CT2 system uses battery efficient frequency shift keying and use 32 Kbps data rate for data transmission.

The world's first cellular system was implemented by the Nippon Telephone and Telegraph Company (NTT) in Japan. The system, extend in 1979, uses 600 FM duplex channels (25 kHz for each one-way link) in the 800 MHz band.

1.4. EXAMPLES OF WIRELESS COMMUNICATION SYSTEMS

Most people are familiar with a number of mobile radio communication system used in everyday life. Garage door openers, remote controllers for home entertainment equipment, cordless telephones, hand-held walkie-talkies, pagers (also called *paging receivers* or "beepers"), and cellular telephones are consider as mobile radio communication systems. However, the cost, complexity, performance, and types of services offered by each of these mobile systems are vastly different.

Table 1.4 shows definitions and terms used to describe elements of wireless communication systems.

Table 1.4: Wireless Communications System Definitions

Base station	A fixed station in a mobile radio system used for radio communication with mobile stations. Base stations are located at the center or on the edge of a coverage region and consist of radio channels and transmitter and receiver antennas mounted on a tower.
Control channel	Radio channel used for transmission of call setup, call request, call initiation, and other beacon or control purposes.
Forward channel	Radio channel used for transmission of information from the base station to the mobile.
Full duplex systems	Communication systems which allow simultaneous two-way communication. Transmission and reception is typically on two different channels (FDD) although new cordless/PCS systems are using TDD.
Half Duplex systems	Communication systems which allow two-way communication by using the same radio channel for both transmission and reception. At any given time, the user can only either transmit or receive information.
Handoff	The process of transferring a mobile station from one channel or base station to another.
Mobile station	A station in the cellular radio service intended for use while in motion at unspecified locations. Mobile stations may be hand-held personal units (portables) or installed in vehicles (mobiles).
Mobile Switching Centre	Switching centre which coordinates the routing of calls in a large service area. In a cellular radio system, the MSC connects the cellular base stations and the mobiles to the PSTN. An MSC is also called a *mobile telephone switching office* (MTSO).
Page	A brief message which is broadcast over the entire service area, usually in a simulcast fashion by many base stations at the same time.
Reverse channel	Radio channel used for transmission of information from the mobile to base station.
Roamer	A mobile station which-operates in a service area (market) other than that from which service has been subscribed.
Simplex systems	Communication systems which provide only one-way communication.
Subscriber	A user who pays subscription charges for using a mobile communications system.
Transceiver	A device capable of simultaneously transmitting and receiving radio signals.

Frequency division duplexing (FDD) provides at same time radio transmission channels for the subscriber and the base station, so that they both may constantly transmit while at the same time receiving signals from one another. At the subscriber unit, however, a single antenna is used for both transmission to and reception from the base station, and a device called a duplexer is used within the subscriber unit to enable the same antenna to be used for simultaneous transmission and reception. To facilitate FDD, it is necessary to distinct the transmit and receive frequencies by about 5% of the nominal RF frequency.

Time division duplexing (TDD) uses the reality that it is possible to share a single radio channel in time, so that a portion of the time is used to transmit from the base station to the mobile station and the remaining time is used to transmit from the mobile to the base station for communication.

1.4.1 Paging Systems

Paging systems are communication systems that send brief messages to a subscriber. Depending on the type of service, the message may be either a numeric message, an alphanumeric message, or a voice message or a picture message.

In modern paging systems, news headlines, stock quotations and faxes may be sent. A message is sent to a paging subscriber by using paging system access number (usually a toll-free telephones number) with a telephone keypad or modem. The issued message is called *a page.* The paging system then transmits the page in every point of service area using base stations which broadcast the page on a radio carrier for communication.

Paging systems vary widely in their complexity and coverage area and the simple paging systems may cover a limited range of 2 to 5 km, or may even be confined to with in individual buildings, wide area paging systems can provide world wide coverage so the wide area paging system are use widely.

Wide area paging systems consist of a network of telephone lines, many base station transmitters, and large radio towers that simultaneously broadcast a page, from each base station (this is called *simulcasting).* Simulcast transmitters may be placed with in the same service area or in different cities or countries.

Paging systems are designed to provide reliable communication to subscribers wherever they are; whether in a building, driving on a highway, or flying in an air plane. This necessitates large transmitter powers (consider in kilowatts) and low data rates (a couple of thousand bits per second) for maximum coverage from each base station. Fig. 1.1 shows a diagram of a wide area paging system and showing how the paging system comminute.

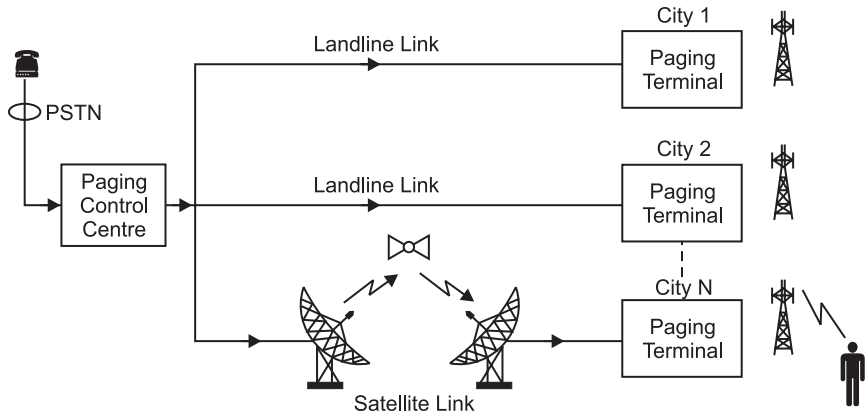


Fig. 1.1: A wide area paging system. The paging control centre dispatches pages received from the PEST throughout several cities at the same time.

1.4.2 Cordless Telephone Systems

Cordless telephone systems are full duplex communication systems that use radio to connect a portable handset to a dedicated base station, which is then connected to a dedicated telephone line with a peculiar telephone number on the Public switched telephone network (PSTN).

Fig. 1.2 illustrates a cordless telephone system with PSTN and base station.

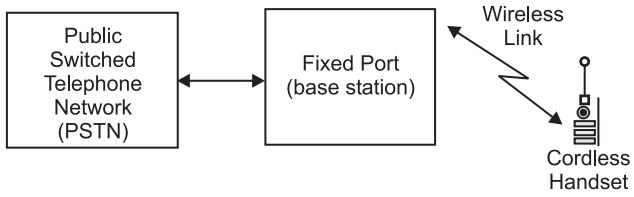


Fig. 1.2: A cordless telephone system

In first generation cordless telephone systems (manufactured in the 1980's), the portable unit communicates only to the dedicated base unit and only over distance of a few tens of meters. Early cordless telephones operate singly as extension telephones to a transceiver connected to a subscriber line on the PSTN and are primarily for in-home use purpose.

Cordless telephone systems provide the user with limited range and mobility, as it is usually not possible to maintain a call if the user move outside the range of the base station. Typical second generation base station provide coverage ranges up to a few hundred meters to a communication system.

1.4.3 Cellular Telephone Systems

A cellular telephone system provides a wireless connection to the PSTN for any user location within the radio range of the system. Cellular systems accommodate a large number of users over large geographic area, within limited frequency spectrum for communication point of view.

Cellular radio systems provide high quality service that is often comparable to that of the landline telephone systems. High capacity is obtained by limiting the coverage of each base station transmitter to a small geographic area called a cell, so that the same radio channels may be reused by another base station located some distance away for transmit the information.

A sophisticated switching technique called a *handoff enables* a call to proceed uninterrupted when the user moves from one cell to another cell with in a system.

Fig. 1.3 shows a basic cellular system which consists of mobile stations, base stations and a Mobile switching centre (MSC). *The mobile* switching centre is sometimes refer as a Mobile telephone switching office (MTSO), since it is responsible for connecting all mobiles to the PSTN in a cellular system. Each mobile communicates with radio with one of the base stations and may be handed-off to any number of base stations throughout the duration of a call and it is shown in Fig. 1.3.

The mobile station contains a transceiver, an antenna, and control circuitry, and may be mounted in a vehicle or used as a portable hand-held unit. The base stations consist of several transmitters and receivers which at same time handle full duplex communications and generally have towers which support several transmitting and receiving antennas. The base station serves as a bridge between all mobile users in the cell and connects the related mobile calls via telephone lines or microwave links to the MSC within a fixed region.

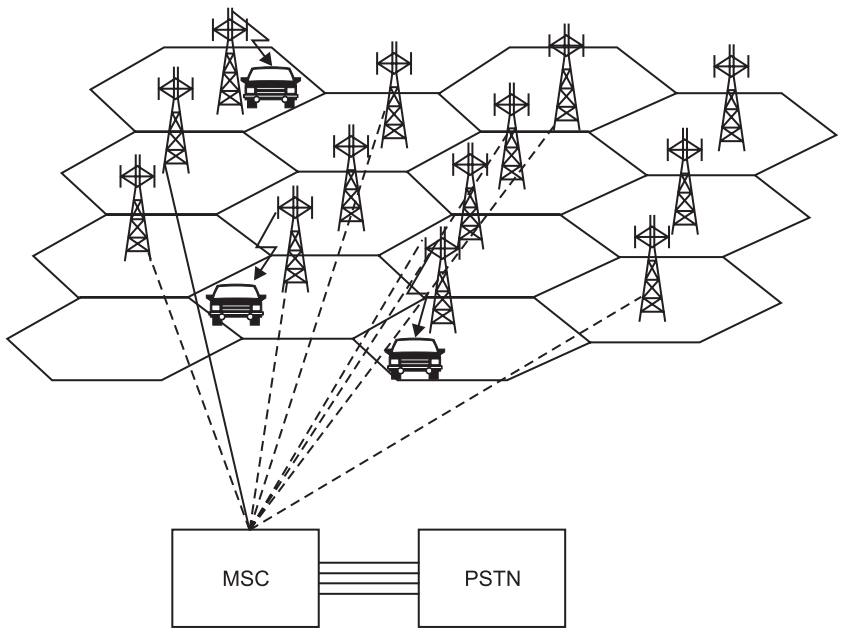


Fig. 1.3: A cellular system. The towers represent base stations which provide radio access between mobile users and the mobile switching center (MSC).

Communication between the base station and the mobiles is defined by a standard Common air interface (CAT) that specifies four different channels. The channels used for voice transmission from the base station to mobiles are called Forward voice channels (FVC) and the channels support the voice tranismission from mobiles to the base station are called Reverse voice channels (RVC). True two channels responsible for initiating mobile calls are the forward control channels (FCC) and Reverse control channels (RCC) between a set of one transmitter and receiver.

1.5 COMPARISON OF MOBILE COMMUNICATION SYSTEM

Table 1.5 and 1.6 give the comparison between mobile communication system.

Table 1.5: Comparison of Mobile Communication Systems—Mobile Station.

Service	*Coverage Range*	*Required Infrastructure*	*Complexcity*	*Hardware Cost*	*Carrier Frequency*	*Functionality*
TV remote control	Low	Low	Low	Low	Infrared	Transmitter
Garage door opener	Low	Low	Low	Low	< 100 MHz	Transmitter
Paging system	High	High	Low	Low	< 1 GHz	Receiver
Cordless phone	Low	Low	Moderate	Low	< 1 GHz	Transceiver
Cellular phone	High	High	High	Moderate	< 2GHz	Transceiver

Table 1.6: Comparison of Mobile Communication Systems—Base Station

Service	*Coverage Range*	*Required Infrastructure*	*Complexcity*	*Hardware Cost*	*Carrier Frequency*	*Functionality*
TV remote control	Low	Low	Low	Low	Infrared	Transmitter
Garage door opener	Low	Low	Low	Low	< 100 MHz	Transmitter
Paging system	High	High	High	High	< 1 GHz	Receiver
Cordless phone	Low	Low	Low	Moderate	< 1 GHz	Transceiver
Cellular phone	High	High	High	High	< 2 GHz	Transceiver

HIGHLIGHTS

1. In 1897, Marconi first invented and demonstrated radio's ability to provide continuous contact with ships sailing the English channel. Beginning with this invention, the field of wireless communication has experienced various advanced developments or invention.
2. Wireless communication is experiencing its fastest growth period in history. This has been possible because of enabling technologies that permit wide spread deployment or development.
3. To provide wireless communication to the whole world was a dream before the development of cellular concept by bell laboratories in the 1960s to 1970s.
4. With the development of highly reliable, miniature, solid state radio frequency hardware in the 1970s, the wireless communication era come into existence.
5. There are various reason for developing a cellular mobile telephone system as:
   (*i*) Limitation of conventional mobile telephone system.
   (*ii*) Spectrum efficiency considerations or defining.
   (*iii*) Technology, feasibility and service affordability.
6. There are following three limitations of conventional mobile telephone system:
   (*i*) Limited service capability
   (*ii*) Poor service performance
   (*iii*) Inefficient frequency spectrum utilization.
7. Radio transmission can define by using many different frequency bands. Each frequency band exhibits certain advantage and disadvantages properties.
8. Wireless communication system that have been in deployment for sometime are those of the first generation and second generation. The third generation systems are also currently under deployment but continue to evolve or explain.
9. The most popular example of wireless communication systems are paging system, cordless telephone system and cellular telephone system.
10. The wireless systems are divided into three generation systems: first generation, second generation and third generation.
11. The first generation technology use the analog base device and the data will be transmit through channel by using FDMA technique.
12. The second generation technology use the digital based device and the information transmit through channel by using TDMA technology.
13. In the third generation technology, the world wide access is possible and any person can communicate at any time and at any place.
14. A basic cellular system define the following three parts:
    (*i*) A mobile unit
    (*ii*) A cell site
    (*iii*) A mobile telephone switching office and the basic cellular system provide the connection between them.

REVIEW QUESTIONS

1.1. Explain the introduction of Wireless Communication System.

1.2. What do you mean by Mobile Radio System arround the world?

1.3. Explain in brief the major mobile radio standard in North America and Europe.

1.4. Give the example of Wireless Communication System (1.5). What is the basic principle of paging system?

1.6. What are the reason for developing a Cellular Mobile Telephone System ?

1.7. What are major three approaches to achieve the ideal Mobile Telephone System ?

1.8. Explain in brief the basic mechanism of Cordless Telephone System.

1.9. What is the basic difference between Mobile Telephone System and Cordless Telephone System ?

1.10. Give an comparison between example of Wireless Communication System.

1.11. Why do paging system need to provide low data rate old how does a low dats rate provide a larger coverage area?

1.12. Between a pager, a cellular phone, and a cordless phone, which device will have

(*i*) the largest battery life between charging? Why?

(*ii*) the shortest battery life between charging? Why?

1.13. Define Page.

1.14. What is Paging System?

2 Chapter

MODERN WIRELESS COMMUNICATION SYSTEMS

2.1. *Introduction.* **2.2.** *Second Generation (2G) Cellular Networks.* **2.3** *Third Generation (3G) Wireless Networks.* **2.4.** *Fourth Generation (4G) Wireless Networks* **2.5.** *Comparison with 3G* **2.6.** *Fouture Trends in Wireless Communication* **2.7.** *Requirements In Wireless Communications*

2.1 INTRODUCTION

Since the mid 1990's the cellular communications industry has witnessed explosive growth. Wireless communications networks have become much more pervasive than anyone could have imagined when the cellular concept was first developed in the 1960's and 1970's for explain the Modern wireless communication system.

As shown in Fig. 2.1, the worldwide cellular and personal communication subscriber base surpassed 600 million users in late 2001, and the number of individual subscribers is projected to reach 2 billion by the end of 2006. This growth of cellular telephone subscriber throughout the world is shown in Fig. 2.1.

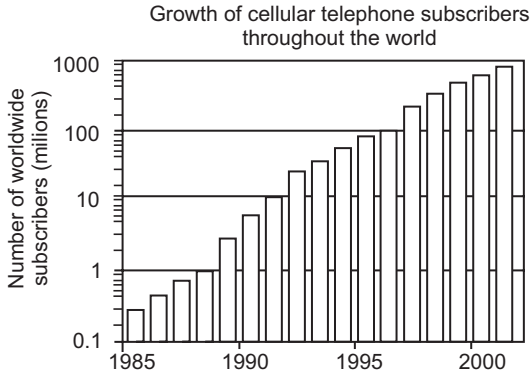


Fig. 2.1: Growth of cellular telephone subscribers throughout the world.

The wide spread adoption of wireless communications was accelerated in the mid 1990's, when Governments throughout the world provided increased competition and new radio spectrum licenses for Personal communications services (PCS) with 1800-2000 MHz frequency bands.

The wide-spread success of cellular has led to the development of newer wireless systems and standards for many other types of telecommunication traffic besides mobile voice telephone calls.

The next generation cellular networks are being designed to facilitate high speed data communications traffic in addition to voice calls. New standards and technologies are being implemented to support wireless networks to replace fiber optic or copper lines between fixed points several kilometers apart (fixed wireless access).

Similarly, wireless networks have been increasingly used as a replacement for wires within homes, building, and office setting through the extend of Wireless local area networks (WLANS). The evolving Bluetooth modem standard promises to replace troublesome appliance communication cords with invisible wireless connections within a person's personal workspace where he work.

Used primarily buildings, WLANS and Bluetooth use low power levels and generally do not require a licence for spectrum use. These licence-free networks provide an interesting dichotomy in the wireless market, since ad-hoc high data rate networks are being extend by individuals within buildings without a licence, whereas wireless carriers who own the spectrum licences for mobile cellular telephone service have focused on support outdoor voice coverage and have been slow to provide reliable in-building coverage and high data rate services to their cellular subscribers.

This chapter highlights the key developments, technical details, and standards nimbleness of the major modern Wireless communication systems throughout the world.

2.2 SECOND GENERATION (2G) CELLULAR NETWORKS

Most of today's ubiquitous cellular networks use, what is commonly called *second generation* or *2G technologies* which conform to the second generation cellular standards. But in first generation cellular systems that relied exclusively on FDMA/FDD and analog FM Second generation standards use digital modulation formats and TDMA/FDD and CDMA/FDD multiple access techniques used for transmission.

The most popular second generation standards include three TDMA standards and one CDMA standard:

(*a*) Global System Mobile (GSM) which supports eight time slotted users with each' 200 kHz radio channel.

(*b*) Interim standard 136 (IS-136), also known as *North American Digital Cellular* (NADC) or *US Digital Cellular* (USDC), which supports three time slotted users with each 30 kHz radio channel.

(*c*) Pacific Digital Cellular (PDC), a Japanese TDMA standard that is similar to IS-136 with more than 50 million users.

(*d*) The most popular 2G Interim Standard 95 Code Division Multiple Access (IS-95), also known as CDMA one, which supports up to 64 users

that are orthogonally coded and at the same time transmitted on each 1.25 MHz channel.

Fig. 2.2 presents how the world subscriber base was divided between the major 1G and 2G technologies as of late 2001. Table 2.1 highlights the key technical specifications of the dominant GSM, CDMA, and IS-136/PDC second generation standards which is adopted from.

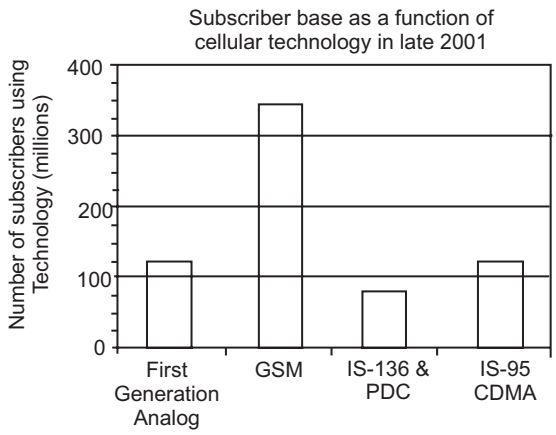


Fig. 2.2: World wide subscriber base as a function of cellular technology in late 2001.

Table 2.1: Key Specifications of Leading 2G Technologies

	CDMA One, IS-95, ANSI J-STD-008, PDC	*GSM, DCS-1900, ANSI J-STD-007,*	*NADC, IS-54/IS-136, ANSI J-STD-0I1, PDC*
Uplink Frequencies	824-849 MHz (US Cellular) 1850-1910 MHz(US PCS)	1890-915 MHz (Europe) 1850-1910 MHz (US PCS)	800 MHz, 1500 MHz (Japan) 1850-1910 MHz (US PCS)
Downlink Frequencies	869-894 MHz (US Cellular) 869-894 MHz (US Cellular) 1930-1990 MHz (US PCS)	935-960 MHz (Europe) 1930-1990 MHz (US PCS)	1930-1990 MHz (US PCS) 800 MHz, 1500 MHz (Japan)
Duplexing Multiple Access	FDD CDMA	FDD TDMA	FDD TDMA
Technology Modulation	BPSK with Quadrature Spreading	GMSK with BT= 0.3	$\pi/4$ DQPSK
Carrier Separation	1.25 MHz	200 kHz	30 kHz (IS-136) (25 kHz for PDC)
Channel Data Rate	1.2288 Mchips/sec	270.833 Kbps	48.6 Kbps (IS-136) (42 Kbps for PDC)

Voice channels per carrier	64	8	3
Speech Coding	Code Excited Linear Prediction (CELP) @ 13 Kbps, Enhanced Variable Rate Codec (EVRC) @8 Kbps	Residual Pulse Excited Long Term Prediction (RPE-LTP) @ 13 Kbps	Vector Sum Excited Linear Predictive Coder (VSELP) @ 7.95 Kbps

Since all 2G technologies offer atleast a three-times increase in spectrum efficiency (and thus atleast a *3x* increase in overall system capacity) as compared to first generation analog technologies, the need to meet a fastly growing customer base justifies the gradual, ongoing change out of analog to digital 2G technologies in any growing wireless network systems.

Most other carriers throughout the world had already committed to adopting a 3G standard based on either GSM or CDMA prior to 2001. Decisions like these have set the stage for the unavoidably of two universal and competing third generation (3G) cellular mobile radio technologies, one based on the philosophy and backward compatibility of GSM, and the other based on the philosophy and backward compatibility of CDMA.

2.2.1 Evolution to 2.5 G Mobile Radio Networks

Before the mid 1990's, the 2G digital standards have been widely extend by wireless carriers for cellular and PCS, even though these standards were designed before the widespread use of the internet.

As a result, 2G technologies use circuit-switched data modems that limit data users to a single circuit-switched voice channel. Data transmissions in 2G are thus generally limited to the data through rate of an particular user and this rate is of the same order of magnitude of the data rate of the designated speech coders given in Table 2.1 with transmitter and receiver parameter.

From inspection of Table 2.1, it can be seen that all 2G networks, as originally developed, only support single user data rates on the order of 10 kilobits per second which is too slow for quick email and internet browsing applications on networks.

Even with relatively small user data rates, 2G standards are able to support limited internet browsing and sophisticated short messaging capabilities using a circuit switched approach. Short Messaging Service (SMS) is a popular feature of GSM and support subscribers to send short, real time message to other subscribers in the same network by simply dialing a recipient's cellphone number on phone.

In an effort to retrofit the 2G standards for compatibility with increased throughout data rates that are required to support modern internet applications, new data-centric standards have been developed that can be overlaid upon

existing 2G technologies. These new standards represent 2.5G technology and allow existing 2G equipment to be modified and supplemented with new base station add-ons and subscriber unit software upgrades to support higher data rate transmissions for web browsing, e-mail traffic, mobile commerce (m-commerce) and location-based mobile services and web page development.

The 2.5G technologies also support a popular new web browsing format language, called *Wireless Application Protocol* (WAP), that support standard web pages to be viewed in a compressed format specifically designed for small, portable hand held wireless devices.

The appropriate 2.5G upgrade path for a particular wireless carries must match the original 2G technology choice made earlier by the same carrier. A wide range of 2.5G standards have been developed to support each of the major 2G technologies (GSM, CDMA and IS-136) to be upgraded incrementally for faster internet data rates in a communication system.

Fig. 2.3 illustrates the various 2.5G and 3G upgrade paths for the major 2G technologies. Table 2.2 describes the required changes to the network infrastructure and the subscriber terminals for the various upgrade options for 2.5G and 3G with example of technology which are used in 2G, 2.5G and 3G.

Table 2.2: Current and Emerging 2.5G and 3G Data Communication Standards

Wireless Data Technologies	*Channel BW*	*Duplix*	*Infrastructure Change*	*News requires Spectrum*	*Requires New Handsets*
HSCSD	200 kHz	FDD	Requires software upgrade at base station.	No	Yes New HSCSD handsets provide Kbps on HSCSD networks, and 9.6 Kbps on GSM networks withdual mode phones GSM only phones will not work in HSCSD networks.
GPRS	200 kHz	FDD	Requires new packet overlay including routers and gateways.	No	Yes New GPRS handsets work on GPRS networks at 171.2 Kbps. 9.6 Kbps on GSM networks withdual nes. GSM-only phones will not work in GPRS networks.

EDGE	200 kHz	FDD	Requires new transceiver at base station. Also, software upgrades to the base station controller and base station.	No	Yes New handsets work on EDGE networks at 384 Kbps. GPRS networks at 144 Kbps, and GSM networks at 9.6 Kbps with tri-mode phones GSM and GPRS-only phones will not work in EDGE netwoks.
W-CDMA	5 MHz	FDD	Requires completely new base stations.	No	Yes New W-CDMA handsets will work on W-CDMA at 2 Mbps. EDGE networks at 384 Kbps, GPRS networks at 144 Kbps, GSM networks at 9.6 Kbps. Older handsets will not work in W-CDMA.
IS-95B	1.25 MHz	FDD	Requires new software in base station controller.	No	Yes New handsets will work on IS-95B at 64 Kbps and IS-95A at 14.4 Kbps. Cdma One phones can work in IS-95B at 14.4 Kbps.
CDMA 2000 1 × RTT	1.25	FDD	Requires new software in back bone and new channel cards at base station. Also need to build a new packet service node.	No	Yes New handsets will work on 1 x RTT at 144 Kbps, IS-95B at 64 Kbps. IS-95A at 14.4 Kbps. Older handsets can work in l × RTT but at lower speeds.
1 × RTT CDMA 2000 1 × EV (DO and DV)	1.25 MHz	FDD	Requires software and digital card upgrade on 1 × RTT network.	No	Yes New handsets will work on 1 × EVA at 2.4 Mbps, 1 × RAT at 144 Kbps. IS-95B at 64 Kbps, IS-95A at 14.4 Kbps. Older handsets can work in 1 × EVA but at lower speeds.

CDMA 2000 3 × RTT	3.75 MHz	FDD	Requires backbone modifications and new channel cards at base station.	May be	Yes New handsets will work on 95A at 14.4 Kbps, 95B at 64 Kbps. 1 × RTT at 144 Kbps, 3 × RTT at 2 Mbps. Older can work in 3 × but at lower speeds.

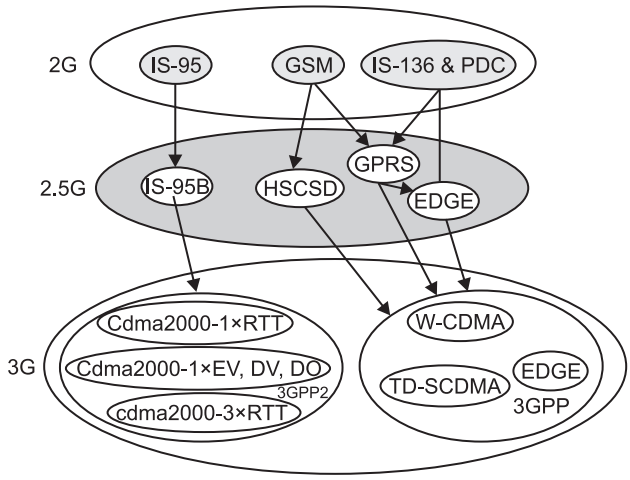


Fig. 2.3: Various upgrade paths for 2G technologies.

2.2.2 Evolution for 2.5G TDMA Standards

Three different upgrade paths have been developed for GSM carriers and two of these solutions also support IS-136. The three TDMA upgrade options are HSCSD, GPRS and EDGE:

(*a*) High Speed Circuit Switched Data (HSCSD);

(*b*) General Packet Radio Service (GPRS) and

(*c*) Enhanced Data Rates for GSM Evolution (EDGE).

These options provide significant improvements in Internet access speed over today's GSM and IS-136 technology and support the development of new Internet-ready cell phone.

2.2.2.1 HSCSD for 2.5G GSM

As the name implies, High Speed Circuit Switched Data is a circuit switched technique that allows a single mobile subscriber to use continuous uses time slots in the GSM standard. That is, instead of limiting each user to only one specific time slot in the GSM TDMA standard.

HSCSD support individual data users to commandeer consecutive time slots in order to obtain higher speed data access on the GSM network to access internet service.

HSCSD relaxes the error control coding algorithms originally specified in the GSM standard for data transmissions and increases the available applications data rate to 14,400 bps, as compared to the original 9,600 bps in the GSM specification parameters.

By using up to four consecutive time slots, HSCSD is able to provide a raw transmission rate of up to 57.6 Kbps to individual users, and this enhanced data offering can be billed as a premium service by the carrier in the network.

HSCSD is ideal for dedicated streaming Internet access or real time interactive web sessions and simply requires the service provider to implement a software change at existing GSM base station for communication.

2.2.2.2 GPRS for 2.5G GSM and IS-136

General Packet Radio Service is a packet-based data network, which is well suited for non-real time internet usage, including the recovering of email, faxes, and asymmetric web browsing, where the user down loads much more data than it uploads on the Internet networks.

Unlike HSCSD which dedicates circuit switched channels to specific users. GPRS supports multiuser network sharing of individual radio channels and time slots. Thus GPRS can support many more users than HSCSD, but in a bursty style technique.

Similar to the Cellular Digital Packet Data (CDPD) standard developed for the North Amercian AMPS systems in the early 1990's, the GPRS standard provides a packet network on dedicated GSM or IS-136 radio channels for transmit the data from one point to another.

GPRS retains the original modulation formats specified in the original 2G TDMA standards, but uses a completely redefined air interface in order to better handle packet data access on internet network.

GPRS subscriber units are automatically instructed to tune to dedicated GPRS radio channels and particular time slots for "always on" access to the network services.

When all eight time slots of a GSM radio channel are dedicated to GPRS an individual user is able to achieve as much as 171.2 Kbps (eight time slots channels multiplied by 21.4 Kbps of raw uncoded data throughput).

GPRS has been installed in markets serving over 100 million subscribers, and is to be supported the most popular near-term packet data solution for 2G TDMA based technologies. The dedicated peak 21.4 Kbps per channel data rate specified by GPRS works well with both GSM and IS-136 and has successfully been implemented using some extra special services.

2.2.2.3 EDGE for 2.5 GSM and IS-136

EDGE, which stands for Enhanced Data rates for GSM (or Global) Evolution is a more advanced upgrade to the GSM standard and requires the addition of new hardware and software standard at base stations.

EDGE was developed from the desire of both GSM and IS-136 operators to have a common technology path for eventual 3G high speed data access, but the initial impetus came from the GSM user group.

EDGE introduces a new digital modulation format 8-PSK (Octal phase shift keying) which is used in addition to GSM's standard GMSK modulation format.

EDGE allows for nine different (autonomously and rapidly selectable) air interface formats known as Multiple Modulation and Coding Schemes (MCS) with varying degrees of error control protection.

EDGE is sometimes referred to us Enhanced GPRS or EGPRS for application point of view.

EDGE uses the higher order 8-PSK modulation and a family of MCSs for each GSM radio channel time slot, so that each user connection may adaptively determine the best MCS setting for the particular radio propagation state and data access requirements of the user.

When EDGE uses 8-PSK modulation without any error protection and all eight times slots of a GSM radio channel are dedicated to a single user, a maximum throughput data rate of 547.2 Kbps can be provided to a user.

In practice, the slotting schemes used in EDGE, when combined with practical network connection issues and error control coding requirements. Limits practical raw data rates to about 384 kilobits per second for a particular or fixed user on a single GSM channel.

By combining the capacity of different radio channels (*e.g.,* using multicarrier transmissions), EDGE can provide up to several megabits per second of data throughput to individual data users to transmit the data.

2.2.3 IS-95B for 2.5G CDMA

Unlike the several GSM and IS-136 evolutionary paths to high speed data access, CDMA (often called cdma one) has a single upgrade path for eventual 3G operation purpose.

Like GPRS, IS-95B is already being deployed world wide, and provides high speed packet and circuit switched data access on a common CDMA radio channel by dedicating multiple orthogonal user channels (Walsh functions) for a particular users and a distinct purposes.

Each IS-95 CDMA radio channel supports up to 64 different user channels with in a network.

The original IS-95 throughput rate specification of 9600 bps was not implemented in practice, but was improved to the current rate of 14,400 bps as specified in IS-95 A for better transmission.

The 2.5G CDMA solution, IS-95B, supports Medium data rate (MDR) service by allowing a dedicated user to command up to eight different user Walsh codes at a same time and in parallel for an instantaneous throughput of 115.2 Kbps per user (8 × 14.4 Kbps).

IS-95B also specifies hard hand-off procedures that allow subscriber units to search distinct radio channels in the network without instruction from the switch so that subscriber units can rapidly tune to different base stations to maintain link quality system.

Prior to IS-95B, the link quality experienced by each subscriber had to be reported back to the switch through the serving base station several hundreds of times per second, and at the suitable moment, the switch would initiate a soft-hand-off between the subscriber and a specified base stations.

2.3 THIRD GENERATION (3G) WIRELESS NETWORKS

3G systems promise unparalleled wireless access in ways that have never been possible before. Multimegabit Internet access, communications using Voice Over Internet Protocol (VOIP) voice-activated calls, unique the network capacity and ubiquitous always-on" access are just some of the advantages being touted by 3G developers.

Companies developing 3G equipment envision users having the ability to receive live music, conduct interactive web sessions, and have same time capability of voice and data access with multiple parties at the same time using a single mobile handset, whether driving, walking, or standing still in an office setting or in running condition.

The eventual 3G evolution for 2G CDMA systems leads to CDMA 2000. Several ariants of CDMA 2000 are currently being developed, but they all are based on the primary of IS-95 and IS-95B technologies.

The eventual 3G evolution for GSM, IS-136 and PDC systems leads to wide band DMA (W-CDMA) also called Universal Mobile Telecommunications Service (UMTS). W-CDMA is identify by the network fundamentals of GSM, as well as the merged versions of GSM and IS-136 through EDGE technique.

It is fair to say that these two major 3G technology camps, cdma 2000 and W-*CDMA,* will remain popular throughout the early part of the 21st century.

Countries throughout the world are currently define the new radio spectrum bands) accommodate the 3G networks that will likely be deployed in the 2004-05 time frame. ITU's 2000 World Radio Conference to set and fixed the 2500-2690 MHz, 1710-1885 MHz and 806-960 MHz bands as candidates for 3G.

2.3.1 3G W-CDMA (UMTS)

The Universal Mobile Telecommunications System (UMTS) is a visionary air interface standard that has developed since late 1996 under the auspices of the European Telecommunication Standards Institute (ETSI) organization.

UMTS was submitted by ETSI to ITU's EVIT-2000 body in 1998 for consideration as a world standard. At that time, UMTS was known as UMTS Terrestrial Radio access (UTRA) and it is explain in Table 2.3 and was designed to provide a high capacity upgrade path for GSM system.

UMTS, or W-CDMA, assures backward compatibility with the second

generation GSM, IS-136 and PDC TDMA technologies, as well as all 2.5G TDMA technologies.

The network structure and bit level packaging of GSM data is reserved by W-CDMA, with additional capacity and bandwidth provided by a new CDMA air interface.

Fig. 2.3 illustrates how the various 2G and 2.5G TDMA technologies will evolve to a unified W-CDMA standard with different technologies.

The 3G W-COMA air interface standard had been designed for "always-on" packet-based wireless service, so that computers, entertainment devices, and telephones may all share the same wireless network and be connected to the Internet, anytime, anywhere and any purpose.

W-CDMA will support packet data rates up to 2.048 Mbps per user (if the user is stationary), thereby allowing high quality data, multimedia, streaming audio, streaming video and broadcast-type services to consumers with conference facility also.

Table 2.3: Leading IMT-2000 Candidate Standards as of 1998

Air Interface	*Mode of Operation*	*Duplixing Method*	*Key Features*
CDMA 200 US TIA TR 45.5	Multi-carrier and Direct Spreading DS-CDMA a N = 1.2288 Mcps with N =1, 3, 6, 9, 12	FDD and TDD Modes	• Backward compatibility with IS-95A and IS-95B. Downlink can be implemented using either Multi-Carrier or Direct Spreading. Uplink can support a simultaneous combination of Multi-Carrier or Direct Spreading. • Auxiliary carriers to help with downlink channelestimation in forward link beam-forming.
UTRA (UMTS) Terrestrial Radio Access) ESTI SMG2 W-CDMA/NA Wideband (CDMA) (North America) USA TIPI-ATIS W-CDMA/Japan (wideband CDMA) Japan ARIB CDMA II South Korea TTA WIMS/W-CDMA USA TIA TR 46.1	DS- CDMA at Rates of N × 0.960 Mcps with N = 4, 8, 16	FDD and TDD Modes	• Wideband DS-CDMA System. • Backward compatibility with GSM/DSC-1900. • Up to 2.048 Mbps on Downlink in FDD Mode. Minimum forward channel • bandwidth of 5 GHz. The collection of proposed • standards represented here each exhibit unique features, but support a common set of chip rates, 10 ms frame structure, with 16 slots per frame. Connection-dedicated pilot • bits assist in downlink beam forming.

CDMAI South Korea TTA	Ds-CDMA at N × 0.9216 Mcps with N = 1,4, 16	FDD and TDD Modes	• Up to 512 Kbps per spreading code, code aggregation up to 2.048 Mbps.
UWC-136 (Universal Wirless Communications Consortium) USA TIA TR 45.3	TDMA-Up to 722.2 Kbps (outdoor/ vehicular), Up to 5.2 Mbps (Indoor office)	FDD (Outdoor/ Vehicular TDD Indoor Office)	• Backward compatibility and upgrade path forboth IS-136 and GSM. • Fits into existing IS-136 and GSM. • Explicit plans to support adaptive antenna technology.
TD-SCDMA China Academy of Telecomm. Technology (CATT)	DS-CDMA 1.1136 MCps	TDD	• RF channel bit rate up to 2.227 Mbps. • Use of smart antenna technology is fundamental (but not strictly required) in TD-SCMA.
DECT ETSI Project (EP) DECT	1150-3456 kbps TEMA	TDD	• Enhanced version of 2G DECT technology.

Future versions of W-CDMA will support stationary user data rates in excess of 8 Mbps.

W-CDMA provides public and private network features, as well as Video conferencing and virtual home entertainment (VHE) facility.

W-CDMA designers contemplate that broadcasting mobile commerce (m-commerce), games, interactive video, and virtual private networking will be possible throughout the world, all from a small portable wireless device.

W-CDMA demand a minimum spectrum allocation of 5 MHz, which is an important distinction from the other 3G standards.

Although W-CDMA is designed to support backward compatibility and interpretability for all GSM, IS-136/PDC, GPRS, and EDGE switching equipment and applications. It is clear that the wider air interface bandwidth of W-CDMA requires a complete change out of the RF equipment at each base station in a network.

With W-CDMA, data rates from as low as 8 Kbps to as high as 2 Mbps will be carried simultaneously on a single W-CDMA 5 MHz radio channel, and each channel will be able to uphold between 100 and 350 simultaneous voice calls at once, depending on antenna sectoring, propagation conditions, user velocity and antenna polarizations.

Because W-CDMA will require costly new base station equipment, the installation of W-CDMA will likely be slow and gradual throughout the world. In this way the evolutionary path to 3G will require dual mode or tri-mode cell phones that can automatically switch between the incumbent 2G TDMA technology, EDGE, or W-CDMA service where its available.

By 2010, it is likely that W-CDMA will be completely installed, eliminating the need for backward compatibility with GSM/GPRS, IS-136, PDC and EDGE technologies.

2.3.2 3G CDMA 2000

The CDMA 2000 vision provides a seamless and evolutionary high data rate upgrade path for current users of 2G and 2.5G CDMA technology. With the help of a building block approach that centers on the original 2G CDMA channel bandwidth of 1.25 MHz per radio channel for transmission.

Based upon the original IS-95 and IS-95A (CDMA One) CDMA standards, as well as the 2.5G IS-95B air interface the cdma 2000 3G standard supports wireless carriers to introduce a family of new high data rate Internet access capabilities in a gradual manner with in existing systems, while assuring that these upgrades maintain backward compatibility with existing cdma One and IS-95B subscriber equipment or instruments.

So, current CDMA operators may seamlessly and selectively introduce 3G capabilities at each cell without having to change out entire base stations or reallocate spectrum system.

The cdma 2000 standard is being developed under the auspices of working group 45 of the Telecommunications Industry Association (TIA) of the US, and involves as a member of the International technical community through the 3G PP2 working group society.

cdma 2000 1X supports an instantaneous data rate of up to 307 Kbps for a user in packet mode, and yields typical throughput rates of up to 144 Kbps per user, depending on the quantity of users, the velocity of a user, and the propagation conditions.

cdma 2000 1X can also support up to twice as many voice user as the 2G CDMA standard and also support the subscriber unit with upto two times the standby time for longer lasting battery life.

cdma 2000 is being developed for both FDD (Mobile radio) and TDD (In-building cordless) applications.

With the help of cdma 2000 1X EV technology, individual 1.25 MHz channels may be installed in CDMA base stations to provide specific high speed packet data access within selected cells in a network.

cdma 2000 supports both voice and data users, and can offer usable data rates upto 144 kilobits per second with about twice as many voice channels as IS-95B technology.

The cdma 2000 standard uses three adjacent 1.25 MHz radio channels that are used together to provide instantaneous packet data throughput speeds in increase of 2 Mbps per user, although actual throughput depends upon cell loading, vehicle speed, and propagation conditions and atmosphere effects.

With peak user data rates in excess of 2 Mbps, it is clear that cdma 2000 3X has a very similar user data rate throughput target when compared to W-CDMA (UMTS).

Advocates of cdma 2000 claim their standard gives a wireless service provider a much more seamless and less expensive upgrade path when compared to W-CDMA since cdma 2000 allows the same spectrum, bandwidth, RF equipment, and air interface framework to be used at each base station as the 3G upgrades are introduced overtime.

2.3.3 Characteristics of Third Generation Systems

Provide world wide web according to requirement.

(*i*) Accesses information through broadband Integrated Service Digital Network (ISDN) and PSTN also.

(*ii*) Information may be in form of voice data and video format.

(*iii*) Enables a person of communicate with anyone at any time at any place.

(*iv*) Global roaming facility to the user and network.

(*v*) Provide more reliable service feature.

2.3.4 Examples of 3G Systems

(*i*) IMT-2000

(*ii*) UMTS

(*iii*) MBS

(*iv*) WLAN.

2.3.5 3GTD-SCDMA

In China, GSM is the most popular wireless air interface standard, and the wireless subscriber growth in China is unmatched anywhere in the world. For example, in late 2001 more than eight million cell phone subscribers were added in just one month in China alone! Given the huge potential market for wireless services in China, and China's desire to craft its own wireless vision, The China Academy of Telecommunications Technology (CATT) and Siemens Corporation jointly submitted an IMT-2000 3G standard proposal in 1998, based on Time Division-Synchronous Code Division Multiple Access (TD-SCDMA). This proposal was adopted by ITU as one of the 3G options in late 1999.

TD-SCDMA relies on the existing core GSM infrastructure and allows a 3G network to evolve through the addition of high data rate equipment at each GSM base station. TD-SCDMA combines TDMA and TDD techniques to provide a data-only overlay in an existing GSM network. Up to 384 kbps of packet data is provided to data users in TD-SCDMA.

The radio channels in TD-SCDMA are 1.6 MHz in bandwidth and rely on smart antennas, spatial filtering, and joint detection techniques to yield several times more spectrum efficiency than GSM. A 5 millisecond frame is used in

TD-SCDMA, and this frame is subdivided into seven time slots which are flexibly assigned to either a single high data rate user or several slower users. By using TDD, different time slots within a single frame on a single carrier frequency are used to provide both forward channel and reverse channel transmissions. For the case of asynchronous traffic demand, such as when a user downloads a file, the forward link will require more bandwidth than the reverse link, and thus more time slots will be dedicated to providing forward link traffic than for providing reverse link traffic. TD-SCDMA proponents claim that the TDD feature allows this 3G standard to be very easily and inexpensively added to existing GSM systems.

2.4 FOURTH GENERATION (4G) WIRELESS NETWORKS

4G stands for the Fourth generation of cellular wireless standards. It is a successor to 3G and 2G families of standards. Speed requirements for 4G service set the peak download speed at 100 Mbit/s for high mobility devices (such as cellphones) and lGbit/s for low mobility devices (such as standalone wireless modems).

A 4G system is expected to provide a comprehensive and secure all-IP based solution where facilities such as **ultra-broadband** internet access, **IP telephony,** gaming services, and streamed multimedia may be provided to users. The wireless telecommunications industry as a whole has generally adopted the term 4G as a short hand way to describe those advanced" wireless technologies that, among other things, are based on or employ wide channel OFDM technology and all-IP based architecture. In the US, sprint's 4G technology meets these parameters (LTE does too).

"In 4G the user has freedom and flexibility to select any desired service with reasonable *QOS* and affordable price, anytime and anywhere."

2.4.1 Features of 4G

4G is described as **MAGIC**–Mobile multimedia, Any time anywhere, Global mobility support, Integrated wireless solution, and customised personal service. Some key features of 4G mobile networks are as follows:

(*i*) *High usability* : anytime, anywhere and with any technology.

(*ii*) Support for multimedia services at low transmission cost.

(*iii*) Personalization

(*iv*) Integrated

Figure 2.4 below shows the architecture of the 4G network

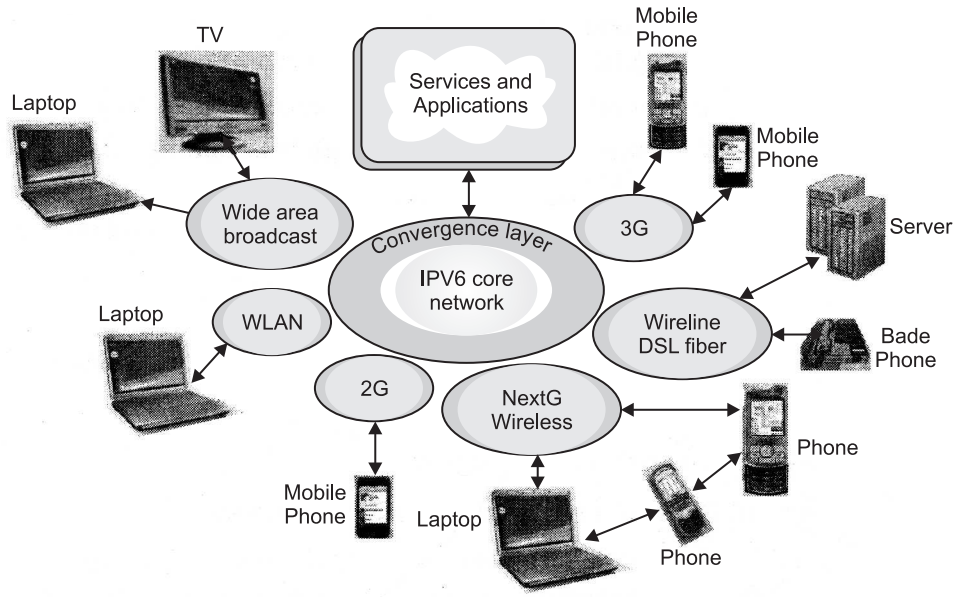


Fig 2.4: An architecture of 4G network.

2.5 COMPARISON WITH 3G

Attribute	*3G*	*4G*
Major Characteristic	Predominantly voice-data as add-on	Converged data and VolP
Network Architecture	Wide area cell based	Hybrid-integration of Wireless Lan (WiFi), Blue Tooth Wide Area
Frequency Band	16-25 GHz	2-8 GHz
Component Design	Optimized antenna multiband adapters	Smart antennas, SW multi-band wideband radios
Bandwidth	5-20 MHz	100 + MHz
Data Rate	385 kbps-2 Mbps	100 Mbps-1 Gbps
Access	WCDMA/CDMA2000	MC-CDMA or OFDM
Forward Error Correction	Convolution code 1/2, 1/3, turbo	Concatenated Coding
Switching	Circuit/Packet	Packet
Mobile top Speed	200 kmph	200 kmph
IP	Multiple versions	All IP (IPV6)
Operational	2003	2010

2.6 FUTURE TRENDS IN WIRELESS COMMUNICATION

In modern years, according to customer demands and working standards, considerable research and development is underway. Some innovative technologies showing the future trends in wireless communication, are given as follows and explain also with application:

1. **Bluetooth**: Bluetooth is an open standard that is used to enable various devices to communicate. With one another within a nominal 10m range in a communication network.
   (*i*) Bluetooth have their places in today's offices and homes, it works for setting up networks, printing or transferring presentations and files.
   (*ii*) Bluetooth is acceptable for situations when two or more devices are connected closely to each other and don't require high bandwidth. Since bluetooth uses short range radio frequencies.
2. **Wi-Fi**: It is known as wireless-fidelity. It is a LAN technology based on the IEEE 802.11 standards for communication. Wi-Fi covers not only office-based LAN, but also home based LAN, Wi-Fi is better suited for developing the networks as it enables faster connection and better security and better performance.
3. **Wi-Max**: Wi-Max is similar to Wi-Fi but has extra range capability with extra security feature.
   (*i*) Wi-Max has a range of 40–50 km while Wi-Fi range is limited to several hundred meters.
   (*ii*) Therefore Wi-Max provide a wireless alternative to cable and wired communication with dedicated connectivity.
   (*iii*) Wi-Max operation is based on IEEE 802.16 standards for communication.
4. **Mobile-Fi**: It is similar to mobile version of Wi-Max in terms of technology and support the mobile communication.
   (*i*) The objective with Mobile-Fi is to provide internet access to mobile user at data rates even higher than the data rate in home broadband links. Thus a mobile user could enjoy broadband internet access while travelling in a moving car or train or walking.
   (*ii*) Mobile-Fi is based on IEEE 802.20 standards specifications.
5. **Ultra wideband**: It is quite different than other technology, it is enable to transfer bulk data as well as small amount of data also.
   (*i*) Ultra wideband enables the moment of message files at high data rates over short distance. Audio and video data can also be transmit by using ultra wideband.

2.7 REQUIREMENTS IN WIRELESS COMMUNICATIONS

Wireless communication is based on specific protocols. Although each protocol has different specifications and criteria but there are general characteristic and requirements that each protocol tries to achieve a fixed target or goal.

Given below some general requirements :

(*i*) Guarantee of delivery the information.

(*ii*) Unlimited roaming and range coverage.

(*iii*) Dependability of delivery of message.

(*iv*) Notification by the user and networks.

(*v*) Millions of users connect.

(*vi*) Priority alerts.

(*vii*) Host mobility.

HIGHLIGHTS

1. The current state of the art of broadband milimeter wave systems were presented.
2. Along with the vision for wireless LMDS broadband networks as a compliment to the existing fiber backbone.
3. The first generation systems were based on analog technology and FDMA.
4. The third generation system mobile world wide access and enable to communicate with a person at any time at any place.
5. To obtain the ideal mobile telephone system, there are three approaches are used:
   (*i*) The single side band (SSB), and it divide the allocated frequency band into maximum number of channels.
   (*ii*) The cellular, and it reuse the allocated frequency band in different location.
   (*iii*) Frequency hopped or spread spectrum system in which generates many codes over a wide frequency band.
6. Smart antenna, software radio, together with advanced base station are the key technique to achieve adaptability of wireless access point to diverse terminals *i.e.* to make ratio system and air networks reconfigurable.
7. The mobile application should refer to a user's profile so that it can be delivered in a way most preferred by the subscriber, such as context based personalized services.
8. The starting time for first generation system is 1985, for second generation system is 1992, for third generation system is 2002.
9. The first generation technology used the 2.4k to 30k bits per message bandwidth.
10. The second generation technology used the 9.6k to 14.4k bits per message bandwidth.
11. The 2.5 generation technology used the 171k to 384k bits per message bandwidth.
12. The third generation technology used the 2m to 5m bits per message bandwidth.
13. The driven technique used in 1G is analogue processing, in 2G is digital signal processing, in 2G is packet switching and in 3G is intelligent signal processing.
14. The fouth generation technology uses flexible chanel bandwidth, between 5 and 20 MHz.

REVIEW QUESTIONS

2.1. What was the first "radio voice link"? In which year it is developed.
2.2. Write down the history of radio communication in brief?
2.3. Write down the classification of mobile radio system?
2.4. How MTS is different from MTS?
2.5. Explain the meaning of modern wireless communication system?
2.6. Explain in brief:
(*i*) 1G cellular networks
(ii) 2G cellular networks
(*iii*) 3G cellular networks.
2.7. What is the basic difference between 1G and 2G technology?
2.8. Why the 3G technology are developed?
2.9. In your home, how many modern wireless communication networks are available to you?
2.10. Write short note on future trends in wireless communication?
2.11. Explain in brief-HSCSD for 2.5G GSM.
2.12. What do you mean by EDGE for 2.5 GSM and IS-136?
2.13. What are the basic requirements in wireless communications?
2.14. Explain in brief – Universal Mobile Telecommunication Systems (UMTS).
2.15. Give some example of 3G technology?
2.16. List major technical difference between CON A2000 and w-cdma.
2.17. Compare 19 1G, 2G, 3G, 4G systems.
2.18. What data rates may be offered by 4G systems.

3 Chapter

MOBILE RADIO PROPAGATION: LARGE-SCALE PATH LOSS, SMALL-SCALE FADING AND MULTIPATH

3.1. *Introduction to Radio Wave Propagation* **3.2.** *Free Space Propagation Model* **3.3.** *The Three Basic Propagation Mechanisms* **3.4.** *Reflection* **3.5.** *Diffraction* **3.6.** Scattering **3.7.** *Practical Link Budget Design Using Path Loss Models* **3.8.** *Outdoor Propagation Models* **3.9.** *Indoor Propagation Models* **3.10.** *Small Scale Multipath Propagation* **3.11.** *Small Scale Multipath Measurement* **3.12.** *Parameters Of Mobile Multipath Channels* **3.13.** *Type Of Small Scale Fading* **3.14.** *Two-Ray Rayleigh Fading Model* **3.15.** *Multipath Shape Factor* **3.16.** *Fading Rate Variance Relationships* **3.17.** *Comparison to omnidirectional Propagation* **3.18.** *Second-Order Statistics using Shape Factors* **3.19.** *Revisiting Classical Channel Models with Shape factors.*

3.1 INTRODUCTION TO RADIO WAVE PROPAGATION

We know that a mobile radio channel have various limitations on the performance of wireless or mobile communication system.

The two fundamental part of any communication system are transmitter and receiver.

In wireless communication, the transmission path between the transmitter and the receiver can vary from simple line of sight and that is severely obstructed by buildings, mountains and foliage.

Further, basically, there are following two types of communication channels:

(*i*) Wired channels and

(*ii*) Radio or wireless channels.

The wired channels are stationary and predictable but the radio channels are extremely random and do not offer easy analysis. Even speed of motion affects how rapidly the signal level fades as a mobile terminal moves in space. So, it has been more difficult to model a radio channel in mobile radio system design. However, the model of a radio channel in mobile radio system design is obtained in a statistical manner which is based upon the measurement observed for a desired communication system or spectrum allocation.

3.2 FREE SPACE PROPAGATION MODEL

The free space propagation model is used to predict received signal power when the transmitter and receiver have a clear, unobstructed line of sight path

between them for example satellite communication system and microwave line of sight radio links typically undergo free space propagation.

Consider a large scale radio wave propagation models, the free space radio wave propagation model predicts that the received power decays as a function of the transmitter and receiver separation distance raised to same power.

Let us consider an isotropic source radiating a total power denoted by P_T, calculate in watts. The radiated power passes uniformly through a sphere of surface area where d is distance (in meter) from the source. So, the power density (P_D) at any point on the surface of the sphere is given by

$$P_D = \frac{P_T}{4\pi d^2} \text{watts/m}^2 \quad ...(3.1)$$

Equation (3.1) show that the power density varies inversely as the square of the distance from a point source and it directly proportional to the transmitted power.

Effective Isotropic Radiated Power (EIRP)

The EIRP is defined as the product of the transmitted power P_T and the power gain of the transmitting antenna G_T *i.e.* show by the following equation.

Effective isotropic radiated power

$$\text{EIRP} = P_T G_T \text{ (watts)} \quad ...(3.2)$$

The EIRP represents the maximum radiated power available from a transmitter in the direction of maximum antenna gain, as compared to an isotropic radiator.

Effective Aperture

Effective aperture represent the ratio of the power available at the antenna terminal to power per unit area of the suitable polarized incident electromagnetic wave, and it is denoted by A_{eff}.

$$A_{eff} = \frac{\lambda^2}{4\pi} G \quad ...(3.3)$$

where λ is wavelength

$\therefore \lambda = c/f$ where c is speed of light (3×10^8 m/s) and/is frequency.

Friss Free-Space Equation

Now let us assume a transmitting antenna with an EIRP defined in equation (3.2), the power density of the transmitting antenna as the following equation:

$$P_D = \frac{\text{ELRP}}{4\pi d^2} \quad ...(3.4)$$

where d is the distance between the receiving and transmitting antennas.

Now, the power P_R, absorbed by the receiving antenna is the product of this power density and the antenna's effective area denoted by A_{eff} *i.e.* shown by following equation:

$$P_R = \left(\frac{\text{EIRP}}{4\pi d^2}\right) A_{eff} \quad ...(3.5)$$

or $$P_R = \frac{P_T G_T . A_{eff}}{4\pi d^2} \text{ watts} \quad ...(3.6)$$

$\therefore$ Because EIRP = $P_T G_T$ from equation (3.2)

The effective area A_{eff} from equation (3.2) can be express for receiving antenna as the following equation:

$$A_{eff} = \frac{\lambda^2}{4\pi} G_R \quad ...(3.7)$$

So from Eqn. (3.6),

$$P_R = P_T G_T G_R \left(\frac{\lambda}{4\pi d}\right)^2 \quad ...[3.8(a)]$$

where,

P_T = Total power radiated by an isotropic source,
G_T = Transmitting antenna gain in dB,
G_R = Receiving antenna gain in dB,
d = Distance between transmitting and receiving antennas,
X = Wavelength of carrier = *clf in* meter,
c = 3×10^8 m/s, velocity of light in meter/second,
f = Carrier frequency in Hz.

Eqn. [3.8 (*a*)] is called free-space equation.

Eqn. [3.8 (*a*)] can be written as

$$\frac{P_R}{P_T} = \frac{\lambda^2 G_T G_R}{16\pi^2 d^2} \quad ...[3.8\ (b)]$$

The gain and attenuation are usually expressed in dB rather than directly as power ratios; the distance between transmitter and receiver is more likely to be measure in kilometers than meters and the frequency in MHz so the equation [3.8 (*b*)] can be written as

$$\frac{P_R}{P_T}(dB) = [G_T(dB_i)] + [G_R(dB_i)] \quad ... (3.9)$$

$$= [32.44 + 20 \log d + 20 \log f]$$

The term dBi indicates that the antenna gains are given with respect to an isotropic radiator in dB.

G_T (dB$_i$) = 10 log (Power density in the direction of the receiver from the transmitting antenna divided by power density in the some direction from the isotropic radiator with same input power) and other quantity are defined in previous discussion.

Equation (3.9) is expressed as decibel gain between transmitting and receiving antenna but the received signal is weaker than the transmitted signal so this gain is always negative. Negative gain are more commonly and it is called losses and

equation (3.9) can be written as a loss. The loss thus *called free space loss or path loss* and can be written as

$$L_{fs} = 32.44 + [20 \log d(\text{km})] + [20 \log f(\text{MHz})] - [G_T \text{dB}_i)] - [G_R(\text{dB}_i)] \quad ...(3.10)$$

where $$L_{fs} = 10\log \frac{P_T}{P_R}$$

and everything else is as shown in equation (3.9).

The far-field or Fraunhofer region of a transmitting antenna is defined as the region beyond the far-field distance *dp* which is related to the largest linear dimension of the transmitter antenna aperture and the carrier wavelength, so the distance is defined as

$$d_f = \frac{2D^2}{\lambda} \quad ...(3.11)$$

where D is the largest physical linear dimension of antenna. Additionally to be in the far-field region, *df* must satisfy the following conditions:

$$d >> D \quad ...(3.12)$$

$$d_f >> \lambda \quad ...(3.13)$$

The received power $P_r(d)$, at any distance $d > d_0$ may be related to P_r at d_0. The value $P_r(do)$ may be predicted from equation [3.8 *(a)*].

The reference distance must be chosen such that it lies in the far-field region, that is $d_0 > df$ and d_0 is chosen to be smaller than any practical distance used in the mobile communication system. Then equation [3.8 (*a*)] can be as, at distance greater than d_0 then

$$P_r(d) = P_r(d_0)\left(\frac{d_0}{d}\right)^2$$

$$d \geq d_0 \geq d_f \quad ...(3.14)$$

Equation (3.14) may be expressed in units of dBm or dBw by simply taking the logarithm of both sides and multiplying by 10 for example if P_r is in units of dBm, the received power is expressed as

$$P_r(d)dBm = m10\log\left[\frac{P_r(d_0)}{0.001W}\right] + 201\log\left(\frac{d_0}{d}\right) d \geq d_0 \geq d_f \quad ...(3.15)$$

where $P_r(d_0)$ is in watts or measure in watts.

The reference distance d_0 for practical system using low gain antennas in the 1–2 GHz region is typically chosen to be lm in indoor environments and 100 m or 1 km in outdoor environments. These are the range for indoor and outdoor environments.

3.3 THE THREE BASIC PROPAGATION MECHANISMS

Reflection, diffraction and scattering are the three basic propagation mechanisms which collision propagation in a mobile communication system.

Received power is generally is the most important parameter predicted by large scale propagation model based on the physics of reflection, diffraction and scattering mechanisms.

Reflection occurs when a propagation electromagnetic wave impinges upon an object which has very large dimensions when compared to the wavelength of the propagating wave. Reflection occurs from surface to earth and from buildings, walls and other object etc.

Diffraction occurs when the radio path between transmitter and receiver is obstructed by a surface that has sharp edges. The secondary waves finally from the obstructing surface are present throughout the space and even behind the obstacle, given rise to a bending of waves around the obstacle even when a line of sight path does not lie between transmitter and receiver.

Scattering occurs when the medium through which the wave travels and consist of objects with dimensions that are small compared to the wavelength and where the number of obstacles per unit volume is large. Scattered wave are produce by rough or hard surface, small objects or by other irregularities in the channel between transmitter and receiver.

3.4 REFLECTION

Reflection occurs when a propagating electromagnetic wave strike upon an object which has very large dimensions when compared to the wavelength of the propagating wave.

When a radio wave propagating in one medium strike upon another medium having different electrical property, the wave is partially reflected and partially transmitted from the medium.

If the plane wave is incident on a perfect dielectric, part of energy is transmitted into the second medium and part of the energy is reflected back into the first medium, and there is no loss of energy in absorption. If the second medium is perfect conductor, then all incident energy is reflected back into the first medium without loss of energy.

The electric field intensity of the reflected and transmitted waves may be related to the incident wave in the medium of origin through the using Fresnel reflection coefficient = (Γ).

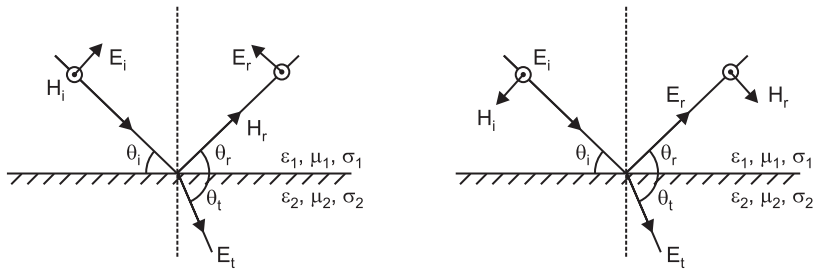
The reflection coefficient is a function of the material property and generally depend on the wave polarization, angle of incidence and the frequency of the propagating wave in a medium.

3.4.1 Reflection from Dielectrics

Fig. 3.1 shows an electromagnetic wave incident at an angle θ_i with the plane of the boundary between two dielectric media.

As shown in the Fig. part of the energy is reflected back to the first medium and at an angle θ_r and part of energy is transmitted (refracted) into the second

medium at an angle θ_i. The nature of reflection varies with the direction of polarization of the E-field is shown in following Figures.



(a) E-field in the plane of incidence. (b) E-field normal to the plane of incidence.

Fig. 3.1: Calculating the reflection coefficients between two dielectrics.

In Fig. 3.1 (*a*), the E-field polarization is parallel with the plane of incidence and in Fig. 3.1 (*b*), the *E*-field polarization is perpendicular to the plane of incidence with the direction of E_i, E_r and E_t.

In Fig. 3.1, the subscripts *i, r, t* refer to the incident, reflected and transmitted fields respectively and parameter E_1, μ_1, σ_2 and E_2, μ_2, σ_2 represent the permittivity, permeability and conductance of the two media respectively as shown in Fig. 3.1.

Often the dielectric constant of a perfect dielectric is related to a relative value of permittivity E_r, such that $\varepsilon = \varepsilon_0\varepsilon_r$, where σ_0 is a constant given by 8.85×10^{-12} F/m and it is a standard value.

If a dielectric material is lossy, it will absorb power and may be described by a complex dielectric constant given by following equations:

$$\varepsilon = \varepsilon_0\varepsilon_r - j\varepsilon' \quad ...(3.16)$$

where,

$$\varepsilon' = \frac{\sigma}{2\pi f} \quad ...(3.17)$$

and σ is the conductivity of material.

Because of superposition, only two orthogonal polarizations need to be considered to solve general reflection problems. The reflection coefficient for the two cases of parallel and perpendicular *E*-field polarization at the boundary of two dielectrics are given by following:

$$\Gamma_{11} = \frac{E_r}{E_i} = \frac{\eta_2 \sin\theta_t - \eta_1 \sin\theta}{\eta_2 \sin\theta_t + \eta_1 \sin\theta} \text{ (}E\text{-field in plane of inicidence)} \quad ...(3.18)$$

$$\Gamma_1 = \frac{E_r}{E_i} = \frac{\eta_2 \sin\theta_i - \eta_0 \sin\theta_t}{\eta_2 \sin\theta_i - \eta_1 \sin\theta_t} \text{ (}E\text{-field normal to the plane of incidence)} \quad ...(3.19)$$

where η_i, is the intrinsic impedance of the *i*th medium (i = 1, 2...) and is given by $\sqrt{\frac{\mu_i}{\varepsilon_i}}$,

the ratio of electric to magnetic field for a uniform plane wave in the particular medium is define the intrinsic impedance.

The velocity of an electromagnetic wave is given by $\frac{1}{\left(\sqrt{\mu\varepsilon}\right)}$, and the boundary conditions at the surface of incidence obey Snell's law which refers to Fig. 3.1, is given by following equation:

$$\sqrt{\mu_1\varepsilon_1 \sin(90-\theta_i)} = \sqrt{\mu_2\varepsilon_2 \sin(90-\theta_t)} \quad ...(3.20)$$

The boundary conditions from Maxwell's equations are used to derive equation (3.18) and (3.19) as well as equation 3.21, 3.22, 3.23.

$$\theta_i = \theta_r \quad ...(3.21)$$

$$E_r = \Gamma E_i \quad ...(3.22)$$

$$E_t = (1 + \Gamma)E_i \quad ...(3.23)$$

where Γ is Γ_{11} or Γ_1 depending on whether the E-field is in (vertical) or normal (horizontal) to the plane of incidence.

When the first medium is free space and $\mu_1 = \mu_2$, the reflection coefficients for the two case of vertical and horizontal polarization can be simplified to as the following equation:

$$\Gamma_{11} = \frac{-\varepsilon_r \sin\theta_i + \sqrt{\varepsilon_r - \cos^2\theta_i}}{\varepsilon_r \sin\theta_i + \sqrt{\varepsilon_r - \cos^2\theta_i}} \quad ...(3.24)$$

$$\Gamma_1 = \frac{\sin\theta_i - \sqrt{\varepsilon_r - \cos^2\theta_i}}{\sin\theta_i + \sqrt{\varepsilon_r - \cos^2\theta_i}} \quad ...(3.25)$$

The superposition may be applied to determine transmitted and reflected waves in a medium.

3.4.2 Brewster Angle

The Brewster angle is the angle at which no reflection occurs in the medium **of** origin. It occurs when the incident angle θ_B is such that the reflection coefficient $\Gamma_{\|}$ is equal to zero means ($\Gamma_{\|} = 0$).

The Brewster angle is given by value of θ_B, which satisfies the following:

$$\sin(\theta_B) = \sqrt{\frac{\varepsilon_1}{\varepsilon_1 + \varepsilon_2}} \quad ...(3.26)$$

For the case, when the first medium is free space and second medium has a relative permittivity ε_r, so equation (3.26) can be written as the –

$$\sin(\varepsilon_B) = \sqrt{\frac{\varepsilon_r - 1}{\varepsilon_r^2 - 1}} \quad ...(3.27)$$

The Brewster angle occurs only for vertical (parallel) polarization.

3.4.3 Reflection from Perfect Conductors

Since electromagnetic energy cannot pass through a perfect conductor a plane wave incident on a conductor has all of its energy reflected.

As the electric field at the surface of the conductor must be equal to zero at all times in order to follow Maxwell's equations, the reflected wave must be equal in magnitude to the incident wave.

For the case when is-field polarization is in the plane of incidence, the boundary conditions requires that the following condition:

$$\theta_i = \theta_r \quad ...(3.28)$$

and

$$E_t = E_r \quad (E\text{-field in place of incidence}) \quad ...(3.29)$$

Similarly, for the case when E-field is horizontally polarized, the boundary conditions requires that the following condition:

$$\theta_i = \theta_r \quad ...(3.30)$$

and

$$E_i = E_r \ (E \text{ is-field normal to plane of incidence}) \quad ...(3.31)$$

Refer to equation 3.28 to 3.31, we see that for a perfect conductor, $\Gamma_{\|} = 1$, and $\Gamma = -1$ regardless of incident angle θ_i.

3.5 DIFFRACTION

Diffraction allows radio signals to propagate around the curved surface of the earth, beyond the horizon, and to propagate behind obstructions (like buildings, trees etc.).

The phenomenon of diffraction can be explained by Huygen's principle, which states that all points on a wavefront can be considered as point sources for the production of secondary wavelets and that these wavelets combine to generate a new wavefront in the direction of wave propagation.

Diffraction is caused by the propagation of secondary wavelets into a shadowed region.

The field strength of a diffracted wave in the shadowed region is the vector sum of the electric field components of all the secondary wavelets in the space around the obstacle like as building, trees etc.

3.5.1 Fresnel Zone Geometry

Consider a transmitter and receiver separated in free space as shown in Fig. 3.2 (a). Let an obstructing screen of effective height h with infinite width be placed between them at a distance d_1 from the transmitter and d_2 from receiver so total distance between transmitter and receiver is $d_x + d_2$.

It is apparent that the wave propagating from transmitter to the receiver via the top of the screen travels a longer distance than if a direct line of sight path existed between them.

Assuming $h << d_{1'}\ d_2$ and $h >> \lambda$, then the difference between the direct path and the diffracted path, called the excess path length (Δ), can be obtained from Fig. 3.2 *(b)* as the following equation:

$$\Delta = \frac{h^2(d_1 + d_2)}{2d_1d_2} \quad ...(3.32)$$

The corresponding phase differences is given by

$$\phi = \frac{2\pi\Delta}{\lambda} \approx \frac{2\pi}{\lambda}\frac{h^2}{2}\frac{(d_1 + d_2)}{d_1d_2} \quad ...(3.33)$$

and when tan $x \approx x$, then $\alpha = \beta + \gamma$ from Fig. 3.2 (*c*) and the a will be given as

$$\alpha = h\left(\frac{d_1 + d_2}{d_1\, d_2}\right)$$

Equation (3.33) is often normalized using the dimensionless Fresnel Kirchoff diffraction parameter v which is given by following equation:

$$v = h\sqrt{\frac{2(d_1 + d_2)}{\lambda d_1 d_2}} = \alpha\sqrt{\frac{2d_1d_2}{\lambda(d_1 + d_2)}} \quad ...(3.34)$$

where a has units of radians and is shown in Fig. 3.2 *(b)* and 3.2 (c). The parameter is convenient because it allows ϕ to be as the following:

$$\phi = \frac{\pi}{2}v^2 \quad ... (3.35)$$

From the above equation it is clear that the phase difference between a direct line of sight path and diffracted path is a function of height and position of the obstruction as well as the transmitter and receiver location or positions.

The concept of diffraction loss as a function of the path difference around an obstruction is explained by Fresnel zones concept.

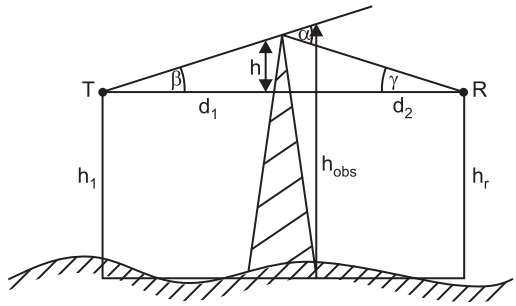
Fresnel zones represent successive regions where secondary waves have a path length from transmitter to receiver which are $\frac{n\lambda}{2}$ greater than the total path length of a line of sight path.

Fig. 3.3 demonstrates a transparent plane located between the transmitter and receiver. The concentric circles on the plane represent the loci of the origins of secondary wavelets which propagates to the receiver such that the total path length increases by $\frac{\lambda}{2}$ for successive circles. These circles are called Fresnel zones.

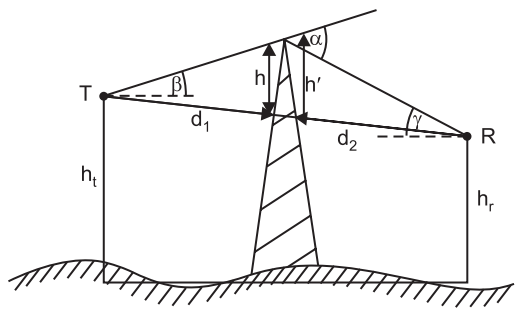
The successive Fresnel zones have the effect of alternately providing constructive and destructive interference to the total received signal. The radius of the nth Fresnel **zone** circle is denoted by r_n and can be expressed in terms of n, λ, d_1 and d_2 by

$$r_n = \sqrt{\frac{n\lambda d_1 d_2}{d_1 + d_2}} \quad ...(3.36)$$

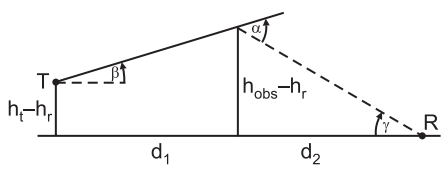
This approximation is valid for $d_1 + d_2 >> r_n$



(*a*) Knife-edge diffraction geometry. The point *T* denotes the transmitter and *R* denotes the receiver, with an finite knife-edge obstruction blocking the line-of-sight path.



(*b*) Knife-edge geometry when the transmitter and receiver are not at the same height. Note that and if a and p are small and $h << d_1$ and d_2 then *h* and *h'are* virtually identical and the geometry may be redrawn as shown in Fig. (*c*).



(*c*) Equivalent knife-edge geometry where the smallest height (in this case h_r) is subtracted from all other height.

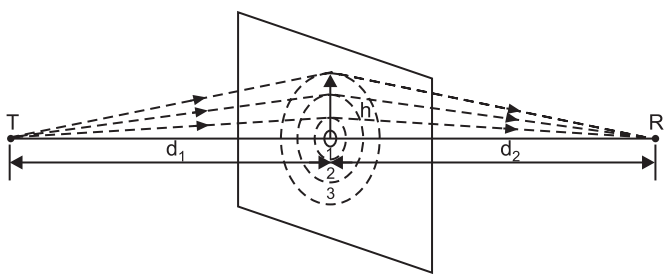


Fig. 3.3: Concentric circles which define the boundaries of successive Fresnel zones.

In mobile communication systems, diffraction loss occurs from the blockage of secondary wave such that only a portion of energy is diffracted arround an obstacle. That is, an obstruction causes a block of energy from some of the Fresnel zones, thus allowing only some of the transmitted energy to reach the

receiver. Depending on the geometry of the obstruction, the received energy will be a vector sum of energy contributions from all unobstructed Fresnel zones concept.

3.6 SCATTERING

Scattering occurs when the medium through which the wave travels consists of objects with dimensions that are small compared to wavelength and where the number of obstacles per unit volume is larger more.

Scattering produce by rough surface or hard surface.

Surface roughness is often tested using the Rayleigh criterion which defines a critical height (h_c) of surface protuberances for a given angle of incidence θ_i given by following equation:

$$h_c = \frac{\lambda}{8\sin\theta_i} \quad ...(3.37)$$

A surface is considered smooth if its minimum to maximum protuberance h is less than h_c and is considered rough if the protuberance is greater than h_c.

For rough surface, the flat surface reflection coefficient need to be multiplied by a scattering loss factor P_s, to account for the diminished reflected field so the loss factor

$P_{s,}$ can be written as

$$P_s = \exp\left[-8\left(\frac{\pi\sigma_h \sin\theta_i}{\lambda}\right)^2\right] \quad ...(3.38)$$

where σ_h is the standard deviation of the surface height about the mean surface height. The loss factor P_s can be modified by Boithias as the following:

$$P_s = \exp\left[-8\left(\frac{\pi\sigma_h \sin\theta_i}{\lambda}\right)^2\right] I_0\left[8\left(\frac{\pi\sigma_h \sin\theta_i}{\lambda}\right)^2\right] \quad ...(3.38)$$

where I_0 is the bassel function of the first kind and zero order.

The reflected field E-field for $h >> h_c$ can be solved for rough surface using a modified reflection coefficient given as the following:

$$\Gamma_{\text{rough}} = P_s \Gamma \quad ...(3.40)$$

3.6.1 Radar Cross Section Model

In radio channel where large, distance objects induce scattering knowledge of the physical location of such object can be used to accurately predict scattered signal strength.

The radar cross section (RCS) of a scattering object is defined as the ratio of the power density of the signal scattered in the direction of the receiver to the power density of the radio wave incidence upon the scattering object, and has units of square meters.

For urban mobile radio system, model based on the bistatic radar equation may be used to compute the received power due to scattering in far-field.

The bistatic radar equation describe the propagation of a wave travelling in free space which impinges on a distance scattering object and is than reradiated in the direction of receiver, given by

$$P_R\,(\text{dBm}) = P_T(\text{dBm}) + G_T\,(\text{dB}_i) + 20\log(\lambda) + \text{RCS}\ [\text{dBm}^2] - 30\log(4\pi) - 20\log d_T - 20\log d_R \quad ...(3.41)$$

where d_T and d_R are the distance from the scattering object to the transmitter and receiver.

In equation (3.41) the RCS is given in units of dBm and can be approximated by the surface area of the scattering object, measured in dB with respect to a one square meter reference.

Equation (3.41) may be applied to scatterers in the far-field of both the transmitter and receiver and is useful for predicting receiver power which scatters off large objects, such as building, which are for both the transmitter and receiver.

For medium and large size buildings located 5–10 km away RCS value were found to be in the range of 14.1 to 55.7 dBm^2.

3.7 PRACTICAL LINK BUDGET DESIGN USING PATH LOSS MODELS

Most radio propagation models are derived using a combination of analytical and empirical method.

The empirical approach is based on fitting curve or analytical expression that recreate a set of measured data. This has the advantage of implicitly taking into account all propagation factor, both known and unknown, through actual field measurement.

However, the validity of an empirical model at transmission frequency or environments other than those used to drive the model can only be established by additional measured data in the new environment at the required transmission frequency.

By using path loss models to estimate the received signal level as a function of distance, it become possible to predict the SNR for a mobile communication system.

3.7.1 Log-Distance Path Loss Model

Both theoretical and measurement based propagation models indicate that average received signal power decreases logarithmically with distance, whether in outdoor or indoor radio channels.

Such model has been used extensively in the literature.

The average large scale path loss for an arbitrary *T-R* separation is expressed as a function of distance by using a path loss exponent, *n*

$$\overline{PL}(d)\alpha\left(\frac{d}{d_0}\right)^n \quad ...(3.42)$$

or
$$\overline{PL}(dB) = \overline{PL}(d_0) + 10n\log\left(\frac{d}{d_0}\right) \quad ...(3.43)$$

where n is the path loss exponent which indicate the rate at which the path loss increases with distance d_0 is the close in reference distance which is determined from measurements close to the transmitter and d is the T-R separation distance.

The bars in equation (3.42) and (3.43) denote the ensemble average of all possible path loss values for a given value of d.

The value of n depends on the specific propagation environment.

It is important to select a free space reference distance that is appropriate for the propagation environment. In large coverage cellular systems, 1 km reference distance are commonly used, whereas in microcellular systems, much smaller distances (100 M or 1 M) are used.

The reference distance should always be in the far-field of the antenna so that near field effects do not alter the reference path loss.

The reference loss (path) is calculated using the free space path loss formula given by equation [3.8 (b)] or through field measurement at distance d_0.

Table 3.1 lists typical path loss exponents obtained in various mobile radio environments with different value.

Table 3.1: Path Loss Exponents for Different Environments.

Environment	*Path Loss Exponent, n*
Free space	2
Urban area cellular radio	2.7 to 3.5
Shadowed urban cellular radio	3 to 5
In building line-of-sight	1.6 to 1.8
Obstructed in building	4 to 6
Obstructed in factories	2 to 3

3.8 OUTDOOR PROPAGATION MODELS

A number of propagation models are use to predict path loss over irregular terrain. While all these models aim to predict signal strength at a particular receiving point or in a specific local area (called a sector), the methods very widely in their approach, complexity and accuracy.

Most of these models are based on a systematic interpretation of measurement data obtained in the service area.

Some of the outdoor propagation models are –

3.8.1 Longley-Rice Model

The longley-Rice Model is applicable to point to point communication system in the frequency range from 40 MHz to 100 GHz over different kinds of terrain. The longley-Rice Model is also reffered to as the ITS irregular terrain model.

The longley-Rice Model is also available as a computer program to calculate large scale medium transmission loss relative to free space loss over irregular terrain for frequencies between 20 MHz and 10 GHz.

For a given transmission path, the program takes as its input the transmission frequency, path length, polarization, antenna heights, surface refractivity, effective radius of earth, ground conductivity, ground dielectric constant and climate. The program also operates on path specific parameters such as horizon distance of the antennas, horizon elevation angle, angular trans-horizon distance, terrain irregularity and other specific inputs.

The longley-Rice method operates in two modes–

(*i*) Point-to-point mode prediction.

(*ii*) Area mode prediction.

In point-to-point mode prediction, when a detailed terrain path profile is available, the path-specific parameters can be easily determined and the prediction is called a point-to-point mode prediction. On the other hand, if the terrain path profile is not available, the Longley-Rice method provides techniques to estimate the path specific parameters, and such a prediction is called an area mode prediction.

One shortcoming of the Longley-Rice Model is that it does not provide a way of determining corrections due to environmental factor in the immediate vicinity of the mobile receiver or consider correction factor to amount for the effects of buildings and foliage.

3.8.2 Durkin's Model-A Case Study

Durkin's model provides an interesting perspective into the nature of propagation over irregular terrain and the losses caused by obstacles in a radio path.

The execution of the Durkin path loss simulator consist of two parts. The first part accesses a topographic data base of a proposed service area and reconstructs the ground profile information along the radial joining the transmitter to the receiver. The assumption is that the receiving antenna receives all off its energy along that radial and therefore experiences no multipath propagation. The effect of this assumption is that the model is somewhat pessimistic in narrow valleys, although it identifies isolated weak reception areas rather well. The second part of the simulation algorithm calculated the expected loss along that radial.

The topographical data base can be thought of as a two-dimensional array. Each array element corresponds to a point as the service area map while the actual content of each array element contain the elevation above sea level data as shown in Fig. 3.4.

These type of digital elevation model (DEM) are readily available from the United States Geological Survey (USGS). By using the quantized map of service area heights, the program reconstructs the ground profile along the radial that combined the transmitter and receiver.

Before the radial may not always be discrete data points and interpolation method are used to calculate the approximate height that may be defined when looking around the radial.

Fig. 3.4 (*b*) define the topographic grid with arbitrary location of transmitter and receiver, the radial between transmitter and receiver and Fig. 3.4 (c) show a typical reconstructed radial terrain profile might look like.

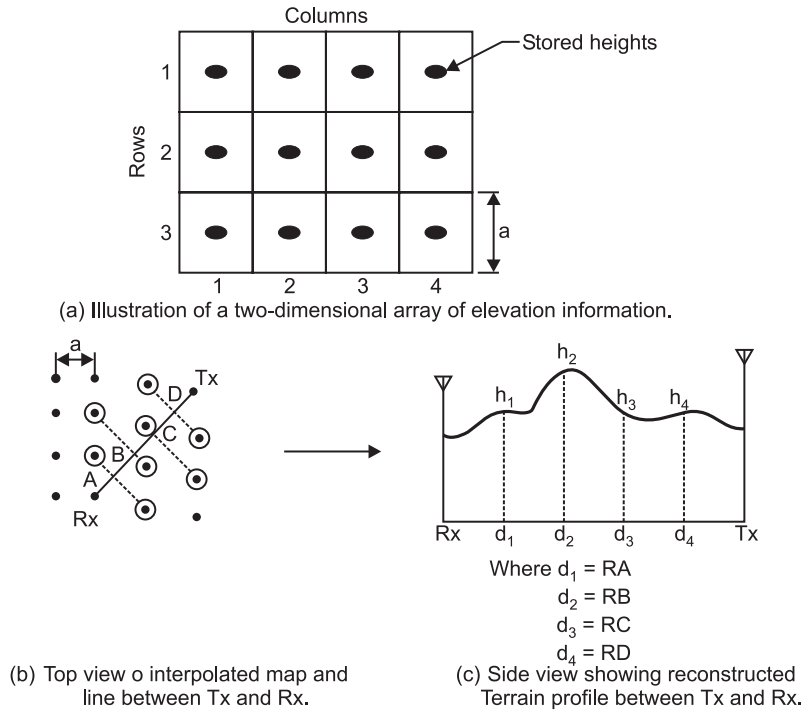


Fig. 3.4: Illustration of terrain profile reconstruction using diagonal interpolation.

The Edwards and Durkin algorithm use the Epstein and Peterson method to define the loss associated or consult with two diffraction edges. In other words it is the sum of two attenuations.

This method is used to read the digital elevation map and perform a site-specific propagation compution on the elevation data.

This method can also be used to develope the signal strength with in a few dB.

This method has some disadvantage that it can not adequately predict propagation effect due to foliage, buildings and from man-made structures and it does not support multipath propagation.

3.8.3 Okumura Model

The Okumura model is used for signal prediction in urban area with frequency range 150 MHz to 1920 MHz and distances of 1 km to 100 km. This model also define the heights ranging from 30 m to 1000 m.

Okumura model produce a set of curves define the mediam attenuation relative to free space (A_{mu}) with effective base station antenna height (h_{te}) of 200 m and a mobile antenna height (h_{re}) of 3 m.

To obtain the path loss by using Okumura model, firstly determine free space path loss between the points of interest and then added the value of A_{mu} (f, d) to it along with correction factor to account for the type of terrain as –

$$L_{50}(dB) = L_F + A_{mu}(f, d) - G(h_{te}) - G(h_{re}) \ldots G_{AREA} \quad \ldots(3.44)$$

where $L_{50}(dB)$ define the 50th percentile value of propagation path loss, L_F define the free space propagation loss, A_{mu} define median attenuation relative to free space, $G(h_{te})$ is gain factor of base station antenna height, $G(h_{re})$ is gain factor of mobile antenna height and G_{AREA} is gain from environment.

Okumura model can also be used to find the $G(h_{te})$ varies at a rate of 20 dB/decode and $G(h_{re})$ varies with rate 10 dB/decade for height less than 3 m.

$$G(h_{te}) = 20\log\left(\frac{h_{te}}{200}\right) 1000 \text{ m} > h_{te} > 30 \text{ m} \quad \ldots[3.45\ (a)]$$

$$G(h_{re}) = 10\log\left(\frac{h_{re}}{3}\right) \quad h_{re} \leq 3 \text{ m} \quad \ldots[3.45(b)]$$

$$G(h_{re}) = 20\log\left(\frac{h_{re}}{3}\right) 10 \text{ m} > h_{re} > 3 \text{ m} \quad \ldots [3.45\ (c)]$$

Some of other important terrain related parameters can also be measured by Okumura model like as terrain undulation height (Ah), isolated ridge height, average slop of the terrain and mixed land sea parameter.

Okumura model is totally based on measured data and it never provide the analytical explanation.

Okumura model is best and simplest method to measure the accuracy in path loss prediction for mature cellular and land mobile radio systems in environment.

The main drawback of this method is that it has very slow response to rapid change in terrain.

3.8.4 Hata Model

Hata model is the solution over the Okumura drawback that it provide empirical formulation of the graphical path loss data which is valid from 150 MHz to 1500 MHz.

The standard formula for median path loss by Hata model is given by in urban areas

$$L_{50}(\text{Urban})(\text{dB}) = 69.55 + 26.16 \log f_c - 13.82 \log h_{te} - a(h_{re}) + (44.9 - 6.55 \log h_{te}) \log d \quad \ldots(3.46)$$

where f_c represents the frequency in MHz and its range from 150 MHz to 1500 MHz, h_{te} represent the transmitter antenna height in meter and its range from 30 m to 200 m, h_{re} represent receiver antenna height in meter and its range

from 1 m to 10 m, and d represents the transmitter-receiver distance in km, and $a(h_{re})$ is a correction factor which represent the function of the size of coverage area and it can be written as –

$$a(h_{re}) = (1.1 \log f_c - 0.7)h_{re} - (1.56 \log f_c - 0.8)dB \quad ...(3.47)$$

For obtain the path loss, the Hata formula from equation (3.46) can modify as –

$$L_{50}(dB) = L_{50}(\text{Urban}) - 2[1\log (f_c/28)]^2 - 5.4 \quad ...(3.48)$$

and

$$L_{50}(dB) = L_{50}(\text{Urban}) - 4.78(\log f_c)^2 + 18.33 \log f_c - 40.94 \quad ...(3.49)$$

From above discussion, the Hata model have some drawback like that it does not have any path specific corrections but the Okumura's model have.

Hata model is suitable only for large cell mobile system not for personal mobile communication system.

3.9 INDOOR PROPAGATION MODELS

The indoor radio channel is different from the traditional mobile radio channel and it can be explain by following two aspects :

(*i*) The distances covered are much smaller and

(*ii*) The variability of the environment is more greater for a small distance range between transmitter and receiver.

The indoor propagation model is used for determine the pathloss within the building.

Reflection, diffraction and scattering, these are three important mechanisms to dominated the indoor propagation models.

In the early 1980's, British Telecom were firstly study the indoor path loss in and around a large numbers of homes and office building.

The indoor channels can be classified into two way either as line of sight (LOS) or obstructed (OBS) with varying degrees of clutter and some features of indoor propagation models are described in following topics:

3.9.1 Partition Losses (Same Floor)

The buildings consist a wide variety of partitions and obstacles which are the base to form the internal and external structure.

The partition which are due to building, houses, offices etc. can be classified into two parts-hard partition and soft partition.

The hard partition is occur due to the building structure but the soft partition that may be moved and which do not span to the ceiling.

The partitions vary widely in their physical and electrical characteristic and these characteristic making it difficult to apply general models to specific indoor installations.

3.9.2 Partition Losses Between Floors

This topics is consult with the losses between floors of a buildings which are determined by the external dimensions and materials of the building with type of constructions used to made the floors and the external surroundings.

The numbers of windows and presence of tinting which attenuates the radio energy and can also impact the loss between floors.

The partition losses between floors can be determine by using floor attenuation factor (FAF), and different FAF is shown in following Table 3.2.

3.9.3 Long Distance Path Loss Model

Equation number (3.50) define the path loss for long distance

$$PL(dB) = PL(d_0) + 10n\log\left(\frac{d}{d_0}\right) + x\sigma \quad ...(3.50)$$

In this equation n is depend on surroundings and building type, and x_a define a normal random variable in dB.

Table 3.2: Total Floor Attenuation Factor and Standard Deviation a(dB) for Three Buildings. Each Point Represents the Average Path Loss Over a 20λ Measurement Track.

Building	*915 MHz FAF(dB)*	*s(dB)*	*Number of locations*	*1900 MHzFAF(dB)*	*s(dB)*	*Number of locations*
Walnut Creek						
One Floor	33.6	3.2	25	31.3	4.6	110
Two Floors	44.0	4.8	39	38.5	4.0	29
SF Pac Bell						
One Floor	13.2	9.2	16	26.2	10.5	21
Two Floors	18.1	8.0	10	33.4	9.9	21
Three Floors	24.0	5.6	10	35.2	5.9	20
Four Floors	27.0	6.8	10	38.4	3.4	20
Five Floors	27.1	6.3	10	46.4	3.9	17
San Ramon						
One Floor	29.1	5.8	93	35.4	6.4	74
Two Floors	36.6	6.0	81	35.6	5.9	41
Three Floors	39.6	6.0	70	35.2	3.9	27

3.10 SMALL SCALE MULTIPATH PROPAGATION

The small scale fading effects are produce by the multipath in the radio channel and the important effects are—

(*i*) Rapid changes in the strength of the signal in a small travel distance or time interval.

(*ii*) Random frequency modulation due to change in Doppler shift.

(*iii*) Time dispersion produce by multipath propagation delay.

In the urban areas, the height of surrounding structures is greater than the height of mobile antennas so there is not a single line of sight path to the base station. But when a line of sight exists, the multipath produce due to reflection from the ground and surrounding structures.

The received signal by the mobile in space may have a large number of plane waves in which randomly distributed amplitudes, phase and angle of arrival is present and these multipath components consist vectorially at the receiver antenna and can cause the signal to distort or fade but if the mobile receiver is stationary then the received signal may be distorted due to moment of surrounding objects in the channel.

If the object is static in the radio channel, then the fading is purely spatial phenomenon.

The spatial variations in the final received signal may be temporal as it move through the multipath field.

Due to the relative motion between the mobile station and base station, there is frequency shift in each multipath wave is possible. This shift is called the Doppler shift and it is directly proportional to the velocity and direction of motion of the mobile with respect to the arrival direction of received multipath wave.

3.10.1 Factor Influencing Small Scale Fading

The small scale fading can be effected by many physical factors in the radio propagation channel. Some factor are described here.

(*i*) **Multipath Propagation:** The random amplitudes and phase of different multipath components produce variation in the strength of the signal due to the small scale fading and signal distortion both will be possible.

The multipath propagation often extend the time required for base band signal at the receiver which can cause signal smearing due to inter symbol interference.

(*ii*) **Speed of Mobile:** The speed of mobile is the most important factor that influencing the small scale fading.

The speed of mobile with reference base station produce frequency modulation by different Doppler shift on each of the multipath component.

The Doppler shift may be positive or may be negative and it is depend on the moment of the receiver from the base station.

(*iii*) Speed of Surrounding Object: In a radio channel, if the object is in the motion condition, then a time varying Doppler shift is produce on multipath component.

If the surrounding object of the mobile are move at a higher rate then the mobile, then this effect dominates the small scale fading, otherwise moment in the surrounding object may be ignored and only mobile speed need be consider.

(*iv*) ***Transmission Bandwidth of the Signal*:** If the bandwidth of the multipath channel is less than the transmitted radio signal bandwidth, then the received signal will be distorted but the strength of the received signal will not fade over a local area.

If the transmitted signal has a narrow bandwidth as compared to channel bandwidth than the signal amplitude will change rapidly but the signal will not be distorted in time.

3.10.2 Doppler Shift

From Fig. 3.5, let us assume, a mobile moving at a constant velocity with v, with path segment length d between point x and y, a remote source s is used for receive signal.

So from source to mobile, the path length difference by the wave at point x and >> can be $\Delta l = d \cos \theta = v\Delta t \cos \theta$

where Δt is the time required to mobile to move from x *to* y and θ is the assumed to the same at point x and y.

Due to the path length difference, the change phase in received signal is

$$\Delta\phi = \frac{2\pi\Delta l}{\lambda} = \frac{2\pi v\Delta t}{\lambda}\cos\theta \qquad ...(3.51)$$

and the Doppler shift frequency f_d is

$$f_d = \frac{1}{2\pi}\cdot\frac{\Delta\varphi}{\Delta t} = \frac{v}{\lambda}\cdot\cos\theta \qquad ...(3.52)$$

Equation (3.52) combined the Doppler shift to the mobile velocity and the spatial angle between the arrival direction of wave and direction of motion of mobile.

From equation (3.52), the Doppler shift will be positive, if mobile is moving toward the arrival direction of wave, and if mobile is moving away from the arrival direction of wave, then the Doppler shift will be negative.

3.11 SMALL SCALE MULTIPATH MEASUREMENT

The multipath structure play an important role to determine the small scale fading effect. There are various type of measurement techniques which are used like as direct pulse measurement, spread spectrum sliding correlator measurement and swept frequency measurement and it is shown in Fig. 3.6.

The direct RF pulse system is widely used technique for small scale multipath measurement.

This technique allow or permit to the engineers to determine to the power delay profile of any channel.

In this system, system transmits a repetitive pulse of width T_{bb} and use a receiver with a band pass filter whose bandwidth is $\frac{2}{T_{bb}} Hz$.

The signal is amplified and detected by using an envelope detector and then stored and display on a CRO.

The main advantage of this system is that this system is the lack of complexity and the main disadvantage of this system is that it is subject to noise and interference from the reason of wide pass band filter required for multipath time resolution.

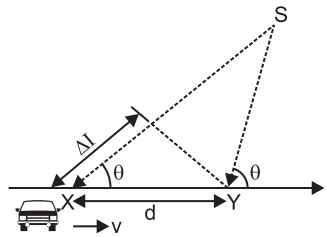


Fig. 3.5: Illustration of Doppler Effect.

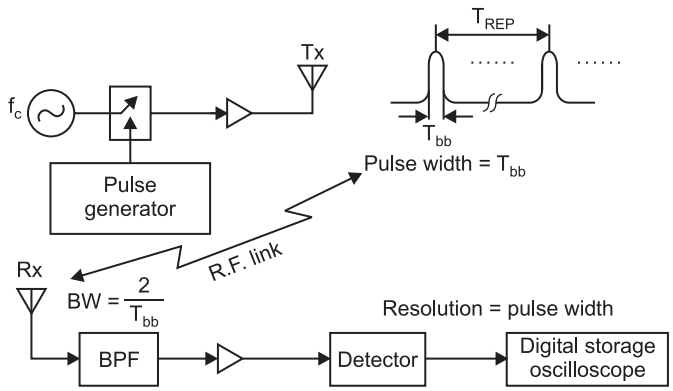


Fig. 3.6: Direct RF channel impulse response measurement system.

3.12 PARAMETERS OF MOBILE MULTIPATH CHANNELS

The parameters of mobile multipath channels play an important role for processing the signal, transmit the signal and receive the signal.

There are three types of parameters are used for mobile multipath channels and these are time dispersion parameter, coherence bandwidth, Doppler spread and coherence time.

3.12.1 Time Dispersion Parameters

The mean excess delay $(\overline{\tau})$ and *rms* delay spread (σ_τ) are used to quantified the time dispersive properties of wide band multipath channel.

The mean excess delay $(\overline{\tau})$ and *rms* delay spread (σ_τ) can be determined by using a power delay profile.

The rms delay spread is the square root of second moment of power delay profile and it is defined as

$$\overline{\tau} = \frac{\sum_k a_k^2 \tau_k}{\sum_k ak^2} = \frac{\sum_k P(\tau_k)\tau_k}{\sum_k P(\tau_k)} \quad \text{... (3.54)}$$

The *rms* delay spread is the square root of second moment of power delay profile and it is defined as –

$$\sigma_\tau = \sqrt{\overline{\tau}^2 - (\overline{\tau})^2} \quad ...(3.54)$$

where,

$$\overline{\tau}^2 = \frac{\sum_k a_k^2 \tau_k^2}{\sum_k ak_k^2} = \frac{\sum_k P(\tau_k)\tau_k^2}{\sum_k P(\tau_k)} \quad ...(3.55)$$

these **delay are** measured at the receiver a $\tau_0 = 0$.

The typical of *rms* delay spread in outdoor mobile radio channel is measured in microseconds and in indoor mobile radio channel, it is measured in nanoseconds.

The maximum excess delay is defined as $\tau_x, - \tau_0$ where τ_0 represent the first arriving signal and τ_x is the maximum delay within *xdB* of the strongest arriving multipath signal.

The maximum excess delay (*x*dB) define the temporal extent of the multipath that is above a particular threshold. Sometime the value of τ_x is called excess delay spread of a power delay profile.

In practically, the value of $\overline{\tau}, \overline{\tau^2}$ and σ_x is depend on noise threshold, if the noise threshold $\overline{\tau}, \overline{\tau^2}$ and σ_τ is set at very low value than the values of x , $_T^2$ and c_t are taken the artificially high.

Analogous to the delay spread parameters in the time domain, coherence bandwidth is used to characterize the channel in frequency domain.

The coherence bandwidth and *rms* delay spread, both are inversely proportional to one another.

3.12.2 Coherence Bandwidth

The coherence bandwidth is defined that it is the range of frequency over which two frequency component have a strong potential for amplitude correlation.

Coherence bandwidth is a statistical measure of range of frequency over which the channel is consider to be flat.

If the coherence bandwidth is defined as the bandwidth over which the frequency correlation function is above 0.9, the coherence bandwidth can be written as –

$$B_c \approx \frac{1}{50\sigma_\tau} \quad ...(3.56)$$

If the frequency correlation function is above 0.5, than the coherence bandwidth can be –

$$B_c \approx \frac{1}{5\sigma_\tau} \quad ...(3.57)$$

Equation (3.56) and (3.57) represent the "ball park estimates".

Generally, the spectral analysis technique and simulation are required to determine the exact impact that time varying multipath consist a particular transmitted signal. Hence, accurate multipath channel models must be used to design the specific modems for wireless applications.

3.12.3 Doppler Spread and Coherence Time

Doppler spread and coherence time are the parameter which are used to describe the time varying nature and time dispersive nature of the channel in a small scale region but they do not provide the information about time varying nature of the channel produced by the relative motion between the mobile and base station or by any movement of object in the channel.

Doppler spread B_D is the measure of a spectral broading produce by change in time rate of mobile radio channel and B_D define a range of frequency in which the received Doppler spectrum is non zero.

When a pure sinusoidal tone of frequency F_C is transmitted than the received signal spectrum is called Doppler spectrum and this Doppler spectrum consist the component range from $(f_c - f_d)$ to $(f_c + f_d)$ where f_d is Doppler shift.

If the B_D is less than base band signal bandwidth, the effect of Doppler spread are negligible at the receiver. This type of channel is called slow fading channel.

The coherence time T_c define the time varying nature of the frequency dispersiveness of the channel in the time domain.

The Doppler spread and coherence time, both are inversely proportional to each other.

$$T_c \approx \frac{1}{f_m} \qquad ...[3.58(a)]$$

where f_m is maximum Doppler shift.

The coherence time is the time duration over which two received signal have a strong potential for amplitude correlation.

If the coherence time is defined as the time over which the time correlation function is more than 0.5 then

$$T_c \approx \frac{9}{16\pi f_m} \qquad ...[3.58(b)]$$

where $f_m = \frac{v}{\lambda}$.

The geometric mean of equation 3.58 (a) and 3.58 (b) define the coherence time like as

$$T_c = \sqrt{\frac{9}{16\pi f_m^2}} = \frac{0.423}{f_m} \qquad ... [3.59(c)]$$

So that the coherence time definition show that two signals arriving with a time separation greater than T_c are differently by the channel.

If a digital system is used, than the symbol rate is greater than $\frac{1}{T_c}$.

3.13 TYPE OF SMALL SCALE FADING

Depending upon the relation between the parameters of signals (like as bandwidth, symbol period etc.) and channels (like as rms delay spread and Doppler spread) will define or introduce the different types of fading.

The time dispersion and frequency dispersion mechanism in a mobile radio channel introduce for different possible effect which are depend upon the nature of the transmitted signal, channel and the velocity.

The multipath delay spread introduce the time dispersion and frequency selective fading while the Doppler spread introduce the frequency dispersion and time selective fading.

Fig. 3.7 show a tree of four different type of fading.

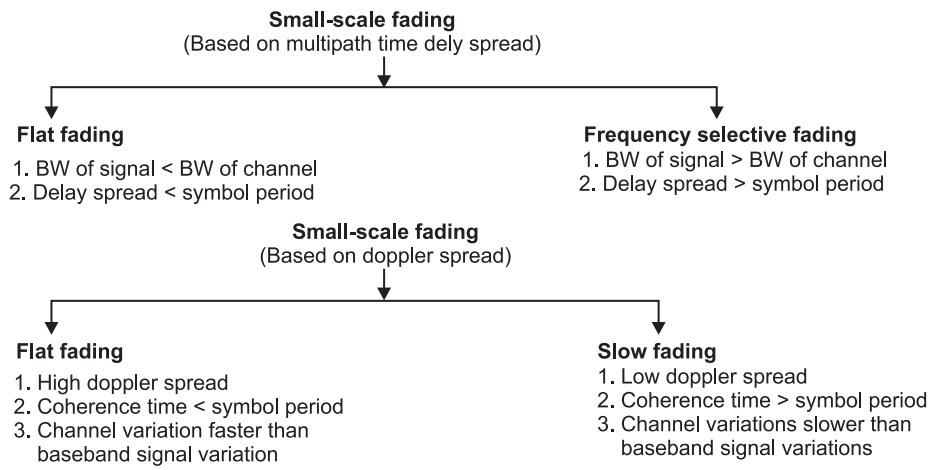


Fig. 3.7: Types of small-scale fading.

3.13.1 Fading Effects Due to Multipath Time Delay Spread

The multipath time delay spread introduce the two type of fading one is flat fading and second is frequency selective fading.

3.13.1.1 Flat Fading

If mobile radio channel consist linear phase and constant gain over a bandwidth which is greater than the transmitted signal bandwidth, then the received signal will represent the flat fading.

In the flat fading, the multipath structure of the channel is such that the spectral characteristics of a transmitted signal are protected at the receiver but the strength of the received signal varying with time and it is due to change in gain of the channel caused by multipath.

The flat fading channel's characteristics shown in Fig. 3.8.

From Fig. 3.8, in a fixed time duration, varies channel gain produce the change in amplitude of received signal but the spectrum of transmission is safe.

In a flat fading channel, the reciprocal bandwidth of transmitted signal is greater than the multipath time delay spread of the channel.

Flat fading channel also called as amplitude varying channel and some time it is known as narrow band channel in which applied signal bandwidth is narrow as compared with flat fading channel bandwidth.

A signal represent the flat fading if

$$B_s << B_c \quad \text{... (3.59)}$$

and

$$T_s >> \sigma_\tau \quad \text{... (3.60)}$$

where T_s is reciprocal bandwidth and B_s is bandwidth of transmitted modulation, o_x is rms delay spread and B_c is coherence bandwidth of the channel.

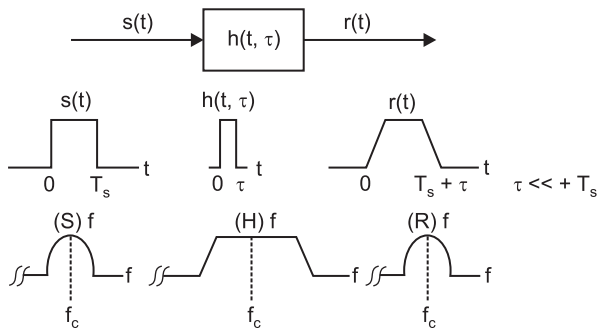


Fig. 3.8: Flat fading channel characteristics.

3.13.1.2 Frequency Selective Fading

If the mobile radio channel consist linear phase response and constant gain over a bandwidth which is smaller than the transmitted signal bandwidth, than the signal represent the frequency selective fading.

The frequency selective fading is occur due to time dispersion of the transmitted signal with in the channel. Thus the channel produce inter symbol interference (ISI).

The modeling of frequency selective fading channel are much more difficult than flat fading channel.

Fig. 3.9 show the characteristics of a frequency selective fading channel.

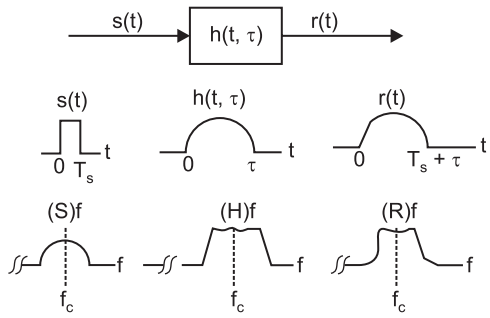


Fig. 3.9: Frequency selective fading channel characteristics.

For the frequency selective fading, the transmitted signal spectrum *s(t)* consist a bandwidth which is greater than the coherence bandwidth B_c of the channel. The frequency selective fading is produce by multipath delays which increase the symbol period of transmitted symbol.

The frequency selective fading channel is also refer as wide band channel in which the bandwidth of the signal $s(t)$ is larger than the bandwidth of the channel impulse response.

A signal represent frequency selective fading if

$$B_S < B_C \quad ...(3.61)$$

and

$$T_S > \sigma_C \quad ...(3.62)$$

where B_s is transmitted signal bandwidth and B_c is coherence bandwidth, T_s is reciprocal bandwidth and a_T is rms delay spread.

3.13.2 Fading Effects Due to Doppler Spread

The Doppler spread introduce two types of fading, one is fast fading and second is slow fading.

3.13.2.1 Fast Fading

In a fast fading channel, the impulse response of the channel varies within the symbol duration and the coherence time of the channel is smaller than the period of the transmitted signal.

The Doppler spreading causes the frequency dispersion as well as signal distortion.

A signal represent the fast fading if

$$T_S > T_C \quad ...(3.63)$$

and

$$B_S < B_D \quad ...(3.64)$$

where T_c and T_s are coherence time and reciprocal bandwidth and B_s and B_D are transmitted signal bandwidth and Doppler spread.

The fast fading only consult with the rate of change of the channel due to motion.

Practically, fast fading occurs only for very low data rate.

3.13.2.2 Slow Fading

In a slow fading channel, the channel impulse response varies at a rate much slower than the transmitted baseband signal $s(t)$. In this condition, the channel may be consider as static over a fixed bandwidth intervals.

A signal represent slow fading if

$$T_s << T_c \quad ...(3.65)$$

and

$$B_s >> B_D \quad ...(3.66)$$

where T_c is coherence time, T_s is reciprocal bandwidth, B_s is transmitted signal bandwidth and B_D is Doppler spread.

The velocity of mobile and the base band signalling determine whether the signal represent slow fading or fast fading.

3.14 TWO-RAY RAYLEIGH FADING MODEL

In the modern mobile communication system with high data rates, it is required to design the fading as well as effect of multipath delay spread. For this purpose a two ray Rayleigh fading model is develope and a model is shown in Fig. 3.10.

The impulse response of the model can be represent as

$$h_b(t) = \alpha_1 \exp(j\phi_1)\delta(t) + \delta(t) + \alpha_1 \exp(j\phi_2)\delta(t-\tau) \quad ... (3.67)$$

where α_1 and α_1 both are independent and 1 both are Rayleigh distributed, ϕ_1 and α_2 are also independent and uniformly distribute d over 0 to 2π, and τ represent the time delay between two ray.

If the σ_2 is equal to zero, then the impulse response can be

$$h_b(t) = \alpha_1 \exp(j\phi_1)\delta(t) \quad ...(3.68)$$

By changing the value of x, than a wide range of frequency selective fading effects will be possible to produce.

Fig. 3.10 show a block diagram of two Ray Rayleigh fading model.

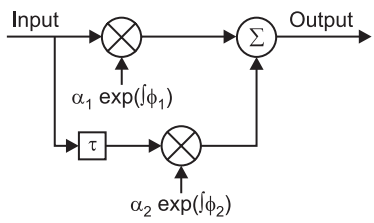


Fig. 3.10: Two Ray Rayleigh Fading Model.

3.15 MULTIPATH SHAPE FACTOR

The shape factors is basically used for explain the statistics of received signal fluctuations in a fading multipath channel.

The multipath shape factor can also support to the quantitative analysis of any non-omni directional multipath wave which is distribute in a local area.

The three basic principle shape factors– the angular spread, the azimuthal direction of maximum fading and angular constriction are used to define the average rate at which a received signal is fade.

The representation of multipath power $P(\theta)$, includes the antenna gains and polarization mismatch the effects.

The shape factors are based on the complex Fourier coefficients of P(θ)

$$F_n = \int_0^{2\pi} P(\theta)\exp(-jn\theta)\,d\alpha \quad ...(3.69)$$

where F_n is the nth complex Fourier coefficient.

(*i*) Angular spread (Δ): The shape factor angular spread (Δ) is a measure of how multipath concentrates about a single azimuthal direction of arrival, than

$$\Delta = \sqrt{1 - \frac{|F_1|^2}{F_0^2}} \qquad ...(3.70)$$

where F_0 and F_1 are defined by the equation (3.69).

(*ii*) Angular Constriction (γ): The shape factor *angular constriction* (γ) is define that how multipath concentrates about two azimuthal directions, than

$$\gamma = \frac{|F_0 F_2 - F_1^2|}{F_0^2 - (F_1)^2} \qquad ...(3.71)$$

where F_0 F_1, and F_2 are defined by the equation (3.69).

(*iii*) Azimuthal Direction of maximum fading(Qmax): The third shape factor, azimuthal direction of maximum fading 6^, may be define as a directional or orientation parameter, than

$$\theta_{max} = \frac{1}{2}\arg\{F_0 F_2 - F_1^2\} \qquad ...(3.72)$$

The θ_{max} value corresponds to in which a mobile user would move in order to experience the maximum fading rate in the local area.

3.16 FADING RATE VARIANCE RELATIONSHIPS

Complex received voltage, received power, and received envelope are the three basic stochastic processes that are studied in small-scale fading. In order to understand how these stochastic processes evolve over space, it is useful to study the position derivatives or rate-of-changes of the three processes. Since the mean derivative of a stationary process is zero, the mean-squared derivative is the simplest statistics that measures the fading rate of a channel. In fact, a mean-squared derivative of a stationary process is actually the **variance** of the rate-of-change.

(*i*) Complex Received Voltage $\widetilde{V}(r)$: The complex received voltage $(\widetilde{V}(r))$, is a baseband representation of the summation of numerous multipath waves that have impinged upon the receiver antenna and have excited a complex voltage component at the input of a receiver. The rate variance, $\sigma_{\tilde{v}}^2$, for the complex voltage of a receiver traveling along the azimuthal direction θ:

$$\sigma_{\tilde{v}}^2(\theta) = E\left\{\left|\frac{d\widetilde{V}(r)}{dr}\right|^2\right\} = \frac{2\pi^2\Lambda^2 P_T}{\lambda^2}(1 + \gamma\cos[2(\theta - \theta_{max})]) \qquad ...(3.73)$$

where λ is the wavelength of the carrier frequency, P_T is the local average mean-squared received power (units of volts-squared).

(*ii*) Received Power (*P*(*r*)): Received power *P(r)* is equal to the magnitude-squared of complex voltage $\widetilde{V}(r)$.

The mathematical operation of taking the squared magnitude of a complex quantity is a nonlinear operation, so in order to derive a rate variance relationship

for received power, we will assume that the channel is **Rayleigh fading.** This assumption, however is unnecessary for the derivation of equation (3.73). The rate variance σ_ρ^2, for the power of a receiver traveling along the azimuthal direction 0:

$$\sigma_\rho^2(\theta) = E\left\{\left(\frac{dP(r)}{dr}\right)\right\} = \frac{4\pi^2\Lambda^2 P_T^2}{\lambda^2}(1+\gamma\cos[2\theta-\theta_{\max}]) \qquad ...(3.74)$$

***(iii)* Received Envelope (*R*(*r*)):** Received envelope, $R(r)$ is equal to the magnitude of complex voltage $\tilde{V}(r)$. Once again, we assume that the channel is **Rayleigh fading** to calculate the mean-squared fading rate. The envelope rate variance,

$$\sigma_R^2(\theta) = E\left\{\left(\frac{dR(r)}{dr}\right)^2\right\} = \frac{4\pi^2\Lambda^2 P_T}{\lambda^2}(1+\gamma\cos[2\theta-\theta_{\max}]) \qquad ...(3.75)$$

Again, equation (3.75) depends on Λ, γ and $\theta_{\max}$. The physical significance of σ_R^2 is that it describes the average **envelope fluctuations** in a local area where the signal envelope undergoes Rayleigh flat fading.

3.17 COMPARISON TO OMNIDIRECTIONAL PROPAGATION

Applying the three shape factors, Λ, γ and $\theta_{\max}$ to the classical omni-directional propagation model, we find that there is not a bias in eitherone or two directions of angle-of-arrival, leading to the maximum angular spread ($\Lambda = 1$) and minimum angular constriction ($\Upsilon = 0$). The statistics of omnidirectional propagation are **isotropic,** exhibiting no dependence on the azimuthal direction of receiver travel, θ.

If the rate variance relationships of equations (3.73)–(3.75) are normalized against their values for omnidirectional propagation, then they reduce to the following form:

$$\bar{\sigma}^2(\theta) = \frac{\sigma_{\tilde{V}}^2(\theta)}{\sigma_{\tilde{V}}^2(\text{omni})} = \frac{\sigma_\rho^2(\theta)}{\sigma_\rho^2(\text{omni})}\frac{\sigma_{\tilde{V}}^2(\theta)}{\sigma_{\tilde{V}}^2(\text{omni})} = \frac{\sigma_\rho^2(\theta)}{\sigma_\rho^2(\text{omni})} \qquad ...(3.76)$$

$$= \frac{\sigma_R^2(\theta)}{\sigma_R^2(\text{omni})} = \Lambda^2(1+\gamma\cos[2(\theta-\theta_{\max})])$$

where $\bar{\sigma}^2$ is a normalized fading rate variance.

Equation (3.76) provides a convenient way to analyze the effects of the shape factors on the second-order statistics of small-scale fading.

$$\frac{1}{2}[\bar{\sigma}^2(\theta)+\bar{\sigma}^2(\theta+\pi/2)] = \Lambda^2 \qquad ...(3.77)$$

Equation clearly shows that the average fading rate within a local area decreases with respect to omnidirectional propagation as multipath power becomes more and more concentrated about a single azimuthal direction.

From equation (3.76), fading rate variance $\bar{\sigma}^2$ will change as a function of direction of receiver travel θ, but will always fall within the following range:

$$\sqrt{1-\gamma} \leq \frac{\bar{\sigma}(\theta)}{\Lambda} \leq \sqrt{1+\gamma} \quad ...(3.78)$$

The upper limit of equation (3.78) corresponds to a receiver traveling in the azimuthal direction of maximum fading ($\theta = \theta_{max}$) while the lower limit corresponds to travel in a perpendicular direction ($\theta = \theta_{max} + \pi/2$). Equation (3.78) clearly shows that the variability of fading rates within the same local area increases as the channel becomes more and more constricted. In fact, equation (3.76) provides mathematical criteria for a multipath channel that may be treated as **"pseudo-omnidirectional"**.

$$|F_1|, \ |F_2 << F_0 \quad ...(3.79)$$

under the condition of equation (3.79), angular spread becomes approximately 1 and angular constriction becomes approximately 0. Thus, the second-order statistics of the channel behave nearly identical to the classical omnidirectional channel.

3.18 SECOND-ORDER STATISTICS USING SHAPE FACTORS

With an understanding of how shape factors describe fading rate variances, it is possible to re-derive many of the basic second-order statistical measures of fading channels. Level-crossing rates, average fade duration, spatial autocovariance, and coherence distance expressions that were originally derived under the assumption of omnidirectional multipath propagation will now be cast in terms of the angular spread, the angular constriction, and the azimuthal directional of maximum fading. A Rayleigh fading signal is one whose envelope K, follows a Rayleigh PDF, $P_R(r)$ given by

$$P_R(r) = r\sqrt{\frac{2}{P_T}}\exp\left(\frac{-r^2}{P_T}\right) \quad r \geq 0 \quad ...(3.80)$$

where P_T is the average total power received in a local area (units of volts-squared).

(*i*) **Level-crossing rates and average fade duration:** The general expression for a level-crossing rate is given by

$$N_R = \int_0^\infty \dot{r}\, p(R,\dot{r})\, d\dot{r} \quad ...(3.81)$$

where R is the threshold level and $p(R,\dot{r})$ is the joint PDF of envelope and its time derivative. For a Rayleigh-fading signal, the level-crossing rate of the envelope process is

$$N_R = \frac{\rho\sigma_{\tilde{V}}}{\sqrt{\pi P_T}}\exp(-\rho) \quad ...(3.82)$$

By substituting Equation (3.73) into equation (3.82) we arrive at an exact expression for the level-crossing rate in a Rayleigh fading channel with any arbitrary spatial distribution of multipath power and any direction of mobile receiver travel θ:

$$N_R = \frac{\sqrt{2}\pi v \Lambda \rho}{\lambda}\sqrt{1+\gamma\cos[2(\theta-\theta_{\max})]\exp(-\rho^2)} \quad ...(3.83)$$

The average fade duration, $\bar{\tau}$ is defined to be

$$\bar{\tau} = \frac{1}{N_R}\int_0^R f_R(r)\,dr \quad ...(3.84)$$

Substitution of the Rayleigh PDF of Equations (3.83) and (3.85) into equation (3.67) yields

$$\bar{\tau} = \frac{\lambda[\exp((\rho^2)-1)]}{\sqrt{2\pi}v\Lambda\rho\sqrt{1+\gamma\cos[2(\theta-\theta_{\max})]}} \quad ...(3.85)$$

Equation (3.83) and (3.85) are useful tools for studying small-scale fading statistics in the presence of non-omnidirectional multipath, and yield identical results to those when a omnidirectional channel is assumed.

(*ii*) **Spatial Autocovariance:** Another important second-order statistic is the spatial autocovariance of received voltage envelope. The autocovariance function determines the correlation of received voltage envelope as a function of change in receiver position and is useful for studies in spatial diversity. The approximate expression for the spatial autocovariance function of envelope based on shape factors is given by

$$p(r,\ \theta) = \left[-23\Lambda^2(1+\lambda\cos[2(\theta-\theta_{\max}])\left(\frac{r}{\lambda}\right)^2\right] \quad ...(3.86)$$

Equation (3.86) allows us to estimate the envelope correlation between two points in space separated by a distance r along an azimuthal direction θ.

(*iii*) **Coherence Distance: Coherence Distance (D_c)** is the separation distance in space over which a fading channel appears to be unchanged. For mobile receivers, a similar parameters called **coherence time T_c,** is the elapsed time over which a fading channel appears to be constant. For the case of a static channel, the coherence time of a mobile receiver may be calculated from the coherence distance ($T_C = D_C/v$ where v is the speed of the mobile). Definitions for coherence distance may be based on the envelope autocovariance function. A convenient definition for the coherence distance D_c, is the value that satisfies the equation $\rho(D_c) = 0.5$. The classical value for coherence distance in an omnidirectional Rayleigh channel is given by

$$D_c \ » \ \frac{9\lambda}{16\pi} \quad ...(3.87)$$

using the generalized autocovariance function of equation (3.86) and evaluating $\rho(r,\ \theta)|_{r\ =\ D_c} = 0.5$ leads to a new definition of coherence distance :

$$D_c = \frac{\lambda\sqrt{\ln 2}}{\Lambda\sqrt{23(1-\gamma\cos[2(\theta-\theta_{\max})])}} \quad ...(3.88)$$

For omnidirectional propagation, equation (3.88) differs from equation (3.70)′ by only –3.0%. Furthermore, equation (3.88) captures the behavior of non-omnidirectional multipath. As angular spread, A decreases the coherence distance in a local area increases. As the angular constriction γ, increases the coherence distance develops a strong dependence on orientation θ.

3.19 REVISITING CLASSICAL CHANNEL MODELS WITH SHAPE FACTORS

The first case corresponds to a narrowband receiver operating in a local area with multipath arriving from all directions, such that the angular distribution of power $p(\theta)$ is a constant. The receiver antenna is assumed to be an omnidirectional whip, oriented perpendicular to the ground. Due to the vertical electric-field polarization of the whip antenna, this propagation scenario is referred to as the E_z-case.

The second two-cases correspond to the same narrowband receiver in the same omnidirectional multipath channel, but with a small loop antenna mounted at top of the receiver such that the plane of the loop is perpendicular to the ground. The antenna pattern of the small loop antenna attenuates the arriving multipath such that the angular distribution of power becomes

$$p(\theta) = A \sin^2 \theta \qquad ...(3.89)$$

where A is some arbitrary gain constant. Unlike the omnidirectional E_z-case, the statistics of the propagation scenario will depend on the direction of travel by the receiver. The H_x-case will refer to a receiver traveling in a direction perpendicular to the main lobes of the loop antenna pattern (9 = 0). The H_y-case will refer to a receiver traveling in a direction parallel to the main lobes (9 = 7i/2). Figure 3.11 illustrates the E_z, H_x and H_y cases for the modeled receiver antennas. The spatial parameters for the E_z-case are A, y and $9_{max} = 0$. Since this case is omnidirectional, the angular spread is at a maximum ($\Lambda = 1$) and the angular constriction is at a minimum ($\gamma = 0$). For the H_{x^-} and H_{y^-} cases, the spatial parameters are $\Lambda = 1$, $\gamma = 1/2$ and $\theta_{max} = \pi/2$. Since the impinging multipath have no clear bias in one direction, the angular spread is at a maximum just like the E_z-case. However, there is clearly a bias in two directions, resulting in an increased angular constriction of $\gamma = 1/2$

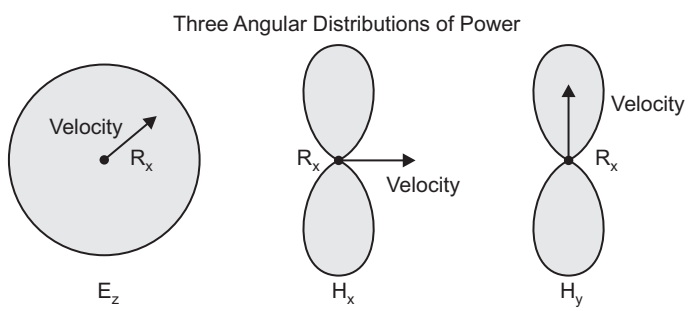


Fig 3.11: Three different multipath-induced mobile-fading scenarios.

After substitution of these parameters into equation (3.83) along with the apprpriate direction of mobile travel, the level-crossing rates for the three cases become

$$E_z: \qquad N_R = \frac{\sqrt{2\pi v\rho}}{\lambda}\exp(-\rho^2) \qquad ...(3.90)$$

$$H_x: \qquad N_R = \frac{\sqrt{\pi v\rho}}{\lambda}\exp(-\rho^2) \qquad ...(3.91)$$

$$H_y: \qquad N_R = \frac{\sqrt{3\pi}\, v\rho}{\lambda}\exp(-\rho^2) \qquad ...(3.92)$$

The corresponding average fade durations are

$$E_z: \qquad \bar{\tau} = \frac{\lambda}{\sqrt{2\pi}\, vp}[\exp(\rho^2)-1] \qquad ...(3.93)$$

$$H_x: \qquad \bar{\tau} = \frac{\lambda}{\sqrt{\pi}\, vp}[\exp(\rho^2)-1] \qquad ...(3.94)$$

$$H_y: \qquad \bar{\tau} = \frac{\lambda}{\sqrt{3\pi}\, vp}[\exp(\rho^2)-1] \qquad ...(3.95)$$

These expressions exactly match the original solution presented by clarke. Now substitute the channel shape factors into the approximate spatial autocovariance functions in equation (3.86).

The results for the three cases are:

$$E_z: \qquad \rho(r) = \exp\left[-23\left(\frac{r}{\lambda}\right)^2\right] \qquad ...(3.96)$$

$$H_x: \qquad \rho(r) = \exp\left[-11.5\left(\frac{r}{\lambda}\right)^2\right] \qquad ...(3.47)$$

$$H_y: \qquad \rho(r) = \exp\left[-34.5\left(\frac{r}{\lambda}\right)^2\right] \qquad ..(3.48)$$

Note that all three model the spatial autocovariance function consistent with the approximation made in the derivation of equation 3.86. The behavior is nearly exact for values of r equal to or less than a correlation distance. The use of spatial parameters to find level crossing rate, average fade duration, and spatial autocovariance is quite simple when compared to the full analytical solutions of the E_Z, H_x and H_y cases. The proposed solution is also more comprehensive.

SOLVED EXAMPLES

Examples 3.1. *Find the far-field distance for an antenna with maximum dimension of 1m and operating frequency of 900 MHz?*

Solution. Given:

Largest dimension of antenna D = lm

Operating frequency f = 900 MHz, λ = ?

$$\lambda = \frac{c}{f} = \frac{3\times10^8\,\text{m/s}}{900\times10^6\,\text{Hz}}$$

Using equation, far-field distance is obtained as

$$df = \frac{2(1)^2}{0.33} = 6\text{m}$$

Examples 3.2. *If a transmitter produces 50 W of power, express the transmit power in units of*

(*a*) dB_m

(*b*) dB_w

If 50 W is applied to a unity gain antenna with a 900 MHz, carrier frequency, find the received power in dBm at a free space distance of 100 m from the antenna. What is Pr (10 km) ? Assume unity gain for receiver antenna.

Solution. Given:

Transmitter power, P_t = 50 W

Carrier frequency, f_c = 900 MHz

Using equation

(*a*) Transmitter power

$$p_t(\text{dBm}) = 10\log\left[\frac{P_t(\text{mW})}{(1\,\text{mW})}\right]$$

$$= 10 \log [50 \times 10^3] = 47.0 \text{ dBm}$$

(*b*) Transmitter power

$$p_t\,(\text{dBw}) = 10\log\left[\frac{P_t(\text{mW})}{(1\,\text{W})}\right]$$

$$= 10\ 1\text{og}\ [50] = 17.0 \text{ dBw}$$

The received power can be determined as

$$P_r = \frac{P_1G_1G_r\lambda^2}{(4\pi)^2d^2L} = \frac{50(1)(1)\left(\frac{1}{3}\right)^2}{(4\pi)^2(100)^2(1)}$$

$$= (8.5 \times 10^{-6})w = 3.5 \times 10^{-3}\ mw$$

$$Pr(\text{dBm}) = 10 \log P_r\ (\text{mW}) = 10 \log (3.5 \times 10\text{–}3 \text{ mW})$$

$$= -24.5 \text{ dBm}$$

The received power at 10 km can be expressed in terms of dBm where,

$$d_0 = 100 \text{ m}$$

$$d = 10 \text{ km}$$

$$P_r(10 \text{ km}) = P_r(100) + 20 \log\left[\frac{100}{10000}\right]$$
$$= -24.5 \text{ dBm} - 40 \text{ dB}$$
$$= -64.5 \text{ dBm}$$

Examples 3.3. *Assume a receiver is located 10 km from a 50 W transmitter. The carrier frequency is 900 MHz, free space propagation is assumed, G_t = 1 and G_r = 2, find*

(*a*) *The power at the receiver*

(*b*) *The magnitude of the E-field at the receiver antenna*

(*c*) *The rms voltage applied to the receiver input assuming that the receiver antenna has a purely real impedance of 50 ohm and is matched to the receiver.*

Solution: Given:

Transmitter power, P_t = 50 W

Carrier frequency, f_c = 900 MHz

Transmitter antenna gain, G_t = 1

Receiver antenna gain, G_r = 2

Receiver antenna resistance = 50 Ω

(*a*) The power received at distance d = 10 km

$$P_r(d) = 10\log\left(\frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2}\right)$$

$$= 10 \log\left(\frac{50 \times 1 \times 2 \times \left(\frac{1}{3}\right)^2}{(4\pi)^2 \times (10000)^2}\right)$$

$$= -91.5 \text{ dBw} = -61.5 \text{ dBm}$$

(*b*) The magnitude of the received *E*-field is

$$|E| = \sqrt{\frac{P_r(d)120\pi}{A_e}} = \sqrt{\frac{P_r(d)120\pi}{G_r \frac{\lambda_2}{4\pi}}} = \sqrt{\frac{7 \times 10^{-10} \times 120\pi}{\frac{2 \times 0.33^2}{4\pi}}}$$

$$= 0.0039 \text{ V/m}$$

(*c*) The open *ckt* rms voltage at the receiver input is

$$V_{ant} = \sqrt{P_r(d) \times 4R_{ant}} = \sqrt{7 \times 10^{-10} \times 4 \times 50} = 0.374 \text{ mV}$$

Example 3.4. *Calculate the Brewster angle for a wave impinging on ground having a permittivity of E_r = 4.*

Solution: The Brewster angle can be found by substituting the values for E_r in equation.

$$\sin(0_i) = \frac{\sqrt{E_r - 1}}{\sqrt{(E_r)^2 - 1}}$$

$$\sin(0_i) = \frac{\sqrt{(4)-1}}{(4)^2-1} = \sqrt{\frac{3}{15}} = \sqrt{\frac{1}{5}}$$

$$0_i = \sin^{-1}\sqrt{\frac{1}{5}} = 26.56°$$

Example 3.5. *A mobile is located 5 km away from a base station and uses a vertical* $\frac{\lambda}{4}$ *monopole antenna with a gain of 2.55 dB to receive cellular radio signals. The E-field at 1 km from the transmitter is measured to be* 10^{-3} *V/m. The carrier frequency used for this system is 900 MHz.*

(*a*) *Find the length and the effective aperture of the receiving antenna ?*

(*b*) *Find the received power at the mobile using the two-ray ground reflection model assuming the height of the transmitting antenna is 50 m and the receiving antenna is 1.5 m above ground*?

Solution. Given:

T-R separation distance = 5 km?

E-field at a distance of 1 km = 10^{-3} V/m

Frequency of operation, f = 900 MHz

$$\lambda = \frac{c}{f} = \frac{3\times10^8}{900\times10^6} = 0.333 \text{ m}$$

(*a*) Length of antenna, $L = \frac{\lambda}{4} = \frac{0.333}{4} = 0.8833 \text{ m} = 8.33 \text{ m}$. Effective aperture of $\frac{\lambda}{4}$ monopole antenna can be obtained.

Effective aperture of antenna = 0.016 m^2

(*b*) Since $d >> \sqrt{h_t h_r}$, the electric field is given by

$$E_r(d) = \frac{2E_0 d_0}{d} \times \frac{2\pi h_t h_r}{\lambda d} = \frac{k}{d^2} \text{ V/m}$$

$$= \frac{2\times10^{-3}\times1\times10^{-3}}{5\times10^3}\left[\frac{2\pi(50)(1.5)}{0.333(5\times10^3)}\right]$$

$$= 113.1\times10^{-6} \text{ V/m}$$

The received power at a distance d can be obtained

$$P_r(d) = \frac{(113.1\times1\times10^{-6})}{377}\left[\frac{1.8(0.333)^2}{4\pi}\right]$$

$$P_r(d) = 5 \text{ km} = 5.4 \times 10^{-13} \text{ w}$$

$$= -122.68 \text{ dBw or } -92.68 \text{ dBm}$$

Example 3.6. *Find the median path loss using Okumura's model for d = 50 km, h_{re} = 100 m, h_{re} = 10 m in a suburban environment? If the base station transmitter radiates an EIRP of 1 kW at a carrier frequency of 900 MHz, find the power at the receiver (assume a unity gain receiving antenna).*

Solution. The free space path loss L_F can be calculated using equation as

$$L_F = 10\log\left[\frac{\lambda^2}{(4\pi)^2 d^2}\right]$$

$$= 10\log\left[\frac{\left(\frac{3\times10^8}{900\times10^6}\right)^2}{(4\pi)^2\times(50\times10^3)^2}\right] = 125.5 \text{ dB}$$

From the Okumura curves

A_{mu} [900 MHz (50 km)] = 43 dB and

G_{AREA} = 9 dB

We have

$$G(h_{te}) = 20\log\left(\frac{h_{te}}{200}\right) = 20\log\left(\frac{100}{200}\right)$$

$$= -6 \text{ dB}$$

$$G(h_{re}) = 20\log\left(\frac{h_{re}}{3}\right) = 20\log\left(\frac{10}{3}\right)$$

$$= 10.46 \text{ dB}$$

The total mean path loss is

$$L_{50}(\text{dB}) = L_F + A_{mu}(f, d) - G(h_{te}) - G(h_{re}) - G_{\text{AREA}}$$

$$= 125.5 \text{ dB} + 43 \text{ dB} - (-6)\text{dB} - 10.46 \text{ dB} - 9 \text{ dB}$$

$$= 155.04 \text{ dB}$$

Therefore, the median received power is

$$P_r(d) = \text{EIRP(dBm)} - L_{50}(\text{dB}) + G_r(\text{dB})$$

$$= 60 \text{ dBm} - 155.04 \text{ dB} + 0 \text{ dB}$$

$$= -95.04 \text{ dBm}$$

Example 3.7. *Consider a transmitter which radiates a sinusoidal carrier frequency of 1850 MHz. For a vehicle moving 60 mph, compute the received carrier frequency if the mobile is moving*

(*a*) *Directly toward the transmitter*

(*b*) *Directly away from the transmitter*

(*c*) *In a direction which is perpendicular to the direction of arrival of the transmitted signal.*

Solution. Given:

Carrier frequency f_c = 1850 MHz

Wavelength $\lambda = \frac{c}{f_c} = \frac{3\times10^8}{1850\times10^6} = 0.162 \text{ m}$

Vehicle speed v = 60 mph = 26.82 m/s

(*a*) The vehicle is moving directly toward the transmitter. The Doppler shift in this case is positive and the received frequency is given by equation

$$f = f_c + f_d = 1850 \times 10^6 + \frac{26.82}{0.162}$$

$$= 850.00016 \text{ MHz}$$

(*b*) *The* vehicle is moving directly away from the transmitter the Doppler shift in this case is negative and hence the received frequency is given by

$$= 1849.99934 \text{ MHz}$$

(*c*) The vehicle is moving perpendicular to the angle of arrival of the transmitted signal.

In this case, $\theta = 90°$, $\cos\theta = 0$

And there is no Doppler shift.

The received signal frequency is the same as the transmitted frequency of 1850 MHz.

Example 3.8. *Assume a discrete channel impulse response is used to model urban RF radio channels with excess delays as large as 100 μs and microcellular channels with excess delays no larger than 4 μs. If the number of multipath bins is fixed at 64, find*

(*a*) $\Delta\tau$

(*b*) *The maximum RF bandwidth which the two models can accurately represent.*

Repeat the exercise for an indoor channel model with excess delays as large as 500 ns. As SIRCIM and SMRCIM are statical channel models based.

Solution. The maximum excess delay of the channel model is given by $t_N = N_{ot}$.

Therefore, for τ_N = 100 (μs, and N= 64 we obtain $dt = \frac{\tau_N}{N}$ = 1.5625 μs. The maximum bandwidth that the SMRCIM model can accurately represent is equal to

$$\frac{2}{\Delta\tau} = \frac{2}{1.5625} = 1.28 \text{ MHz}$$

For the SMRCIM urban microcell model,

$\tau_N = 4\mu_s, \Delta\tau = \frac{\tau_N}{N} = 62.5 \text{ ns}.$ The maximum RF bandwidth that can be represented as

$$\frac{2}{\Delta\tau} = \frac{2}{62.5 \text{ ns}} = 32 \text{ MHz}$$

Similarly, for indoor channels

$$\Delta\tau = \frac{\tau_N}{N} = \frac{500 \times 10^{-9}}{64}$$

$$= 7.8125 \text{ ns}$$

The maximum RF bandwidth for the indoor channel model is

$$\frac{2}{\Delta\tau} = \frac{2}{7.8125 \text{ ns}} = 256 \text{ MHz}$$

Example 3.9. *Assume a mobile traveling at a velocity of 10 m/s receives two multipath components at a carrier frequency of 1000 MHz. The first component is assumed to arrive at* $\tau = 0$ *with an initial phase of 0° and a power of –70 dBm, and the second component which is 3 dB weaker than the first component is assumed to arrive at* $\tau = 1$ *μs, also with an initial phase of 0°. If the mobile moves directly toward the direction of arrival of the first component and directly away from the direction of arrival of the second component. Compute the narrow band instantaneous power at time intervals of 0.1s from 0s to 0.5s. Compute the average narrowband power received over this observation interval. Compare average narrowband and wideband received powers over the' interval, assuming the amplitude of the two multipath components do not fade over the local area.*

Solution. Given:

$$V = 10 \text{ m/s}$$

Time intervals of 0.1 s correspond to spatial intervals of 1 m. The carrier frequency is given to be 1000 MHz. Hence the wavelength of the signal is

$$\lambda = \frac{c}{f} = \frac{3\times10^8}{1000\times10^6} = 0.3 \text{ m}$$

The narrowband instantaneous power can be computed as,

Note. 70 dB_m = 100 PW. At time $t = 0$, the phases of both multipath components are 0°, hence the narrowband instantaneous power is equal to

$$|r(t)| = \left|\sum_{i=0}^{N-1} a_i \exp[j\theta_i(t,\tau)]\right|^2$$

$$= \left|\sqrt{100\text{PW}} \times \exp(0) + \sqrt{50\text{PW}} + \exp(0)\right|^2$$

$$= 291 \text{ PW}$$

Now as the mobile moves, the phase of the two multipath components changes in opposite directions.

At $t = 0.1$ s, the phase of the first component is

$$\theta_i = \frac{2\pi d}{\lambda} = \frac{2\pi vt}{\lambda} = \frac{2\pi\times10(\text{m/s})\times0.1s}{0.3 \text{ m}}$$

$$= 20.94 \text{ rad} = 2.09 \text{ rad} = 120°$$

Since the mobile moves toward the direction of arrival of the first component, and away from the direction of arrival of the second component, θ_1 is positive and θ_2 is negative.

Therefore, at t = 0. 1s, θ_1 = 120° and θ_2 = –120° and the instantaneous power is equal to

$$|r(t)|^2 = \left|\sum_{i=0}^{N-1} a_i \exp[j\theta_i(t,\tau)]\right|^2 = \left|\sqrt{100\text{ PW}} \times \exp(j120) + \sqrt{50\text{PW}} \times \exp(-j120)\right|^2$$

$$= 79.3 \text{ PW}$$

Similarly, at t = 0.2 s, θ_1 = 240° and θ_2 = –240° and the instantaneous power is equal to

$$|r(t)^2 = \left|\sum_{i=0}^{N-1} a_i \exp[j\theta_i(t,\tau)]\right|^2$$

$$= \left|\sqrt{100\text{ PW}} \times \exp(j120°) + \sqrt{50\text{PW}} \times \exp(-j120°)\right|^2$$

$$= 79.3 \text{ PW}$$

Similarly, at t = 0.3s, θ_1 = 360° = 0° and θ_1 = 0° and the instantaneous power is equal to

$$|r\ (t)|^2 = \left|\sum_{i=0}^{N-1} a_i \exp[j\theta_i(t,\tau)]\right|^2$$

$$\left|\sqrt{100\text{ PW}} \times \exp(j0°) + \sqrt{50\text{PW}} \times \exp(-j0°)\right|^2$$

$$= 291 \text{ PW}$$

It follows that at t = 0.4s, $|r\ (t)^2$ = 79.3 PW and at t = 0.5s, $|r(t)|^2$ = 79.3 PW. The average narrowband received power is equal to

$$\frac{(2)(291)+(4)(79.3)}{6}\text{PW} = 149 \text{ PW}$$

The wideband power is given by

$$E_{a,\theta}\ [P_{w,\ B}] = 100 \text{ PW} + 50 \text{ PW} = 150 \text{ PW}$$

As can be seen, the narrowband and wideband received power are virtually identical when averaged over 0.5s. While the (w signal fades over the observation interval, the wideband signal power remain constant over the same spatial interval.

Example 3.10. *Compute the RMS delay spread for the following power delay profile:*

(*a*) *P*(τ)

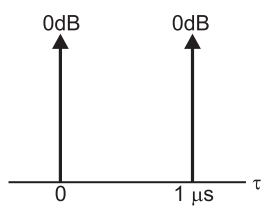


Fig. 3.12

(*b*) *If BPSK modulation is used, what is the maximum bit rate that can be sent through the channel without needing an equalizer?*

Solution. (*a*)

$$\bar{\tau} = \frac{(1)(0)+(1)(1)}{1+1} = \frac{1}{2} = 0.5\ \mu s$$

$$\bar{\tau^2} = \frac{(1)(0)+(1)(1)^2}{1+1} = \frac{1}{2} = 0.5\ \mu s^2$$

$$\sigma_\tau = \sqrt{\bar{\tau^2} - (\bar{\tau})^2} = \sqrt{0.5-(0.5)^2}$$

$$= \sqrt{0.25} = 0.5\ \mu s$$

(*b*)

$$\frac{\sigma_\tau}{T_s} \le 0.1$$

$$T_s \ge \frac{\sigma_\tau}{0.1}$$

$$T_s \ge 5\ \mu s$$

$$R_s = \frac{1}{T_s} = 0.2\times 10^6 \text{sps} = 200 \text{ ksps}$$

$$R_b = 2000 \text{ Kbps}$$

Example 3.11. *For a Rayleigh fading signal, compute the positive-going level crossing rate for P = 1, when the maximum Doppler frequency (fj is 20 Hz. What is the maximum velocity of the mobile for this Doppler frequency if the carrier frequency is 900 MHz?*

Solution. The number of zero level crossing

$$N_R = \sqrt{2\pi}(20)(1)e^{-1} = 18.44 \text{ crossings per second}$$

The maximum velocity of the mobile can be obtained using the Doppler relation $f_{d.\max} = \frac{V}{\lambda}$

Therefore velocity of the mobile at f_m = 20 Hz

$$v = f_d\lambda = 20 \text{ Hz}\left(\frac{1}{3 \text{ m}}\right) = 6.66 \text{ m/s} = 2 \text{ km/hr}$$

Example 3.12. *Find the average fade duration for threshold levels P = 0.01, P = 0.1 and P = 1, when the Doppler frequency is 200 Hz.*

Solution: Average fade duration can be found by substituting the given values as

For P = 0.01

$$\bar{\tau} = \frac{e^{0.01}-1}{(0.01)200\sqrt{2\pi}} = 19.9\ \mu s$$

For P = 0.1

$$\bar{\tau} = \frac{e^{0.1^2}-1}{(0.1)200\sqrt{2\pi}} = 200\ \mu s$$

For $P = 1$ $$\bar{\tau} = \frac{e^{1^2} - 1}{(1)200\sqrt{2\pi}} = 3.43\ \mu s$$

R_b = 200 Kbps

HIGHLIGHTS

1. These are three ranges: transmission range, detection range and interference ranges are used to define propagation of radio signal.
2. These are three mechanism are responsible for electromagnetic propagation, which are reflection, diffraction and scattering.
3. The reflection mechanism occurs when the signal strike with an object, whose dimension is large compared with wavelength of signal.
4. The diffraction mechanism occurs when the radio wave signal is strike at the edge of object and propagate in the different direction.
5. The scattering mechanism occurs when the dimension of the obstacle is in the order of the wavelength of signal or less.
6. There are two type of radio channels are used or communication channels to transmit the information.
   (*i*) Wired channels
   (*ii*) Wireless channel
7. The received power strength at a particular location is predict by the free space propagation model when the transmitter and receiver are connected with clear, unobstructed line of sight path between them.
8. The effective isotropic radiated power (EIRP) is defined by the product of transmitted power (P_t) and transmitting at antenna gain (G_t).
   EIRP = $P_r G_t$
9. Generally the radio waves can exhibit the three fundamental propagation behaviours depending on their frequency
   (*iii*) Ground wave (L2 MHz)
   (*ii*) Sky wave (2 – 30 MHz)
   (*iii*) Line of sight (> 30 MHz)
10. Generally the propagation models are divided into two parts:
    (*i*) Outdoor propagation model
    (*ii*) Indoor propagation model.
11. The outdoor propagation model related with the outside environment and predict signal strength at a particular receiving point or in a specific local area.
12. There are various type of outdoor propagation models are used to define the path loss by the irregular terain line presence of trees, buildings and other obstacles-like as longley-rice models, Durkin's models, Okumura's model, Hata models and Walfisch of Bertoni's model.
13. The indoor propagation model characterize the radio propagation inside the buildings nd this model is also influenced by the construction material.
14. In general, the indoor channel may be classified either as line of sight or obstructed (OBS) with varying degree of clutter.

15. There are various type of indoor propagation model are used to define the variation in signal strength line as long distance path loss model, Ericsson multiple breakpoint model and attention Factor model.

16. The effective aperture defined the ratio of the power available at the antenna terminal to the power per unit area of the appropriately palarized incident electromagnetic wave.

$$A_e = \frac{\lambda^2}{4\pi} \cdot G$$

and the wavelength define the ratio between speed of light ($c = 3 \times 10^8$ m/s) and frequency $\lambda = \frac{c}{f}$

17. The fading is occurs when the amplitude of received signal level is fluctuates at receiver due to corruptive element in propagation medium.

18. The multipath fading is also known as Rayleigh fading and it is produce by the multipath propagation.

19. When a radio wave propagating from one medium to another medium then the wave is partially reflected and partially transmitted if wave have different electrical property.

20. The small scale fading based on multipath time delay spread is divided into two parts, one is flat fading in which the bandwidth of signal is less than bandwidth of the channel, and second is frequency selective fading in which the bandwidth of signal is greater than the bandwidth of channel.

21. The small scale fading is also divided into part which is based on Doppler spread, one is fast fading which has high Doppler spread and second is slow fading, the coherence time is greater than the symbol period.

22. In the fast fading, the channel variation are faster than the base band signal variation but in case of slow fading the channel variations are slower than base band signal variations.

23. In the flat fading, the delay spread is less than the symbol period but in lose of frequency fading, the delay spread is greater than the symbol period.

24. The time selective channels are time varying channels, for an example wireless terminal moving through the environment and undergoing Rayleigh fading.

25. The frequency channels have a flat response that cannot be assumed to be flat over the bandwidth of the signal.

26. The frequency selectivity is due to multipath that has significant delay spread relative the symbol period of the transmission.

PROBLEMS

3.1. If P_t = 10 W, G_t = 0 dB, G_r = 0 dB and f_c = 900 MHz, find P_r in watts at a free space distance of 1 km.

3.2. Assume a receiver is located 10 km from a 50 W transmitter. The carrier frequency is 6 GHz and free space propagation is assumed G_t – 1 and G_r = 1.

(*a*) Find the power at the receiver

(*b*) Find the magnitude of the E-field at the receiver antenna.

(*c*) Find the rms voltage applied to the receiver input, assuming that the receiver antenna has a purely real impedance of 50 Ω and is matched to the receiver.

3.3. Fraunhofer distance: Calculate the gain, half power beamwidth (HPBW), and Fraunhofer distance for a uniformly illuminated horn antenna at 60 GHz with dimensions of 4.6 cm x 3.5 cm.

Hint: HPBW for the horn antenna can be estimated as HPBW = 51 λ/a where, a is the aperture width.

3.4. Free space propagation: Assume the transmitter power is 1W at 60 GHz fed into the transmitter antenna. Using the horn antenna from problem 3.3 at both the transmitter and receiver.

(*a*) Calculate the free space path loss at 1m, 100 m and 1000 m.

(*b*) Calculate the received signal power at these distance.

(*c*) What is the rms voltage received at the antenna. If the receiver antenna has purely real impedance of 50 £1 and is matched receiver?

3.5. For the Knife- edge geometry in Figure 3.13, show that

(*a*) $\phi = \frac{2\pi\Delta}{\lambda} = \frac{2\pi}{\lambda}\left[\frac{h^2}{2}\left(\frac{d_1 + d_2}{d_1\, d_2}\right)\right]$

(*b*) $V = \alpha\sqrt{\frac{2d_1\, d_2}{\lambda(d_1 + d_2)}}$ and $\frac{v^2\pi}{2} = \phi, d_1, d_2 > h, h >> \lambda$ and $\Delta = P_1 + P_2 - (d_1 + d_2)$

3.6. In the following cases, tell whether the two-ray model could be applied, and explain why or why not:

h_t = 35 m, h_r = 3m, d – 250 m

h_t = 30m, h_r = 1.5 m, d = 450 m

3.7. Assuming a receiver is located 10 km from a 50 W transmitter. The carrier frequency is 1900 MHz, free space propagation is assumed, $G_t = l$, $G_r = 2$, find:

(*a*) The power at the receiver.

(*b*) The magnitude of the E-field at the receiver antenna.

(*c*) The open-circuit rms voltage applied to the receiver input assuming that the receiver antenna has a purely real impedance of 50 Ω and is matched to the receiver.

(*d*) Find the received power at the mobile using the two-ray ground reflection model assuming the height of the transmitting antenna is 50 m, receiving antenna is 1.5 m above the ground, and the ground flection is –1.

3.8. Diffraction: From the knife-edge defloration model. Show how the diffracted power depends on frequency. Assume $d_1 = d_2$ = 500 m and h = 10 m in figure 3.12, *Hint:* You may need to calculate Fresnel's integral or use Lee's approximation.

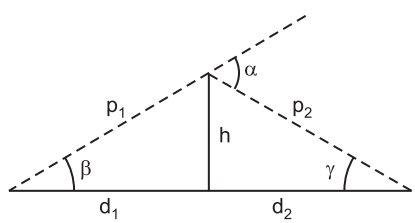


Fig. 3.13: Knife-edge geometry for problem 3.5.

3.9. If P_t = 10 W, G_t = 10 dB, G_r = 3 dB and L = 1 dB at 900 MHz, compute the received power for the knife-edge geometry shown in figure 3.14. Compare this value with the theoretical free space received power if an obstruction did not exist. What is the path loss due to diffraction for this case?

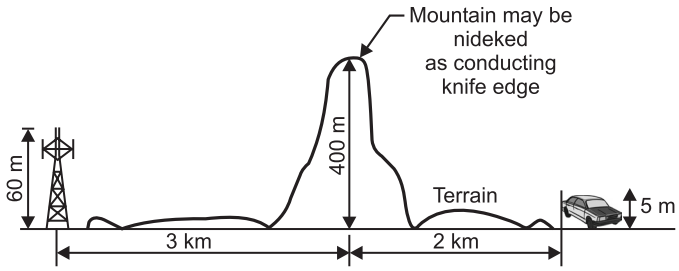


Fig. 3.14: Knife-edge geometry tor Problem 3.9.

3.10. Assume that local average signal strength field measurements were made inside a building, and past processing revealed that the measured data fit a distant-dependent mean power law model having a log-normal distribution about the mean. Assume the mean power law was found to be $P_r\ (d) \propto= d^{-3.5}$. If a signal of 1 MW was received at d_0 = 1 m from the transmitter and at a distance of 10 m 10% of the measurements ware stronger than – 25 dBm, define the standard deviation, σ, for the path loss model at d = 10 m.

3.11. If the received power at a reference distance d_0 = 1 km is equal to 1 microwatt. Find the received powers at distance of 2 km, 5 km, 10 km and 20 km from the same transmitter for the following path loss model.

(*a*) Free space

(*b*) h = 3

(*c*) h = 4

(*d*) Two-ray ground reflection using the exact expression,

(*e*) Extended Hata model for a large city environment.

Assume f = 1800 MHz, h_t = 40 m, h_r = 3m, G_t = G_r = 0 dB.

Plot each of these models on the same graph over the range of 1 km to 20 km. Comment on the differences between these five models.

3.12. A transmitter provides 15 W to an antenna having 12 dB gain. The receiver antenna has a gain of 3 dB and the receiver handwidth is 30 kHz. If the receiver system noise figure is 8 dB and the carrying frequency is 1800 MHz. Find the maximum *T-R* separation that will ensure that a SNR of 20 dB is provided for 95% of the time. Assume h = 4, o = 8 dB and d_0 = 1 km.

3.13. Assume a SNR of 25 dB is desired at the receiver. If a 900 MHz cellular transmitter has an ERP of 100 W and the AMPS receiver uses a 0 dB gain antenna and has a 10 dB noise figure, find the percentage of time that the desired SNR is achieved at a distance of 10 km from the transmitter. Assume h = 4, σ = 8 dB and d_0 = 1 km.

3.14. Four received power measurements were taken at distance of 100 m, 200 m, 1 km and 2 km from a transmitter. The measured values at these distances are – 0 dB_m, – 25 dB_m, – 35 dB_m, and – 38 dB_m respectively. It is assumed that the path loss for these measurements follows the model.

$$PL(d)[dB] = \overline{PL}(d) + x_\sigma = \overline{PL}(d_0) + 10n\log\left(\frac{d}{d_0}\right) + x_n$$

where, d_0 = 100m

(*a*) Find the minimum mean square error (MMSE) estimate for the path loss exponent *n*.

(*b*) Calculate the standard deviation of shadowing about the mean value.

(*c*) Estimate the received power at d = 2 km using the resulting model.

(*d*) Predict the likelihood that the received signal level at 2 km will be greater than -35 dBm express your answer as a *Q*-function.

3.15. Determine the maximum and minimum spectral frequencies received from a stationary GSM transmitter that has a center frequency of exactly 1950.000000 MHz, assume that the receiver is traveling at speeds of

(*a*) 1 km/hr (*b*) 5 km/hr

(*c*) 100km/hr (*d*) 1000 km/hr

3.16. Describe all the physical circumstances that relate to a stationary transmitter and a moving receiver such that the Doppler shift at the receiver is equal to

(a) 0 Hz (*b*) $f_{d\max}$

(*c*) $-f_{d\ \max}$ (*d*) $f_{d\max/2}$

3.17. If a particular modulation provides suitable BER performance whenever $\sigma_\tau/T_s \leq 0.1$ determine the smallest symbol period T_s (and thus the greatest symbol rate) that may be sent through RF channels shown in figure without using an equalizer.

3.18. For a mobile receiver operating at frequency of 860 MHz and moving at l00 km/hr

(*a*) Sketch the Doppler spectrum if *a cw* signal is transmitted and indicated the maximum and minimum frequencies.

(*b*) Calculate the level crossing rate and average fade duration if $P = -20$ dB.

3.19. Derive the RF Doppler spectrum for a 5/8*d*, vertical monopole receiving a *cw* signal using the models by Clarke and Gans. Plot the RF Doppler spectrum and the corresponding base-band spectrum out of an envelope detector. Assume isotropic scattering and unit average received power.

3.20. A flat Rayleigh fading signal at 8 Hz × 6 GHz is received by a mobile traveling at 80km/hr.

(*a*) Determine the number of positive-going zero crossings about the rms value that occur over a 5s interval.

(*b*) Determine the average duration of a fade below the rms level.

(*c*) Determine the average duration of a fade at a level of 20 dB below the rms value.

3.21. Show that Brewster angle (case where $T_{11} = 0$) is given by θ_1, where

$$\sin\theta_1 = \sqrt{\frac{\varepsilon_r^2 - \varepsilon_r}{\varepsilon_r^2 - 1}}$$

REVIEW QUESTIONS

3.1. Explain the mobile radio propagation in terms of large scale path loss and small scale fading.

3.2. Explain the types of communication channels with wired and wireless channel.

3.3. Explain in brief the free space propagation model.

3.4. Define the following terms - EIPP, effective aperture.

3.5. Give a comparison between reflection, diffraction and scattering.

3.6. Define the following in terms of reflection:

(*i*) Reflection from dielectric

(*ii*) Brewster angle

(*iii*) Reflection from perfect conductor.

3.7. Explain the Freshel zone geometry in terms of diffraction.

3.8. Explain the scattering process with radar cross section model.

3.9. What do you mean by practical link budget design using path loss model.

3.10. Explain the following in terms of outdoor propagation model:

(*i*) Longley-Rice Model

(*ii*) Durkin's Model

(*iii*) Okumura Model

(*iv*) Hata Model

3.11. Explain the partition losses at floors and between floors.

3.12. Explain the long distance path less model.

3.13. What do you mean by small scale multipath propagation.

3.14. Explain the factor which effect the small scale fading.

3.15. Explain the parameter of mobile multipath channel.

3.16. Explain the following fading mechanism:

(*i*) Flat fading

(*ii*) Frequency selective fading

(*iii*) Slow fading

(*iv*) Fast fading

3.17. Explain the small scale multipath measurement.

3.18. Explain in brief two ray ray-leigh fading model.

3.19. What do you mean by multipath shape factor.

3.20. Define the three multipath shape factors and give the fading rates variance relationships in terms of the shapes factors.

3.21. Which pdf are applicable for describing the mobile channel? Explain in brief.

3.22. What is Doppler spread? Explain.

3.23. Give the equation for signal to undergo flat fading?

3.24. What do you understand by Rayleigh fading distribution?

3.25. What is meant by choreance time?

4
Chapter

SPREAD SPECTRUM MODULATION TECHNIQUES

4.1. *Introduction* **4.2.** *Advantage of Spread Spectrum Techniques* **4.3.** *Classification of Spread Spectrum System* **4.4.** *Spread Spectrum Modulation Techniques* **4.5.** *Pseudo-Noise (PN) Sequences* **4.6.** *Direct Sequence Spread Spectrum (DS-SS)* **4.7.** *Frequency Hopped Spread Spectrum (FH-SS)* **4.8.** *Performance of Direct Sequence Spread Spectrum* **4.9.** *Performance of Frequency Hopping Spread Spectrum* **4.10.** *Modulation Performance in Fading and Multipath Channels.*

4.1 INTRODUCTION

As the name "spread spectrum" suggests, this technique supports spreading the bandwidth to transmit data with more security with in a system.

Spread spectrum technique was developed for military applications to avoid jamming and interception becasue a signal spread over a wider bandwidth made jamming. More difficult to transmit via channel.

Bandwidth expansion is achieved by modulating again the modulated signal with a spreading code. This code is generated by pseudo-noise or pseudo-random number generator system also provide the security.

4.2 ADVANTAGE OF SPREAD SPECTRUM TECHNIQUES

Advantage of spread spectrum technique over narrow band systems include the following:

1. Reduced interference with the other systems.
2. Highly secure communication between two terminals.
3. Low density power spectrum for signal hiding.
4. Antijamming capability of the network.
5. Increased capacity of system.
6. Low cost.
7. Immunity to noise and multipath distortion or error.
8. Several users can transmit the data simultaneously or at a time.
9. Increased spectral efficiency which is used to represent the performance of a system.

4.3 CLASSIFICATION OF SPREAD SPECTRUM SYSTEM

Spread spectrum system has been classified by their architecture and modulation technique into the following categories as shown in Fig. 4.1 in four parts:

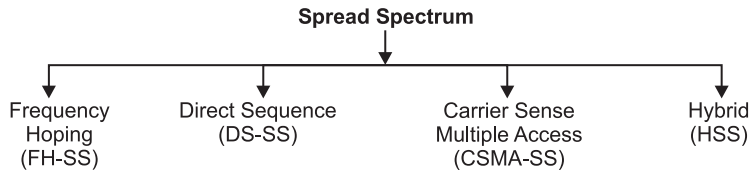


Fig. 4.1

4.4 SPREAD SPECTRUM MODULATION TECHNIQUES

Spread spectrum techniques employ a transmission bandwidth that is several orders of magnitude greater then the minimum required signal bandwidth.

Spread spectrum techniques is very bandwidth inefficient for a single user or may be multi-users.

The advantage of spread spectrum is that many users can simultaneously use the same bandwidth without significantly interfering with one another.

Independently from using a very large bandwidth, spread spectrum signals are pseudo-random and have noise like properties when compared with the digital information data or analog information data.

The spreading waveform is controlled by a pseudo-noise (PNS) sequence or pseudo-noise code, which is a binary sequence that appears random but can be regenerate in a deterministic manner by intended receivers.

Spread spectrum signals are demodulated at the receiver through cross-correlation with a locally generated version of the pseudorandom carrier at the receiver to get the original information.

Since each user is assigned a unique PN code which is approximately orthogonal to the codes of other users, the receiver can separate each user based in their codes, even though they use the same spectrum at all times, this implies that, up to a certain number of users, interference between spread spectrum signals using the same frequency is negligible or nill.

Spread spectrum system are not only resistant to multipath fading, but they can also exploit the delayed multipath component to improve the performance of the system. This can be achieve by using RAKE receiver which anticipates multipath propagation delays of the transmitted spread spectrum signal and combines the information obtain from several resolvable multipath components to form a stronger version of signal for transmit the information from one source to another.

4.5 PSEUDO-NOISE (PN) SEQUENCES

Pseudo-Noise Sequence can simply be defined as a series of bits (1'S and 0's) that seems to be random but repeat itself after n numbers (orbits) where n represent length of the pseudo-noise sequence (or PN code).

The major work of pseudo-noise sequence in wireless communication is –

(*i*) Spread the bandwidth of user signal in a channel.

(*ii*) Distinguish between different users signals sharing the same bandwidth by changing the PN sequence in wireless communication.

A pseudo-noise sequence, it is deterministic, has many characteristic that are similar to those of random binary sequence, such as having a nearly equal number of 0's and 1's, very low correlation between changed versions of the sequence, very low cross correlation between any two sequence etc.

The PN sequence is usually generated using sequential logic circuits.

A feedback shift register, which is shown in Fig. 4.2, consist of consecutive stages of two state memory devices and feedback logic with clock which is used for defining the state.

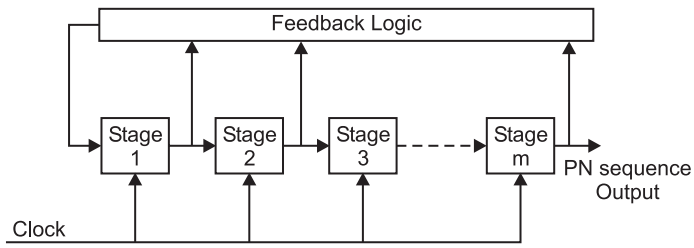


Fig. 4.2: Block diagram of a generalized feedback shift register with *m* stages.

Binary sequence are shifted through the shift register in response to clock pulse, and the out of the various stages are logically combined and feedback as the input to the first stage and last stage given the PN sequence output.

When the feedback logic consists of exclusive OR-gates, which is usually the case, the shift register is called a linear PN sequence generator.

The initial contents of the memory stages and the feedback logic circuit define the successive contents of the memory of the system.

If a linear shift register reaches zero state at some time, it would always remain in the zero state and the output would subsequently be all 0's until the state is changed.

Since there are exactly $2^m - 1$ non zero states for an *m*-stage feedback shift register, the period of a PN sequence produced by a linear *m*-stage shift register never exceed $2^m - 1$ symbols.

A sequence of period $2^m - 1$ generated by a linear feedback register is called a maximal length (ML) sequence.

The application of PN sequence in transmittor and receiver side is given in following paragraph with working function.

At transmitter side: The correlator yields an encrypted digital representation of the original signal using PN sequence.

This encrypted signal is than spread over a very wide frequency spectrum and then transmit to the receiver.

At receiver side: The signal is demodulated back to a narrow bandwidth and the then feel to a decorrelator.

This decorrelator uses its unique PN code (chipping code) to extract only the information intened for it from the received signal at the receiver.

A signal correlated with a given chipping sequence and decorrelated with the same chipping sequence returns the original signal.

Decorrelating the signal with the wrong PN sequence would result in the pure noise or distortion and which represent the wrong transmission and receiving.

4.6 DIRECT SEQUENCE SPREAD SPECTRUM (DS–SS)

A direct sequence spread spectrum (DS–SS) system spreads the baseband data by directly multiplying the baseband data pulses with a pseudo-noise sequence that is produced by a pseudo-noise code generators.

A single pulse or symbol or the PN wave form is called a chip which play an important role for transmit or receiving the information.

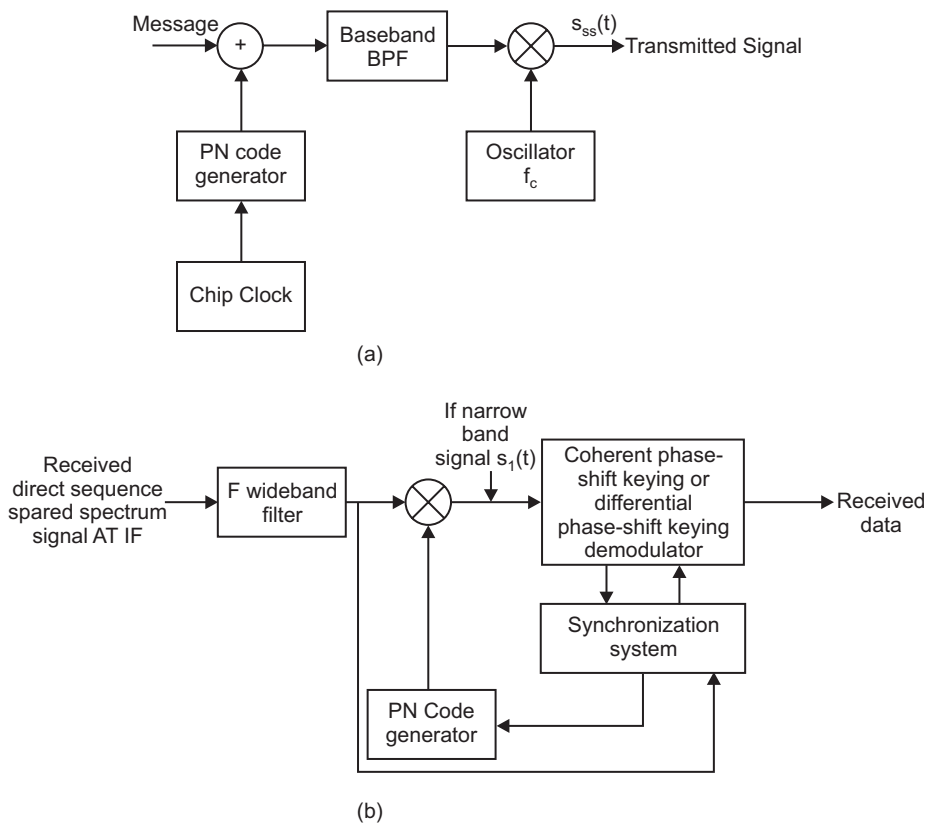


Fig. 4.3: Block diagram of a DS-SS system with binary phase modulation: (*a*) transmitter, and (*b*) receiver.

Fig. 4.3 shows a functional block diagram of a DS system with binary phase modulation. This system is one of the most widely used direct sequence spread

spectrum system. Synchronized data symbols which may be information bits or binary channel code symbols, are added in modulo-2 fashion to the chips before being phase modulated during the starting of transmission.

A coherent or differentially coherent phase-shift keying (PSK) demodulation may be used in the receiver.

The received spread spectrum signal for a single user can be represented as the following:

$$S_{ss}(t) = \sqrt{\frac{2E_s}{T_s}}\, m(t)p(t)\cos(2\pi f_c\, t + \theta) \qquad ...(4.1)$$

where $m(t)$ is the data sequence, $p(t)$ is the PN spreading sequence, f_c is the carrier frequency, and θ is the carrier phase angle at $t = 0$ and S_{ss} is spread signal.

The data waveform is a time sequence of non-overlapping rectangular pulses, each of which has an amplitude equal to +1 or –1. Each symbol in $m(t)$ represents a data symbol and has duration T_s (mean time duration T_s).

Each pulse *in* $p(t)$ represents a chip, is usually rectangular with an amplitude equal to +1 or –1, and has a duration of T_c.

The transition of the data symbols and chips coincide hence the ratio T_s to T_c in an integer of duration.

If B_{ss} is the bandwidth of $S_{ss}(t)$ and B is the bandwidth of a conventionally modulated signal $m(t)\ cos(2\pi\ f_c t)$ the spreading due to $p(t)$ gives $B_{ss} >> B$ means B_{ss} is greater than B.

Fig. 4.3 illustrates a DS receiver. Assuming that code synchronization has been achieved at the receiver, the received signal passes through the wideband filter and is multiplied by a local replica of the PN code sequence $p(t)$. If $p(t) = \pm 1$, then $P_2(t) = 1$, and this multiplication yields the despread signal $s(t)$ given by following equation.

$$s_1(t) = \sqrt{\frac{2E_s}{T_s}}\, m(t)\cos(2\pi f_c t + \theta) \qquad ...(4.2)$$

at the input of the demodulator. Because $S_1(t)$ has the form of a BPSK signal, the corresponding demodulation draw out by $m(t)$.

Fig. 4.4 shows the received spectra of the desired spread spectrum signal and the interference at the output of the receiver wideband filter. Multiplication by the spreading waveform produces the spectra of Fig. 4.3 at the demodulator input and Fig. 4.4 also represent the correlator output after despreading.

The signal bandwidth is reduced to B, while the interference energy is spread over an RF bandwidth exceeding B_{ss}. B_{ss} means bandwidth of spread signal $S_{ss}(t)$

The filtering action of the demodulator removes most of the interference spectrum that does not overlap with the signal spectrum. Thus, most of the original interference energy is eliminated by spreading and minimally affects the desired receiver signal at the receiver.

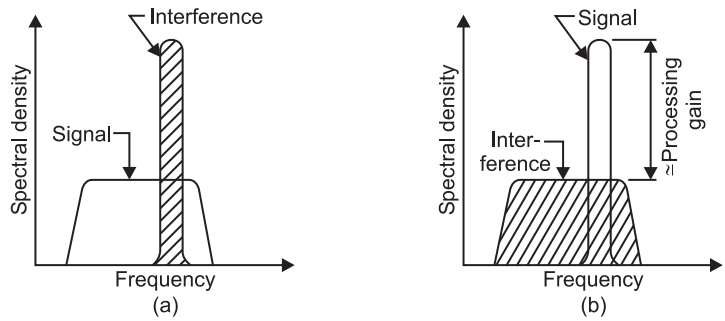


Fig. 4.4. Spectra of desired received signal with interference: (*a*) wideband filter output and (*b*) correlator output after despreading.

An approximate measure of the interference rejection capability is given by the ratio $\frac{B_{ss}}{B}$, which is equal to the processing gain defined as following equation

$$PG = \frac{T_s}{T_c} = \frac{R_c}{R_s} = \frac{B_{ss}}{2R_s} \quad ...(4.3)$$

The greater the processing gain of the system, the greater will be its ability to suppress inband interference.

4.7 FREQUENCY HOPPED SPREAD SPECTRUM (FH–SS)

Frequency hopping involves a periodic change of transmission frequency during the transmission.

A frequency hopping signal may be regarded as a sequence of modulated data bursts with time-varying pseudorandom carrier frequencies at the transmitter.

The set of possible carrier frequency is called the hopset.

Hopping occurs over a frequency band that includes a number of channels. Each channels is expalin as a spectral region with a central frequency in the hopset and a bandwidth large enough to include most of the power in a narowband modulation burst (usually FSK) having the corresponding carrier frequency for transmission purpose.

The bandwidth of a channel used in the hopset is called the instantaneous bandwidth. The bandwidth of the spectrum over which the hopping occurs in called the total hopping bandwidth or hopped bandwidth.

Data is sent by hopping the transmitter carrier to seemingly random channels which are known only to the desired receiver. On each channel, small bursts of data are sent using conventional narrowband modulation before the transmitter hops again to the receiver.

If only a single carrier frequency (single channel) is used on each hop, digital data modulation is called single channel modulation.

Fig. 4.5 shows a single channel FH–SS system with single channel modulation.

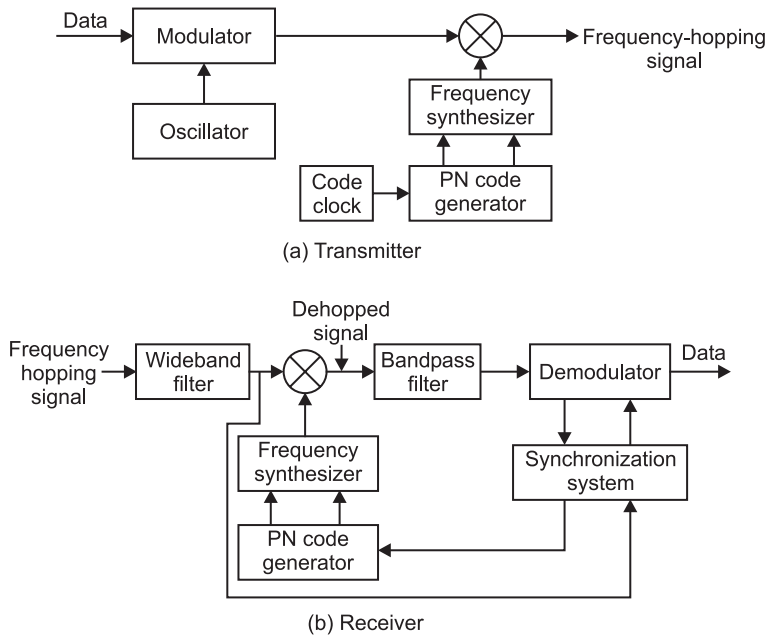


Fig. 4.5: Block diagram of frequency hopping (FH) system with single channel modulation.

The time duration between hops is called the hop duration or the hopping period and is denoted by T_h.

The total hopping bandwidth and the instantaneous bandwidth are denoted by B_{ss} and B respectively of a FH system.

The processing gain $\frac{B_{ss}}{B}$ for FH systems.

After frequency hopping has been removed from the received signal, the resulting signal is said to be the dehopped. If the frequency pattern produced by the receiver synthesizer in Fig. 4.5 is synchronized with the frequency pattern of the received signal, then the mixer output is a dehopped signal at a fixed difference frequency. Before demodulation, the dehopped signal is applied to a conventional receiver for getting the transmitted information.

In FH, whenever an undesired signal occupies a particular hopping channel, the noise and interference in that channel are translated in frequency so that they enter into, the demodulator. So that it is possible to have collisions in an FH system where an undesired user for transmission.

Frequency hopping may be classified as fast or slow hopping.

(*i*) Fast frequency hopping occurs if there is more than one frequency hop during each transmitted symbol. Thus fast frequency hopping implies that the hopping rate equals or exceeds the information symbol rate means $H_r \geq I_r$.

(*ii*) Slow frequency hopping ocurs if one or more symbols are transmitted in the time interval between frequency hops.

The frequency channel occupied by a transmitted symbol is called the transmission channel. The channel that would be occupied if the alternative symbol were transmitted is called the complementary channel or alternative channel.

The frequency hop rate of an FH–SS system is determined by the frequency agility of receiver synthesizers, the type of information being transmitted, the amount of redundancy used to code against collisions, and the distance to the nearest potential interferer.

4.8 PERFORMANCE OF DIRECT SEQUENCE SPREAD SPECTRUM

A direct sequence spread spectrum system with k multiple access users is shown in Fig. 4.6. In which 4.6 (*a*) shows model of k-users in a CDMA spread spectrum system and 4.6 (*b*) show receiver structure for user 1.

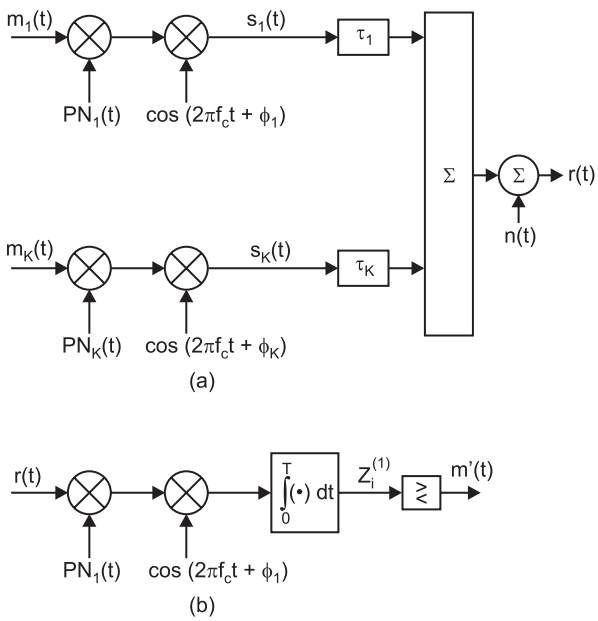


Fig. 4.6: A simplified diagram of a DS-SS system with k-users, (*a*) Model of K users in a CDMA spread spectrum system (*b*) receiver structure for User 1.

Assume each user has a PN sequence with N chips per message symbol period T such that $NT_C = T$.

The transmitted signal of the kth user can be expressed by following equation:

$$S_k(t) = \sqrt{\frac{2E_s}{T_s}} m_k(t) p_k(t) \cos(2\pi f_c t + \phi_k) \qquad ...(4.4)$$

where $p_k(t)$ is the PN code sequence of the kth user, and $m_k(t)$ is the data sequence of the kth users in the system.

The received signal will consist of the sum of k different transmitted signals (one desired user and k–1 undesired users) plus additive noise or distortion.

Reception is accomplished by correlating the received signal with the approapriate signature sequence to produce a decision variable. The decision variable for the ith transmitted bit of user 1 is represent by following equation:

$$Z_i^{(1)} = \int_{(i-1)T+\tau_1}^{iT+\tau_1} r(t)p_1(t-\tau_1)\cos[2\pi f_c(t-\tau_1)+\phi_1]\,dt \quad ...(4.5)$$

If $m_{1,i} = -1$, then the bit will be received in error if $Z_i^{(1)} > 0$. The probability of error can now be calculated as $P_r\,[Z_i^{(1)} > 0 \mid m_{1i} = -1]$. Since the received signal $r(t)$ is a linear combination of sign also plus additive noise or distortion.

Equation (4.5) can be rewritten as the following:

$$Z_i^{(1)} = I_1 + \sum_{k=2}^{k} I_k + \xi \quad ...(4.6)$$

where I_1 can be written as

$$I_1 = \int_0^T s_1(t)\,p_1(t)\cos(2\pi f_c t)\,dt = \sqrt{\frac{E_s T}{2}} \quad ...(4.7)$$

is the response of the receiver to the desired signal from user 1.

$$\xi = \int_0^T n(t)\,p_1(t)\cos(2\pi f_c t)\,dt \quad ...(4.8)$$

is a Gaussian random variable representing noise with mean zero and variance

$$E\,|\,\xi\,|^2 = \frac{N_0 T}{4}$$

$$I_k = \int_0^T S_k(t-\tau_k)p_1(t)\cos(2\pi f_c t)\,dt \quad ...(4.10)$$

represents the multiple access interference from user k with in a system.

Assuming that I_k is cumposed of the cumulative effects of N random chips from the kth interferer over the integration period T of one bit.

Since there are $k - 1$ users which serve as identically distributed interferers, the total multiple access interference $I = \sum_{k=2}^{k} I_k$ may be approximated by a Gaussian random variable or represented by a Gaussian random variable.

Gaussian approximation assumes that each I_k is independent, but they are not. The Gaussian approximation yields a convenient expression for the average probability of bit error given by the following equation:

$$P_e = Q\left[\frac{1}{\sqrt{\frac{k-1}{3N} + \frac{N_0}{2E_b}}}\right] \quad ...(4.11)$$

For a single user, $k = 1$, this expression reduces to the BER expression for BPSK modulation. For the interference limited case where thermal noise is not a factor $\frac{E_b}{N_0}$ tends to infinity, and the BER expression has a value equal to

$Q\left(\sqrt{\frac{(3N)}{(k-1)}}\right)$ which is represent by the following equation:

$$P_e = Q\left(\sqrt{\frac{(3N)}{(k-1)}}\right)$$

This is the irreducible error floor due to multiple access interference and assumes that all interferer provide equal power, the same as the desired user, at the DS-SS receiver system.

The near-far problem presents difficulty for DS-SS systems.

Without careful power control of each mobile user, one close-in user may dominate the received signal energy at a base station, making the Gaussian assumption inaccurate or Gaussian assumption in error.

For a large number of users, the bit error rate is limited more by the multiple access interference than by thermal noise or any other noise.

4.9 PERFORMANCE OF FREQUENCY HOPPING SPREAD SPECTRUM

In FH-SS systems, several users independently hop their carrier frequencies while using BFSK modulation technique.

If two users are not simultaneously utilizing the same frequency band, the probability of error for BFSK can be given by the following equation:

$$P_e = \frac{1}{2}\exp\left(-\frac{E_b}{2N_0}\right)$$

However, if two users transmit simultaneously in the same frequency band, a collision, or "hit", occurs. In this case, it is reasonable to consider that the probability of error is 0.5. Thus, the overall probability of bit error can be modeled as the following equation:

$$P_e = \frac{1}{2}\exp\left(-\frac{E_b}{2N_0}\right)(1-p_h)+\frac{1}{2}p_h \qquad ...(4.14)$$

where p_h is the probability of a hit, which must be determined for define the performance of a system.

If there are M possible hopping channels (called slots), there is a $\frac{1}{M}$ probability that a given interferer will be present in the desired user's slot and slot may be frequency or time.

If there are $k-1$ interfering users, the probability that at least one is present in the desired frequency slot is equal to one mm *vs* the probability of no hits, given as the following equation number (4.15).

$$P_h = 1-\left(1-\frac{1}{M}\right)^{k-1} \approx \frac{k-1}{M} \qquad ...(4.15)$$

Assuming M is large substituting this in equation (4.14), then the equation (4.14) will be

$$P_e = \frac{1}{2}\exp\left(-\frac{E_b}{2N_0}\right)\left(1-\frac{k-1}{M}\right)+\frac{1}{2}\left[\frac{k-1}{M}\right] \quad ...(4.16)$$

Now consider the following special cases. If $k = 1$, the probability of error reduces to equation (13), the standard probability of error for BFSK. Also if $\frac{E_b}{N_0}$ approaches E_h infinity or $\frac{E_b}{N_0}$ tends to infinity.

$$\lim_{\frac{E_b}{N_0}\to\infty} (P_e) = \frac{1}{2}\left[\frac{k-1}{M}\right] \quad ...(4.17)$$

which illustrates the irreducible error rate due to multiple access interference in a communication system.

The previous analysis assumes that all users hop their carrier frequencies synchronously. This is called slotted frequency hopping. This may not be a realistic scenario for many FH-SS systems or for a particular FH-SS systems.

Even when synchronization can be achieved between individual user clocks, radio signals will not arrive synchronously to each user due to the various propogation delays. As described by Geraniotis, the probability of a hit in the asynchronous case is

$$p_h = 1-\left\{1-\frac{1}{M}\left(1+\frac{1}{N_b}\right)\right\}^{k-1} \quad ...(4.18)$$

where N_b represent number of bits per hop.

Comparing equation (4.18) to (4.15), we see that for the asynchronous case, the probability of a hit is increased (this would be expected) or extend.

Using equation (4.18) in equation (4.14), the probability of error for the asynchronous FH-SS case is represents as:

$$p_e = \frac{1}{2}\exp\left(1-\frac{E_b}{N_0}\right)\left\{1-\frac{1}{M}\left(1+\frac{1}{N_b}\right)\right\}^{k-1}$$
$$+\frac{1}{2}\left[1-\left\{1-\frac{1}{M}\left(1+\frac{1}{N_b}\right)\right\}^{k-1}\right] ...(4.19)$$

FH-SS an advantage over DS-SS in that it is not as susceptible to the near-far problem. Because signals are generally not utilizing the same frequency simultaneously, the relative power levels of signals are not as critical as in DS-SS. The near-far problem is not totally avoided, however since there will be some interference caused by the stronger signals bleeding into weaker signals due to imperfect filtering of adjacent channels in the system.

To combat the occasional hits error-correction coding is required on all transmissions. By applying strong Reed-Solomon or other burst error correcting codes, performance can be increased dramatically, even with an occasional collision.

4.10 MODULATION PERFORMANCE IN FADING AND MULTIPATH CHANNELS

The mobile radio channel is characterized by various implements such as fading multipath and doppler spread technique.

In order to study the effectiveness of any modulation scheme in a mobile radio environment, it is required to evaluate the performance of modulation scheme over such channel conditions, or channel environment.

Although bit error rate (BER) evaluation gives a good indication of the performance of a particular modulation scheme, it does not provide information about type of error or distortion.

In a fading mobile radio channel, it is likely that a transmitter signal will suffer deep fades which can lead to outage or a complete loss of signal or information.

Evaluating the probability of outage is another means to judge the effectiveness of the signaling scheme in a mobile radio channel.

Bit error rate and probability of outage for various modulation scheme under various type of channel impairment can be evaluated through analytical techniques or through simulations. While simple analytical technique for computing bit error rate in slow flat fading channels exist, performance evaluation in frequency selective channels and computations of outage probabilities are often made through computer simulations.

Computer simulations are based on convalving the input bit stream with a suitable channel impulse response model and counting the bit error at the output of the receiver decision circuit.

Before a study of the performance of various modulation scheme in multipath fading channels is made, it is imperative that a thorough understanding of the channel characteristics be obtained or channel properties be obtained.

SOLVED EXAMPLES

Example 4.1. *A zero mean sinusoidal message is applied to a transmitter that radiates an AM signal with 10 kW power. Compute the carrier power if the modulation index is 0.6. What percentage of the total power is in the carrier? Calculate the power in each sideband.*

Solution. We have

$$P_c = \frac{P_{AM}}{1+\frac{k^2}{2}} = \frac{10}{1+\frac{0.6^2}{2}} = 8.47 \text{ kW}$$

Percentage power in the carrier is

$$\frac{P_c}{P_{AM}} \times 100 = \frac{8.47}{10} \times 100 = 84.7\%$$

Power in each sideband is given by

$$\frac{1}{2}(P_{AM} - P_c) = 0.5 \times (10 - 8.47) = 0.765 \text{ kW}$$

Example 4.2. *A sinusoidal modulating signal, $m(t)$ = 4 cos 2n4 × 10$^{3'}$–t, is applied to an FM modulator that has a frequency deviation constant gain of* 10 *kHz/V. compute*

(*a*) The peak frequency deviation.

(b) The modulation index.

Solution. Given:

Frequency deviation constant k_f = 10 kHz/V

Modulating frequency, f_m = 4 kHz

(*a*) The maximum frequency de 0.5 × (10 – 8.47) = 0.765 kW viation will occur when the instantaneous value of the input signal is at its maximum. For the given $m(t)$, the maximum value is 4V, and hence the peak deviation is equal to maximum value is 4V, and hence the peak deviation equal to

$$\Delta f = 4\text{V} \times 10 \text{ kHz/V} = 40 \text{ kHz}$$

(*b*) The modulation index is given by

$$\beta_f = \frac{k_f A_m}{f_m} = \frac{\Delta_f}{f_m} = \frac{40}{4} = 10$$

Example 4.3. *An 880 MHz carrier signal is frequency modulated using a 100 kHz sinusoidal modulating waveform. The peak deviation of the FM signal is 500 kHz. If this FM signal is received by a superheterodyne receiver having an IF frequency of 5 MHz. Determine the IF bandwidth necessary to pass the signal.*

Solution. Given:

Modulating frequency, f_m = 100 kHz

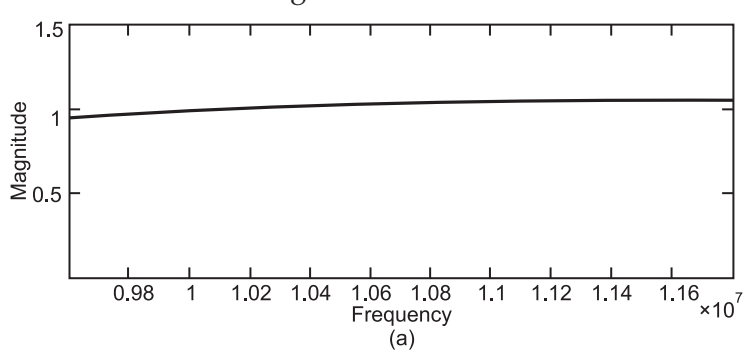
Frequency deviation, Δ_f = 500 kHz

Therefore modulation index $\beta_f = \dfrac{\Delta f}{f_m} = \dfrac{500}{100} = 5$

Using Carson's rule, the bandwidth occupied by the FM signal is given by

$$B_T = 2(\beta_f + 1)\, f_m = 2(5 + 1)\ 100 \text{ kHz}$$
$$= 1200 \text{ kHz}$$

The IF filter at the receiver needs to pass all the components in this bandwidth hence the IF filter should be designed for a bandwidth of 1200 kHz.



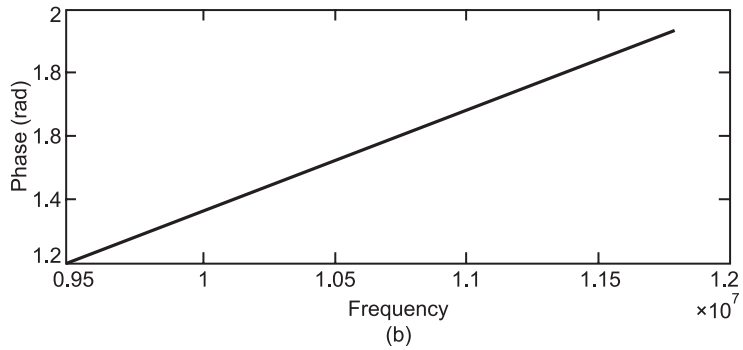


Fig. 4.7: (*a*) Magnitude response of the designed phase-shift network; (*b*) Phase response of the designed phase-shift network.

Example 4.4. *Design an RLC network that implements an IF quadrature FM detector with f_c = 10.7 MHz, and a 500 kHz symmetrical bandpass spectrum. Also plot the transfer function of the designed network to verify that it will work.*

Solution. A quadrature detector is represented by the block diagram in Fig. 4.8 and the phase-shift network is implemented by the RLC circuit shown in Fig. 4.8. Here the phase shift is 90° instead of –90° for $f = f_c$ with reference to figure.

$$\frac{V_q(\omega)}{V_f(\omega)} = \frac{V_q(\omega)}{V_f(\omega)}$$

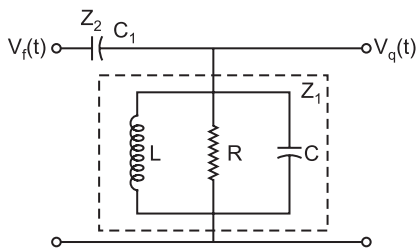


Fig. 4.8: Circuit diagram of an RLC phase-shift network.

Multiplying and dividing by $\frac{1}{(Z_1 Z_2)}$, we get

$$\frac{V_q(\omega)}{V_f(\omega)} = \frac{Y_2}{Y_1 + Y_2} = \frac{j\omega C_1}{j\omega C + \frac{1}{R} + \frac{1}{j\omega L} + j\omega C_1}$$

$$= \frac{j\omega R C_1}{1 + jR\left(\omega(C + C_1) - \frac{1}{\omega L}\right)}$$

Let $\omega_c^2 = \frac{1}{(L(C_1 + C))}$ for the overall circuit, then

$$Q = \frac{R}{\omega_C L} = R\omega_C(C_1 + C)$$

$$\frac{V_q}{V_t} = \frac{j\omega RC_1}{1 + jQ\left(\frac{\omega}{\omega_c} - \frac{\omega_c}{\omega}\right)}$$

So, for $\omega = \omega_c$

$$\frac{V_q}{V_f} = j\omega_c RC_1$$

This provides the desired 90° phase shift at ω_c. At IF frequencies, the phase shift introduced by the network may be expressed as

$$\phi(\omega_i) = \frac{\pi}{2} + \tan^{-1}\left[Q\left(\frac{\omega_i}{\omega_c} - \frac{\omega_c}{\omega_i}\right)\right] = 90° + \eta$$

For a good system we need $-5° < \phi\ (\omega_1)\ (\omega_i) < 5°$ (approximately) therefore, for $f_c = 10.7$ MHz and $B = 500$ kHz, at the largest IF frequency $f_i = f_c + 250$ kHz. Thus we require.

$$Q\left(\frac{10.7\times10^6 + 250\times10^3}{10.7\times10^6} - \frac{10.7\times10^6}{10.7\times10^6 + 250\times10^3}\right) = \tan 5°$$

Therefore, $Q = 1.894$

Using $Q = 1.894$, one may verify the phase shift at the smallest IF frequency $f_i = f_c - 250$ kHz.

$$\tan^{-1}\left[1.894\left(\frac{10.45}{10.7} - \frac{10.7}{10.45}\right)\right] = -5.12° = -5°$$

We have verified that a circuit with $Q = 1.894$ will satisfy the phase shift requirements.

Now, to compute the values of L, R, C and C_1

Choose $L = 10$ μH. Using the first part of equation, the value of R can be computed as 1.273 kΩ.

Using the second part of equation

$$C_1 + C = \frac{Q}{R_{\omega_c}} = \frac{1.894}{(1.273\times10^3)\,2\pi\,(10.7\times10^6)} = 22.13 \text{ pF}$$

Assuming $C_1 = 12.13$ pF = 12 pF, we get $C = 10$ pF.

The magnitude transfer function of the designed phase shift network is given by

$$|H(f)| = \frac{2\pi fRC}{\sqrt{1 + Q^2\left(\frac{f}{f_c} - \frac{f_c}{f}\right)}} = \frac{97.02\times10^{-9} f}{\sqrt{1 + 3.587\left(\frac{f}{10.7 + 10^6} - \frac{10.7\times10^6}{f}\right)^2}}$$

and the phase transfer function is given by

$$<H(f) = \frac{\pi}{2} + \tan^{-1}\left[Q\left(\frac{f}{f_c} - \frac{f_c}{f}\right)\right]$$

$$= \frac{\pi}{2} + \tan^{-1}\left[1.894\left(\frac{f}{10.7\times10^6} - \frac{10.7\times10^6}{f}\right)\right]$$

It is clearly seen from the above equation that the transfer function satisfies the requirements of the phase-shift network, thus allowing detection of FM.

Example 4.5. *How much bandwidth is required for an analog frequency modulated signal that has an audio bandwidth of 5 kHz and a modulation index of three? How much output SNR improvement would be obtained if the modulation index is increased to five? What is the trade-off bandwidth for this improvement*?

Solution. From Carson's rule, the bandwidth is

$$B_T = 2(B_f + 1)\, f_m = 2\ (3 + 1)\ 5 \text{ kHz} = 40 \text{ kHz}$$

From equation, the output SNR improvement factor is approximately

$$3B_f^3 + 3B_f^3$$

Therefore, for $B_f = 3$, the output SNR factor is $= 3(3)^3 + 3(3)^3 = 108$

$$= 20.33 \text{ dB}$$

for $B_f = 5$, the output SNR factor is $= 3(5)^3 + 3(5)^2 = 450$

$$= 26.53 \text{ dB}$$

The improvement in output SNR by increasing the modulation index from three to five is therefore 26.53 – 20.33 = 6.2 dB.

This improvement is achieved at the expense of bandwidth, for $B_f = 3$, a bandwidth of 40 kHz is needed, while $B_f = 5$ requires a bandwidth of 60 kHz.

Example 4.6. *If the SNR of a wireless communication link is 20 dB and the RF bandwidth is 30 kHz determine the maximum theoretical data rate that can be transmitted. Compare this rate to the US digital cellular standard described in chapter 1.*

Solution. Given:

S/N = 20dB =100

RF Bandwidth B = 30000 Hz

Using Shannon's channel capacity formula, the maximum possible data rate

$$C = B\log_2\left(1 + \frac{S}{N}\right)$$

$$= 30000 \log_2 (1 + 100)$$

$$= 199.75 \text{ Kbps}$$

The VSDC data rate is 48.6 Kbps, which is only about one fourth the theoretical limit under 20 dB SNR conditions.

Example 4.7. *What is the theoretical maximum data rate that can be supported in a 200 kHz channel for SNR = 10 dB, 30 dB. How does this compare to the GSM standard described in chapter 1?*

Solution. For SNR = 10 dB = 10

$$B = 200 \text{ kHz}.$$

Using Shannon's channel capacity formula, the maximum possible data rate

$$= 200000 \log_2 (1 + 10) = 691.886 \text{ Kbps}$$

The GSM data rate is 270.833 Kbps, which is only about 40% of the theoretical limit for 10 dB SNR conditions

For $\text{SNR} = 30\text{dB} = 1000$

$B = 200$ kHz.

The maximum possible data rate

$$C = B \log_2 \left(1+\frac{S}{N}\right) = 200000 \log_2 (1 + 1000)$$

$$= 1.99 \text{ Mbps}$$

Example 4.8. *Find the first zero-crossing RF bandwidth of a rectangular pulse which has T_s = 41.06 us. Compare this to the bandwidth of a raised cosine filter pulse with T_s = 41.06 us and α = 0.35.*

Solution. The first zero-crossing (null-to-null) bandwidth of a rectangular pulse is equal to

$$\frac{2}{T_s} = \frac{2}{(41.06\mu s)} = 48.71 \text{ kHz}$$

and that of a raised cosine filter with α = 0.35 is

$$\frac{1}{T_s}(1+\alpha) = \frac{1}{41.06\ \mu s}(1+0.35) = 32.88 \text{ kHz}$$

Example 4.9. *Using the $\frac{\pi}{4}$ QPSK signal, demonstrate how the received signal is detected properly using a baseband differential detector.*

Solution. Assume the transmitter and receiver are perfectly phase locked, and the receiver has a front-end gain of two. The phase difference between the three transmitter phases yield $(x_1, y_1) = (-0.707, -0.707)$,

$$(x_2, y_2) = (0.707, -0.707)$$

$$(x_3, y_3) = (0.707, 0.707)$$

Applying the decision rules of equation and the detected bit stream is

$$(S_1, S_2, S_3, S_4, S_5, S_6) = (0, 0, 1, 0, 1, 1).$$

Example 4.10. *Find the 3-dB bandwidth for a Gaussian law pass filter used to produce 0.25 GMSK with a channel data rate of R_b –270 Kbps. What is the 90% power bandwidth in the RF channel? Specify the Gaussian filter parameter α.*

Solution: From the problem statement

$$T = \frac{1}{R_b} = \frac{1}{270\times10^3} = 3.7\ \mu s$$

Solving for B, where BT= 0.25

$$B = \frac{0.25}{T} = \frac{0.25}{3.7\times10^{-6}} = 67.567 \text{ kHz}$$

Thus the 3-dB bandwidth is 67.567 kHz. To determine the 90% power bandwidth, use table to find that 0.57 R_b is the desired value. Thus, the occupied RF spectrum for a 90% power bandwidth is given by

$$RF\ BW = 0.57\ R_b$$

$$= 0.57 \times 270 \times 10^3 = 153.9 \text{ kHz}.$$

HIGHLIGHTS

1. Spread spectrum is a technique which convert narrow band signal into wide band signal.
2. Main advantage of spread spectrum system is anti jamming capability and secure communication with reduced interference.
3. Pseudo-noise code (PN code) is a code word which is used to spread the bandwidth of a signal.
4. A signal correlated with a given PN sequence and decorrelated with the same PN sequence returns the original signal or recover the original signal.
5. Spread spectrum systems are divided into two main parts
   (*i*) Direct-Sequence Spread Spectrum (DS-SS)
   (*ii*) Frequency Hopping Spread spectrum (FH-SS).
6. DS-SS systems use the total available bandwidth and FH-SS systems uses only a portion of the total bandwidth available at any time.
7. Spread spectrum may be defined as a technique in which a transmitted signal used a bandwidth which is kept much larger than that which is required by a base band information signal at the transmitter end.
8. A combination of these techniques is also available for use. Pseudo-random (PN) codes are used to achieve spreading of the spectrum. Before this explain the methods of generating these codes in detail, let us first consider the techniques for widening the spectrum of a signal.
9. One of the most common method to extend the spectrum of a signal is simply to multiply *(i.e.,* modulate) it by wide-band signal. We will observe that the spreading signal must have properties which aid in acquiring and tracking the signal at any time.
10. The spread spectrum communication, spreads the transmitted (RF) signals over a bandwidth which is as larger as in practical. So that the RF bandwidth which is 10 to 100 times the information signal bandwidth. Frequency hopping and direct sequence modulation are two principle spread spectrum techniques which are used.
11. If the PN generator sequenced through all its states, then the number of states will be an even number and the sequence available at any Q or $\overline{Q}$.
12. There are three techniques used to explain an idea to the energy in bandwidth B and spread it over the wider bandwidth $B_{RF,}$ and these techniques are–
    (*i*) Direct Sequence Pseudo-noise (DSPN)
    (*ii*) Frequency Hopping (FH)
    (*iii*) Time Hopping (TH)
13. The second method of increasing the message signal is explain by transmitting each bit or subinteger (bit) chip on a different carrier which is selected from a wide range of frequency.
14. The third method of increasing the message signal spectrum is time hop spreading.
15. A receiver must have the following information for detect and demodulate the received signal, which are:
    (*i*) PN code rate (*ii*) Modulation type

(*iii*) Signal carrier frequency (*iv*) Propagation delay time

(*v*) Existence of the signal mixed-up in the noise signal.

16. The main advantage of spread spectrum communication system is to reject the interference whether it be the unintentional interference by another user same time attempting to transmit through the channel.

17. Spread spectrum modulation was originally developed for military applications, where resistance of interference is of major concern.

18. The spread spectrum modulation system can be used to provide multipath rejection in a ground based mobile radio environment.

19. The spread spectrum modulation system can be used in multiple access communication in which a number of independent users are required to share a common channel without an external synchronizing mechanism.

20. In DS-SS system, the spreading system is complex but in FH-SS system the spreading system is simple.

21. In a direct sequence spread spectrum (DS-SS), the rejection capabilities to cancel out interference and noise are good but in lose frequency hopped spread spectrum (FH-SS) the interference rejection capabilities are poor.

22. The spread spectrum are classified on the basis of their architecture and modulation concept into the following categories–

(*i*) Frequency Hopping Spread Spectrum (FH-SS)

(*ii*) Direct Sequence Spread Spectrum (DS-SS)

(*iii*) Carrier Sequence Multiple Access Spread Spectrum (CSMA-SS)

(*iv*) Hybrid Spread Spectrum (HSS)

23. The main work of pseuodo-noise sequence in wireless communication is–

(*i*) Spread the bandwidth of the user signal.

(*ii*) Differentiate between different user signals shairing the source bandwidth by changing the *P-N* sequence.

24. *AP-N* generator, produce a random series of binary ones and zeros which represent a *P-N* sequence. This *P-N* sequence can be repeats but appears to be in random.

PROBLEMS

4.1. Compare the channel spectral efficiencies of IS-54, GSM, PDC and IS-95. What are the theoretical spectral efficiencies for each of these standards if they operate at 20 dB SNR?

4.2. If 63 users share a CDMA system and each user has a processing gain of 511, determine the average probability of error for each user. What assumptions have you made in determine your result?

4.3. Assume that a synchronous frequency hopping multiple access system is used. Binary frequency shift keying having a bit error probability of–

$$P_e = \frac{1}{2} e^{-\frac{E_b}{2N_0}}$$

is used by each user. Prove that if there are M possible hopping channels and k total users, that the irreducible error rate due to multiple access interference is given by

4.4. (*a*) In a DS-SS multiple user system, how many simultaneous users may be supported such that an average bit error rate of less than 10-3 is maintained for each user? Assume all users employ power control such that the received powers of each user are maintained at an average $\frac{E_b}{N_0} = 10\,dB$, and assume each user has a PN code that is produced from an 11 bit shift register.

(*b*) Using your answer in (*a*), what is the resulting average probability of error for a user if one more simultaneous user is diode? Would this be noticeable to all the other users? Why or why not?

4.5. A FH-SS system uses 50 kHz channels over a continuous 20 MHz spectrum. Fast frequency hopping is used, where two hops occur for each bit. If binary FSK is the modulation method used in this system, determine.

(*a*) The number of hops per second if each user transmits at 25 Kbps.

(*b*) The probability of error for a single user operating at an $\frac{E_b}{N_0} = 20\,dB$.

(*c*) The probability of error for a user operating at $\frac{E_b}{N_0} = 20\,dB$ with 20 dB other EH-SS users which are independently frequency hopped.

(*d*) The probability of error for a user operating at $\frac{E_b}{N_0} = 20\,dB$ with 200 other FH-SS users which are independently frequency hopped.

4.6. A frequency hopping WLAN system is proposed for Bluetooth using binary FSK. Assume fifty 1-MHz channels are available and *k* devices are to be connected in the same room. If an $\frac{E_b}{N_0}$ of 20 dB can be supplied to each user, plot the average probability of errors as a function of the number of devices, where Granges from 1 to 100. What would you deem to be the maximum acceptable number of devices?

4.7. Prove that a Rayleigh fading signal has an exponential pdf for the power of the signal.

REVIEW QUESTIONS

4.1. What do you understant by 'spread spectrum technique'? What are the benefits in this technique?

4.2. What is direct sequence spread spectrum?

4.3. What is the basic principle of frequency hopping spread spectrum? How many types of frequency hopping are possible?

4.4. Draw the block diagram of FH-SS transmitter system and explain each block in detail.

4.5. Draw the block diagram of a basic spread spectrum digital communication system and explain its working detail.

4.6. What is the concept of spreading and despreading? Explain in brief.

4.7. What do you understand by pseudo-noise sequence?

4.8. Explain the difference between DS-SS systems and FH-SS systems.

4.9. Explain the performance of DS-SS systems.

4.10. Explain the performance of FH-SS systems.

4.11. What are the advantages of spread spectrum systems over narrow band spectrum systems?

4.12. Compare similarity and differences between basic concepts of DS-SS versus FH-SS systems. Analyze the performance and a synchronous frequency hopping spread spectrum system and show that the irreducible error rate due to multiple access interference is equal to $\frac{1}{2}\left[\frac{k-1}{M}\right]$ where k is number of MA uses and M is number of possible hopping slots.

5 Chapter

EQUALIZATION AND DIVERSITY TECHNIQUES

5.1. *Introduction* **5.2.** Fundamentals of Equalization **5.3.** *Equalizers in a Communications Receiver* **5.4.** *Survey of Equalization Techniques* **5.5.** *Linear Equalizers* **5.6.** *Non-Linear Equalization* **5.7.** *Diversity Technique* **5.8.** *Rake Receiver*

5.1 INTRODUCTION

Equalization and diversity are two technique which can be used independently or in tandom to improve received signal quality and link performance over small scale times and distance or large scale.

Equalization compensates the inter symbol interference (ISI) created by multipath within time dispersive channels.

An equalizer within a receiver compensates for the average range of expected channel amplitude and delay characteristics of the channel.

Equalizers must be adaptive since the channel is generally unknown and time varying.

Diversity is another technique used to compensate for fading channel impairment and is usually implemented by using two or more receiving antennas at the receiver.

The evolving 3G common air interface also use transmit diversity where base stations may transmit multiple duplicate of the signal on spatially separated antennas or frequencies either receiver side or transmitter side.

As with an equalizer, diversity improves the quality of a wireless communication link without altering the common air interface and without increasing the transmitted power or bandwidth. Nevertheless while equalization is used to counter the effect of time dispersion (ISI), diversity is usually employed to reduce the depth and duration of the fades experienced by a receiver in a local area which are due to motion of the things by which fades are produced.

Diversity techniques are often employed at both base station and mobile receiver means at the transmitter as well as at the receiver.

The most common diversity techniques are spatial diversity, polarization diversity, frequency, and time diversity.

The two general techniques of equalization and diversity are used to improve radio link performance (*i.e.* to minimize the bit error rate), but the approach, cost complexity and effectiveness of each technique varies widely in practical wireless communication systems or experimentry wireless communication systems.

5.2 FUNDAMENTALS OF EQUALIZATION

It was observed that intersymbol interference (ISI) is a major obstacle to high speed data transmission over wireless channels. Intersymbol interference is reputation or esteem by multipath reception within time dispersive channels. If the modulation bandwidth exceeds the coherence bandwidth of radio channel then ISI occurs and hence the modulation pulses are spread in time into adjacent (nearby) symbols. A variety of equalizer (adaptive) can be used to suppress the interference.

The term equalization can be used to describe any signal processing during the operation that minimizes ISI.

An equalizer is usually implemented at base band or at IF stage in a receiver to minimize the interference.

Adaptive equalizer works on two operating modes, that are training and tracking operating'modes.

A fixed length training sequence is sent by the transmitter so that the receiver's equilizer may receive the signal with minimum bit error rate (BER). This sequence is a pseudo-random binary signal followed by the user data or information.

Adaptive equilizer at receiver side utilizes a receiver algorithm to compensate distortion or noise.

The training sequence permits the receivers equalizer to get the proper filter coefficients in the worst case (*i.e.* at fades, delays, distortion, noise, error etc.).

At the time user data is received, the adaptive algorithm of the equalizer tracks the changing channel. Therefore, we can conclude that adaptive equalizer is continually changing its filter characteristics over time duration.

Fig. 5.1, shows a block diagram of a communication system with an adaptive equalizer in the receiver and it is simplified communication systems.

If $x(t)$ is the original information signal and $f(t)$ is the combined complex baseband impulse response of the transmitter, channel and the RF/IF sections of the receiver, the signal received by the equalizer may be expressed in the following equation:

$$y(t) = x(t) \otimes f^*(t) + n_b(t) \qquad ...(5.1)$$

where $f^*(t)$ denotes the complex conjugate of $f(t)$, $n_b(t)$ is the baseband noise at the input of the equalizer and $\otimes$ denotes the convolution operation of the system. If the impulse response of the equalizer is $h_{eq}(t)$, then the output of the equalizer will be as:

$$\begin{aligned} \hat{d}(t) &= x(t) \otimes f^*(t) \otimes 8 > h_{eq}(t) + n_b(t) \otimes h_{eq}(t) \qquad ...(5.2) \\ &= x(t) \otimes g(t) + n_b(t) \otimes h_{eq}(t) \end{aligned}$$

where $g(t)$ is the combined impulse response of the transmitter, channel, RF/IF section of the receiver. The complex baseband impulse response of a transversal filter equalizer is given by the following equation:

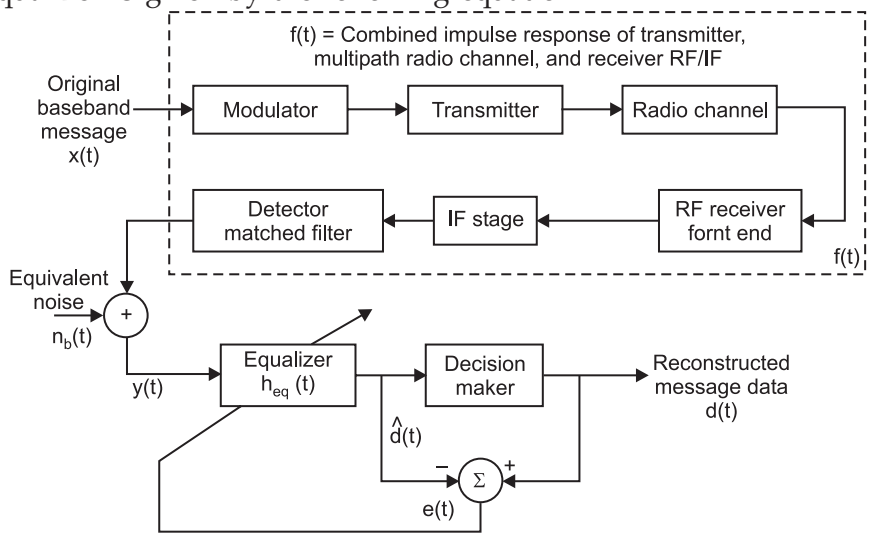


Fig. 5.1: Block diagram of a simplified communications system using an adaptive equalizer at the receiver.

$$h_{ea}(t) = \sum_n C_n \delta(t - nT) \qquad ...(5.3)$$

where C_n are the complex filter coefficient of the equalizer.

The desired output of the equalizer is $x(t)$, the original source data and assume that $n_b(t) = 0$ then in order to force $\hat{d}(t) = x(t)$ in equation (5.2), $g(t)$ must be equal to $f^*(t) \otimes h_{eq}(t)$ and it is given as:

$$g(t) = f^*(t) \otimes h_{eq}(t) = \delta(t) \qquad ...(5.4)$$

The goal of equalization is to satisfy equation (5.4) so that the combination of the transmitter, channel and receiver appear to be an all pass channel.

In the frequency domain equation (5.4) can be expressed as the following equation:

$$H_{eq}(f)F^*(-f) = 1 \qquad ...(5.5)$$

where $H_{eq}(t)$ and $F(f)$ are Fourier transforms of $h_{eq}(t)$ and $f(t)$ respectively.

Equation (5.5) indicates that an equalizer is actually an inverse filter of the channel or communication link between transmitter and receiver.

If the channel is frequency selective, the equalizer enhances the frequency components with small amplitudes and attenuates the strong frequencies in the received frequency spectrum in order to support a flat, composite, received frequency response and linear phase response.

For a time-varying channel, an adaptive equalizer is designed to track the channel variations so that equation (5.5) is approximately satisified to track the channel variations.

5.3 EQUALIZERS IN A COMMUNICATIONS RECEIVER

Equalizer in a communication receiver means how the equalizer fits and works into the wireless communication link to remove the ISI.

The basic purpose and working of an equalizer have been explained in previous section, now we will study how the equalizer in communication link works.

Fig. 5.1, show that there is always a possibility of ISI and tracking error at receiver and therefore the instantaneous combined frequency response will not always be flat, this causes some finite prediction error is produce or present.

Because adaptive equalizers are implemented using digital logic, it is most convenient to represents all time signals in discrete form.

Let T is used to show some increment of time between successive observations of signal state. Letting $t = t_{en}$, where n is an integer that represent time $t_n = nT$, time waveform may be corresponding expressed as a sequence on n in the discrete domain, using this notation, equation (5.2) may be expressed as the following equation number (5.6):

$$\hat{d}(t) = x(n) \otimes g(n) + n_b(n) \otimes h_{eq}(n) \quad ...(5.6)$$

and the prediction error is

$$e(n) = d(n) - \hat{d}(t) = dn - [x(n) \otimes g(n) + n_b(n) \otimes h_{eq}(n)] \quad ...(5.7)$$

The mean squared error $E[|e(n)^2]$ is one of the most important measure of how well an equalizer works in a communication system.

$E[|e(n)|^2]$ is the expected value of the squared prediction error $|e(n)|^2$ but time averaging can be used if $e(n)$ is ergodic. But practically ergodicity is impossible to prove and algorithms are developed and implemented using time averages instead of ensemble averages. This prove to be highly effective and in general, better equalizers provide smaller values of $E[|e(n)|^2]$ to prove and to implementation.

Minimizing the mean square error tends to reduce the bit error rate. This can be understood with a simple intuitive explanation of BER.

Suppose $e(n)$ is Gaussian distributed with zero mean. Then $E[|e(n)|^2]$ is the variance of the error signal. If the variance is minimized then there is less chance of disturbing the output signal $d(n)$. Thus the decision device is likely to detect $d(n)$ as the transmitted signal $x(n)$ (see Fig. 5.1). As a result, there is a smaller probability of error when $E[|e(n)|^2]$ is minimized. For wireless communication link, it would be best to minimize the instantaneous probability of error (P_e) instead of the mean squared error with minimize the effect of error on the system or in the system.

5.4 SURVEY OF EQUALIZATION TECHNIQUES

Equalization techniques can be subdivided into two general categories:

(*i*) Linear equalization categories.

(*ii*) Non-linear equalization categories.

These categories are determined from how the output of an adaptive equalizer is used for subsequent control (feedback) of the equalizer.

The main difference in linear and non-linear equalization is of feedback it reconstructed message data $d(t)$ is not used in the feedback path to adapt the equaliser then the equalization in linear and if $d(t)$ is feedback to change the subsequent outputs of the equalizer, then the equalization is non-linear, this is the main difference between linear and non-linear equalization.

Fig. 5.2 provides a general categorization of the equalization techniques according to the types, structures and algorithms used by it.

The most common equalizer structure is a linear transversal equalizer (LTE). A linear transversal filter is made up of tapped delay lines, with the tapping spaced a symbol period (T_s), apart as shown in Fig. 5.3. Fig. 5.3 also show the effect of clock pulse in a LTE with delay element.

Assuming that the delay elements have unity gain and delay T_s, the transfer function of a linear transversal equalizer can be written as a function of the delay operator $\exp(-j\omega T_s)$ or z^{-1} which is direct related to input $y(t) + n_b(t)$.

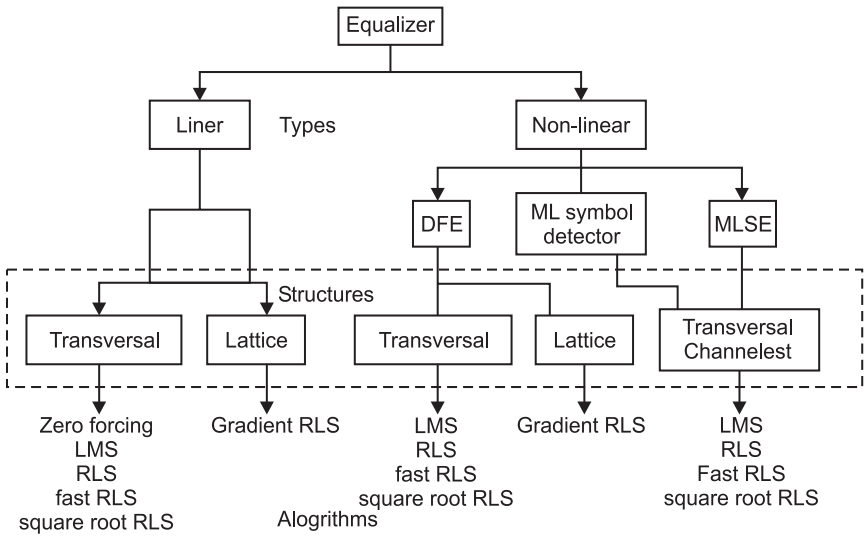


Fig. 5.2: Classification of equalizers

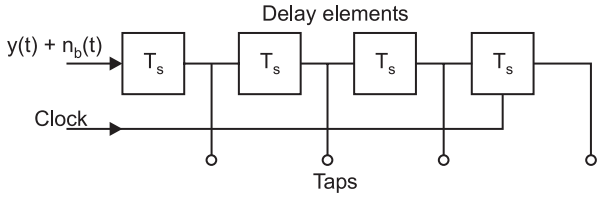


Fig. 5.3: Basic linear transversal equalizer structure.

The simplest LTE uses only feed forward taps and the transfer function of the equalizer filter is a polynomial in z^{-1}. This filter has many years but pole only at $z = 0$, and is called a finite impulse response (FIR) filter or simply a transversal filter means both filter similar (FIR and LTF).

If the equalizer have both feed forward and feedback taps, its transfer function is a rational function of z^{-1} and is called an infinite impulse response (IIR) filter with poles and zeroes in both direction, either forward or backword.

Fig. 5.4 shows a tapped delay line filter with both feed forward and feedback taps with a summing device, one with feedback taps and one with feed forward tops.

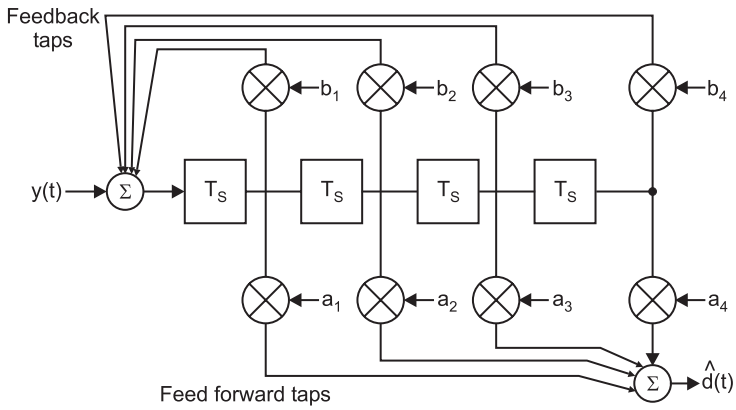


Fig. 5.4: Tapped delay line filter with both feed forward and feedback taps.

Since IIR filters tends to be unstable when used in channels where the strongest pulse arrives after an echo pulse (*i.e.* leading echoes), they are rarely used in the system.

5.5 LINEAR EQUALIZERS

Linear equalizers is further divided into two parts:

(*i*) Lineartransversal equalizers categories.

(*ii*) Linear lattice equalizers categories.

(*i*) **Linear Transversal Equalizers:** These equalizers are simplest to construct and have current and past value of received signal for equalization.

It can be implemented by an FIR filter.

In this equalizer, the correct and past values of the received signal are linearly weighted by the filter coefficient and summed to produce the output, as shown in Fig. 5.5 and summer output is feeded to the threshold detector for obtain the output.

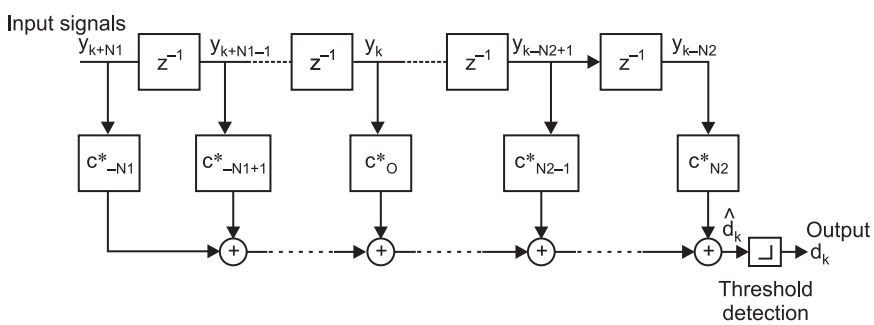


Fig. 5.5: Structure of linear transversal equalizer.

If the delays and the tap gains are analog, the continuous output of the equalizer is sampled at the symbol rate and samples are applied to the decision device for gain of the final sample.

The implementation is, however, usually carried out in the digital domain where the samples of the received signal are stored in a shift register as a memory element.

The output of this transversal filter before a decision is made (threshold detection) is given by the following equation:

$$\hat{d}_k = \sum_{n=-N_1}^{N_2} (C_n^*) y_{k-n} \qquad ...(5.8)$$

where C^*_n represents the complex filter coefficient or tap weights, $\hat{d}_k$ is the output at time index k, y_i is the input received signal at time $t_0 + iT$, t_0 is the equalizer starting time and $N = N_1 + N_2 + 1$ is the number of taps. The value of N_1 and N_2 denote the number of taps used in forward and reverse position of the equalizer, means N_1 for forward and N_2 for reverse position.

The minimum mean squared error $E[|e(n)|^2]$, that is a linear transversal equalizer can be written as–

$$E[|e(n)|^2] = \frac{T}{2\pi}\int_{-\frac{\pi}{T}}^{\frac{\pi}{T}} \frac{N_0}{|F(e^{j\omega T})^2 + N_0|} d\omega \qquad ...(5.9)$$

where $F(e^{j\omega r})$ is the frequency response of channel, and N_0 is noise power spectral density of an linear transversal equalizer.

(*ii*) **Linear Lattice Eqaualizer:** Linear lattice equalizer is implemented using a lattice filter from Fig. 5.6.

In a lattice filter, the input signal y_k is transformed into a set of N intermediate forward and backward error signals $f_n(k)$ and $b_n(k)$, respectively, which are used as input to the top multipliers and are used to calculate the updated coefficients (showing in Fig. 5.6).

Fig. 5.6 show the structure of a lattice equalizer.

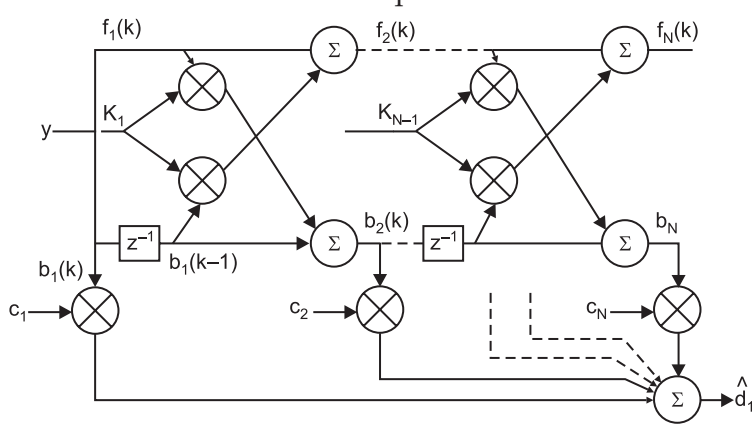


Fig. 5.6: The structure of a lattice equalizer.

Each stage of the lattice is then characterized by the following equation:

$$f_1(k) = b_1(k) = y(k) \quad ...(5.10)$$

$$f_n(k) = y(k) - \sum_{i=1}^{n} K_i y(k-i) = f_{n-1}(k) + K_{n-1}(k) b_{n-1}(k-1) \quad ...(5.11)$$

$$b_n(k) = y(k-n) - \sum_{i=1}^{n} K_i y(k-n+i) \quad ...(5.12)$$

$$b_n(k) = b_{n-1}(k-1) + K_{n-1}(k) f_{n-1}(k) \quad ...(5.13)$$

where $K_n(k)$ is the reflection coefficient for the nth stage of the lattice.

The backward error signals $b_{n'}$ are then used as inputs to the tap weights, and the output of the equalizer is given by the following equation:

$$\hat{d}_k = \sum_{n=1}^{N} C_n(k) b_n(k) \quad ...(5.14)$$

This output is given for k-stage length equalizer. The length of this equalizer is dynamic in nature, it means the length of the linear lattice equalizer can be increased by the algorithm without stoping the operation of the equalizer or during the operation.

Linear lattice equalizer has some advantages like as the –

(*i*) Faster convergence.

(*ii*) Numerical stability.

(*iii*) Unique structure of the lattice filter.

(*iv*) Dynamic assignment of the length.

5.6 NON-LINEAR EQUALIZATION

The main function of non-linear equalization can be explain as that the –

Non-linear equalizers are used in application where the channel distortion is too high for a linear equalizer to handle and are common place in practical wireless system.

Linear equalizers do not perform well on channels which have deep spectral nulls in the passband so that non-linear equalizer is used to perform well on channels for equalization.

Two very effected non-linear methods have been developed which offer improvements over linear equalization techniques and are used in most 2G and 3G systems or 2G and 3G technologies –

(*i*) Decision Feedback Equalization (DFE).

(*ii*) Maximum Likelihood Squence Estimation (MLSE).

5.6.1 Decision Feedback Equalization (DFE)

It is based on feedback system.

The basic idea behind decision feedback equalization is that once an information symbol has been detected and decided upon, the ISI that it induces

on future symbols can be estimated and subtracted out before detection of subsequent symbols are possible to final.

The DFE can be realized in either the direct transversal form or as a lattice filter means in any form.

The direct form of DFE is shown in Fig. 5.7. It consists of a feed forward filter (FFF) and a feedback filter (FBF) with input y_{k+nl}.

The FBF is driven by decision on the output of the detector and its coefficient can be adjusted to cancel the ISI on the current symbol from past detected symbols, ISI means inter symbol interference.

The equalizer has $N_1 + N_2 + 1$ taps in the feed forward filter and N_3 taps in the feedback filter and its output can be expressed as by the following equation Fig. 5.7. Decision feedback equalizer (DFE).

$$\hat{d} = \sum_{n=-N_1}^{N_2} C_n^* y_{k-n} + \sum_{i=1}^{N_3} F_i^* d_{k-i} \qquad ...(5.15)$$

where C_n^* and y_n are tap gain and the inputs, respectively to the forward filter, F^*_t are tap gain for the feedback filter, and $d(i < k)$ is the previous decision made on the detected signal or received signal.

That is once $\hat{d}_k$ is obtained using equation 5.15, d_k is decided from it, then d_k along with previous decision d_{k-1}, d_{k-2} are feedback into the equalizer and $\hat{d}_{k+1}$ is obtained using equation 5.15 which is used to represent the output of an equalizer.

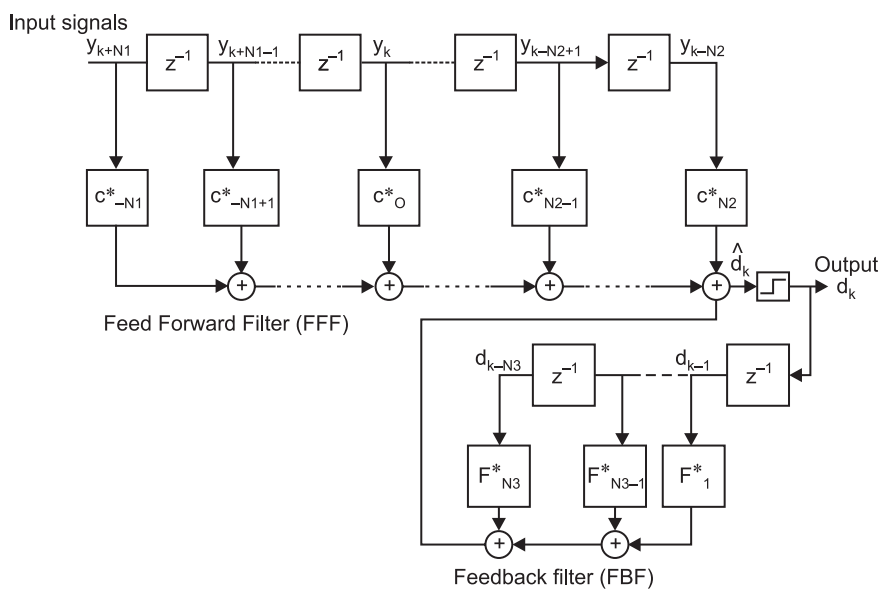


Fig. 5.7: Decision Feedback Equlizer (DFE).

The minimum mean squared error of DFE can be achieve by the following:

$$E[e(n)\,|^2]_{\min} = \exp\left[\frac{T}{2\pi}\int_{-\frac{\pi}{T}}^{\frac{\pi}{T}} \ln\left[\frac{N_0}{|F(e^{j\omega T})|^2 + N_0}\right] d\omega\right] \quad ...(5.16)$$

It can be shown that minimum MSE for DFE in equation (5.16) is always smaller then that of an LTE in equation (5.9) unless |F(e'to7)| is a constant {*i.e.* when adaptive equalization is not needed during the operation).

If there are null in |F(e/'<n7)|, a DFE has significantly smaller minimum MSE than an LTE there, an LTE is well behaved when the channel spectrum is comparatively flat, but if the channel is severly distorted or exhibits nulls in the spectrum, the performance of an LTE deteriorates and the mean square error of a DFE is much better than a LTE means MSE and performance is much better of a DFE then LTE.

The lattice implementation of the DFE is equivalent to a transversal DFE having a feed forward filter of length N_1 and a feedback filter of length N_2 where $N_1 > N_2$ means N_1 is greater than N_2.

5.6.2 Maximum Likelihood Sequence Estimation (MLSE)

The MSE based linear equalizers described previously are optimum with respect to the criterion of minimum probability of symbol error when the channel does not introduce any amplitude distortion. Yet this is precisely the condition in which an equalizer is needed for a mobile communication link. This limitation on MSE based equalizer lead researchers to investigate optimum nonlinear structure. These equalizer use various forms of classical. Maximum likelihood receiver structure. Using a channel impulse response simulator within the algorithm, the MLSE tests all possible data sequences and choose the data sequence with the maximum probability as the output.

MLSE needs match filter, delay element and channel estimator for its implementation or operation.

This equalizer is based on an algorithm known as "Viterbi algorithm". This algorithm was recognized to be a maximum likelihood sequence estimator of the state sequences of a finite state (Markov) process observed in memory less noise. This algorithm has recently been implemented successfully for equalizers used in mobile channels.

The block diagram of a MLSE receiver based on the DFE is shown in Fig. 5.8 with match filter, channel estimator, and delay element.

As shown in Fig. 5.8 the channel output is given to a match filter which operates on continuous time signal. The output of this filter is given in discrete form to MLSE. The estimated data sequence from MLSE is feedback to channel estimator for error reduction. This is achieved by getting some delay by using delay element and the summed with the output of channel estimator. After summing point, the signal is given to match filter and again channel estimator for further estimation.

The main advantage of this equalizer is that it minimizes the probability of sequence error or reduce the probability of sequence error.

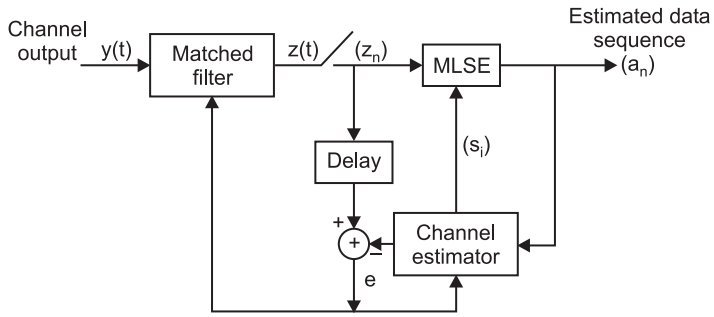


Fig. 5.8: The structure of a Maximum Liklihood Sequences Estimator (MLSE) with an adaptive matched filter.

5.7 DIVERSITY TECHNIQUE

Diversity is a powerfull communication receiver technique that provide wireless link improvement at relatively low cost means quality of connection.

Unlike equalization, diversity requires no training overhead since a training sequence is not required by the transmitter for transmission.

Diversity exploits the random nature of radio propagation by finding independent signal path for communication between two terminal.

The diversity concept can be explained that if one radio path undergoes a deep fade, and other independent path may have a strong signal by having more than one path to select from both the instantaneous and average SNR at the receiver may be improved often by as much as 20 dB to 30 dB and it is fixed with in this range.

Diversity is changed from equalization in terms of compensation factor.

Equalization technique was used to reduce ISI where as diversity is used to reduce the effect of fading on wireless communication. Both technique have the separate application or role in a wireless communication.

There are various type of diversity techniques used in microscopic diversity, space diversity, time diversity, frequency diversity, polarization diversity etc.

5.7.1 Derivation of Selection Diversity Improvment

Before discussing the many diversity techniques that are used, it is suitable to quantitatively determine the advantage that can be achieved using diversity technique.

Consider M independent Rayleigh fading channels available at a receiver. Each channel is called a diversity branch further assume that each branch has the same average SNR given by following as a function of E_b and N_b:

$$\textbf{SNR} = \Gamma = \frac{E_b}{N_d}\overline{\alpha^2} \qquad ...(5.17)$$

where we assume $\overline{\alpha^2} = 1$

If each branch has an instantaneous SNR = *vi*, then the *Pdf of* v_i is

$$P(v_i) = \frac{1}{\Gamma} e^{-\frac{-v_i}{\Gamma}} \quad v_i \geq 0 \qquad ...(5.18)$$

where Γ is the mean SNR of each branch of diversity technique.

The probability that a single branch has an instantaneous SNR less than some threshold v is

$$P_r\,[v_i \leq v] = \int_0^v P(v_i)dv_i = \int_0^v \frac{1}{\Gamma} e^{\frac{-v_i}{\Gamma}} dv_i = 1 - e^{\frac{v}{\Gamma}} \qquad ...(5.19)$$

Now the probability that all M independent diversity branches receive the signals which are at the same time less then some specific SNR threshold v is

$$P_r[v_i \;....\; v_m \leq v] = \left(1 - e^{\frac{-v}{\Gamma}}\right)^M = P_M(v) \qquad ...(5.20)$$

$P_M(v)$ in equation (5.20) is the probability of all branches failing to achieve an instantaneous SNR = v. If a single branch achieves SNR > v then the probability that SNR > v for one or more branches is given by

$$P_r\,[v_i - v] = 1 - P_M\,(v) = 1 - \left(1 - e^{\frac{-v}{\Gamma}}\right)^M \qquad ...(5.21)$$

equation (5.21) is an expression for the probability of exceeding a level of threshold when selection diversity is used and is plotted in Fig. (5.9).

To determine average signal to noise ratio of the received signal when diversity is used, it is first necessary to find out the *Pdf of* fading signal. For selection diversity, the average SNR is found by first computing the derivative of the CDF $P_M\,(V)$ in order to find the *Pdf of v,* the instaneneous SNR when M branches are used. Proceeding along these lines can be written as–

$$P_M(v) = \frac{d}{dv} P_M(v) = \frac{M}{\Gamma}\left(1 - \frac{-v}{\Gamma}\right)^{M-1} e^{\frac{-v}{\Gamma}} \qquad ...(5.22)$$

Then the mean SNR, $\overline{v}$ may be expressed as the following:

$$\overline{v} = \int_0^\infty v P_M(v)\,dv = \Gamma \int_0^\infty M_x (1 - e^{-x})^{M-1} e^{-x} dx \qquad ...(5.23)$$

where $x = \frac{v}{\Gamma}$ is the average SNR for a single branch, equation (5.23) is evaluated to field the average SNR improvement offered by selection diversity which is given by the

$$\frac{\overline{v}}{\Gamma} = \sum_{k=1}^{M} \frac{1}{k} \qquad ...(5.24)$$

From equation (5.24), it can be shown that the average SNR in the branch which is selected using selection diversity naturally increases, since it is always guaranteed to be above the specified threshold. Thus selection diversity offers an average improvement in the link margin without requiring additional transmitter power or sophisticated receiver circuitry.

Selection diversity is easy to implement because all that is needed is a side monitoring station and an antenna switch at the receiver. However, it is not an optimal diversity technique because it does not use all of the possible branches simultaneously.

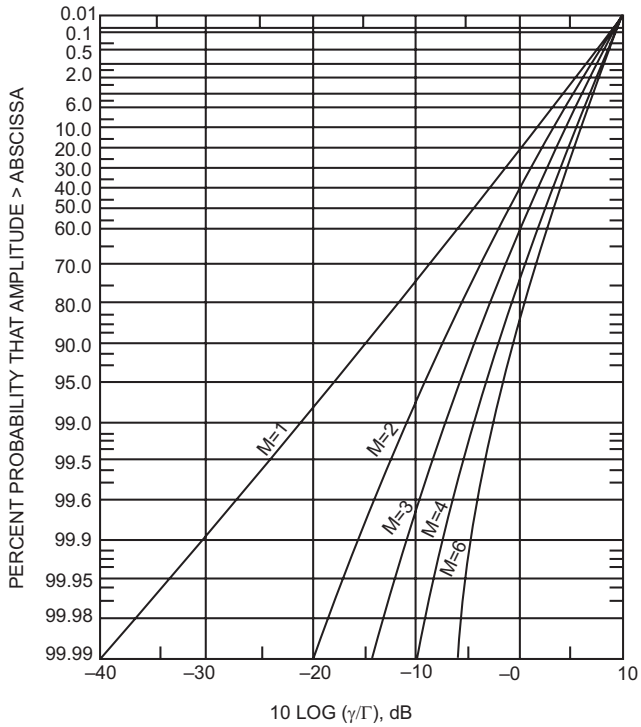


Fig. 5.9: Graph of probability distributions of SNR = λ threshold for *M* branch selection diversity. The term Γ represents the mean SNR on each branch.

5.7.2 Derivation of Maximal Ratio Combining Improvement

In Maximal ratio combining, the voltage signal r_i from each of the *M* diversity branches are co-phased to provide coherent voltage addition and are particular weighted to provide optimal SNR.

If each branch has gain G_i, then the resulting signal envelope applied to the detector is given by the following equation:

$$r_M = \sum_{i=1}^{M} G_i r_i \qquad ...(5.25)$$

Assuming that each branch has the same average noise power *M*, the total noise power N_T applied to the detector is simply the weighted sum of the noise in each branch. Thus we can write that.

$$N_T = N\sum_{i=1}^{M} G_i^2 \qquad ...(5.26)$$

Which results in an SNR applied to the detector, v_M, is given by following equation with noise power N_T

$$N_T = \frac{r_M^2}{2N_T} \quad ...(5.27)$$

Using Chebychev's inequality, v_M is maximized when $G_i = \frac{r_i}{N}$, which leads to

$$V_M = \frac{1}{2}\frac{\sum\left(\frac{r_N^2}{N}\right)^2}{N\sum\left(\frac{r_i^2}{N^2}\right)}\frac{1}{2}\sum_{i=1}^{M}\frac{r_i^2}{N} = \sum_{i=1}^{M} v_i \quad ...(5.28)$$

Thus the SNR out of the diversity combiner is simply the sum of the SNRs in each branch.

The value for v_i is $\frac{r_i^2}{2N}$, where r_i is the equal to *r(t),* so the received signal envelope *2N′* for a fading mobile radio signal can be modeled from two independent Gaussian random variables T_c and T_s each having zero mean and equal variance σ^2. That is

$$v_i = \frac{1}{2N}r_i^2 = \frac{1}{2N}(T_c^2 + T_s^2) \quad ...(5.29)$$

Hence v_M is a chi-square distribution of $2M$ Gaussian random variable with variance $\frac{\sigma^2}{(2N)} = \frac{\Gamma}{2}$, where r is define in equation (5.17).

The resulting *Pdf* for V_M can be shown to be

$$P(V_M) = \frac{V_M{}^{M-1}e^{\frac{-v_M}{\Gamma}}}{\Gamma^M(M-1)!} \text{ for } v_M \geq 0 \quad ...(5.30)$$

The probability that v_m is less then some SNR threshold v is

$$P_r\{V_M \leq v\} = \int_0^v P(v_M)dv_M = 1 - e^{\frac{-v_M}{\Gamma}}\sum_{k=1}^{M}\frac{\left(\frac{V}{\Gamma}\right)^{k-1}}{(k-1)!} \quad ...(5.31)$$

Equation (5.31) is probability distribution for maximal ratio combining it follows directly from equation (5.28) that the average SNR, $\overline{v}_M$, is simply the sum of the individual $\overline{v}_i$ from each branch.

The control algorithms for setting the gains and phases for maximal ratio combining receivers are similar to those required in equalizers and RAKE receiver.

5.7.3 Practical Space Diversity Consideration

Space diversity, also known as *antenna diversity,* is one of the most popular form of diversity is also used in base station designing in a wireless communication system.

The concept of antenna of space diversity is also used in base station designing in a wireless communication system.

At each cell site multiple base station receiving antenna are used to provide diversity reception. However, since the important scatters are generally on the ground in the neighbour of the mobile, the base station antenna must be spaced considerably for apart to achieve decorrelation to the others.

Space diversity can thus be used at either the mobile station or base station or both base and mobile station.

Fig. 5.10 shows a general block diagram of space or antenna diversity scheme with variable gain and switching logic or demodulators.

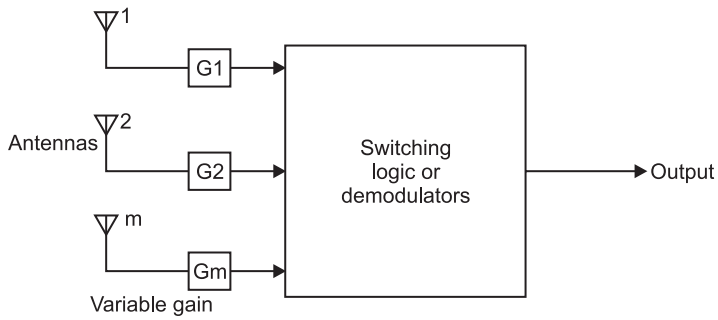


Fig. 5.10: Block diagram of space diversity or antenna diversity.

Space diversity reception method can be classified into four categories:

(*i*) Selection diversity

(*ii*) Feedback diversity

(*iii*) Maximum ratio combining

(*iv*) Equal gain diversity.

5.7.3.1. Selection diversity

In this technique, the receiver branch having the highest instantaneous signal to noise ratio (SNR) is connected to the demodulator circuitry (Fig. 5.10) or switching circuit.

The antenna signals themselves could be sampled and the best signal having high signal power and low noise power is sent to demodulation process by the demodulator.

(S + N) In practice, the branch with the largest $\frac{(S+N)}{N}$ is used, since it is difficult to measure SNR alone means signal and noise seperately.

5.7.3.2. Feedback diversity [scanning diversity]

It is also known scanning diversity technique.

In this diversity technique, instead of always using the best of 'M' signals, the M signals are scanned in a fixed sequence untill one is found to be above a predetermined threshold level.

A simplest block diagram of feedback or scanning diversity is shown in Fig. (5.11) with comparator and controlling system.

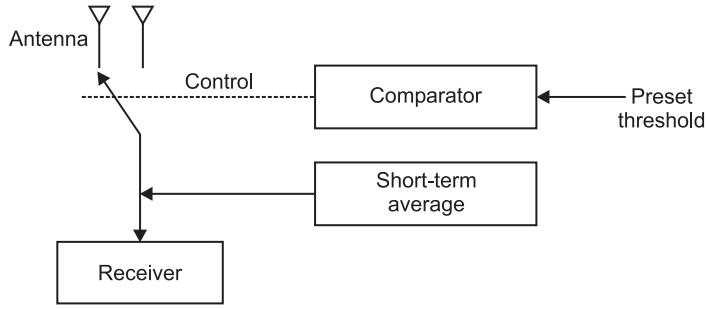


Fig. 5.11: Basic form of scanning diversity.

The main advantage of this technique is that it is very simple to construct since only one receiver is required but the resulting fading statistics of this technique are some what inferior to those obtained by other methods so it may be a drawback.

5.7.3.3 Maximal ratio diversity

This technique is most useful to produce an output with an acceptable SNR even when none of the individual signals are themselves enough to produce a strong (SNR) signal at output of the diversity system.

This technique gives the best statical reduction of fading of any known linear diversity combiner.

A simple block diagram of this technique is shown in Fig. (5.12) with co-phase and sum device, detector with adaptive control.

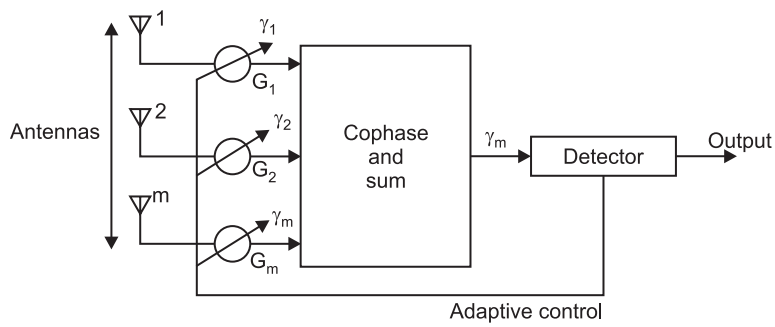


Fig. 5.12: Maximal ratio combiner diversity.

As shown in Fig. (5.12), the signals from all of the *M* branches are weighted according to their individual SNR power ratio's and then summed in summer which are in same phase.

Note that all the signal must be in same phased before being summed in a summer.

Maximal ratio combining produce and output SNR equal to the sum of the individual SNRs which are produce seperately.

5.7.3.4 Equal Gain Diversity

This technique is little different from Maximal ratio combining technique in which variable weighting capability was required for true Maximal ratio combining explaination.

In this case, each branch weight is constant and set to unity but the signals from each branch must be in same phase to provide equal gain combination to show equal gain diversity.

This allows the receiver to exploit the signals that are simultaneously received on each branch of the system.

The over all performance of this diversity technique is superior to selection diversity while less valuable to Maximal ratio combining diversity.

5.7.4 Polarization Diversity

The complexity and high cost of space diversity at base station forced research take interest in another technique known as polarization diversity development. In this technique, the signal is received at base station by a polarization diversity antenna for functionality.

In early days, vertical polarization was used to all subscriber units were mounted in vehicles and use vertical chip antennas but now the subscriber units are portable. Therefore it is reported that the horizontal and verticals, path between a mobile and base station are uncorrected to communicate.

This decorrelation in polarization is due to multipath reflection. Therefore the reflection coefficient for each polarization is different which lead to different amplitude and phase for each reflection with-in the polarization.

To characterize the multipath inside buildings, circular and linear polarized antennas have been used to perform.

The result of practical experiments shows that the polarization diversity is viable diversity reception technique.

5.7.5 Frequency Diversity

In this diversity technique, information is carried on more then one carrier frequency from one point to another.

This diversity could be achieved because the carrier frequency are seperated by more then the coherence bandwidth of the channel therefore these are uncorrected. As a result, they does not experience the same fades or uncorrected information.

Frequency diversity is often used in microwave line of sight link which carry several channel in a frequency division multiplex mode (FDM) in which channels are divided into frequency slots.

The main drawback of frequency diversity is that it requires spare bandwidth and as many receivers as there are channels used.

New FDM modulation and access techniques exploit frequency diversity by providing, at same time modulation signal with error control coding across a

large bandwidth so that if a particular frequency undergoes a fade, the composite signal will still be demodulated by demodulator.

5.7.6 Time Diversity

Time diversity repeatedly transmit information at time spacings that exceed the coherence time of the channel so that multiple repetitions of the signal will be received with independent fading conditions, thereby providing for diversity and TDM technique may be used or not.

Time diversity involves the use of RAKE receiver for spread spectrum CDMA, where multipath channel provides redundancy in the transmitted message or information.

By demodulating several replicas of the transmitted CDMA signal, where each replica experiences a particular multipath delay, the RAKE receiver able to align the replicas in time so that a better estimate of the signal may be formed at the receiver.

5.8 RAKE RECEIVER

A RAKE receiver collect the time shifted versions of the original signal by providing a separate correlation receiver for each of the multipath signals.

There are four 'RAKE' receivers with in each base transceiver and three 'RAKE' receivers with in each mobile phone. These are so called because they resemble a lawn RAKE receiver.

The main function of RAKE receiver at both ends (mobile and base station) is to aggregate the diversity received signals at the receiver.

The direct signal at the RAKE receiver to form the composite signals that is used to process the mobile call.

The multipath signals are additives to the direct signal to obtain the cleanest, strongest signal possible. Therefore we can conclude that the signal to noise ratio (SNR) gets improved by using RAKE receiver.

A RAKE receiver circuitry is shown in Fig. 5.13 with correlator 1 to correlator M and integrator.

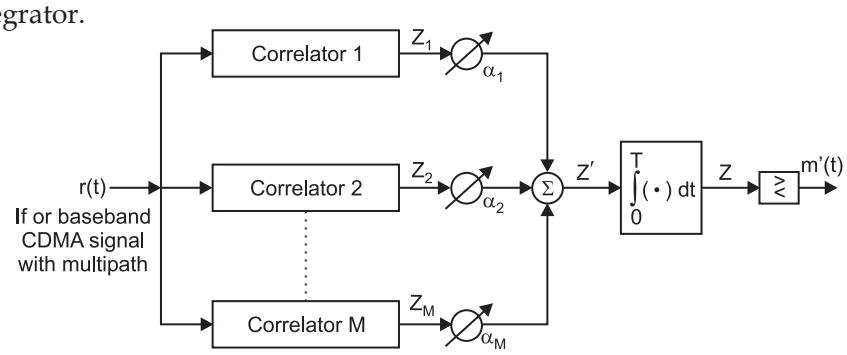


Fig. 5.13: RAKE receiver.

A RAKE receiver perform its function in following step:

It collects the time shifted version of original signal by providing a separate correlation receiver for each of the multipath signals.

Each correlation receiver is adjusted in time delay so that a microprocessor controller can cause different correlation receivers to search in different time windows for significant multipath.

The range of time delays that a particular correlator can search is called a search window. A RAKE receiver having *'M'* correlators separately detect *'M'* strongest multipath components.

The output of each correlator is than weighted to provide a better estimate of the transmitted signals. This is done in weighting network that provide a linear combination of the correlators output.

Decisions based on the combination of the *'M'* separate decision statistics offered by the RAKE receiver provide a form of diversity which can over come fading and there by improve CDMA reception.

From Fig. 5.13, here $r(t)$ is the input signal after multipath reception which is given to *'M'* correlators. The output of *'M'*correlators are denoted by Z_1, Z_2 Z_M. These output are weighted by α_1, α_2 ... α_M respectively. The weighting coefficients are based on the power (SNR) from each correlator output.

Note that SNR (or power) is small then it will assign a small weighting factor. The overall signal Z′ is given by

$$Z' = \sum_{m=1}^{M} \alpha_m Z_m \qquad ...(5.32)$$

where a_m are weighting coefficients which are normalized to the output signal power of the correlator in such a way that the coefficient sum of unity.

$$a_M = \frac{Z_m^2}{\sum_{m=1}^{M} Z_m^2} \qquad ...(5.33)$$

Weighting coefficient can be generated by many ways but choosing weighting coefficient based on the actual outputs of the correlators yields better RAKE receivers.

SOLVED EXAMPLES

Example 5.1. *Consider the design of the US digital cellular equalizer. If f = 900 MHz and the mobile velocity V = 80 km/hr, determine the following :*

(*a*) *The maximum Doppler shift*

(*b*) *The coherence time of the channel*

(*c*) *The maximum number of symbols that could be transmitted without updating the equalizer, assuming that the symbol rate is* 24.3 *k symbols/sec.*

Solution. (*a*) From equation 3.52, the maximum Doppler shift is given by

$$f_d = \frac{V}{\lambda}\frac{\left(\frac{80,000}{3600}\right)m/s}{\left(\frac{1}{3}\right)m} = 66.67 \text{ Hz}$$

(*b*) From equation 3.58 (*c*), the coherence time is approximately

$$T_c = \sqrt{\frac{9}{16\pi f_d^2}} = \frac{0.423}{66.67} = 6.34 \text{ msec}$$

Note that if equation 3.58 (*b*) or 3.58 (*a*) were used, T_c would increase or decrease by factor of 2–3.

(*c*) To ensure coherence over a TDMA time slot, data must be during a 6.34 ms interval. For R_s = 24.3 k symbols/sec, the number of bits that can be sent is

$$N_b = R_S T_C = 24,300 \times 0.00634 = 154 \text{ symbols}$$

Each time slot in the US digital cellular standard has a 6.67 ms duration and 162 symbols per time slot, which are very close to values in this example.

Example 5.2. *Assume four branch diversity is used, where each branch receives an independent Rayleigh fading signal. If the average SNR is 20 dB, determine the probability that the SNR will drop below 10 dB. Compare this with the case of a single receiver without diversity.*

Solution. For this question, the specified threshold γ = 10 dB, Γ = 20 dB and there are four branches. Thus $\frac{\gamma}{\Gamma} = 0.1$ and using equation 5.20

$$P_4(10 \text{ dB}) = (1 - e^{-0.1})^4 = 0.000082$$

When diversity is not used, equation 5.20 may be evaluated using m = 1

$$P_1 (10 \text{ dB}) = (1 - e^{-0.1})\ 1 = 0.095$$

Notice that without diversity, the SNR drops below the specified threshold with a probability that is three orders of magnitude greater than if four branch diversity is used !

Example 5.3. Equalization. *The IS-136 USDC standard specifies the use of decision feedback equalizers (DFEs).*

Diversity

(*i*) *The US AMPS system makes use of spatial diversity.*

(*ii*) *The PACS standard specifies the use of antenna diversity for base station and portable units.*

Solution. Channel coding

(*i*) The IS-95 standard as proposed makes use of a rate 1/3. Constraint length L = 9 convulational code with block interleaving. The interleaver used is a 32* 18 block interleaver.

(*ii*) The AMPS system makes use of a (40, 28) BCH code for the forward control channel and a (48, 30) BCH code for the reverse control channel.

Equalization, diversity and channel coding, as discussed, have the same goal of improving the reliability and quality of the communication service over small-scale channel changes. Each technique has its own advantages and disadvantages. The trade offs to be considered are those of complexity/power/cost versus system performance. Each technique is capable of improving system performance significantly.

HIGHLIGHTS

1. Equalization techniques can be subdivided into two general categories, linear and nonlinear equalization. The most common equalizer structure is a linear transversal equalizer (LTE). Nonlinear equalizers are used in applications where the channel distortion is too high for a linear equalizer to handle.
2. Decision feedback equalization (DFE) and maximum likelihood sequence estimation (MLSE) are the most commonly used non-linear equalization techniques, used for equalization.
3. Diversity improves transmission performance by making use of more than one independetly faded version of the transmitted signal. If various replicas of the signal, carrying the same information, are received over multiple channels that exhibit independent fading with comparable strengths, the changes that all the independently faded signal components experience deep fading simultaneously are greatly reduced to minimize the effect.
4. Angle diversity can be viewed as a special case of space diversity since it also requires multiple antennas for transmission.
5. A Rake receiver can separate the received signal components from different propagation paths by using code correlation and after it combine the signal components constructively.
6. In CDMA, exploring the path diversity, reduces the transmitted power required and increases the system capacity for transmission.
7. Diversity is a powerful communication receiver technique that provide wireless link improvement at relatively low cost.
8. Diversity achieve the random nature of radio propagation by finding independent or at least highly uncorrected signal paths for communications. Transmit diversity (TD) is one of the key contributing technologies to defining the ITU and define 3G system WCDMA and CDMA 2000.
9. Spatial diversity is introduced into the signal by transmitting through multiple antennas or by using multiple antennas.
10. In addition to diversity gain, antenna gain can also be incorporated through channel state feedback. This leads to the categorization of transmit diversity (TD) methods into open loop and closed loop methods. The present paper reviews the concept of transmit diversity and describes the features of selected transmit diversity (TD) techniques or both.
11. In open loop diversity methods, a predetermined form of diversity is introduced using multiple antennas. Open loop transmit diversity methods have following advantages or properties:
    (*i*) Signaling overhead is not needed to achieve this form of diversity.
    (*ii*) The mobile station (MS) unit.

12. Space-time transmit diversity is an open loop technique in which the symbols are spread using multiple Walsh codes. If differs slightly from STTD. The main merit of STS over OTD is that all symbols are transmitted over all antennas. Therefore, it provides the addition of temporal diversity in the form of repetition coding prior to the decoding process.

13. Equalization is a technique used for combat intersymbol interference (ISI). In fact, equalization, diversity and channel coding are three techniques which can be used independently or in tandem to improve received signal quality and link performance over small-scale time and distances between two points.

14. Equalization compensates for intersymbol interference (ISI) created by multipath time dispersive channels. If the modulation bandwidth increase the coherence bandwidth of the radio channel. ISI occurs and modulation pulses are spread in time into adjacent symbols. An equalizer within a receiver compensates for the average range of required channel amplitude and delay characteristics. Equalizers must be adaptive since the channel is generally unknown and time varying.

15. A diversity technique can be classified into the following types:

(*i*) Space or antenna diversity
(*ii*) Microscopic diversity
(*iii*) Polarization diversity
(*iv*) Frequency diversity
(*v*) Time diversity

16. A space diversity or antenna diversity can also be classified into following broad (four) categories:

(*i*) Selection diversity
(*ii*) Feedback diversity
(*iii*) Maximal ratio combining diversity
(*iv*) Equal gain diversity

17. The comparison between equalization and diversity can be explain on the basis of compensation means the equalization is used to compensate intersymbol interference (ISI) but the diversity is used to compensate fading which occurs due to multipath reception.

18. The main function of RAKE receiver at both ends (mobile and base station transceivers) is to aggregate the diversity received signals.

19. The power control is classified into two types:

20. (*i*) Open loop power control and
(*ii*) Closed loop power control.

20. In open loop power control, the transmitter define the channel conditions (gain) and adjusts its transmission power. The channel calculation depend on the transmission on the other link under the consideration that the uplink and downlink channel conditions are closely correlated.

21. In the FDD systems, the uplink and downlink connection use different carrier frequency which define the different Doppler shift and different channel fading conditions.

22. In the TDD system, both the uplink and downlink share the same channel frequency at different time periods.

23. There are three non-linear methods are used over linear equalization technique:
(*i*) Decision feedback equalization (DFE)
(*ii*) Maximum likelihood symbol detection (MLSD)
(*iii*) Maximum likelihood sequence estimation (MLSE).

PROBLEMS

5.1. Consider the two-tap adaptive equalizer shown in Fig. 5.14.
(*a*) Find an expression for MSE in terms of w_0, w_1 and N.
(*b*) If $N > 2$, find the minimum MSE.
(*c*) If $w_0 = 0$, $w_1 = 2$ and $N = 4$ samples/cycle, what is the MSE?
(*d*) For parameters in (*c*) what is the MSE if $d_k = 2\sin\left(\frac{2\pi k}{N}\right)$?

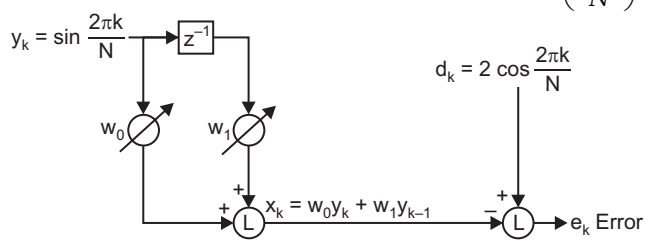


Fig. 5.14: A two-tap adaptive linear equalizer.

5.2. For a equalizer in Fig. 5.14, what weight values will produce arms value of e_k =2? Assume $N = 5$ and express your answer in terms of w_0 and w_1.

5.3. If a digital signal processing chip can perform one million multiplications per second, determine the time required between each interation for the following adaptive equalizer algorithms.
(*a*) LMS
(*b*) Kalmanroot RLS
(*c*) quare root RLS DFE
(*d*) radient lattice DFE.

5.4. Suppose a quick rule of thumb is that an RLS algorithm requires 50 iterations to cover converge. Whereas the LMS algorithm requires 1000 iterations. For a DSP chip that performs 25 million multiplications per second, determine the maximum symbol rate and maximum time interval before retraining if a five tap equalizer requires 10% transmission overhead, and the following Doppler spreads are found in a 1900 MHz channel. In case where 25 million multiplications per second is not fast enough, determine the minimum DSP multiplications per second required for each equalizer implementation and compare maximum DSP multiplication per second required for each equalizer implementation and compare maximum symbol rates for a fixed number of multiplications per second.
(*a*) 100 Hz (*b*) 1000 Hz (*c*) 10,000 Hz
Hint: (Consider the coherence time and its impact on equalizer training).

5.5. Consider a single branch Rayleigh fading signal has a 20% chance of being 6 dB below some mean SNR threshold.

(*a*) Determine the mean of the Rayleigh fading signal as referred to the threshold.

(*b*) Find the likelihood that a two branch selection diversity receiver will be 6 dB below the mean SNR threshold.

(*c*) Find the likelihood that a three branch selection diversity receiver will be 6 dB below the mean SNR threshold.

(*d*) Find the likelihood that a four branch selection diversity receiver will be 6 dB below the mean SNR threshold.

(*e*) Based on your answers above, is there a law of diminishing returns when diversity is used?

5.6. Compare $\frac{\overline{\gamma}}{\Gamma}$ (selection diversity) with $\frac{\overline{\gamma}_m}{\Gamma}$ (maximum ratio combining) for one to six branches, specifically, compare how the average SNR increases for each diversity scheme as a new branch is added. Does this make sense? What is the average SNR improvement offered by six-branch maximal ratio combining as compared to six-branch selection diversity ? If $\frac{\gamma}{\Gamma} = 0.01$, = 0.01, determine the probability that the received signal will be below this threshold for maximal ratio combining and selection diversity (assume six branches are used). How does this compare with a single rayleigh fading channel with the some threshold.

REVIEW QUESTIONS

5.1. What do you understand by equalization? Explain.

5.2. Explain the difference between equalization and diversity.

5.3. Draw the block diagram of a communication system using an equalizer and explain.

5.4. Prove that the equalizer is nothing else but an inverse filter of the channel.

5.5. How bit rate can be improve using equalizer.

5.6. What are different filter used in equalizers? Explain them.

5.7. Write down classification of different diversity techniques used to improve the signal in wireless link.

5.8. Explain decision feedback equalization in detail with a neat diagram.

5.9. What do you mean by Maximal ratio diversity?

5.10. Explain now equal gain diversity is different from maximum ratio diversity.

5.11. What is the difference between frequency diversity and time diversity?

5.12. What do you understand by RAKE receiver? Explain the working of a *M*-branch RAKE receiver.

5.13. Discuss survey of equalization techniques.

5.14. Classify and explain diversity techniques used in wireless communications.

6 Chapter

SPEECH CODING AND QUANTIZATION TECHNIQUES

6.1. *Introduction* **6.2.** *Characteristics of Speech Signals* **6.3.** *Quantization Techniques* **6.4.** *Vocoders* **6.5.** *Linear Predictive Coders (LPC)*

6.1 INTRODUCTION

The basic purpose of using speech coder in wireless communication system is to provide the code for speech or information bits in such a way so as to maximize the channel capacity without increasing the bandwidth of the channel.

Low bit speech coding offers a ways to meet this challenge. At low bit rate, the speech signals are represented by some amount of bits, therefore the information contents per bit increases.

The lower bit rate, make more channel can be compressed within a given bandwidth during an operation.

The need to develop speech coders arises because bandwidth is limited for wireless systems and number of users are continuously increasing. Therefore, such a technique was required which support more number of users in a channel without increasing the bandwidth. This is known as speech coding and the system used for this coding known as speech coder.

The goal of all speech coding systems is to transmit speech with the highest possible quality using the least possible channel capacity. This has to be accomplished while maintaining certain required levels of complexity of implementation and communication delay within the system.

In general, there is a positive correlation between coder bit rate efficiency and the algorithmic complexity required to obtain it. The more complex algorithm is, the more its processing delay and cost of implementation.

The hierarchy of speech coders is shown in Fig. 6.1 with source and waveform coders.

Speech coders differ widely in their approaches to achieving signal compression. Based on the means by which they achieve compression.

Mainly speech coders are divided into two parts-source and Waveform coders.

***(i)* Waveform coders:** This class of coder achieves only moderate economy in transmission bit rate. But the major advantages of waveform coders are less

complexity, robust design for noisy environment, source independent, therefore these are able to code equally well a variety of signals in a communication system.

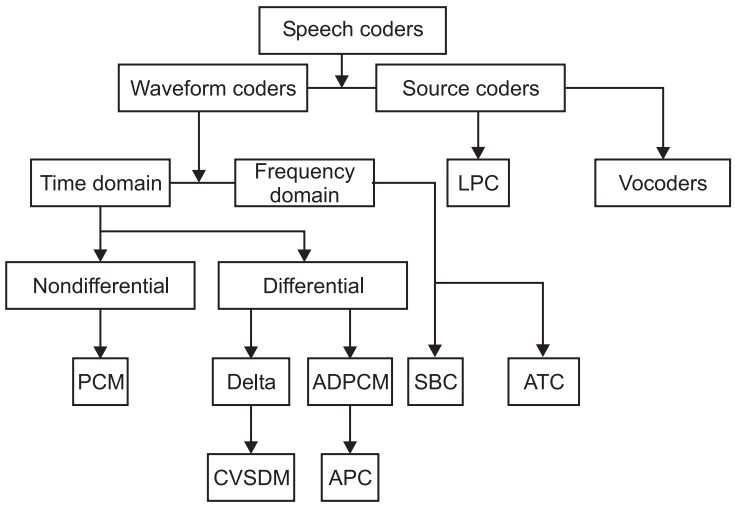


Fig. 6.1: Hierarchy of speech coders.

Here**:**

PCM → Pulse code modulation
CVSDM → Continuously variable slope delta modulation
ADPCM → Adaptive differential pulse code modulation
AP C → Adaptive predictive coding

Waveform coders are again subdivided into two parts– time domain coders and Frequency domain coders (SBC, ATC etc.).

The well known representation of speech signal using time domain waveform coding is the *A* low or *u*. low companded PCM at 64 Kbps.

In frequency domain waveform coders, the input signal undergoes a more or less accurate short time spectral analysis of signal.

The signal is divided into a number of frequency domain sub-bands. The individual sub-band signals are then encoded by using different number of bits to fulfil the quality requirements in a system.

(*ii*) **Source coders:** These are highly economical in transmission of bit rate. It is again subdivided into LPC and vocoders.

Vocoders are based on a prior knowledge of the signal that is required for coded. The main disadvantage of these are high complexity.

All these coders will be explained later on in this chapter.

6.2 CHARACTERISTICS OF SPEECH SIGNALS

Speech waveform have a number of useful properties that can be achieved when designing efficient coders. Some of the property that are most often utilized

in coder design include the non-uniform probability distribution of speech amplitude, the non-zero autocorrelation between successive speech samples, the non-flat nature of speech spectra, the presence of voiced and unvoiced segments in speech and the quasi periodicity of voiced speech signals. The most basic property of speech waveforms that is achieved by all speech coders is that they are band limited. A finite bandwidth means that it can be time-discretized (sampled) at a finite rate and reconstructed completely from its samples, provided that the sampling frequency is greater then twice the highest frequency component in the low pass signal means ($f_c > 2f_M$).

6.2.1 Probability Density Function (Pdf)

The non-uniform probability density function of speech amplitudes is perhaps the next most achieved property of speech. The Pdf of a speech signal is in general characterized by a very high probability of near zero amplitudes, a expressive probability of very high amplitudes and a monotonically decreasing function of amplitudes between these extremes. The accurate distribution, however, depends on the input bandwidth and recording conditions. The two-sided exponential (1aplacian) function given in equation (6.1) provides a good approximation to the long-term Pdf of telephone quality speech signals which can be written as:

$$P(x) = \frac{1}{\sqrt{2\sigma^x}} \exp\left(\frac{-\sqrt{2}\,|x|}{\sigma_x}\right) \qquad ...(6.1)$$

Note that this Pdf shows a distinct peak at zero, which is due to the existence of frequent pauses and low level speech segments. Short time Pdfs of speech segments are also single packed functions and are usually provided as a Gaussian distribution.

6.2.2. Autocorrelation Function (ACF)

Another very useful property of speech signals is that there exists much correlation between adjacent samples of a segment of speech. This signify that in every sample of speech, there is a large component that is easily predicted from the value of the previous samples with a small random error. All differential and predictive coding schemes are based on achieved this property. The autocorrelation function (ACF) gives a quantitative measure of the closeness or similarity between samples of a speech signals as a function of their time separation. This function is defined as the following equation:

$$c(k) = \frac{1}{N} \sum_{n=0}^{n-|k|-1} x(n)x(n+|k|) \qquad ...(6.2)$$

where $x(k)$ represent the kth speech sample.

The autocorrelation function is often normalized to the variance of the speech signal and hence is restrained to have value in the range {–1,1} with $c(1) = 1$ and for simple signal $c(l)$ as high as 0.85 to 0.9.

6.2.3 Power Spectral Density Function (PSD)

The PSD of speech makes it possible to obtain important compression by coding speech in the frequency domain.

Long term averaged PSDs, of speech show the high frequency components contribute very little to the total speech energy. This indicates that coding speech separately in different frequency bands, can guide to significant coding gain.

The maximum coding gain can be obtained by Spectral Flatness Measure (SFM). SFM is defined the ratio of the arithmetic to geometric mean of the samples of the PSD taken at uniform intervals in frequency and it is shown in Fig.

$$\text{SFM} = \frac{\left[\frac{1}{N}\sum_{k=1}^{N} S_k{}^2\right]}{\left[\prod_{k=1}^{N} S_k^2\right]^{\frac{1}{N}}} \qquad ...(6.3)$$

where S_k is the kth frequency sample of the PSD of the speech signal.

For long term SFM, speech signal have a value of 8 and for short term PSD speech signal have the value between 2 to 500. These are standard values.

6.3 QUANTIZATION TECHNIQUES

Quantization is a technique to break a continuous signal into many discrete signal level. It can be understood by Fig. 6.2 in which a quantized signal is shown.

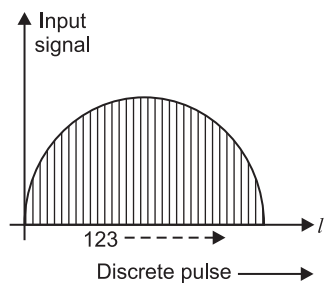


Fig. 6.2: A quantized signal.

As shown in Fig. 6.2 a continuous analog signal is divided into many discrete pulses at a fix interval in between. This process, is known as quantization. Each pulses (quantized signal) has a fixed amplitude which represent a fixed quantization level.

Quantization is divided into many techniques that utilizes the different characteristics of speech signals. Some quantization techniques are given below and described also:

(*i*) Uniform quantization (*ii*) Non-uniform quantization

(*iii*) Adaptive quantization (*iv*) Vector quantization

6.3.1 Uniform Quantization

In uniform quantization, the step size remain same for whole the input range. In uniform quantization, each signal is quantized uniformly and does not depend on whether the signal is strong or weak.

A quantizer that has n bits can have $M = 2^n$ discrete amplitude levels. The distortion produce by any quantization operation is directly proportional to the square of the step size, step size is inversely proportional to the number of levels for a given amplitude range.

The performance of a quantizer is measured in terms of output signal-to-quantization noise ratio (SQNR). The SQNR relation with the total number of bits used for encoding is equal to the

$$(\text{SQNR})_{\text{dB}} = 6.02n + \alpha \qquad ...(6.4)$$

where $\alpha = 4.77$ dB for peak SQNR

$\alpha = 0$ dB for average SQNR

A pulse code modulation (PCM) coder is basically a quantizer of sampled speech amplitudes. PCM coding, using 8 bits per sample at a sampling frequency of 8 kHz, was the first digital coding standard adopted for commercial telephony for coding purpose.

6.3.2 Non-Uniform Quantization

In non-uniform quantization, the step size varies according to the input signal values variations.

Non-uniform quantizers distribute the quantization levels in accordance with the Pdf of the input waveform. For an input signal with a Pdf *P(x),* the mean square distortion is given by the following equation:

$$D = E[\{x - f_Q(x)\}^2] = \int_{-\infty}^{\infty} [x - f_Q(x)]^2 P(x)\,dx \qquad ...(6.5)$$

where $f_Q(x)$ = output of quantizer

$[x - f_Q(x)]^2$ = quantization noise

The logarithmic quantizer is a simple and robust implementation of non-uniform quantizer. It is generally used in commercial telephony.

This quantizer divides the quantization steps into two parts–

1. Fine quantization steps are used for the frequently occuring low amplitudes in speech.
2. Coarser quantization steps are used for less frequent large amplitude excursions. For this, different companding schemes are used at different places of world *e.g.* μ-low companding scheme is used in US while *A*-low companding scheme is used in Europe.

6.3.3 Adaptive Quantization

The signals are generally non-stationary in nature or we can say time-varying. This leads to wide dynamic range. Therefore time varying quantization techniques are used to cover such a huge dynamic range.

In adaptive quantization, the step size of quantization level varies in accordance with the input speech signal power. Its characteristic shrink and expand in time like an accordion.

One simple adoption strategy would be to make the step size Δ_k, of the quantizer at any given sampling instant, proportional to the quantizer output f_Q at the preceding sampling instant.

Since the adoptation follows the quantizer output rather than the input step size information need not be explicitly transmitted, but can be recreated at the receiver.

6.3.4 Vector Quantization

Vector quantization is a delayed-decision coding technique which maps a group of input samples, called a vector, to a code book index.

A code book is made by a finite set of vectors covering the entire range of values. At each quantization interval, the code book is searched and the best match to the input signal frame is selected.

Vector quantization can yield better performance even when the samples are independent of one another. Performance is greatly enhanced if there is strong correlation between samples in the group.

The number of samples in a block (vector) is called the dimension L of the vector quantizer. The rate R of the vector quantizer is defined as

$$R = \text{bits/sample} \qquad ...(6.6)$$

where n is the size of vector quantization code book and L is the number of samples in a block.

Vector quantization is known to be most efficient at very low bit rate (R = 0.5 bits/sample or less).

6.4 VOCODERS

Vocoders are used specially to code the voice signals. These are based on certain properties of human speech production mechanism.

Vocoder are a class of speech coding systems that analyze the voice signal at the transmitter, transmit parameters derived from the analysis and than synthesize the voice at the receiver using those parameters.

A simple speech generation model shown in Fig. 6.3 that is the basis of all vocbding system.

Working Principle of Vocoders

In human speech, the air compressed by longs excites the vocal chords in two typical modes : Voiced and Unvoiced.

When generating unvoiced sounds (*f, s, sh* pronunciations), the vocal chords do not participate and the source acts like a noise generator. These are produced by turbulent air flow through a constriction. The excitation signal is then filtered through a vocal apparatus.

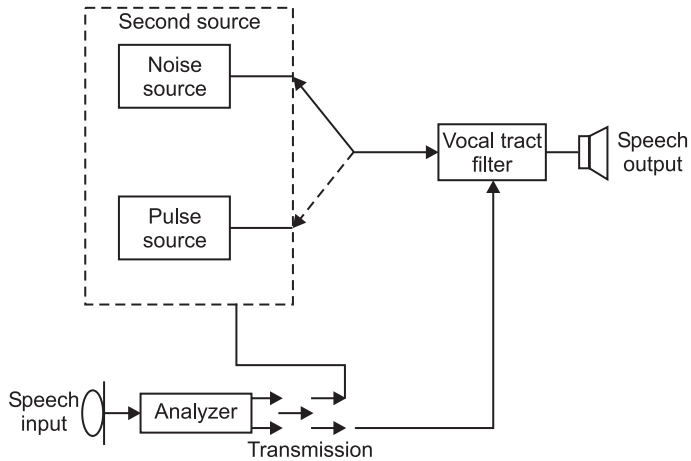


Fig. 6.3: Speech generation model.

When generating voiced sounds *(m, n, v* pronunciations) the vocal chords vibrate and generate quasiperiodic voice sound.

The parameters associated with the model shown in Fig. 6.3 are voice pitch, pole-frequencies of the filter and amplitude parameters.

The pole frequency correspond to the resonant frequencies of the vocal tract and often called the formants of the speech signal. The value of formants for adults centered arround 500 Hz, 1500 Hz, 2500 Hz and 3500 Hz.

Advantage of Vocoders

(*i*) Low bit rate for data transmission.

(*ii*) Can be classified into the frequency domain and time domain subclasses.

Disadvantage of Vocoders

(*i*) Much more complex for designing.

(*ii*) Less robust.

(*iii*) Relatively low synthetic speech quality.

Vocoders classification

Vocoders can be classified into four types or four sub-class–

(*i*) Channel vocoder

(*ii*) Formant vocoder

(*iii*) Cepstrum vocoder

(*iv*) Voice-excited vocoder

6.4.1 Channel Vocoder

Channel vocoder was the first among the analysis synthesis system of speech demonstrated practically.

Channel vocoders are frequency domain vocoders that determine the envelope of the speech signal for a number of frequency bands and then sample, encode and multiplex, these samples with the encoded outputs of the other filters.

The sampling is done synchronously every 10 ms to 30 ms.

Along with the energy information about each band, the voiced/unvoiced decision and the pitch frequency for voiced speech are also transmitted.

6.4.2 Formant Vocoders

The word formant indicates to the pole frequency which correspond to the resonant frequencies of the vocal tract. Theoretically, these vocoders operates at lower bit rate then the channel vocoder, because it uses fewer control signals.

Instead of sending samples of the power spectrum envelope, the formant vocoder attempts to transmit the positions of the peaks (formants) of the spectral envelope.

Typically a formant vocoder must be able to identify at least three formants for representing the speech sounds and must also control the intensities of the formants.

Formant vocoders can reproduce speech at bit rates lower then 1200 bits/sec.

6.4.3 Cepstrum Vocoder

The cepstrum vocoder separates the excitation and vocal tract spectrum by inverse Fourier transforming of the log magnitude spectrum to produce the cepstrum of the signal.

Linear filtering is used to filter the vocal tract cepstral coefficient from the excitation coefficient.

In the receiver side, the cepstral coefficients are Fourier transformed and then convolved with a excitation signal (random noise) to reproduce the original speech.

6.4.4 Voice-Excited Vocoder

Voice-excited vocoders eliminate the need for pitch extraction and voicing detection operations.

This system uses a hybrid combination of PCM transmission for the low frequency band of speech, combined with channel vocoding of higher frequency bands.

In a voice-excited vocoder, a pitch signal is generated at the synthesizer by rectifying, band pass filtering and clipping the signal.

Voice-excited vocoders have been designed for operation at 7200 bits/s to 9600 bits/s and its quality is superior to other traditional pitch excited vocoders.

6.5 LINEAR PREDICTIVE CODERS (LPC)

This class of coders is subdivided into four parts:

(*i*) LPC vocoders

(*ii*) Multipulse-excited LPC

(*iii*) Code-excited LPC

(*iv*) Residual-excited LPC

6.5.1 LPC Vocoders

It belong to the time domain class of vocoder.

These vocoders take the advantage of lower bit rate and are most popular among all vocoders with linear predict coding. It is possible to transmit good quality speech at 4.8 Kbps and poor quality speech at even lower rate.

Fig. 6.4 shows the block diagram of a LPC coding system.

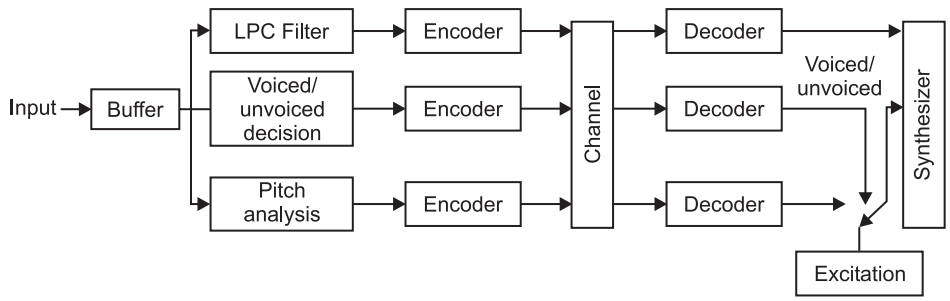


Fig. 6.4: Block diagram of a LPC coding system.

As show in Fig. 6.4, the LPC system transmit only selected characteristics of error signal. These characteristic are the gain factor, pitch information, and the voiced/unvoiced decision information. All these allow approximation of the correct error signal.

At the receiver, the information is decoded and given to a synthesizer filter. Synthesizer filter is designed using the received predictor coefficients.

The linear predictive coding system models are vocal tract as an all pole linear filter with a transfer function described by

$$H(z) = \frac{G}{1+\sum_{k=1}^{M} b_k Z^{-1}} \qquad ...(6.7)$$

where G is the gain of filter and Z^{-1} represents a unit delay operation.

6.5.2 Multipulse-Excited LPC

Multipulse excited LPC works on the principle that excitation should be done by more then one pulse rather than a single pulse per pitch period. Typically, eight pulses are used per pitch period. Therefore, the problem of finding the exact position of pulse gets solved, this reduces the available distortion. As a result better speech quality is achieved.

6.5.3 Code-Excited LPC

This type of code has its own code book of excitation signals and this code book carried by both, the transmitter and receiver.

The transmitter searches through its code book for the best match of the sound signal. This match code is then transmitted. The receiver uses this code to pick the correct excitation signal for its synthesizer filter.

Fig. 6.5 shows the block diagram of code excited LPC code book search.

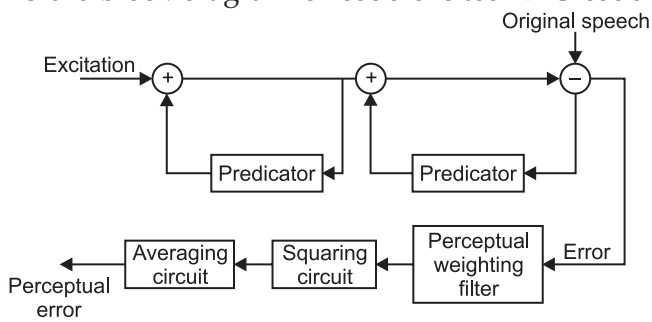


Fig. 6.5: Block diagram illustrating the CELP code book search.

As shown in Fig. 6.5, the samples or excitation are passed sequentially through two filters which introduce voice periodicity.

The generated speech samples at the output of second filter are compared with samples of speech signal (original).

The difference signal represents the objective error in the regenerated speech signal. This is further processed through a linear filter which amplify the perceptually more important frequencies and attenuates the perceptually less important frequency.

The code excited LPC coders are extremely complex and can require more than 500 million multiply and add operation per Second. They can provide high quality even when the excitation is coded at only 0.25 bits per sample.

These coders can achieve transmission bit rates as low as 4.8 Kbps.

6.5.4 Residual Excited LPC

In residual excited LPC, DPCM technique in waveform coding are use.

In this class of LPC coders, after estimating the model parameters (LP co-efficients or related parameters) and excitation parameters (voiced/unvoiced decision pitch gain) from a speech frame, the speech is synthesized at the transmitter and subtracted from the original speech signal to form a residual signal.

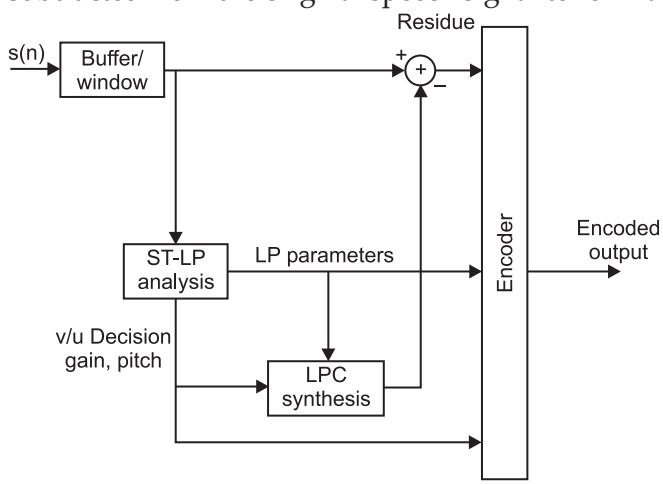


Fig. 6.6: Block diagram of a RELP encoder

Fig. 6.6 shows a block diagram of a simple RELP coder.

The residual signal is quantized, coded and transmitted to the receiver along with the LPC model parameters.

At the receiver, the residual error signal is added to the signal generated using the model parameters to synthesize an approximation of the original speech signal.

SOLVED EXAMPLES

Example 6.1. *Let the input signal to a quantizer have a probability density function (pdf) as shown in Fig. 6.7. Assume the quantization levels to be {1, 3, 5, 7}. Compute the mean square error distortion at the quantizer output and the output signal-to-distortion ratio. How would you change the distribution of quantization levels to decrease the distortion? For what input pdf would this quantizer be optimal?*

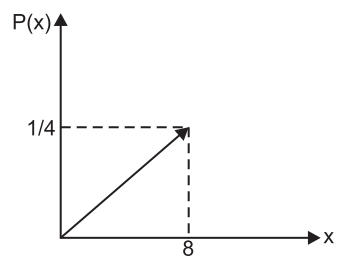


Fig. 6.7: Probability density function of the input signal.

Solution. From Fig. 6.7. The pdf of the input signal can be recognized as:

$$P(x) = \frac{x}{32} 0 \le x \le 8$$

$$P(x) = 0 \quad \text{elsewhere}$$

Given the quantization levels to be [1, 3, 5, 7], we can define the quantization boundaries as [0, 2, 4, 6, 8].

Mean square error distortion = $D = \int_{-\infty}^{\infty} [\{x - f_Q(x)\}^2 P(x) dx]$

$$D = \int_0^2 (x-1)^2 P(x)\, dx + \int_2^4 (x-3)^2 P(x)\, dx + \int_4^6 (x-5)^2 P(x) dx + \int_6^8 (x-7)^2 P(x) dx$$

This expression evaluates to 0.333

Signal power = $E[x^2] = \int_0^8 P(x) x^2 dx = \int_0^8 \frac{1}{32}(x.x^2) dx = 32$

$$\text{Signal to distortion ratio} = 10 \log\left[\frac{E(x^2)}{D}\right]$$

$$= 10 \log\left(\frac{32}{0.333}\right) = 19.82 \text{ dB}$$

To minimize the distortion, we need to concentrate the quantization levels in regions of higher probability. Since the input signal has a greater probability of higher amplitude levels then lower amplitudes we need to place the quantization levels closer *(i.e.,* more quantization levels) at amplitudes close to eight and further *(i.e.* less quantization levels) at amplitudes close to zero.

Since this quantizer has quantization levels uniformly distributed, this would be optimal for an input signal with a uniform *pdf* .

Example 6.2. *In an adaptive PCM system for speech coding, the input speech signal is sampled at 8 kHz, and each sample is represented by 8 bits. The quantizer step size is recomputed every 10 ms, and it is encoded for transmission using 5 bits. Compute the transmission bit rate of such a speech coder. What would be the average and peak SQNR of this system?*

Solution. Given:

Sampling frequency = f_s = 8 kHz

Number of bits per sample = h = 8 bits

Number of information bits per second = 8000 × 8 = 64000 bits.

Since the quantization step size is recomputed every 10 ms, we have 100 step samples to be transmitted every second.

Therefore, the number of overhead bits = 100 × 5 = 500 bits/s

Therefore, the effective transmission bit rate = 64000 + 500 = 64.5 kbps

The signal to quantization noise ratio depends only on the number of bits used to quantize the samples.

Peak signal to quantization noise ratio in dB

$$= 6.02\, n + 4.77 = (6.02 \times 8) + 4.77 = 52.93 \text{ dB}$$

Average signal-to-noise ratio in dB = 6.02 n = 48.16 dB.

Example 6.3. *Consider a sub-band coding scheme where the speech bandwidth is partitioned into four bands. The table below gives the corner frequencies of each band along with the number of bits used to encode each band. Assuming that no side information need be transmitted, compute the minimum encoding rate of this SBC coder.*

Sub-band Number	Frequency Band (Hz)	# of Encoding Bits
1	225 – 450	4
2	450 – 900	3
3	1000 – 1500	2
4	1800 – 2700	1

Solution: Given:

Number of sub-bands = N = 4

For perfect reconstruction of the band-pass signals, they need to be sampled at a Nyquist rate equal to twice the bandwidth of the signal. Therefore, the different sub-bands need to be sampled at the following rates:

Sub-band 1 = 2 × (450 – 225) = 450 samples/s

Sub-band 2 = 2 × (900 – 450) = 900 samples/s

Sub-band 3 = 2 × (1500 – 1000) = 1000 samples/s

Sub-band 4 = 2 × (2700 – 1800) = 1800 samples/s

Now, total encoding rate is

450 × 4 + 900 × 3 + 1000 × 2 + 1800 × 1 = 8300 bits/s = 8.3 Kbps.

Example 6.4. *The output of a speech coder has bits which contribute to signal quality with varying degree of importance. Encoding is done on blocks of samples of 20 ms duration (260 bits of coder output). The first 50 of the encoded speech bits (say type 1) in each block are considered to be the most significant and hence to protect them from channel errors are appended with 10 CRC bits and* 1 *convolutionally encoded with a rate* $\frac{1}{2}$ *FEC coder. The next 132 bits (say type 2) are appended with 5 CRC bits and the last 78 bits (say type 3) are not error protected. Compute the gross channel data rate achievable.*

Solution. Number of type 1 channel bits to be transmitted every 20 ms

(50 + 10) × 2 = 120 bits

Number of type 2 channel bits to be transmitted every 20 ms

132 + 5 = 137 bits

Number of type 3 channel bits to be encoded = 78 bits

Total number of channel bits to be transmitted every 20 ms

120 + 137 + 78 bits = 335 bits

Therefore, gross channel bit rate = $\frac{335}{(20\times10^{-3})} = 16.75\,\text{Kbps}$

HIGHLIGHTS

1. The performance of speech coder determine the quality of the recovered speech and the capacity of communication system so that the speech coders have great importance in communication system.
2. Basically the speech coders are classified into two broad categories–
   (*i*) Waveform coders
   (*ii*) Source coders or vocoders.
3. The waveform coders reproduce the time waveform of speech signals as closely as possible and the waveform coders are designed to be source independent hence they can provide equally code to a variety of signals, and the waveform coders have the advantage that they are robust for a large range of speech characteristics and for noisy environments.
4. The waveforms coders are also classified into the following broad categories:
   (*i*) Pulse code modulation (PCM)
   (*ii*) Differential pulse code modulation (DPCM)
   (*iii*) Adaptive differential pulse code modulation (ADPCM)

(*iv*) Continuously variable slope delta modulation (CVSDM)

(*v*) Adaptive predictive coding (APC).

5. The speech waveform coders have some different and important properties and the properties are:

(*i*) Non-uniform probability distribution of speech amplitude

(*ii*) The non-zero autocorrelation between successive speech samples

(*iii*) The non-flat nature of speech spectra

(*iv*) The existence of voiced and unvoiced segments in speech

(*v*) The quasiperiodicity of voiced speech signals.

6. The probability density function (*pdf*) of a speech signal is defined by a very high probability of near zero amplitudes, the important probability of very high amplitudes and a monotonically decreasing function of amplitudes between these extremes.

7. The autocorrelation function (ACF) define a quantitative measure of the similarity between samples of a speech signal as a function of their time seperation.

8. The power spectral density function (PSD) define a non-flat characteristic which is used to obtain significant compression by coding speech in frequency domain.

9. The vocoders are used to code the signals and these are based on certain properties of human speech production mechanism. The signals may be voiced and unvoiced.

10. The vocoders are classified into following four types:

(*i*) Channel vocoder

(*ii*) Formant vocoder

(*iii*) Cepstrum vocoder

(*iv*) Voice-excited vocoder.

11. The channel vocoder are used to analyze the speech signal and these are based on frequency domain coding but the formant vocoder indicate the pole frequency which correspond to the resonant frequency of the vocal tract and these are operated at lower bit rate than the channel vocoder.

12. The voice excited vocoder quality is superior to other traditional pitch excited vocoders. Here a pitch signal is generated at the synthesizer by rectifying, bandpass filtering and clipping the signal and the operated bit rate are 7200 bits/sec to 9600 bits/sec.

13. The linear predictive coder (LPC) related to the time domain class of coders and the LPC divided into four parts–

(*i*) LPC vocoders

(*ii*) Multipulse-excited LPC

(*iii*) Code excited LPC

(*iv*) Residual excited LPC

14. The LPC system only transmit selected characteristics of error signal, these characteristics are the gain factor, pitch information and the voiced/unvoiced decision information and these system operate at lower bit rate.

15. The multipulse excited LPC works on the principle that the excitation should be done by more than one pulse rather than a single pulse per pitch period and also reduce the audible distortion.

16. The code-excited LPC consist his own code book of excitation signals and this code book is carried by both transmitter and receiver, from this book, the transmitter search the best match of the sound signal and the receiver pick the correct excitation signal from its synthesizer filter.

17. In the residual excited LPC, the residual signal is being transmitted (inspite of the transmitting the original speech and the residual) signal is generated by subtracting the synthesized speech at the transmitter to original speech.

18. The quantization is divided into four main type of techniques that utilizes the different characteristics of speech signals, and these are –

(*i*) Uniform quantization

(*ii*) Non-uniform quantization

(*iii*) Adaptive quantization

(*iv*) Vector quantization.

19. In uniform quantization, each signal is quantized uniformly means the step size remain same throughout the input range and it does not depend on whether the signal is strong or weak, and the performance is measure in terms of output signal to quantization noise ratio (SQNR).

20. The non-uniform quantizer define the quantization level in accordance with the PDF of the input waveform means the step size is varies according to the input signal values and quantizer is generally used in commercial telephony.

21. In the adaptive quantizer, the speech signals are generally non-stationary in nature or time-varying. This quantizer cover a wide dynamic range and the step size is varies accordance with the input speech signal power.

22. The vector quantization is a delayed decision coding technique, it use a code book index, which contain a group of input samples or speech frames and the code book is also contain a finite set of vector covering the entire range of values.

23. The pulse code modulation, differential pulse code modulation, adaptive differential pulse code modulation are not the modulation technique but these are the conversion technique, which convert the analog signal into the digital signal. Delta modulation is the another technique to convert the analog to digital signal.

24. For compress the data, there are two type of compressor technique are used–

(*i*) μ-low companding

(*ii*) A-low companding.

25. The cepstrum vocoder is used to seperated the excitation of the signal and the vocal tract spectrum by using inverse Fourier transform.

PROBLEMS

6.1. For an 8 bit uniform quantizer that spans the range (–1 V, 1V), determine the step size of the quantizer. Compute the SNR due to quantization if the signal is a sinusoid that spans the entire range of the quantizer.

6.2. Derive a general expression that relates the signal-to-noise ratio due to quantization as a function of the number of bits.

6.3. For a *u-law compander* with $\mu = 255$, plot the magnitude of the output voltage as a function of the magnitude of the input voltage. If an input voltage of

0.1 V is applied to the compander, what is the resulting output voltage? If an input voltage of 0.01 V is applied to the input determine the resulting output voltage. Assume the compander has a maximum input of 1V.

6.4. For an *A-law compender* with $A = 90$, plot the magnitude of the output voltage as a function of the magnitude of the input voltage. If an input voltage of 0.1V is applied to the compander. What is the resulting output voltage ? If an input voltage of 0.01V is applied to the input. Determine the resulting output voltage. Assume the compander has a maximum input of 1V.

6.5. A compander relies on speech compression and speech expansion (decompression) at the receiver in order to restore signals to their correct relative values. The expander has an inverse characteristic when compared to the compressor. Determine the proper compressor characteristics for the speech compressors in problem 3 and 4.

6.6. A speech signal has an amplitude? That can be characterized as a zero-mean Gaussian process with a standard deviation of 0.5V. For such a speech signal, determine the mean square error distortion at the output of a 4-bit quantizer if quantization levels are uniformly spaced by 0.25V. Design a non-uniform quantizer that would minimize the mean square error distortion and determine the level of distortion.

6.7. Consider a sub-band speech coder that allocates 5 bits for the audio spectrum between 225 Hz and 500 Hz, 3 bits for 500 Hz to 1200 Hz and 2 bits for frequencies between 1300 Hz and 3 kHz. Assume that the sub-band coder output is then 3 applied to a rate $\frac{3}{4}$ convolutional coder. Determine the data rate out of the channel coder.

6.8. List four significant factors which influence the choice of speech coders in mobile communication system. Elaborate on the trade offs which are caused by each factor. Rank order the factors based on your personal view point, and defend your position.

REVIEW QUESTIONS

6.1. What do you understand by speech coding? How speech coders are beneficial in communication system?

6.2. Give a brief classification of speech coders. Explain the difference between waveform and source coders.

6.3. What are different characteristic of speech signals? How they are used in designing of coders?

6.4. What do you mean by transmission bit rate and quality in coders?

6.5. What are the basic steps that are used in quantization and coding of the signal?

6.6. Explain different quantization techniques in detail.

6.7. Draw a simple model of speech generation and explain it?

6.8. What is the basic principle of vocoders? Write down advantages and disadvantages of vocoders.

6.9. Explain different type of vocoders in brief.

6.10. Explain code excited LPC in detail.

6.11. Derive a general expression that relates the signal to noise ratio due to quantization as a function of the number of bits.

7 Chapter

MULTIPLE ACCESS TECHNIQUES FOR WIRELESS COMMUNICATION

7.1. *Introduction* **7.2.** *Introduction to Multiple Access* **7.3.** *Frequency Division Multiple Access (FDMA)* **7.4.** *Time Division Multiple Access (TDMA)* **7.5.** *Space Division Multiple Access (SDMA)* **7.6.** *Comparison among Different Multiple Access Technology* **7.7.** *ALOHA* **7.8.** *Code Division Multiple Access (CDMA)* **7.9.** *Capacity of Cellular System*

7.1 INTRODUCTION

In wireless communication systems, it is desirable to allow the subscriber to send at the same time information to the base station while receiving information from the base station. For example, in conventional telephone systems, it is possible to talk and listen at the same time, and this effect is called duplexing.

Duplexing may be done using frequency or time domain techniques.

Frequency division duplexing (FDD) provide two distinct band of frequency for every user. The forward band allow traffic from the base station to the mobile, and the reverse band allow traffic from mobile to the base station.

In FDD any duplex channel consist of two simplex channel (a forward and reverse) and a device called a duplexer is used inside each subscriber unit and base station to provide, at the same time bidirectional radio transmission and reception for both the subscriber unit and the base station on the duplex channel pair.

Time division duplexing (TDD) uses time instead of frequency to provide both a forward and reverse link for transmission.

In TDD, multiple users share a single radio channel by taking turns in the time domain. They are allowed to access the channel in alloted time slots and each duplex channel has both a forward time slot and a reverse time slot to facilitate bidirectional communication.

Fig. 7.1 show the FDD and TDD techniques with simplex channel.

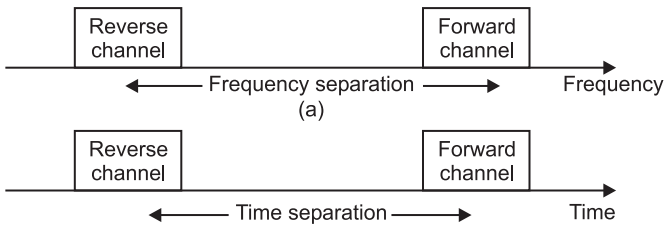


Fig. 7.1: (*a*) FDD provides two simplex channels at the same time, (*b*) TDD provides two simplex time slots on the same frequency.

7.2. INTRODUCTION TO MULTIPLE ACCESS

Frequency division multiple access (FDMA), time division multiple access (TDMA) and code division multiple access (CDMA) are the three major access techniques used to share the available bandwidth in a wireless communication system. These techniques may be grouped as narrowband and wideband systems, depending upon how the available bandwidth is allocated to the users.

7.2.1. Narrowband System

In a narrowband multiple access system, the available radio spectrum is splited into a large number of narrowband channels. The channels are usually operated using FDD technique.

In narrowband FDMA, a user is assigned a particular channel which is not shared by other users and if FDD is used (that is each duplex channel has a reverse and forward simplex channel), then the system is called FDMA/FDD system.

Narrowband TDMA, support to users to share the same radio channel but proivde a unique time slot to each user in a channel, thus separating a small number of users in time on a single channel.

For narrowband TDMA system, there are a large number of radio channels provide using either FDD or TDD and each channel is shared using TDMA, such system are called TDMA/FDD or TDMA/TDD access systems.

7.2.2. Wideband systems

In wideband systems, the transmission bandwidth of a single channel is much larger then the coherence bandwidth of the channel.

In wideband multiple access systems a large number of transmitter are support to transmit on the same channel.

TDMA allocates time slot to the many transmitters on the same channel and allows only one transmitter to access the channel at any instant of time, whereas spread spectrum CDMA support to all the transmitter to access the channel at the same time.

TDMA and CDMA systems may use either FDD or TDD multiplexing techniques for transmission.

7.3. FREQUENCY DIVISION MULTIPLE ACCESS (FDMA)

Frequency division multiple access (FDMA) assigns individual channels to individual users. As shown in Fig. 7.2, that each user is allocated a unique frequency band or channel. These channels are allocated on demand to users who request service. During the period of call no other user can share the same channel.

In FDD systems, the users are alloted a channel as a pair of frequency, one frequency is used for forward channel, with the other frequency is used for the reverse channel.

The features of FDMA are explain as follows:

(*i*) The FDMA channels carries only one phone circuit at a time.

(*ii*) If an FDMA channel is not in use, then it sits idle and can not be used by other users to increase or share capacity. It is wasted resource of the system.

Table 7.1: Multiple Access Techniques Used in Different Wireless Communication Systems.

Cellular system	Multiple access technique
Advanced mobile phone system (AMPS)	FDMA/FDD
Global system for mobile (GSM)	TDMA/FDD
US digital cellular (USDC)	TDMA/FDD
Pacific digital cellular (PDC)	TDMA /FDD
CT2 (cordless telephone)	FDMA/TDD
Digital European cordless telephone (DECT)	FDMA/TDD
US narrowband spread spectrum (IS–95)	CDMA/FDD
W-CDMA (3GPP)	CDMA/FDD CDMA/TDD
cdma2000 (3GPP2)	CDMA/FDD CDMA/TDD

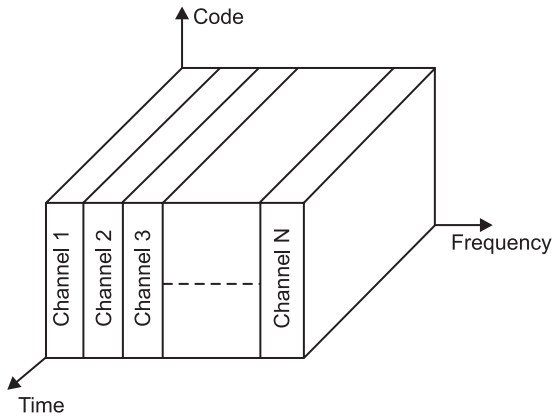


Fig. 7.2: FDMA where different channels are assigned different frequency bands.

(*iii*) After the assignment the voice channel, the base station and mobile transmit at the same time and continuously.

(*iv*) The bandwidth of FDMA channels are retatively narrow (30 kHz) as each channel allow only one circuit per carrier.

(*v*) The symbol time of a narrowband signal is large as compared to the average delay spread. This signify that the amount of inter symbol interference is low and thus little or no equalization is required in FDMA narrowband systems.

(*vi*) The complexity of FDMA mobile systems is lower when compared to TDMA system so FDMA is a simple technique.

(v*ii*) Since FDMA is a continuous transmission scheme, fewer bits are needed for overhead purposes (such as synchronization and framing bits) as compared to TDMA.

(*viii*) FDMA system have higher cell site system costs as compared to TDMA system means FDMA system is a cheap system.

(*ix*) The FDMA mobile unit uses duplexers since both transmitter and receiver operates at the same time this result in an increase in the cost FDMA subscriber units and base stations.

(*x*) FDMA requires tight RF filtering to minimize neighbouring channel interference.

7.3.1 Non-Linear Effect in FDMA

In a FDMA system, many channel share the same antenna at the base station. The power amplifiers or the power combiners, when operated at or near saturation for maximum power efficiency are non-linear. The non-linearities produces signal spreading in the frequency domain and generates intermodulation (IM) frequency. IM is undesired RF radiation which can interfere with other channels in the FDMA systems. Spreading of the spectrum results in bordering channel interference. Intermodulation is the generation of undesirable harmonics. Harmonics generated outsize the mobile radio band cause interference to adjacent services, while those present inside the band produce interference to other users in the wireless system.

7.3.2 Number of Channel in FDMA

FDMA decides a limit on the maximum numbers of channels that can be simultaneously supported. If Ware the maximum number of channels then the number of channels in FDMA can be express as –

$$N = \frac{B_t - 2B_g}{B_c} \qquad ...(7.1)$$

where,

B_t = Complete spectrum

B_g = Guard band at the edge of spectrum

B_c = Channel bandwidth

7.4 TIME DIVISION MULTIPLE ACCESS (TDMA)

TDMA systems divide the radio spectrum into time slots, and in each slot only one user is allowed to either transmit or receive, it can be seen from Fig. 7.3 that each user hold a cyclically repeating time slot so a channel may be thought of as a particular time slot reoccurs every frame, where *N*-time slots comprise a frame as shown in Fig. 7.3.

TDMA systems transmit data in a buffer and burst method, thus the transmission for any user is non-continuous. This implies that, unlike a FDMA systems which correspond to analog FM digital data and digital modulation must be used with TDMA system.

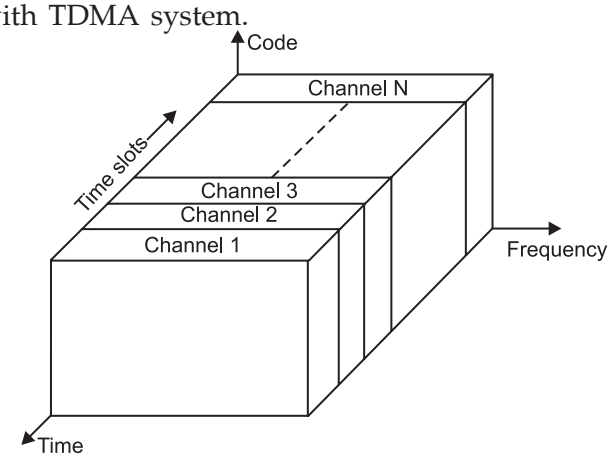


Fig. 7.3 : TDMA scheme where each channel occupies a cyclically repeating time slot.

The transmission from various user is interlaced into a repeating frame structure as shown in Fig. 7.4.

From Fig. it can be seen that a frame consists of a number of slots, each frame is made up of a preamble, an information message and the tail bits with sync bit and guard bits.

In TDMA/TDD, half of the time slots in the frame information message would be used for the forward link channels and half would be for reverse link channels.

In TDMA/FDD systems an similar frame structure would be used single for either forward or reverse transmission, but the carrier frequency would be distinct for forward and reverse links.

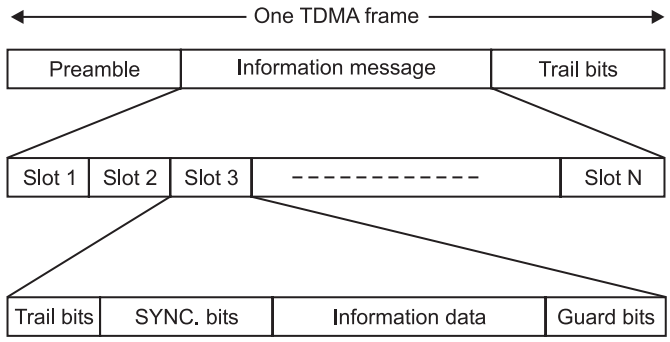


Fig. 7.4: TDMA frame structure. The frame is cyclically repeated over time.

In a TDMA frame, the preamble contain the address and synchronization information that both the base station and subscriber use to identify each other guard times are utilized to support synchronization of the receivers between different slots and frames.

The features of TDMA include the following important things:

(*i*) TDMA shares a single carrier frequency with several users, where each user makes use of non-overlapping time slots. The number of time slots per frames depends on various factors such as modulation technique, bandwidth etc.

(*ii*) Data transmission for users of a TDMA is not continuous but produce in bursts. This result low battery consumption.

(*iii*) Because of discontinuous transmission in TDMA, the handoff process is much simpler for a subscriber unit.

(*iv*) TDMA uses distinct time slots for transmission and reception, thus duplexers are not required. Even if FDD is used, a switch rather than a duplexer inside the subscriber unit is all that is demanded to switch between transmitter and receiver using TDMA.

(*v*) Adaptive equalization is usually necessary in TDMA system, since the transmission rates are generally very high as compared to FDMA channels.

(*vi*) In TDMA, the guard time should be minimized or minimum.

(*vii*) High synchronization overhead is required in TDMA systems because of burst transmission.

(*viii*) TDMA has an advantage in that it is possible to allocate different number of time slots per frame to distinct users. Thus bandwidth can be supplied on demand to different users by concatenating or reassigning time slot based on priority.

7.4.1 Efficiency of TDMA

The efficiency of a TDMA system is a measure of the percentage of transmitted data that contains information as opposed to providing overhead for the access scheme. The frame efficiency n_f is the percentage of bits per frame which contain transmitted data.

The number of overhead bits per frame is expressed as –

$$b_{0H} = N_r b_r + N_t b_p + N_t b_g + N_r b_g \qquad ...(7.2)$$

where,

N_r = Number of reference bursts per frame

N_t = Number of traffic bursts per frame

b_r = Number of overhead bits per reference burst

b_p = Number of overhead bits per preamble in each slot

b_g = Number of equivalent bits in each guard time interval

The total number of bits per frame b_T is equal to

$$b_T = T_f R \quad ...(7.3)$$

where,

T_f = frame duration

R = channel bit rate

The frame efficiency $\eta_f = 1 - \frac{b_{0H}}{B_T} \times 100\%$ ...(7.4)

7.4.2 Number of Channel in TDMA

Maximum number of channels slots in TDMA systems are found by multiplying the number of time slot per channel by the number of channel available than the number of channel in TDMA are

$$N = \frac{M(B_t - 2B_g)}{B_c} \quad ...(7.5)$$

where,

M = Maximum number of TDMA slots (or user)

B_t = Total bandwidth

B_g = Guard bands (one at lower end and other at upper band)

B_c = Channel bandwidth

7.5 SPACE DIVISION MULTIPLE ACCESS (SDMA)

Space division multiple access (SDMA) control the radiated energy for each user in space.

It can be seen from Fig. 7.5 that SDMA serves different users by using spot beam antennas. These distinct area covered by the antenna beam may be served by the same frequency (in a TDMA or FDMA system) or different frequency (in an FDMA system).

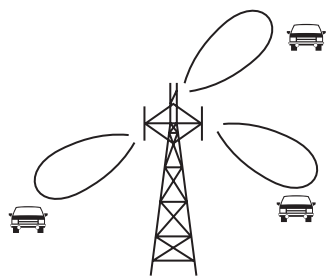


Fig. 7.5: A spatially filtered base station antenna serving different users by using spot beams.

Sectorized antennas may be thought of as a original application of SDMA. In future, adaptive antenna will likely be used to simultaneously steer energy in the direction of many users at once and appear to be best suited for TDMA and CDMA base station architectures.

The reverse link presents the most difficulty in cellular systems for several reasons.

First, the base station has complete control over the power of all the transmitted signals on the forward link. However because of different radio propagation paths between each user and the base station, the transmitted power from each subscriber unit must be dynamically controlled to prevent any single user from driving up the interference level for all other users. Second transmit power is limited by battery consumption at the subscriber unit therefore there are limit on the degree to which power may be controlled on the reverse link.

If the base station antenna is made to spatially filter each desired user so that more energy is detected from each subscriber, then the reverse link for each user is improved and less power is required.

To simply understand the SDMA technology, let us consentrate on the example of car driving on a highway. At highway, each driver had his/her lane at which he/she is moving on. Driver just can't enter into the other's lane. But this may be a waste of space also because at a time suppose lane 1 is empty and lane 2 had very high traffic even then the traffic at lane 2 can't use lane 1. Fig. 7.5 show the SDMA technology.

The basis of SDMA technology is formed by cells and sectorized antennas which constitute the infrastructure for SDMA technology.

The sectorized antennas cover its area using spot beams. These different areas covered by the antennas beam may be served by the same frequency (in TDMA or CDMA) or different frequency (in FDMA systems).

The main advantage of this technology is that all users in the system using SDMA technology would be able to communicate at the same time using the same channel.

7.6. COMPARISON AMONG DIFFERENT MULTIPLE ACCESS TECHNOLOGY

Table 7.2: Shows a Sharp Comparison Among Different Multiple Access Technologies

S.No	Parameter	FDMA	TDMA	CDMA	SDMA
1.	Scaling	Frequency scale is shared by different users	Time scale is shared by different users	Code scale is shared by different users	Space scale is shared by different users.
2.	Segmentation	Segment frequency band into disjoint sub-bands	Segment time into disjoint time-slots	Spread the spectrum using orthogonal codes	Segment space into cells/sectors
3.	Terminals	Every terminal uses its own frequency uninterruptedly	All terminals are active for short periods of time on the same frequency	All terminals can be active uniterrupted at the same time	Only one terminal can be active in one cell or one sector

4.	Signal separation	Filering in the frequency domain	Synchronization in the time domain	Code with special receivers	Call structure directed antennas
5.	Advantages	Simple, established, robust	Very flexible fully digital	Soft hand-off flexible, highly secure	Very simple capacity increases per km^2
6.	Disadvantages	Inflexible, less efficient, more adjacent channel interference	Guard space needed, synchronization difficult	Complex receivers, near-far problem, complicated power control	Inflexible, antennas typically fixed
7.	Systems using technology	AMPS, DECT, TACS etc.	GSM, ISD-54, IS-136 etc.	Used in many 3G systems	Used in combination with TDMA, FDMA or CDMA

7.7. ALOHA

ALOHA is a media used to transmit the data from one station to another. ALOHA is classified into two forms:

(*i*) Pure ALOHA (*ii*) Slotted ALOHA

7.7.1 Pure ALOHA

Pure ALOHA has no restriction or ALOHA with no restriction is called pure ALOHA.

There is more possibility of collision means only 18.4% data is transmit.

When a station or terminal start transmission, anyone of the following three things may happen :

(*i*) There may be an ongoing transmission already in the channel and new transmission may collide with the existing one, thereby affecting both transmission.

(*ii*) The channel is free and the transmission goes on successfully.

(*iii*) The channel is free and to start with, but another station gets ready and start transmission, when this transmission is in progress. The transmission collide with each other and no useful data transfer take place.

These event are shown in Fig. 7.6. The end part of frame *X* is corrupted with *A* means part of frame *A* collide with an ongoing transmission X and the part of frame *B* is corrupted by transmission *C* with the results both the frame are not transmitted successfully. Frame *D* is successfully transmitted and frame *E* and *F* collide again.

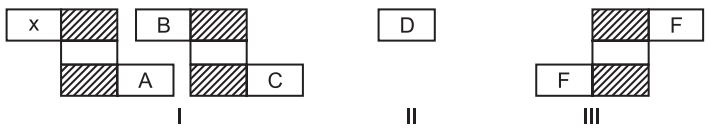


Fig. 7.6: Collision in pure ALOHA access.

The probability of the K frame are generated in time t given by the possion arrival –

$$P_k(t) = \frac{G^K e^{-G}}{\lfloor K} \quad ...(7.6)$$

where,

$P_k(t)$ = Probability that there are K arrival in time t

G = Load offered to the system per frame time

= λt, t is frame time and λ is arrival rate

Consider a frame F being placed on the channel at any time, the frame will be transmitted successfully only if two conditions –

(*i*) There is no ongoing transmission at the time of placement of this frame.

(*ii*) No other station ready for transmission during the transmission of this frame.

Both condition are shown in Fig. 7.7, in which the time duration $2T$ is called vulnerability period.

The probability of zero arrival during vulnerability period from equation (7.6).

$$P_0(2\tau) = \frac{G^0 e^{-2G}}{\lfloor 0} = e^{-2G} \quad ...(7.7)$$

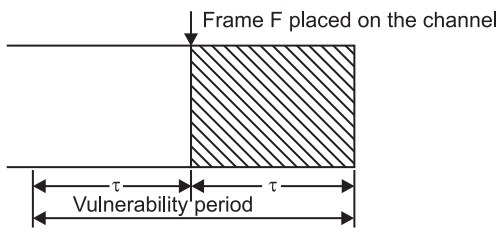


Fig. 7.7: Vulnerability period in pure ALOHA.

Equation (7.7) show that the probability of successful transmission of one packet.

Now if there are G packet are transmit so throughput of the system.

$$S = Ge^{-2G} \quad ...(7.8)$$

From equation (7.8), the maximum throughput occur at G = 0.5, so the maximum throughput in pure ALOHA is

$$S_{max} = \frac{1}{2e} = 18.4\% \quad ...(7.9)$$

So we can say that the only 18.4% data will be transmit through channel.

7.7.2 Slotted ALOHA

The throughput of the pure ALOHA scheme can be doubled by the introducing time slot in the channel and permitting a new transmission to start only at the beginning of a time slot such a scheme is known as slotted ALOHA.

Every station divided the total time or particular time into different slots

and there is small collision and in it only 37% data is transmitted.

A time slot equal to the frame time τ. The vulnerability period for a slotted ALOHA is shown in Fig. 7.8 which is τ.

A frame *F* is placed on the channel at the begining of the time slot period 2-3.

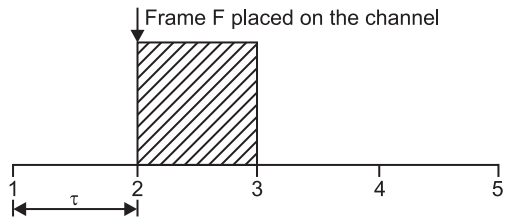


Fig. 7.8: Vulnerability period in slotted ALOHA

The frame will be transmitted successfully only if no other transmission commences. During the begining of same slot and another transmission can exist only if only other station get ready during the slot period 1–2.

So the probability of zero arrival in the time period *x*.

$$P_0(\tau) = e^{-G} \quad ...(7.10)$$

For *G* packet transmission, the throughput *S* is

$$S = Ge^{-G} \quad ...(7.11)$$

The maximum throughput will be obtain when *G* = 1 to be consider than

$$S_{max} = e^{-1} = \frac{1}{e} = 36.0 \approx 37\%$$

So in slotted ALOHA only 37% data will be transmit through channel.

7.8 CODE DIVISION MULTIPLE ACCESS (CDMA)

The code division multiple access system is a system in which the narrowband message signal is multiplied with a large bandwidth signal which is called the spreading signal.

The spreading signal is a pseudonoise code sequence which has a chip rate whose magnitude is greater than data rate of message.

A CDMA system is shown in 7.9 in which all users use the same carrier frequency and transmit the data at the same time.

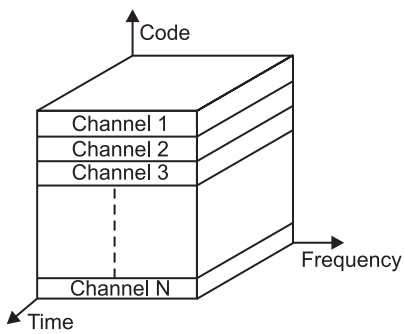


Fig. 7.9: Spread spectrum multiple access in which each channel is assigned a unique PN code which is

In CDMA, each users has its own pseudorandom codeword which is orthogonally to all other codewords.

The receiver support a time correlation operation to detect the original codeword.

Some time the original signal is received with noise so for receive the original signal without noise, receiver needs to know the codeword by the transmitter.

Each user operates independently with no knowledge of other users.

Some important feature of CDMA are explain in following points:

(*i*) Most of the CDMA users use the same frequency. Either FDD or TDD.

(*ii*) Unlike the FDMA or TDMA, the CDMA consist a soft capacity limit. If the number of users in a CDMA system increase then it increase the noise floor in a linear manner hence there is no absolute limit on the number of users in CDMA.

(*iii*) In a CDMA system, multipath fading is reduce because the signal is spread over a large spectrum.

(*iv*) A CDMA system use the very high channel data rate.

(*v*) The CDMA systems uses the co-channels cells and this can also use the macroscopic spatial diversity to provide soft hand-off. The MSC is used to provide the soft hand-off, which can monitor a particular user from two or more base station at the same time.

(*vi*) The CDMA system is suffer by a problem called self jamming. The self jamming occurs from that the spreading sequence of different users are not exactly orthogonal.

(*vii*) The another problem in CDMA is near-far problem which is produce when if an undesired user has a high detected power as compared to the desired user.

7.9. CAPACITY OF CELLULAR SYSTEM

For a radio system, the channel capacity defined the number of channels or users which are provided in a fixed frequency band.

For measure the spectrum efficiency of a wireless system, a parameter named radio capacity is used. This parameter is determine by the channel bandwidth B_c and

carrier-to-interference ratio $\frac{c}{I}$.

There are two type of channel interference possible or occurs between the base station receiver and subscriber –

(*i*) Reverse channel interference

(*ii*) Forward channel interference.

Let us consider the forward channel interference case, let R is the cell radius and D is the distance between co-channel cells, then the co-channel reuse ratio define the relation between D and R, and is given by

$$Q = \frac{D}{R} \qquad ...(7.12)$$

From Fig. (7.10) the carrier to interference rate $\frac{c}{I}$ can be written as

$$\frac{c}{I} = \frac{D_0^{-n_0}}{\sum_{k=1}^{M} D_k^{-n_k}} \qquad ...(7.13)$$

where n_0 is the path loss exponent, D_0 is the distance from desired base station to the mobile, D_k is kth cell distance from mobile and n_k is the path loss exponent to the kth interfering base station.

If only the six closest interfering cells are consider, then the $\frac{c}{I}$ can be written as –

$$\frac{c}{I} = \frac{1}{6} \cdot \frac{D_0^{-n}}{D^{-n}} \qquad ...(7.14)$$

Equation number (7.14) can be written as–

$$\frac{1}{6}\left(\frac{R}{D}\right)^{-n} \geqslant \left(\frac{c}{I}\right)_{\min} \qquad ...(7.15)$$

From equation (7.12), the co-channel reuse factor can be given as

$$Q = \left[6\left(\frac{c}{I}\right)_{\min}\right]^{\frac{1}{n}} \qquad ...(7.16)$$

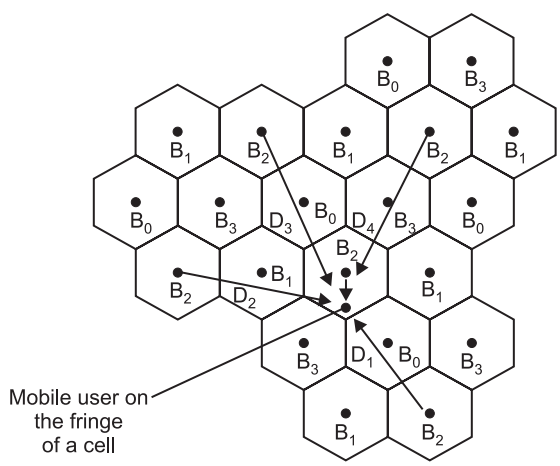


Fig. 7.10: Illustration of forward channel interference.

For a cellular system, the radio capacity is defined as –

where M is radio capacity matrix, B_t is the total allocated specify of the system, B_c is the channel bandwidth and N is the number of cell than the co-channel reuse factor Q.

$$Q = \sqrt{3N} \qquad ...(7.18)$$

Than from equation 7.16, 7.17, 7.18, the radio capacity can be written as

$$M = \frac{B_t}{B_c \frac{Q^2}{3}} = \frac{B_1}{B_c \left[3^{\frac{6}{n}} \left(\frac{c}{I} \right)_{\min} \right]^{\frac{2}{n}}} \qquad ...(7.19)$$

As shown by Lee, when $n = 4$, the radio capacity is given by

$$M = \frac{Bt}{B_c \sqrt{\frac{2}{3}\left(\frac{C}{I}\right)}} \text{ radio channels/cell} \qquad ...(7.20)$$

The value of $\left(\frac{C}{I}\right)_{\min}$ may be lower in digital system when compared to an analog system.

SOLVED EXAMPLES

Examples 7.1: *In the AMPS system, the system bandwidth is* 12.5 *MHz, the channel spacing is* 30 *kHz and the edge guard spacing is* 10 *kHz. The number of channels allocated for control signaling is* 21. *Find :*

(*a*) *The number of channels available for message transmission, and*

(*b*) *The spectral efficiency of FDMA.*

Solution: We have

$$B_s = 12.5 \text{ MHz}$$
$$B_c = 30 \text{ kHz}$$
$$B_g = 10 \text{ kHz}$$

(*a*) We know that the number of available channel is given by

$$N_s = \frac{B_s - 2B_g}{B_c} = \frac{12.5 \times 1000 - 20}{30} = 416 \text{ channels}$$

(*b*) We know that the spectral efficiency of this FDMA system is given by following expression:

$$\eta_{\text{FDMA}} = \frac{30 \times (416 - 21)}{12.5 \times 1000} = 0.948$$

Examples 7.2*: Given a cellular system in which the one-way bandwidth of the system is* 12.5 *MHz, the channel spacing is* 30 *kHz, and the guard band at each boundary of the spectrum is* 10 *kHz. If*

(*i*) *The cell area is* 6 *km*

(*ii*) *The frequency reuse factor (cluster size) is* 7 *and*

(*iii*) 21 *of the available channels are used to handle control signalling, calculate*

(*a*) *The total number of available channels per cluster*

(*b*) *The number of available data channels per cluster*

(*c*) *The number of available data channels per cell*

(*d*) *The system spectral efficiency in units of channels/MHz/km.*

Solution: We allocate all of the available frequencies to one cluster and these frequencies, or channels, are distributed evently among the N cells in the cluster.

(*a*) The total number of available channels in the cluster is

$$N_{\text{ch/cluster}} = \frac{B_s - 2B_g}{B_c} = \frac{12.5 - 2\times 0.01}{0.03} = 416$$

(*b*) The number of available data channels per cluster is

$$N_{\text{data/cluster}} = N_{\text{ch/cluster}} - N_{\text{ctl/cluster}}$$
$$= 416\text{-}2 = 395$$

(*c*) The number of availabe data channels per cell is given by

$$N_{\text{data/cell}} = \frac{N_{\text{data/cluster}}}{N} = \frac{395}{7} = 56$$

(*d*) The overall spectral efficiency of system is

$$\eta = \frac{N_{\text{data/cell}}}{B_s \times A_{cell}} = \frac{56}{12.5\times 6} = 0.746 \text{ channels/MHz/km}^2$$

Examples 7.3*: Given that the system parameter values are the same as those in Example 7.2. In addition, there are the following other specifications:*

The area of the entire cellular system ***3024*** *km.*

The average number of cells per user during a busy hour is 1.5.

The average channel holding time of a call is 180 seconds.

The trunk efficiency, n_t, is 0.95.

Calculate the following parameter values:

(*a*) *The number of cells in the system*

(*b*) *The number of cells per hour per cell*

(*c*) *The average number of users per hour per cell*

(*d*) *The system spectral efficiency in Erlangs/MHz/km^2.*

Solution:

(*a*) The number of cells in the system = $\frac{3024}{6} = 504$

(*b*) The number of calls per hour per cell, $N_{cell'}$ is given by

$$N_{\text{call}} = \frac{N_{\text{data/cluster}}}{N} \times \eta_t \times \text{number of calls per hour}$$

$$N_{\text{call}} = 56 \times 0.95 \times \frac{3600}{180}$$

$$N_{\text{call}} = 1064 \text{ calls/hour/cell.}$$

(*c*) The average number of users per hour per cell, N_{users} is given by

$$N_{\text{users}} = \frac{\text{Number of calls per hour per cell}}{\text{Average number of calls per user hour}}$$

$$N_{users} = \frac{1064}{1.5} = 709 \text{ users/hour/cell}$$

(*d*) The system spectrum efficiency is, from last question

$$\eta = \eta_t \times 0.747$$
$$= 0.710 \text{ Erlangs/MHz/km}^2$$

Example 7.4: *Find the intermodulation frequencies generated if a base station transmits two carrier frequencies at 1932 MHz and 1930 MHz that are amplified by a saturated clipping amplifier. If the mobile radio band is allocated from 1920 MHz to 1940 MHz, designate the 1M frequencies that lie inside and outside the band.*

Solution: Intermodulation distortion products occur at frequencies $mf_1 + nf_2$ for all integer values of m and n, *i.e.*, $-\infty < m, n < \infty$. Some of the possible intermodulation frequencies that are produced by a non-linear device are

$(2n+1) - 2nf_2$, $(2n+2)f_1 - (2n+1)f_2$, $(2n+1)f_1 - 2nf_2$, $(2n+2)f_2 - (2n+1)f_1$, etc. for $n = 0, 1, 2, \ldots\ldots\ldots\ldots$

Table lists several intermodulation product terms.

Table 7.3: Intermodulation Products

$n = 0$	$n = 1$	$n = 2$	$n = 3$
1930	1926	1922	1918
1928	1924	1920	1916
1932	1936	1940	1944*
1934	1938	1942*	1946*

The frequencies in the table marked with an asterisk (*) are the frequencies that lie outside the allocated mobile radio band.

Example 7.5: *If a US AMPS cellular operator is allocated 12.5 MHz for each simplex band, and if B_t is 12.5 MHz, B_{guard} is 10 kHz, and B_t is 30 kHz, find the number of channels available in an FDMA system.*

Solution: The number of channels available in the FDMA system is given as –

$$N = \frac{12.5\times10^6 - 2(10\times10^3)}{30\times10^3} = 416$$

In the US, each cellular carrier is allocated 416 channels.

Example 7.6: *Consider Global system for mobile, which is a TDMA/FDD system that uses 25 MHz for the forward link, which is broken into radio channels of 200 kHz. If 8 speech channels are supported on a single radio channel, and if no guard band is assumed, find the number of simultaneous users that can be accommodated in GSM.*

Solution: The number of simultaneous users that can be accommodated in GSM is given as –

$$N = \frac{25MHz}{\frac{(200\,kHz)}{8}} = 1000$$

Thus, GSM can accommodate 1000 simultaneous users.

Example 7.7: *If GSM uses a frame structure where each frame consist of eight time slots and each time slot contains 156.25 bits, and data is transmitted at 270.833 Kbps in the channel, find*

(a) *The time duration of a bit*

(*b*) *The time duration of a slot*

(*c*) *The time duration of a frame*

(*d*) *How long must a user occupying a single time slot wait between two successive transmissions ?*

Solution:

(*a*) The time duration of a bit,

$$T_b = \frac{1}{270.833\,\text{Kbps}} = 3.692\ \mu\text{s}$$

(*b*) The time duration of a slot,

$$T_{\text{slot}} = 156.25 \times T_b = 0.577\ \mu\text{s}$$

(*c*) The time duration of a frame,

$$T_f = 8 \times T_{\text{slot}} = 4.615\ \mu\text{s}$$

(*d*) A user has to wait 4.615 ms, the arrival time of a new frame, for its next transmission.

Example 7.8: *If a normal GSM time slot consists of six trailing bits, 8.25 guard bits, 26 training bits, and two traffic bursts of 58 bits of data, find the frame efficiency ?*

Solution: A time slot has 6 + 8.25 + 26 + 2(58) = 156.25 bits

A frame has 8 x 156.25 = 1250 bits/frame.

The number of overhead bits per frame is given by

b_{OH} = 8(6) + 8(8.25) + 8(26) = 322 bits

Thus, the frame efficiency

$$\eta_t = \left[1 - \frac{322}{1250}\right] \times 100 \ \times 100$$

$$\eta_t = 74.24\ \%$$

Example 7.9: *Determine the maximum throughput that can be achieved using ALOHA and slotted ALOHA protocols.*

Solution: The rate of arrival which maximizes the throughput for ALOHA is found by taking the derivating of equation 7.8 and equating it to zero.

$$\frac{dT}{dG} = e^{-2G} - 2Ge^{-2G} = 0$$

Maximum throughput achieved by using the ALOHA protocol is found by substituting R_{max} in equation 7.8 and this value can be seen as the maximum throughput in Fig. 7.11.

$$T = \frac{1}{2}e^{-1} = 0.1839.$$

Thus the best traffic utilization one can hope for using ALOHA is 0.184 Erlangs.

The maximum throughput for slotted ALOHA is found by taking the derivative of equation 7.11 and equating it to zero.

$$\frac{dT}{dR} = e^{-G} - Ge^{-G} = 0$$

Maximum throughput is found by substituting R_{max} in equation 7.11 and this value can be seen as the maximum throughput in Fig.

$$T = e^{-1} = 0.3679$$

Notice that slotted ALOHA provides a maximum channel utilization of 0.368 Erlangs, double that of ALOHA.

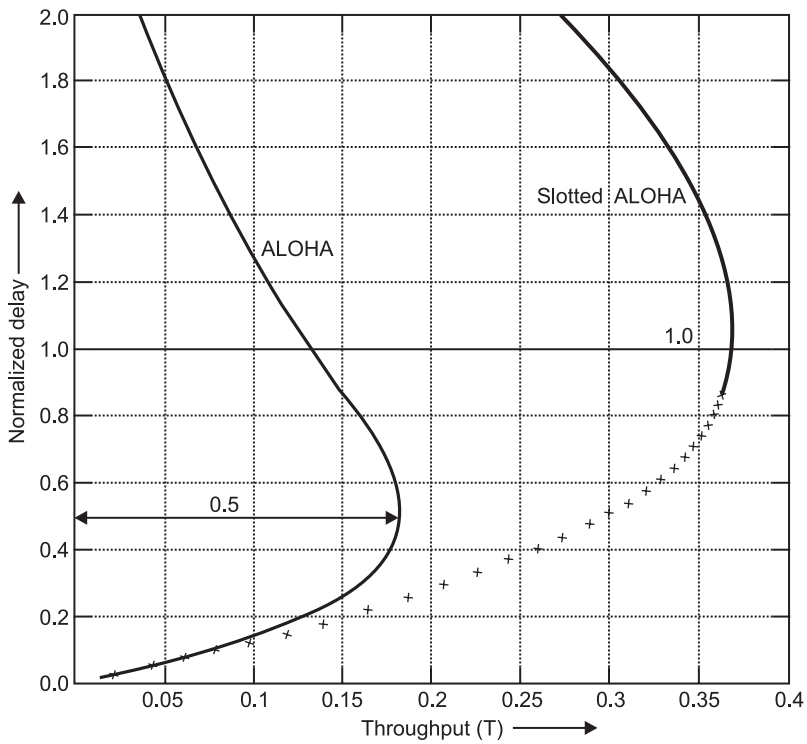


Fig. 7.11: Tradeoff between throughput and delay for ALOHA and slotted ALOHA packet radio protocols.

Example 7.10: *Evaluate four different cellular radio standards, and choose the one with the maximum radio capacity. Assume n = 4 propagation path loss.*

System A : B_c = 30 *kHz,* $(C/I)_{min}$ = 75 *dB*

System B : B_c = 25 *kHz,* $(C/I)_{min}$ = 14 *dB*

System C : B_c = 12.5 *kHz,* $(C/I)_{min}$ = 2 *dB*

System D : B_c = 6.25 *kHz,* $(C/I)_{min}$ = 9 *dB*

Solution: Consider each system for 6.25 kHz bandwidth, and use equation

System *A :* B_c = 6.25 kHz, $(C/I)_{eq} = 18 - 20 \log \frac{6.25}{30} = 31.680$ dB

Systems B : B_C = 6.25 kHz, $(C/I)_{eq}$= 14 – 201og $\frac{6.25}{25}$ = 26 dB

System C: B_c = 6.25 kHz, $(C/I)_{eq}$ = 12 – 201og $\frac{6.25}{12.5}$ = – 4 dB

System D : B_c = 6.25 kHz, $(C/I)_{eq}$ = 9 – 20log $\frac{6.25}{6.25}$ = 9 dB

Based on comparison, the smallest value of $(C/I)_{eq}$ should be selected for maximum capacity in equation system B offers the best capacity.

Example 7.11: *Consider an FDMA system with three channels, each having a bandwidth of 10 kHz and a transmission rate of 10 Kbps. A TDMA system has three time slots, channel bandwidth of 30 kHz, and a transmission rate of 30 Kbps.*

Solution: For the TDMA scheme, the received carrier-to-interference ratio for an single user is measured for 1/3 of the time the channel is in use. For example C'/I' can be measured in 333.3 ms in one second, thus, C'/I' is given by

$$C' = E_b R_b^{'} = \frac{E_b 10^4 bits}{0.33s} = 3R_b E_b = 3C$$

$$I' = I0B'_C = I_0\ 30 \text{ kHz} = 3I$$

It can be seen that the received carrier-to-interference ratio for a user in this TDMA system C'/I' is the same as C/I for a user in the FDMA system. Therefore, for this question, FDMA and TDMA have the same radio capacity and consequently the same spectrum efficiency. However, the required peak power for TDMA is 10 log k higher than FDMA, where k is the number of time slots in a TDMA system of equal bandwidth.

Example 7.12: *If w = 1.25 MHz, R = 9600 bps, and a minimum acceptable E_b / N_0 is found to be 10 dB, determine the maximum number of users that can be supported in a single cell CDMA system using*

(a) Omnidirectional base station antennas and no voice activity detection.

*(b) Three-sectors at the base station and activity detection with **a** = 3/8. Assume the system is interference limited.*

Solution:

(*a*) Using equation $$N = 1 + \frac{\frac{W}{R}}{\frac{E_b}{N_0}} - e\left(\frac{\eta}{s}\right)$$

$$N = 1 + \frac{1.25 \times 10^6 / 9600}{10} = 1 + 13.02 = 14$$

(*b*) For each sector, we can find $N_s = 1 + \frac{1}{\alpha}\left[\frac{\frac{W}{R}}{\frac{E_b}{N_0^{'}}}\right]$

$$N_s = 1 + \frac{1}{0.375}\left[\frac{1.25\times10^6/9600}{10}\right] = 35.7.$$

The total number of users is given by $3N_S$, since three sectors exist within a cell; therefore N = 3 x 35.7 = 107 users/cell.

HIGHLIGHTS

1. The multiple access technology allow a large number ot users to share a common radio channel and any user can gain access to any channel.
2. The multiple access technology are divided into parts on the basis of frequency range –
   (*i*) Narrow band system
   (*ii*) Wide band system.
3. For transmission and reception between base station and user mobile, two types of connectivity is used –
   (*i*) Uplink – from user mobile to base station connection.
   (*ii*) Downlink – from base station to user mobile connection.
4. The multiplexing is basically used to transmit the signal from one end to the another end over a channel while multiple access techniques are used to communicate many user with each other. There are two type of multiplexing technique used in telecommunication networks –
   (*i*) Frequency division multiplexing (FDM)
   (*ii*) Time division multiplexing (TDM).
5. The FDM use the analog modulation system and in FDM, the user use the channel on the base of frequency division but the TDM use the digital modulation system and in TDM, the user use the channel on the basis of time division.
6. The multiple access technology are divided in four major parts –
   (*i*) Time division multiple access technology (TDMA)
   (*ii*) Frequency division multiple access technology (FDMA)
   (*iii*) Space division multiple access technology (SDMA)
   (*iv*) Code division multiple access technology (CDMA).
7. In case of frequency division multiple access, the complete radio spectrum is divided into different channels (frequency bands). These individual channels are assigned to individual users. During the period of the call conversation, no two users can share the same channels.
8. The FDMA architecture is also known as narrow band mobile radio as the bandwidth of the individual data or digitized analog signal (like as voice, fax etc.) is relatively narrow compared with TDMA and CDMA applications.
9. By using FDMA, technique, the communication system can classified as –
   (*i*) Advanced mobile phone system (AMPS)
   (*ii*) Digital European cordless telephone (DECT)
   (*iii*) Total access communication system (TACS).
10. The FDMA decides a limit on maximum number of channels that can be simultaneously supported if N are the maximum number of channels then the number of channels in FDMA can be

$$N = \frac{B_t - 2B_g}{B_c}$$

B_t = complete spectrum
B_c = channel bandwidth
B_g = Guard band

11. ALOHA is a media used to transmit the data from one station to another station and ALOHA can be classified into two parts –
(*i*) Pure ALOHA
(*ii*) Slotted ALOHA

12. The ALOHA with no restriction is called pure ALOHA and the collision possibility in pure ALOHA is more and only 18.4% data is transmit.

13. In the slotted ALOHA, the throughput of the pure ALOHA scheme can be doubled by the introducing time slot in the channel and permitting a new transmission to start only at the begining of a time slot.

14. The collision possibility in slotted ALOHA is less then the pure ALOHA and 36.8% data can be successfully transmitted.

15. The code division multiple access (CDMA) system is a system in which the narrow band message signal is multiplied with a large bandwidth signal, which is called the spreading signal.

16. The spreading signal is a pseudonoise code sequence which has a chip rate whose magnitude is greater than data rate of message and in CDMA system, each user has its own pseudonoise code sequence or word and in CDMA, all users use the same carrier frequency and transmit the data at the same time.

17. Self jamming and near-far problem are two major problems produce in CDMA system. The self jamming occurs when the spreading sequence of different users are not exactly orthogonal and the near-far problem occur when an undesired user has a high detected power as compared to desired user.

18. TDMA systems divided the radio spectrum into time slots and in each slot only one user is allowed to either transmit or receive.

19. The TDMA system transmit the data in a buffer and burst method thus the transmission for any user is non-continuous. This implies that, unlike a FDMA system which correspond to analog FM, digital data and digital modulation must be used with TDMA system.

20. The maximum number of channels slots in TDMA system are found by multiplying the number of time slot per channel by the number of channel available, then the number of channel in TDMA system

$$N = \frac{M(B_t - 2B_g)}{B_c}$$

where,
M = Maximum number of TDMA slot (or user).

21. The space division multiple access (SDMA) control the radiated energy for each user in space and SDMA serves different users by using spot beam antennas. These different areas covered by the antenna beam may be served by the same frequency or different frequency.

22. The SDMA technology is very simple and the capacity will be increased per kilometer but it has two drawbacks, one is inflexibility and second is, the antennas typically fixed, and SDMA used in combination with TDMA, FDMA or CDMA has a high detected power as compared to the desired user.

23. For a radio system, the channel capacity defined the number of channels or users, which are provided in a frequency band.

24. For measuring the spectrum efficiency of a wireless system, a parameter named radio capacity is used. This parameter is determine by the channel bandwidth B_c and carrier to interference ratio (C/T).

25. In wireless communication system, it is desirable to allow the subscriber to send the information at the same time to the base station while receiving information from the base station for example in conventional telephone system, it is possible to talk and listen at the same time, this effect is called duplexing.

26. Duplexing may be done using frequency or time domain techniques.

PROBLEMS

7.1. The GSM TDMA system uses a 270.833 Kbps data rate to support eight users per frame.

(*a*) What is the raw data rate provided for each user ?

(*b*) If guard time, ramp-up time, and synchronization bits occupy 10.1 Kbps, uses a 48.6 Kbps data rate determine the traffic efficiency for each user.

7.2. The US digital cellular TDMA system uses a 48.6 Kbps data rate to support three users per frame. Each user occupies two of the six time slots per frame. What is the raw data rate provided for each user ?

7.3. For problem 2 assume each reverse channel frame contains six time slots with 324 bits per time slot, and within each time slot, assume there are six guard bits, six bits reserved for ramp-up, 28 synchronization bits, 12 control channel bits, 12 bits for supervisory control signals, and 260 data bits.

(*a*) Determine the frame efficiency for the US digital cellular standard.

(*b*) If half-rate speech coding is used, then six users may be supported within a frame. Determine the raw data rate and frame efficiency with half-rate speech coding.

7.4. The pacific digital cellular (PDC) TDMA system uses a 42.0 Kbps data rate to support 3 uses per frame. Each user occupies two of the six time slots per frame.

(*a*) What is the raw data rate provided for each user ?

(*b*) If the frame efficiency is 80% and the frame duration is 6.667 ms, determine the number of information bits sent to each user per frame.

(*c*) If half-rate speech coding is used, six users per frame are accomodated. Determine the number of information bits provided for each user per frame.

(*d*) What is the information data rate per user in half-rate PDC ?

7.5. Assume that a non-linear amplifier is used to broadcast FDMA transmissions for the US AMPS standard. If control channel 352 and voice channel 360 are simultaneously transmitted by a base station, determine all cellular channels on the forward link that might carry interference due to intermodulation.

7.6. If a paging system transmits at 931.9375 MHz and a cellular base station broadcasts on AMPS channel 318 and both antennas are co-located on the same broadcast tower, determine the cellular channels (if any) that may contain intermodulation interference, if detected by a receiver with a non-linear amplifier. In what practical situation might intermodulation of this type be produced ?

7.7. In an unslotted ALOHA system the packet arrival times from a Poisson process having a rate of 10^3 packets/sec. If the bit rate is 10 Mbps and there are 1000 bits/packet, find

(*a*) The normalized throughput of the system

(*b*) The number of bits per packet that will maximize the throughput.

7.8. Repeat problem 7 for a slotted ALOHA system.

7.9. Determine the propagation delay in packet transmission units if a 19.2 kbps channel data rate is used and each packet contains 256 bits. Assume a line of sight radio path exists for a user 10 km away from the transmitter. If slotted ALOHA is to be used, what is the best choice for the number of bits/packet for this system (assuming that 10 km is the maximum distance between the transmitter and receiver)?

7.10. Determine the number of analog channels per cell for the case of $n = 3$ propagation path loss, where the minimum acceptable $C/I = 14$ dB. What is the appropriate cluster size for the system? Assume the channel bandwidth is 30 kHz and the total spectrum allocation is 20 MHz.

7.11. A CDMA cellular system uses SDMA, and multiple cells are used, where each cell shares the same radio channel. Consider propagation path loss exponents of n= 2, 3 and 4 and determine the number of simultaneous users that can be supported at an average probability of error of 10^{-2}

Assume $k = 511$ and 6 dB of directionality is provided by the base station for each user.

7.12. Simulate reverse channel interference in a one-cell reuse CDMA system. To do this, consider seven hexagonal cells and assume each one has equal area of 10 km^2. Randomly place 30 users in each cell and assume, that $d_0 = 1$ m is the close in reference distance for propagation and $n = 4$ is the path loss exponent. If power control is used for the 30 in-cell users, find:

(*a*) The received in-cell interference power

(*b*) The received out of cell interference power

(*c*) The frequency reuse factor.

REVIEW QUESTIONS

7.1. What are various accessing techniques used in mobile networks ?

7.2. Explain frame efficiency in a TDMA system. How many number of channel slots can be provided in a TDMA mobile system whose channel bandwidth is B_c and total spectrum allocated is B_t. The maximum number of TDMA users supported on each radio channel is M and allocated guard band is B_g .

7.3. Discuss the advantages of digital wireless systems over analog wireless systems.

7.4. List out the salient features of FDMA technology along with its advantages and disadvantages.

7.5. Name those communication systems which are using TDMA technology.

7.6. What is the basic principle of SDMA technology. Explain.

7.7. Do a comparison among various multiple accessing technologies.

7.8. What do you mean by narrowband and wideband systems in multiple access technique.

7.9. What is the fundamental difference in concept for calculating the capacity of CDMA cellular system over FDMA cellular system.

7.10. Find out the expression for calculating the number of users in a CDMA cellular system.

7.11. Explain in brief CDMA.

7.12. Explain in brief ALOHA.

7.13. What do you understand by multiple access technique

7.14. Explain packet Radio Protocols – pure ALOHA and Slotted ALOHA.

7.15. Explain TDMA, FDMA and CDMA techniques in detail.

8 Chapter

THE CELLULAR CONCEPT-SYSTEM DESIGN FUNDAMENTALS

8.1. *Introduction* **8.2.** *The Cellular Concept* **8.3.** *Frequency Reuse* **8.4.** *Channel Assignment Strategies* **8.5.** *Handoff Strategies* **8.6.** *Interference and System Capacity* **8.7.** *Power Control For Reducing Interference* **8.8.** *Improving Coverage And Capacity In Cellular Systems* **8.9.** *Comparison Among Cell Splitting Cell Sectoring Micro Cell Zone Concept Increasing Number Of Repeaters*

8.1 INTRODUCTION

To fulfill the requirement of day by day increasing number of mobile users, some techniques are required to use the available channel (set of frequencies or allowed frequency spectrum) more efficiently so as to avoid congestion. The techniques provide us to increase the system capacity of the existing systems. Therefore frequency reuse concept come into evolution. But at the same time this frequency reuse technology faces the challenge from interference between the communication of various radio signal. This chapter concern with frequency reuse strategy, interference, hand-off strategy, cell sectoring and various other technology to improve cellular transmission and system capacity for transmission.

8.2. THE CELLULAR CONCEPT

The cellular concept was a major breakthrough in solving the problem of spectral congestion and user capacity. It support very high capacity in a limited spectrum allocation without any major or minor technological changes.

The cellular concept is a system level idea which calls for replacing a single high power transmitter (large cell) with many low power transmitter (small cells) each support coverage to only a small portion of the service area.

Each base station is allocated a portion of the total number of channels available to the entire system and nearby base station are allocate different group of channels so that all the available channels are assigned to a relatively small number of neighbouring base stations. Neighbouring base station are alloted different group of channels so that the interference between base station is minimized.

By in a regular manner spacing base stations and their channel group throughout a market, the available channels are distributed throughout the geographic region and may be reuse as many times as required so long as the interference between co-channel stations is kept below acceptable levels.

As the demand for service increase (*i.e.* more channel are demanded) the number of base station may be increased (along with a corresponding decrease in transmitted power to avoid added interference), thereby allow additional radio capacity with no additional increase in radio spectrum.

This fundamental principle is the foundation for all modern wireless communication system, since it authorize a fixed number of channels to serve an arbitrarily large number of subscribers by reusing the channels throughout the coverage region or area.

8.3 FREQUENCY REUSE

The cellular system are high capacity land mobile system in which assigned radio spectrum is divided into discrete channels which are assigned in groups to geographic cells covering cellular geographic region. The discrete channels are capable of being reused in different cells through a process known as frequency reused.

The design process of selecting and assigning the channel groups for all of the cellular base station within a system is called frequency reuse of frequency planning concept.

Each cellular base station is allocated a group of radio channels to be used with in a small geographic region called a cell.

Fig. 8.1 show the concept of cellular frequency reuse, where cells adhesive with the same letter use the same group of channels.

The hexagonal cell shape shown in Fig. 8.1 is conceptual and is a simplistic model of the radio coverage for each base station, but it has been universally embrace since the hexagon permits easy and manageable analysis of a cellular system.

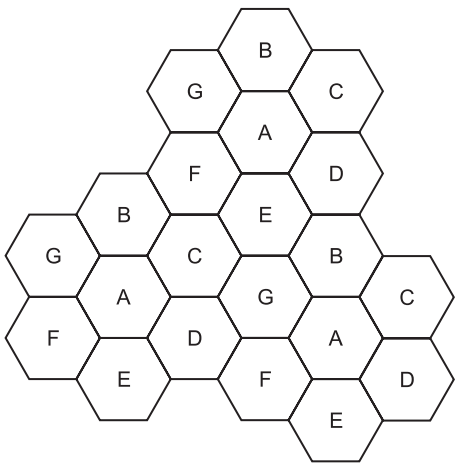


Fig. 8.1: Illustration of the cellular frequency reuse concept. Cells with the same letter use the same set of frequencies. A cell cluster is outlined in bold and replicated over the coverage area. In this example, the cluster size, *N*, is equal to seven, and the frequency reuse factor is 1/7 since each cell contains one-seventh of the total number of available channels.

The actual radio coverage of a cell is called the foot print and is determined from field measurement or propogation prediction models.

So far we have study that a given geographical region coverage is subdivided into hexagonal cells, where each cell has its own set of channels and a number of channels were required for cellular communication, but this is not a practical approach to support different sets of channels to each cell for entire geographical area due to limited range of frequency spectrum for wireless (cellular communication). Therefore to reduce this drawback, a technique is adopted which user same set of channels for different cells, the cells are separated from one another by a distance large enough to make interference levels within tolerable limits. Fig. 8.1 show the reuse concept ($N = 7$ reuse format).

To explain the frequency reuse concept, consider a cellular system which has a total of s duplex channels available for use. If each cell is allocated a group of k-channels ($k < s$) and if the s-channels are splited among N cells into unique and disjoint channel groups which each have the same number of channels, the total number of available radio channels can be expressed as the following equation:

$$s = kN \qquad \text{... (8.1)}$$

The N–cells which collectively use the complete set of available or present frequencies is called a cluster (a group of cells).

If a cluster is replicated M times within the system, the total number of duplex channels C, can be used as a measure of capacity and written as -

$$C = MkN = Ms \qquad \text{... (8.2)}$$

From equation (8.2), the capacity of a cellular system is directly proportional to the number of times a cluster is replicated in a fixed service area or region.

The factor N (cluster size) is typically equal to 4, 7, or 12, if the cluster size N is reduced while the cell size is kept constant, more cluster are required to cover a given region, and hence more capacity is achieved. A large cluster size causes the ratio between the cell radius and the distance between co-channel cells to decrease, leading to weaker co-channel interference, conversely a small cluster size represent the co-channel cells are located much closer together.

From design point of view, the smallest possible value of N is desirable in order to maximize capacity over a given coverage region (*i.e.* to maximize C in equation 8.2).

The frequency reuse factor of a cellular system is given by $1/N$ since each cell within a cluster is only allot the $1/N$ of the total available channels in the system.

The hexagonal geometry of Fig. 8.1, has exactly six *equidistant* neighbours and that the lines joining the centers of any cell and each of its neighbours are distinct by multiple of 60 degree, there are only certain cluster size and cell layouts which are possible to connect without gaps between neighbouring cells, the geometry of hexagons is such that the number of cells per cluster N, can only have value which satisfy equation (8.3) than the equation can be written as –

$$N = i^2 + ji + j^2 \qquad \text{... (8.3)}$$

where i and j are non-negative integers.

To find the nearest co-channel neighbour of a particular cell, one most do the following:

(*i*) Move *i* cells with any chain of hexagons and then

(*ii*) Turn 60 degree counter clockwise and move *j* cells.

This is show in Fig. 8.2 for $i = 3$ and $j = 2$ (example, $N = 19$).

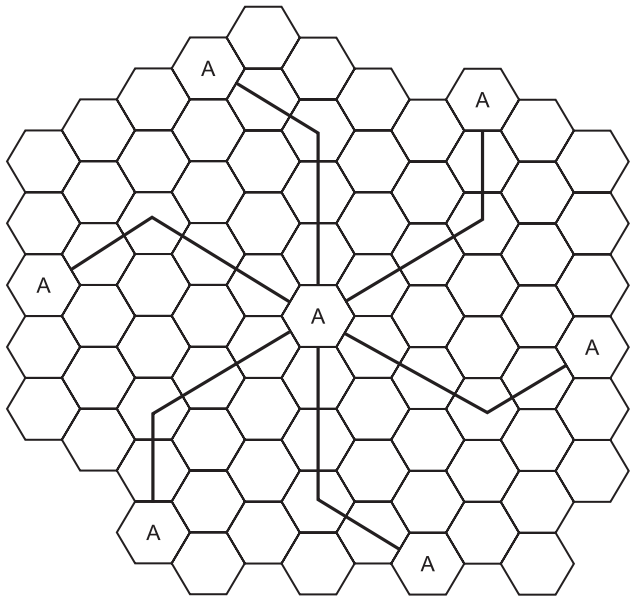


Fig. 8.2: Method of locating co-channel cells in a cellular system.
In this example, $N = 19$ (*i*.e., $i = 3$, $j = 2$)

8.4 CHANNEL ASSIGNMENT STRATEGIES

The main moto of using channel assignment strategies is to optimize utilization of frequency reuse scheme and minimization of interference so that overall channel capacity can be increased by using this strategies.

Channel assignment strategies can be classified as the following:

(*i*) Fixed channel assignment strategy

(*ii*) Dynamic channel assignment strategy

The choice of channel assignment strategy impacts the performance of the system, particularly as to how calls are managed when a mobile user is handed off from one cell to another moment.

8.4.1 Fixed Channel Assignment Strategy

In fixed channel assignment strategy, each cell is allocated a predetermined set of voice channel for transmission.

Maximum number of calls that can be attempted depend on the number of voice channels. If all the channels in a cell are found busy than the call is blocked and no service will be delivered to the subscriber.

In this strategy, an approach is sometime adopted that is called the borrowing approach. In this approach, one cell is allowed to borrow some voice channels from its neighbouring cell under the handle by mobile switching centre (MSC).

In this strategy, the neighbouring cell should not be disturbed or interfered any way by borrower cell or neighbour cell.

8.4.2 Dynamic Channel Assignment Strategy

In a dynamic channel assignment strategy, voice channels are not allocated to different cells permanently. In place of each time a call request is made, the serving base station request a channel from the MSC. Then the switch assigned a channel to the requested cell following an algorithm that takes into account the probability of future blocking within in cell, the frequency of use of the candidate channel, the reuse distance of the channel and other cost function.

The MSC only assign a given frequency if that frequency is not presently is use in the cell or any other cell which fall within the minimum restricted distance of frequency reuse to ignore co-channel interference.

Dynamic channel assignment strategies require the MSC to collect real time data on channel occupancy, traffic distribution, and radio signal strength indications (RSSI) of all channels on a continuous basis.

This growing the storage and computational load on the system but provides the advantage of increased channel utilization and decreased probability of a blocked call with in a system.

8.5 HANDOFF STRATEGIES

When a mobile moves into a different cell while a conversation is in progress, the MSC automatically transfer the call to a new channel related to the new base station. This handoff operation not only involves identifying a new base station but also require that the voice and control signals be allocated to channels related with the new base station.

Handoff should be performed as infrequently as possible and be completely transparent (seamless) to the subscriber *(i.e.* the subscriber cannot discerned that their facility has been switched).

A handoff consist of four stages –

(*i*) **Initiation** : Either the mobile unit or the network determines the need for a handoff and support the necessary network procedures.

(*ii*) **Resource reservation** : Appropriate network procedures, reserve the resources needed to support the handoff *(i.e.* voice and control channel).

(*iii*) **Execution** : The actual transfer of control from one base station to another base station take place or establish.

(*iv*) **Completion** : Unnecessary network resources are relinquished and made available to other mobile units or subscriber.

A connection that is momentarely broken during the cell-to-cell transfer is called a hard handoff. A hard handoff is a break before make process. With

a hard handoff, the mobile unit divides its connection with one base station before establishing voice communication with a new base station. Hard handoff generally produce when a mobile unit is passed between disjointed system with different frequency assignments, air interface characteristic or technologies.

A flowless handoff (*i.e.* no perceivable interruption of service is known as soft handoff and normally takes approximately 200 ms which is imperceptible to voice telephone users, although the delay may be disruptive when transmitting data. By using soft handoff, a mobile unit establishes contact with a new base station before given up its current radio channel by transmitting coded speech signal to two base stations at the same time. Both base station send their received signal to the MTSO, which estimate the quality of the two signals and determines when the transfer should occur. A complementary process is produced in the opposite direction. A soft handoff requires that the two base station operate synchronously with one another stations.

Fig. 8.3 shows how a base station transfer is accomplished when a mobile unit moves from one cell into another (this Fig. 8.3 define a soft handoff). The mobile unit is moving away from base station 1 (*i.e.* toward base station 2). When the mobile unit is at *w* and *x* position then it is well with in the range of base station 1 and very distance from base station 2. However, when the mobile unit reaches *y* position it receive signals from base station 1 and base station 2 at approximately the same power level and the two base station should be establish for a handoff. When the mobile unit crosses from cell 1 into cell 2, the handoff should be executed and completed as shown in following Fig. 8.3.

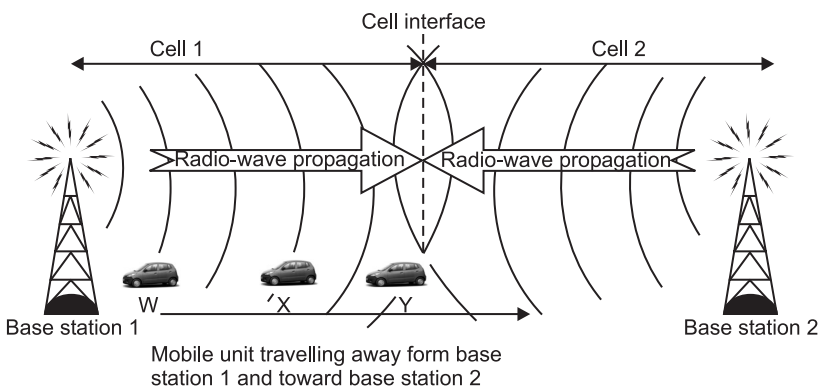


Fig. 8.3

When a call is in progress, the switching center monitors the received signal strength of each user channels information.

Handoff can be initiated when the signal strength (or signal to interference ratio) measured by either the base station or the mobile unit's receiver, falls below a predecide threshold level typically between –90 dB and –100 dB or when a network resource management needs to force a handoff to free resources to place an emergency call from station.

During a handoff, information about the user stored in the first base station is transferred to the new base station. A condition called blocking produce when the signal level drops below a usable level are there are no usable channels available in the target cell to switch to.

By using avoid blocking or loss of a call during a handoff, the system employs a load balancing techniques that trees channels for handoff and sets handoff priorities.

8.5.1 Prioritizing Handoffs

One method for giving priority to handoffs is known as guard channel concept whereby a fraction of the total available channels in a cell is reserved exclusively for handoff requests from ongoing calls which may be handedoff into the cell. This method has the disadvantage of minimizing the total carried traffic, as fewer channels are allocated to originating calls guard channels, however, offer efficient spectrum utilization when dynamic channel assignment strategies, which minimize the number of required guard channels by efficient requirement based allocation are used.

Queuing of handoff requests is another method to reduce the probability of forced termination of a call due to lack of available channels.

8.5.2 Practical Handoff Considerations

In practical cellular systems, many problems arise when attempting to design for a wide range of mobile velocities and in practice it is difficult for cellular service providers to obtain new physical cell site location in urban area or region.

By using different antenna heights (often on the some building or tower) and different power levels, it is possible to provide large and small cells which are co-located at a single location. This technique is known as the umbrella cell approach and is used to provide large area coverage to high speed users while support small area coverage to users travelling at low speeds.

Fig. 8.4 shows an umbrella cell which is co-located with some small microcells. The umbrella cell approach ensures that the number of handoff is minimize for high speed users and arrange additional microcell channels for padestrian users.

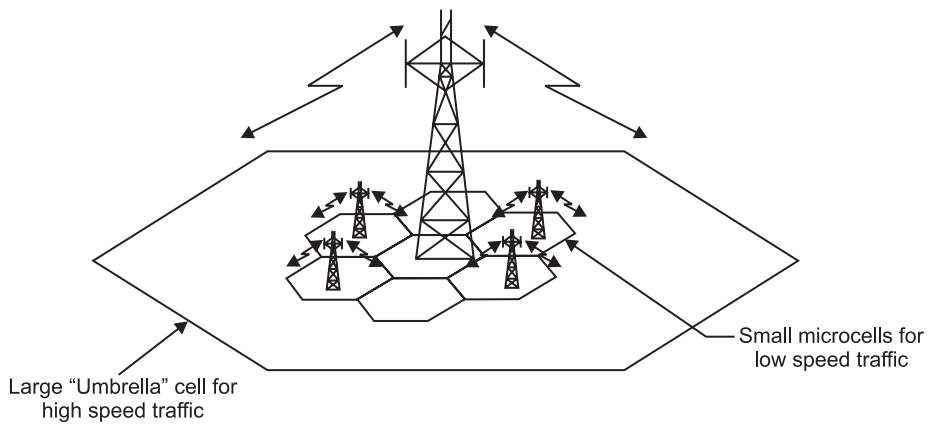


Fig. 8.4: The umbrella cell approach.

Another practical handoff problem in microcell systems is known as dragging Cell dragging results from pedestrian users that support a very strong signal to the base station. As the user travells away from the base station at a very slow speed, the average signal strength does not reduce rapidly. Even when the user has travelled well beyond the designed range of cell. The received signal at the base station may be above the handoff threshold, thus a handoff may not be made. This produces a potential interference and traffic management problem. Since the user has meanwhile travelled deep within a not neighbouring cell. To solve the cell dragging problem, handoff thresholds and radio coverage parameters should be adjusted and handle carefully.

8.6 INTERFERENCE AND SYSTEM CAPACITY

Interference can be understood as interaction of radio signals that cause noise or distortion.

Interference and system capacity both terms are inversely proportionally to each other. If interference of any system increases then its capacity decreases and if interference decrease than capacity increase.

Interference is simply defined as a disturbance during conversation or during establishment of a call in cellular radio systems.

Sources of interference with another mobile in the same cell, a call in progress in a neighbouring cell, other base station operating in the same frequency band or any non-cellular system which neighbouring leaks energy into the cellular frequency band.

Interference on voice channels produce cross talk.

Interference is more serve in urban area, due to the greater RF noise floor and the large number of base stations and mobiles users.

Interaction of signals of one cell may be either due to its reused frequency cell in other cluster or due to its neighbouring cell. Therefore interference is categorized into two parts or in two types –

(*i*) Co-channel interference

(*ii*) Adjacent channel interference

8.6.1 Co-channel Interference and System Capacity

Frequency reuse implies that in a given coverage area, there are several cells that use the same set of frequencies. These cells are known as co-channel cells and the interference between signals from these cells is called co-channel interference.

Co-channel interference can not be combated by simply increasing the carrier power of a transmitter. This is because an growing in carrier transmitted power growing the interference to neighbouring co-channels cells. To reduce co-channel interference, co-channel cells must be physically separated by a minimum distance to allow sufficient isolation due to propagation.

When the size of each cell is closely the same and the base station transmit the same power, the co-channel interference ratio is independent of the transmitted

power and becomes a function of the radius of the cell (R) and the distance between centers of the neighbour co-channel cells (D).

By increasing the ratio of D/R, the spatial separation between co-channel cells relative to the coverage distance of a cell increases, thus interference is minimize from improved isolation of RF energy from the co-channel cells.

The parameter Q, called the co-channel reuse ratio, is related to the cluster size. For a hexagonal geometry, Q can be given as –

$$Q = \frac{D}{R} = \sqrt{3N} \qquad \text{... (8.4)}$$

A small value of Q support large capacity since the cluster size N is small, whereas a large value of Q improve the transmission quality due to a smaller level of co-channel interference.

Now the i_0 be the number of co-channel interfering cells. Then the signal to interference ratio (S/I or S/R) for a mobile receiver which monitors a forward channel can be expressed as the following equation:

$$S = \frac{S}{\sum_{i=1}^{i_0} I_i} \qquad \text{... (8.5)}$$

where S is the desired signal power from the desired base station and I_t is the interference power produce by the i th interfering co-channel cell base station. If the signal levels of co-channel cells are known, then the S/I ratio for the forward link can be defined using equation (8.5).

The average received power P_r at a distance d from the transmitting antenna is approximated by the following equation can be written as –

$$P_r = P_0 \left(\frac{d}{d_0} \right)^{-n} \qquad \text{... (8.6)}$$

$$P_r\ (dBm) = P_0(dBm) - 10n \log\left(\frac{d}{d_0} \right) \qquad \text{... (8.7)}$$

where P_0 is the power received at a close in reference point in the far-field region of the antenna at a small distance d_0 from the transmitting antenna and n is the path loss exponent.

Now consider the forward link where the desired signal is the serving base station and where the interference is due to the co-channel base stations, if D_i is the distance of the ith interferer. from the mobile, the received power at a given mobile due to the ith interfering cell will be proportional to $(D_i)^{-n}$ or depends on the $(D_i)^{-n}$.

When the transmit power of each base station is equal and path loss exponent is the same throughout the coverage area, *S/I for* a mobile can be approximated, written as the following equation –

$$\frac{S}{I} = \frac{R^{-n}}{\sum_{i=1}^{i_0} (D_i)^{-n}} \qquad \text{... (8.8)}$$

If this distance equal to distance D between cell centers, then the equation can be –

$$\frac{S}{I} = \frac{\left(\frac{D}{R}\right)^n}{i_0} = \frac{(\sqrt{3N})^n}{i_0} \quad \text{... (8.9)}$$

Equation (8.9) show the relation between S/I and cluster size N, which in turn design the overall capacity of the system from equation (8.2).

Generally, co-channel cells are located is tiers. Tier is a large hexagon that is formed by connecting nearest co-channel cells. A cell is covered by *6-n* co-channel cells for nth tier (in N = 7 cluster size). Fig. 8.5 shows the first tier and second tier co-channel cells.

8.6.2 Adjacent Channel Interference

Adjacent channels are those channels which show less separation in frequency in desired channel. The interference resulting from signals of adjacent channels is known as "Adjacent Channel" interference.

Adjacent channel interference is produce by the inability of mobile phone to filter out the signals (frequencies) of adjacent channels assigned to nearby cell sites.

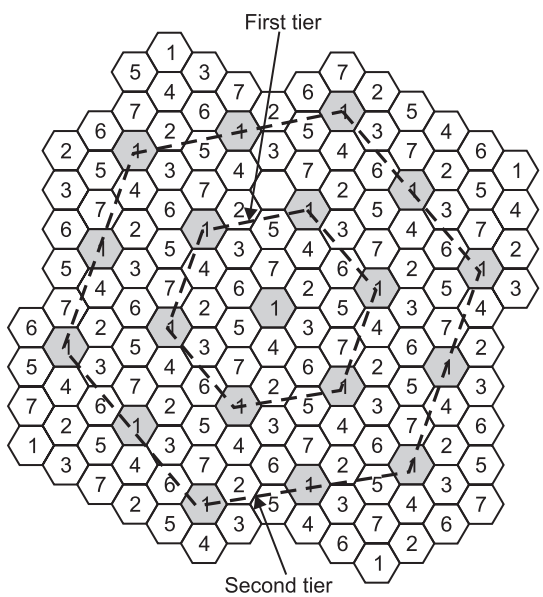


Fig. 8.5: Different co-channel cells forming first and second tier.

If the receiver is not strict to support only its own passband but permit some nearby frequencies also then adjacent channels traffic do interfere in the current cell.

Fig. 8.6 shows adjacent channel and desired main channel and important channel.

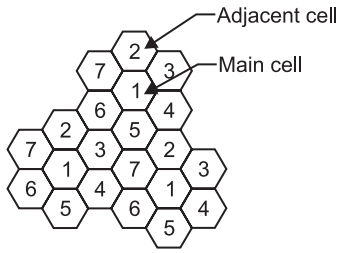


Fig. 8.6: Illustration of adjacent channel cells responsible for adjacent channel Interference.

The main drawback of adjacent channel interference is near-far problem. Near-far problem occurs when a near-far transmitter captures the subscriber's receives that is very close in frequency range while the subscriber attempts to receive a base station on the desired channel.

To avoid adjacent channel interference, three things are important, which are –

1. Careful filtering is required.
2. Frequency separation between each cell should be maximize.
3. Use separate uplink and downlink frequency for transmission.

As we know that cluster size (N) is inversely proportional to frequency reuse factor. For example – if cluster size N= 7, then the frequency reuse factor is 1/7 since each cell contains 1/7th part of total number of available channels.

Therefore for a small value of cluster size, the frequency factor becomes large which indicates insufficient separation between the adjacent channels. It may produce adjacent channel interference to exceed the tolerable limits. Thus overall system capacity get minimize. If there are two mobile units in a cell, first one is close to base station and second one is distance from the base station, then to ignore adjacent channel interference at distance mobile unit, very accurate base station filters are required.

8.7 POWER CONTROL FOR REDUCING INTERFERENCE

In a practical system cellular radio and personal communication systems, the power levels transmitted by every subscriber unit are under constant control by the serving base stations. This is done to confirm that each mobile transmit the smallest power necessary to maintain a good quality link on the reverse channel.

Power control not only helps to prolong battery life for the subscriber unit but also dramatically minimize the reverse channel S/I in the system.

Power control is especially important for emerging CDMA spread spectrum systems that allow every user in every call to use the same radio channel.

8.8 IMPROVING COVERAGE AND CAPACITY IN CELLULAR SYTEMS

As the demand for wireless service increase, the number of channels alloted to a cell eventually becomes insufficient to support the required number of users.

At this point, cellular design techniques are needed to allow more channels per unit coverage area:

The following techniques are used to increase the capacity of cellular systems:

(*i*) Cell splitting

(*ii*) Sectoring

(*iii*) Coverage zone approaches (repeaters for range extension)

(*iv*) Microcell zone concept.

8.8.1 Cell splitting

Cell splitting is the process of subdividing a congested cell into smaller cells, each with its own base station and a corresponding decreasing in antenna height and transmitter power.

Cell splitting increases the capacity of cellular system since it increases the number of times that channels are reduced.

By introducing new cells which have a smaller radius than the original cells and by installing these smaller cells (called microcells) between the existing cells, capacity extend due to the additional number of channels per unit area.

An example of cell splitting is shown in Fig. 8.7. In Fig. the base stations are placed at corners of the cells, and the region served by base station *A* is assumed to be saturated with traffic.

New base stations are therefore needed in the region to increase the number of channels in the area and to minimize the area served by the single base station. Note in Fig. 8.7 that the original base station *A* has been covered by six new microcells.

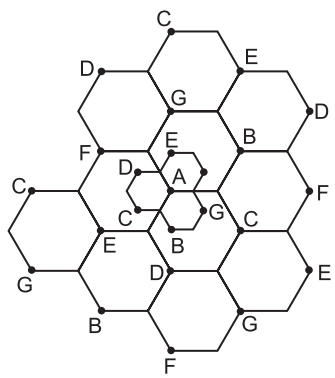


Fig. 8.7: Illustration of cell splitting.

In the example shown in Fig. 8.7, the smaller cells were added in such a way as to preserve the frequency reuse plan of the system. For example, the microcell base station marked *G* was placed half way between two larger stations utilizing the same channel set *G*. This also the case for the other microcells in the Fig. 8.7. Figure shows that cell splitting merely scales the geometry of the cluster. In this case, the radius of each new microcell is half that of the original cell.

For the new cells, which is smaller in size, the transmit power of these cells must be reduced. The transmit power of the new cells with radius half that of the original cells can be determine by examining the received power P_r at the new and old cell boundaries and setting them equal to each other. This is necessary to ensure that the frequency reuse plan for the new microcells behaves exactly as for the original cells. From figure P_r can be written as –

$$\boldsymbol{P_r}\text{ [at old cell boundary] } \alpha\ P_{t1}\ R^{-n} \qquad ...(8.10)$$

$$\text{and } P_r\text{ [}at\text{ new cell boundary] } \alpha P_{t2}\left(\frac{R}{2}\right)^{-n} \qquad ...(8.11)$$

where P_{t1} and P_{t2} are the transmitted power of the larger and smaller cell base station respectively and n is the path loss exponent.

If we take $n = 4$ and set the received power equal to each other then –

$$P_{t2} = \frac{P_{t1}}{16} \qquad ...(8.12)$$

In other words, the transmitted power must be minimize by 12 dB in order to fill in the original coverage area with microcells while maintaining the *S/I* requirement.

Advantage of Cell Splitting

(*i*) Number of channels per unit area increase or maximize.

(*ii*) System capacity increase.

(*iii*) Antenna height is reduced.

(*iv*) Transmitter power is reduced or minimize.

Disadvantage of Cell Splitting

(*i*) Number of base station increases.

(*ii*) Overall lost of the system increases.

8.8.2 Cell Sectoring

In the sectorization technique, neither more base stations are required nor cell division into microcell is done but each cell is splitted into sectors with each sector directional antennas to increases the system capacity by increasing the signal to interference ratio (SIR) or reduce the interference level. SIR is increased by using directional antennas.

As we know that directional antenna radiate in a particular direction not in all directions as done by omidirectional antennas. So the directional antenna reduces interference and therefore increase SIR. Here each cell is splited into more then one sector having one directional antenna at each sector.

The factor by which co-channel interference is reduced depends on the amount of sectoring used.

Fig. 8.8 shows sectoring at different angles. One at 120° and second at 60° sectoring.

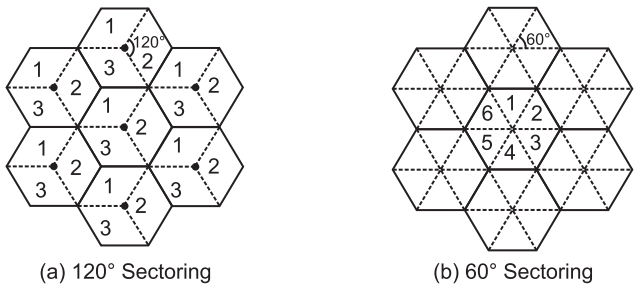


Fig. 8.8: Sectoring at different angles.

Sectoring technique uses more than one antenna per base station or per cell.

The available channels in the cell are subdivided and each division is dedicated to a particular antenna. For example, in 120 sectorization (for cluster size N= 7), each cell is subdivided into three sectors and each sector consists its own directional antenna and channel frequency.

Fig. 8.9, define the 120° sectoring reduces interference from co-channel cells.

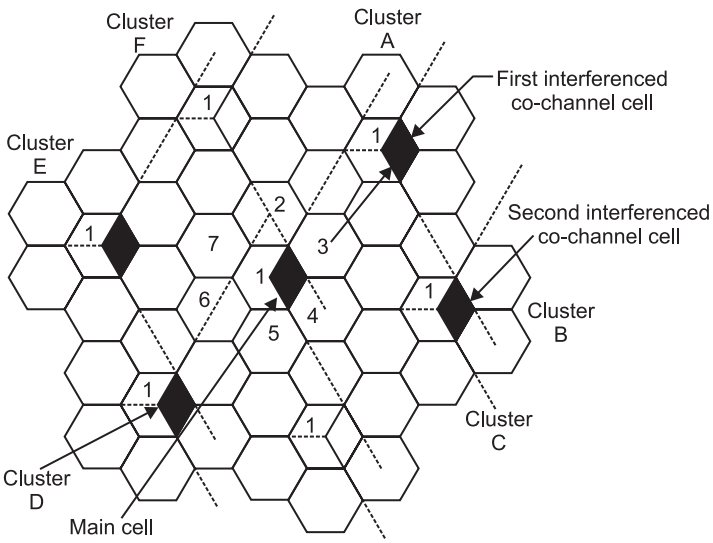


Fig. 8.9: Reduced interference in co-channel cells (showing by no. 1) by using sectorization.

Before sectoring, six co-channel cell interfere with the centre center cell but after sectoring (120°) only two co-channel cells are reserve to interfere.

Co-channel cell (1) of cluster A and B gets interfere due to main cell after sectorization while interference because these are at anti side of sectorization and not lie in the interference region of main cell.

In practice, the reduction in interference enables the designers to reduce the cluster size N, and consequently a small value of N increases the capacity and therefore a small value of N increase the capacity and support an additional degree of freedom assigning channels.

Advantage of Sectoring

(*i*) Reduces co-channel interference with in a system.

(*ii*) Increases capacity by increasing signal to interference ratio (SIR).

(*iii*) Cell radius gets unchanged (as it was half in splitting technique by introducing new microcells in a system).

Disadvantage of Sectoring

(*i*) Number of antennas per cell get increased.

(*ii*) Thinking efficiency decreases due to channel sectoring at the base station because available trunked channel poor is splitted into several smaller pools.

(*iii*) Sectoring makes handoff more typical because mobiles crossing the boundary of sector to another sector within the same cell demanded handoff.

8.8.3 Repeaters for Range Extension

Often a wireless operator need to allow dedicated coverage for hard to reach areas, such as within buildings or in valleys or tunnels, radio transmitters, known as repeaters, are often used to provide such range extension capabilities to the system.

By using repeaters, coverage area can be improved or we can say range can be extended but it does not extend the capacity of the system.

Repeaters are simply "Radio retransmitters" bidirectional in nature, that provide range extension capabilities to the system.

Repeaters receive signals from base station (in forward link) then amplify that signal and retransmit to the specific coverage area.

Repeaters work using over the air signals, so they may be installed anywhere and are capable of repeating an entire cellular.

By using repeaters, radius of cell remain unchanged means constant.

Cost of the system will increase by using repeater.

The main drawback is that it amplify the noise and interference along with the signals. Consequently a case must be taken to properly place the repeaters to minimize the noise and interference.

8.8.4 Microcell Zone Concept

This technique is being adopted in most cellular and personal communication systems. A microcell zone concept is quite similar to sectoring but all the drawback faced during sectoring are removed in microcell zone concept. And this concept is explain as the following.

This proposal is based on microcell concept for seven cell reuse ratio as illustrated in Fig. 8.10 with T_r and R_x.

In this scheme, each of three (or possibly more) zone sites (represented as T_x/R_x in Fig. 8.10) are connected to a single base station and use the same radio

equipment. The zone are connected by coaxial cable or microwave link to the base station.

Multiple zones and a single base station make up a cell. As a mobile travels within the cell, it is provide by zone with the strongest signal.

This approach is superior to sectoring since antenna are placed at the edges of the cell, and any base station channel may be alloted to any zone by the base station.

Each base station has its own channel set that is collectively used by the cell zones in time and space pattern. This means same channel can be reserve by a mobile travelling from one zone to another within the cell, by simply switching the channel to the second zone using zone selector. This eliminate the requirement of handoff from one zone to another.

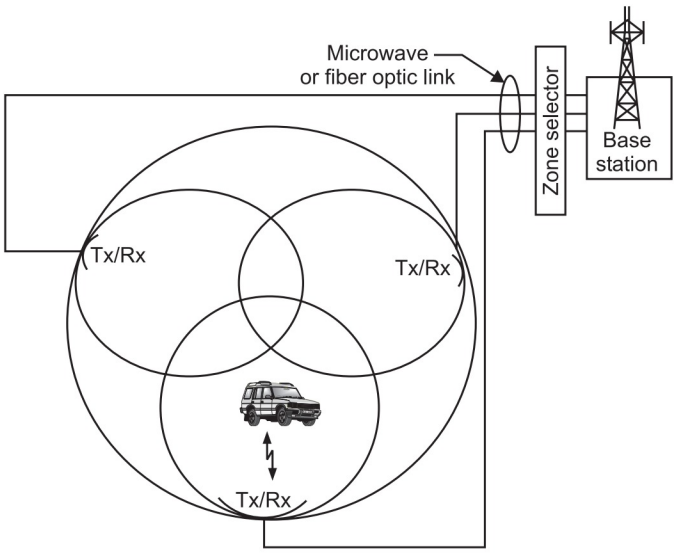


Fig. 8.10: The microcell concept.

This technique is particularly useful along highways or along urban traffic corridors or at both.

Interference is reduced in this technique because a given channel is working only in that zone where the mobile is travelling.

Advantage of Microcell Zone Concept

(*i*) Only one base station can reserve for more than one zone.

(*ii*) Any base station channel may be assigned to any zone by zone selector.

(*iii*) Interference is reduced since a large central base station is replaced by various lower powered transmitter on the edge of the zone.

(*iv*) Co-channel cells may be formed in the normal fassion that uses the same frequency.

(*v*) Improves the signal quality by reducing the co-channel interference.

(*vi*) System capacity increases without any degradation in trunking efficiency because the capacity is directly related to the distance between co-channel cells (and not zones).

8.9 COMPARISON AMONG CELL SPLITTING CELL SECTORING MICRO CELL ZONE CONCEPT INCREASING NUMBER OF REPEATERS

S. No.	Parameters	Cell splitting	Cell sectoring	Micro cell zone concept	Increasing number of repeaters
1.	Process	Each cell is divided into smaller radius cells.	Each cell is divided into sectors of equal area having centred mid of the main cell.	Each cell is formed using many zones.	Repeaters are used to cover more coverage area.
2.	Base station	Each smaller cell requires new base station.	No more base stations are required.	Each cell has only one base station.	Repeaters are required.
3.	Interference	Interference gets increased due to nearer co-channel cells.	Co-channel interference is reduced.	Interference is reduced.	Interference is increased due to amplification.
4.	System capacity	System capacity increases due to increase in the number of cluster.	System capacity increased by increase in the signal to interference ratio (SIR).	System capacity increases without any degradation in trunking efficiency.	System capacity remains unchanged.
5.	Radius of cell	Radius of cell decreases.	Rading of cell remains unchanged.	Radius of cell increases.	No effect at radius of cell.
6.	Cost	Cost increase due to new base stations.	Cost increases due more antennas.	Case decreases due to less base station.	Increases due to repeaters.

SOLVED EXAMPLES

Example 8.1: *If a total of* 33 *MHz of bandwidth is allocated to a particular FDD cellular telephone system which uses two* 25 *kHz simplex channels to provide full duplex voice and control channels, compute the number of channels available per cell if a system uses* (*a*) *four cell reuse* (*b*) *seven cell reuse* (*c*) Twelve *cell reuse.*

Solution: Given total bandwidth = 33 MHz

Channel bandwidth = 25 kHz × 2 simplex channels = 50 kHz/duplex channel

Total available channels = 33,000/50 = 660 channels

(*a*) For $N = 4$

Total number of channels available per cell = 660/4 =165 channels

(*b*) For $N = 7$

Total number of channels available per cell = 660/7 = 95 channels

(*c*) Fori $N = 12$

Total number of channels available per cell = 660/12 = 55 channels.

Example 8.2: *If a signal to interference rati of* 15 *dB is required for satisfactory forward channel performance of a cellular system, what is the frequency reuse factor and cluster size that should be used for maximum capacity if the path loss exponent is* (*a*) $n = 4$, (*b*) $n = 3$. *Assume that there are six co-channel cells in the first tier and all of them are at the same distance from the mobile. Use suitable approximations.*

Solution: (*a*) $n = 4$

First let us consider a seven cell pattern by using equation (8.4) the co-channel reuse ratio $D/R = 4.583$.

By using equation (8.9) the signal to noise interference ratio is given by

$$S/I = 1/6 \times (4.583)^4 = 75.3 = 18.66 \text{ dB}$$

Since this is greater than the minimum required $SII,N = 7$ can be used.

(*b*) $n = 3$

First let us consider a seven cell reuse pattern

Using equation (8.9) : $S/I = 1/6 \times (4.583)^3 = 16.04 = 12.05$ dB

Since this is less than the minimum required S/I, we need to use a larger N, 0.

Using equation (8.3), the next possible value of Wis 12 ($i = j = 2$)

The corresponding co-channel ratio is given by equation (4) as $D/R = 6.0$

Using equation (8.3), the signal to interference ratio is given by

$S/I = 1/6 \times (6)^3 = 36 = 15.56$ dB

Since this is greater then the minimum required $S/I = N = 12$ is used.

Example 8.3: *A hexagonal cell within a four cell system has a radius of* 1.387 *km. A total of* 60 *channels are used within the entire system. If the load per user is* 0.029 *Erlangs and* $\lambda - 1$ *call/hour, compute the following for an Erlang C system that has a* 5% *probability of a delayed cell.*

(*a*) *How many users per square kilometer will this system support ?*

(*b*) *What is probability that a delayed call will have to wait for more than 10 second?*

(*c*) *What is probability that a call will be delayed for more than 10 seconds ?*

Solution: Call radius, $R = 1.387$ km

Area covered per cell is $2.598 \times (1.387)^2 = 5$ sq. km

Number of cell per cluster = 4

Total number of channels = 60

Therefore, number of channels per cell = 60/4 =15 channels

(*a*) From Erlang C chart for 5% probability of delay with $C = 15$ traffic intensity = 9.0 Erlangs.

Therefore, number of users = total traffic intensity/traffic per user

= 9.0/0.029 = 310 users

= 310 users/5 sq. km = 62 users/sq. km

(*b*) Given λ, = 1, holding time

$$H = A/\lambda = 0.029 \text{ hour} = 104.4 \text{ second}$$

The probability that a delayed call will have to wait longer than 10 second is

$$P_r[\text{delay} > t \mid \text{delay} \mid] = exp[-(C - A)\, t/H]$$
$$= \exp[-(15 - 9.0)10/104.4] = 56.29\%$$

(*c*) Given $p_r[\text{delay} > 0] = 5\% = 0.05$

Probability that a call is delayed more than 10 second

$$P_r[\text{delay} > 0] = P_r[\text{delay} > 0]\, P_r\,[\text{delay} > t\ (\text{delay})]$$
$$= 0.05 \times 0.5629 = 2.81\%$$

Example 8.4: *Given a cellular system with a total bandwidth of* 30 *MHz which uses two* 25 *kHz simplex channels to provide full duplex voice and control channels. Assuming that the system uses a nine-cell reuse pattern and* 1 *MHz of the total bandwidth is allocated for control channels.*

(*i*) *Calculate the total available channels.*

(*ii*) *Determine the number of control channels.*

(*iii*) *Determine the number of the voice channel per cell.*

(*iv*) *Determine an equitable distribution of control and voice channel in each cell.*

Solution: Given that

Total bandwidth = 30 MHz

Channel bandwidth = 25 kHz × 2 = 50 kHz/duplex channel

(*i*) Total number of available channels = $\frac{30000}{50} = 600$ = 600

(*ii*) The number of control channels = $\frac{1000}{50} = 20$

(*iii*) The number of voice channel per cell = $\frac{1000}{50} = 20$

(*iv*) Since, only a maximum of 20 channels can be used as control channel for *N=9,* one way to allocate is 7 cells with two control channels and 64 voice channels each, and 2 cells with 3 control channels and 66 voice channels each.

Example 8.5: *Verify that the cell cluster size is* $N = i^2 + ij + j^2$, *where i and j are the integer parameters determine the co-channel cells as shown in Fig.8.11.*

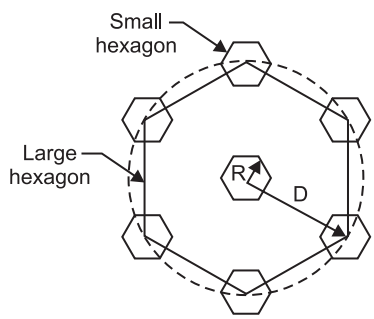


Fig. 8.11: First tier co-channel interfering cells.

Solution: A candidate cell has 6 nearest co-channel cells. By joining the centres of the 6 nearest neighbouring co-channel cells, we form a large hexagon as shown in Fig. 8.11. This large hexagon has radius equal to *D,* which is also the co-channel cell separation.

With the cell radius *R,* we have

$$D = \sqrt{3}RD_{\text{norm}} = \sqrt{3(i^2 + ij + j^2)R}$$

In general the area of a hexagon is proportional to the square of its radius. Let $P(\cong 2.598)$ be the proportional constant. Then the area of large hexagon with radius *D* is

$$A_{\text{large}} = \beta D^2 = \beta[3(i^2 + ij + j^2)R^2]$$

and the area of a cell (the small hexagon) with radius *R* will be

$$A_{\text{small}} = \beta R^2$$

The number of cell in the large hexagon is then

$$\frac{A_{\text{large}}}{A_{\text{small}}} = 3(i^2 + j^2 + ij) \qquad \text{... } (i)$$

On other hand, from the geometry, it can be seen that, in general, **the large** hexagon enclose the center cluster of Af-cells plus 1/3 the number of cells associated with six other peripheral large hexagons. Therefore the total number of cells enclosed by the large hexagon is given by

$$N + 6\left(\frac{1}{3}N\right) = 3N \qquad \text{... } (ii)$$

From equation (*i*) and (*ii*) we get

$$N = i^2 + ij + j^2$$

Hence proved.

Example 8.6: *The acceptable signal to co-channel interference ratio in a certain cellular communication situation is S/I = 20 dB or 100. Also from measurements it is determined that k = 4. What will be the minimum cluster size ?*

Solution: The frequency reuse ratio can be calculated as under

$$q = (6 \times 100)^{1/4} = 4.9492$$

The cluster size is given by

$$N = \frac{q^2}{3} = 8.165 \cong 9$$

In this case, a 9 reuse pattern is required for an *S/I* ratio of at least 20 dB.

Since $\quad q = \dfrac{D}{R}$ or $D = qR$

D can be determined, given the cell radius *R* and vice-versa.

It may be noted that if *N* is less than 9, the *S/I* value would be below the acceptable level of 20 dB.

$$\frac{S}{I} = \frac{R^{-k}}{2(D-R)^{-k} + 2D^{-k} + 2(D+R)^{-k}}$$

Recall that $D/R = q$ with a path loss exponent of $k = 4$ so equation can be written as

$$\frac{S}{I} = \frac{1}{2(q-1)^{-4} + 2q^{-4} + 2(q+1)^{-4}}$$

Example 8.7: *Given a cellular system in which there are a total of 1001 radio channels available for handling traffic. It is also given that the area of a cell is 6 km and the area of the entire system is 2100 km^2.*

(*a*) *Calculate the system capacity if the cluster size is 7.*

(*b*) *How many times would the cluster of size 4 have to be replicated in order to approximately cover the entire cellular area ?*

(*c*) *Calculate the system capacity if the cluster size is 4.*

(*d*) Does decreasing the cluster size increase the system capacity? Explain.

Solution: Given:

The total number of available channels $k = 1001$

Cluster size $N = 7$

Area of cell $A_{\text{cell}} = 6$ km^2

Area of cellular system $A_{sys} = 2100$ km^2

(*a*) Because, the number of channels per cell is given by

$$J = \frac{k}{N}$$

Then, we have

$$j = \frac{1001}{7} = 143 \text{ channels/cell}$$

Also, we have know that the coverage area of a cluster is given by

$$A_{\text{Cluster}} = N \times A_{\text{cell}} = 7 \times 6 = 42 \text{ km}^2$$

The number of times that the cluster has to be replicated to cover the entire system will be

$$M = \frac{A_{\text{sys}}}{A_{\text{cluster}}}$$

$$M = \frac{2100}{42} = 50$$

Hence, the system capacity C will be

$$C = MJN$$

$$C = 50 \times 143 \text{ x } 7$$

$$C = 50{,}050 \text{ channels}$$

(*b*) For $N = 4$, $A_{\text{cluster}} = 4 \times 6 = 24$ km^2

$$M = \frac{2100}{24} = 87.5 = 84$$

(*c*) With $N = 4$, $J = \frac{1001}{4} = 250$ channels/cell

The system capacity is then

$$C = 87 \times 250 \times 4 = 87{,}000 \text{ channels}$$

(*d*) From (*a*) and (*c*), it is obvious that a decrease in N from 7 to 4 is accompanied by an increase in M from 50 to 87, and the system capacity is increased from 50,050 channels to 87,000 channels. Hence, decreasing the cluster size does increase the system capacity.

Example 8.8: *Given a cellular system with 416 radio channels available for handling traffic. Assume that 21 of these channels are designated as control channels. Let the average channel holding time of a call be 3 minutes, the blocking probability during busy hours be 2% and the frequency reuse factor be 9.*

(*a*) *Determine the number of calls per cell per hour.*

(*b*) *Determine the signal-to-co-channel interference ratio, S/I in dB.*

Solution:

(*a*) In a cellular system with frequency reuse, all the available radio channels are allocated to handle services in a single cluster of cells. Therefore, the number of voice channels in a cluster is

$$416 - 21 = 395$$

With a frequency reuse factor of 9, the effective number of voice channels and a 2% blocking probability, the traffic load is 34.683 Erlangs.

The number of calls per unit time = $\frac{\text{Number of calls per cell}}{\text{Average holding time of a call}}$

Therefore, the number of calls per cell per unit time

$$= \frac{34.683}{3} \text{ calls per cell per minute}$$

$$= \frac{34.686}{3} \times 60 = 693 \text{ calls per cell per hour}$$

(*b*) The frequency reuse ratio $q = \sqrt{3N} = \sqrt{3 \times 9} = 5.1962$.

For $k = 4$, the signal-to-co-channel interference ratio will be given by

$$\frac{S}{I} = \frac{1}{6} \times q^k$$

$$\frac{S}{I} = \frac{1}{6} \times (5.1962)^2$$

$$= 121.5 \text{ or } 20.845 \textbf{ dB}$$

Example 8.9: *An urban area has a population of two million residents. Three competing trunked mobile networks (system A, B and C) provide cellular service in*

this area. System A has 394 cells with 19 channels each, system B has 98 cells with 57 channels each, and system C has 49 cells, each with 100 channels. Find the number of users that can be supported at 2% blocking if each user averages two calls per hour at an average call duration of three minutes. Assuming that all three trunked systems are operated at maximum capacity, compute the percentage market penetration of each cellular provider.

Solution: System *A*

Given:

Probability of blocking = 2% = 0.02

Number of channels per cell used in the system, C = 19

Traffic intensity per user, $A_u = \lambda H = 2 \times \left(\frac{3}{60}\right) = 0.1$ Erlangs

For GOS = 0.02 and C = 19, from the Erlang B chart, the total carried traffic, A is obtained as 12 Erlangs.

Therefore, the number of users that can be supported per cell is

$$U = \frac{A}{A_u} = \frac{12}{0.1} = 120$$

Such there are 394 cells, the total number of subscribers that can be supported by system A is equal to 120 × 394 = 47280

System *B*

Given:

Probability of blocking = 2% = 0.02

Number of channels per cell used in the system, C = 57

Traffic intensity per user, $A_U = \lambda H = 2 \times \left(\frac{3}{60}\right) = 0.1$ Erlangs

For GOS = 0.02 and C = 57, from the Erlang B chart, the total carried traffic, A is obtained as 45 Erlangs.

Therefore, the number of users that can be supported per cell is

$$U = \frac{A}{A_u} = \frac{45}{0.1} = 450$$

Since there are 98 cells, the total number of subscribers that can be supported by system B is equal to 450 × 98 = 44,100.

System *C*

Given:

Probability of blocking = 2% = 0.02

Number of channels per cell used in the system C = 100

Traffic intensity per user, $A_u = \lambda H = 2 \times \left(\frac{3}{60}\right) = 0.1$ Erlangs

For GOS = 0.02 and C = 100, from the Erlang B chart, the total carried traffic, A is obtained as 88 Erlangs.

Therefore, the number of users that can be supported per cell is

$$U = \frac{A}{A_u} = \frac{88}{0.1} = 880$$

Since there are 49 cells, the total number of subscribers that can be supported by system C is equal to 880 × 49 = 43,120.

Therefore, total number of cellular subscribers that can be supported by these three system are 47,280 + 44,100 + 43,120 = 134,500 users.

Since there are two million residents in the given urban area and the total number of cellular subscribers in system A is equal to 47280.

The percentage market penetration is equal to

$$\frac{47,280}{2,000,000} = 2.36\%$$

Similarly, market penetration of system B is equal to

$$\frac{44,100}{2,000,000} = 2.205\%$$

And the market penetration of system C is equal to

$$\frac{43,120}{2,000,000} = 2.156\%$$

The market penetration of the three systems combined is equal to

$$\frac{134,500}{2,000,000} = 6.725\%$$

Example 8.10: *A certain city has an area of* 1,300 *square miles and is covered by a cellular system using a seven-cell reuse pattern. Each cell has a radius of four miles and the city is allocated* 40 *MHz of spectrum with a full duplex channel bandwidth of* 60 *kHz. Assume a GOS of* 2% *for an Erlang B system is specified. If the offered traffic per user is 0.03 Erlangs, compute:*

(*a*) *The number of cells in service area*

(*b*) *The number of channels per cell*

(*c*) *Traffic intensity of each cell*

(*d*) *The maximum carried traffic*

(*e*) *The total number of users that can be served for* 2% *GOS*

(*f*) *The number of mobiles per unique channel (where it is understood that channels are reused*)

(*g*) *The theoretical maximum number of users that could be served at one time by the system.*

Solution:

(*a*) Given:

Total coverage area = 1300 miles, and cell radius = 4 miles

The area of a cell (hexagon) can be shown to be 2.5981 R^2, thus each cell covers 2.5981 × $(4)^2$ = 41.57 sq. meter

(*b*) The total number of channels per cell (*C*)

$$= \frac{\text{Allocated spectrum}}{(\text{Channel width} \times \text{Frequency reuse factor})}$$

$$= \frac{40{,}000{,}000}{(60{,}000 \times 7)} = 95 \text{ channels/cell}$$

(*c*) Given:

$$C = 95, \text{ and GOS} = 0.02$$

From the Erlang *B* chart, we have traffic intensity per cell *A* = *84* Erlangs/cell

(*d*) Maximum carried traffic = number of cells x traffic intensity per cell

= 34 x 84 = 2604 Erlangs

(*e*) Given traffic per user = 0.03 Erlangs

$$\text{Total number of users} = \frac{\text{Total traffic}}{\text{Traffic per user}}$$

$$= \frac{2604}{0.03} = 86{,}800 \text{ users}$$

(*f*) Number of mobiles per channel = $\frac{\text{Number of users}}{\text{Number of channel}}$

$$= \frac{86{,}800}{666} = 130 \text{ mobiles/channel}$$

(g) The theoretical maximum number of served mobiles is the number of available channels in the system (all channels occupied)

$= C \times N_C = 95 \times 31 = 2945$ users, which is 3.4% of the customer base.

Example 8.11: *A hexagonal cell within a four-cell system as a radius of 1.387 km. A total of 60 channels are used within the entire system. If the load users is 0.029 Erlangs and* λ *= 1 call/hour, compute the following an Erlang C system that has a 5% probability of a delayed call.*

(*a*) *How many users per square kilometer will system support ?*

(*b*) *What is the probability that a delayed call will have to wait for more than 10 seconds?*

(*c*) *What is the probability that a call will be delayed for more than 10 seconds?*

Solution: Given:

Cell radius, R= 1.387 km

Area covered per cell is 2.598 × (1.387)2 = 5 sq. km

Number of cells per cluster = 4

Total number of channels = 60

Therefore, number of channels per cell = $\frac{60}{4} = 15$ channels

(*a*) From Erlang *C* chart, for 5% probability of delay with $C = 15$, traffic intensity = 9.0 Erlangs.

$$\text{Therefore, number of users} = \frac{\text{Total traffic intensity}}{\text{Traffic per user}}$$

$$\frac{9.0}{0.29} = 310 \text{ users}$$

$$= \frac{9.0}{0.29} = 310 \text{ users}$$

$$= 310 \text{ users/5 sq.} = 62 \text{ users/sq. km}$$

(*b*) $\lambda = 1$, holding time

$$H = \frac{A_u}{\lambda} = 0.029 \text{ hour} = 104.4 \text{ seconds}$$

The probability that a delayed call will have to wait longer than 10 seconds

$$P_r\,[(\text{delay} > t)\text{delay}] = \exp\left[\frac{-(C-A)t}{H}\right]$$

$$= \exp\left[\frac{-(-15-9.0)10}{104.4}\right] = 56.29\%$$

(*c*) Given P_r [delay > 0] = 5% = 0.05

Probability that a call is delayed more than 10 seconds

$$P_r\,[\text{delay} > 10] = P_r\,[\text{delay} > 0]\; P_r\,[\text{delay} > t \mid \text{delay}]$$

$$= 0.05 \times 0.5629 = 2.81\%$$

HIGHLIGHTS

1. A large geographical area is subdivided into smaller cells by using frequency reuse concept.
2. In frequency reuse concept, a frequency can be reused by some other co-channel cell that is may be sufficiently apart (or distant) from the home cell to avoid interference or distortion.
3. By using hexagonal shape for a cell, a fewer number of cells can cover a geographic region because the hexagons have the largest area over circle, triangle or square shape of cells.

4. The number of cells which collectively use the complete set of allocated frequencies is called a 'cluster'.
5. Large cluster size minimizes co-channel interference while small cluster size increase capacity.
6. Frequency reuse ratio is defined as the following:

$$Q = \frac{D}{R}\sqrt{3N}$$

where D is the distance between two co-channel cells. R is the radius of the cell and N is cluster size.

7. In fixed channel assignment strategy, a fixed set of frequencies is allocated to each cell.
8. In dynamic channel assignment strategy, no fixed set is defined but whole voice channel is dynamic in nature *i. e.* Allocated to a cell when there is a request for call establishment or call arrangement.
9. When mobile moves from one cell coverage area to other cell then hand-off is required. Hand-off transfers the call from one base station (BIS1) to another base station (BIS2) or one MSC to other MSC.
10. The increase in system capacity comes from the use of smaller cells, reuse of frequencies and antenna is sectoring. The amplification of frequency reuse and the gain obtainable from cell splitting and antenna sectoring are explained and discussed in detail.
11. A wider spectrum will allow for greater sharing of the spectral width to support more users and increase the through out and system capacity can be interpreted as the largest number of users that can be supported during any one use of the channel and system.
12. Any cellular radio system mainly depends upon an intelligent assignment and reuse of channels (*i*.e. may be frequency) throughout a coverage regions. Every cellular base station is assigned a group of radio channels which are used with in a small geographic area called a cell.
13. Cells which use the same set of frequencies which may be reffered to as co-channel interference. Frequencies or channels represent radio resources. Here, we shall use the terms frequencies and channels interchangeably.
14. If the different cells in the entire cellular system were to use different sets of frequencies, intercell interference would be kept at a minimum. However, the system capacity would be limited. In fact, deployment of frequency reuse is necessary to enlarge the system capacity.
15. Generally, there are two important questions for define connection–
    (*i*) How successfully can a new user get a connection established?
    (*ii*) After connection establishment, how successfully will the connection be maintained as the user moves from one cell to another?
16. The capacity of a cellular system can be enlarged through frequency reuse. The capacity can also be improve based on cellular layout and antenna design. Basically, there are following three popular mechanisms to increase capacity of a cellular system –

(*i*) Cell splitting
(*ii*) Antenna sectoring
(*iii*) Microcell zone concept.

17. The cellular concept was a major breakthrough in solving the problem of spectral congestion and user capacity. It support very high capacity in a limited spectrum allocation without any major or minor technological changes.

18. As the demand for service increase *(i.e.* more channel are demanded), the number of base stations may be increased (along with a corresponding decreases in transmitted power to avoid added interference), thereby allow additional radio capacity with no additional increase in radio spectrum.

19. The design process of selecting and assigning the channel groups for all of the cellular base station within a system is called frequency reuse or frequency planning concept.

20. The frequency reuse factor of a cellular system is given by $1/N$ since each cell with in a cluster is only allot the l/N of the total available channels in the system.

21. The main function of using channel assignment strategies is to optimize utilization of frequency reuse scheme and minimization of interference so that the overall channel capacity can be increased and this strategies can be classified into two parts –

(*i*) Fixed channel assignment strategy
(*ii*) Dynamic channel assignment strategy.

22. The following technique are used to increase the capacity of cellular system–

(*i*) Cell splitting
(*ii*) Sectoring
(*iii*) Coverage zone approaches (repeaters for range extension) (v) Microcell zone concept.

23. Cell splitting is the process of subdividing a congested cell into smaller cells, each with its own base station and a corresponding decreasing in antenna height and transmitted power.

24. In the sectorization technique, neither more base stations are required nor cell division into microcell is done but each cell is splitted into sectors with each sector directional antennas to increases the system capacity by increasing the signal to interference ratio (SIR) or reduce the interference level. SIR is increased by using directional antennas.

25. By using repeaters, coverage area can be improved or we can say range can be extended but it does not extend the capacity of the system. And repeaters are simply radio retransmitters and bidirectional in nature and that provide range extension capabilities to the system.

26. Microcell zone concept is being adopted in most cellular and personal communication system and it is similar to sectoring but all the drawback faced during sectoring are removed in microcell zone concept.

27. The system capacity of the system is increased by cell splitting, cell sectoring and microcell zone concept but the system capacity remain unchanged by increasing the number repeaters.

PROBLEMS

8.1. Show that the frequency reuse factor for a cellular system is given by k/s, when k is the average number of channels per cell and s is the total number of channels available to the cellular service provider.

8.2. If 20 MHz of total spectrum is allocated for a duplex wireless cellular system and each simplex channel has 25 kHz RF bandwidth find –

(*a*) The number of duplex channels

(*b*) The total number of channels per cell site, if N = 4 cell reuse is used.

8.3. A cellular service provider decides to use a digital TDMA scheme which can tolerate a signal to interference ratio of 15 dB in the worst case. Find the optimal value of N for –

(*a*) Omnidirectional antennas

(*b*) 120° sectoring

(*c*) 60° sectoring

Should sectoring be used ? If so, which case (60° or 120°) should be used ? (Assume a path loss exponent of N= 4 and consider trunking efficiency).

8.4. What is the average distance (in terms of R) between the mobile on the fringe of the serving cell and the first tier of co-channel cells ? (These cells are called the "nearest neighbours"). How many cells are located in the first tier ? Solve for the case of $2N$= 1, N = 3, N= 4, N = 7 and N = 12 cluster sizes. How does the average distance compare to the value of $D = QR$ where $Q = \sqrt{3N}$.

8.5. What is the average distance (in terms of./?) between the mobile on the fringe of the serving cell and the second and third tier of co-channel cells and how many cells are in the second and third tier of co-channel cells for the cases of N = 1, N = 3, N = 4, N = 7 and N = 12 cluster sizes?

8.6. Determine the forward link $C/2$ for the following frequency reuse designs: N = 1, N=3, N = 4, N =7 and N = 12. Assume that the propagation path loss exponent is four, and evaluate the S/I contribution due to just the first tier and then due to additional outer tiers of co-channel cells. Indicate the number of tiers at which there is a diminishing contribution to the interference at the mobile.

8.7. Repeat part 6 except now consider a line-of-sight path loss exponent of n = 2. Notice the huge impact that the propagation path loss exponent has on C/I. What can you say about the cluster size, path loss exponent and the C/I values which result ? How would this impact practical wireless system design?

8.8. Suppose that a mobile station is moving along a straight line between base stations BS_1 and $BS_{2'}$ as shown in Fig. 8.12. The distance between the base stations is D = 2000 m. For simplicity, assume small scale fading is neglected and the received power (in dB_m) at base station i, from the mobile station, *is* modeled as a function of distance on the reverse link.

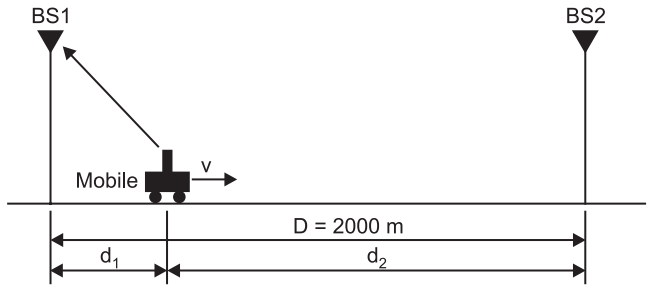


Fig. 8.12: Cellular system with two base stations.

$$P_{ri}(d) = P_0 - 10n \log_{10}\left(\frac{d_i}{d_0}\right)(\text{dB}_m)\, i = 1,2$$

where d_t is the distance between the mobile and the base station i, in meters. P_0 is the received power at distance d_0 from the mobile antenna. Assume that $P_0 = 0$ dB$_m$ and d_0 = 1 m. Let n denote the path loss which is assumed to be equal to 2.9.

Assume the minimum usable signal level for acceptable voice quality at the base station receiver is $P_{r,\min} = -88$ dB$_m$ and the threshold level used by the switch for handoff initiation is $P_{r\ H0}$. Consider that the mobile is currently connected to BS_1 and is moving toward a handoff (time required to complete a handoff, once that received signal level reaches the handoff threshold $P_{r,\ H0}$ = is Δt 4.5 seconds).

(*a*) Determine the minimum required margin $\Delta = P_{rHO} - P_{r,}$ min to assure that calls are not lost due to weak signal condition during handoff. Assume that the base station antenna heights are negligible compared to the distance between the mobile and the base stations.

(*b*) Describe the effects of the margin $\Delta = P_{rHO} - Pr,$ min on the performance of cellular systems.

8.9. If an intensive propagation measurement compaign showed that the mobile radio channel provided a propagation path loss exponent of $n = 3$ instead of four, how would your design decisions in problem 3 change ? What is the optimal value of N for the case of $n = 3$?

8.10. A total of 24 MHz of bandwidth is allocated to a particular FDD cellular telephone system that uses two 30 kHz simplex channels to provide full duplex voice and control channels.

Assume each cell phone user generates 0.1 Erlangs of traffic. Assume Erlang B is used.

(*a*) Find the number of channels in each cell for a four-cell reuse system.

(*b*) If each cell is to offer capacity that is 90% of perfect scheduling. Find the maximum number of users that can be supported per cell where omnidirectional antennas are used at each base station.

(*c*) What is the blocking probability of the system in (*b*) when the maximum number of users are available in the user pool ?

(*d*) If each new cell now uses 120° sectoring instead of omnidirectional for each base station. What is the new total number of users that can be supported per cell for the same blocking probability as in (C) ?

(*e*) If each cell covers five square kilometers, then how many subscribers could be supported in an urban market that is 50 km × 50 km for the case of omnidirectional base station antennas ?

(*f*) If each cell covers five square kilometers, then how many subscribers could be supported in an urban market that is 50 km × 50 km for the case of 120° sectored antennas ?

8.11. A certain area is covered by a cellular radio system with 84 cells and a cluster size N= 300 voice channels are available for the system. Users are uniformly distributed over the area covered by the cellular system and the offered traffic per user is 0.04 Erlang. Assume that blocked calls are cleared and the designated blocking probability is P_b = 1%.

(*a*) Determine the maximum carried traffic per cell if cluster size N= 4 is used. Repeat for cluster sizes N = 7 and 12.

(*b*) Determine the maximum number of users that can be served by the system for a blocking probability of 1% and cluster size N= 4. Repeat for cluster sizes N = 7 and 12.

8.12. A receiver in an urban cellular radio system detects a 1 mw signal at $d = d_0$ = 1 meter from the transmitter. In order to mitigate co-channel interference effects, it is required that the signal received at any base station receiver from another base station transmitter which operates with the same channel must be below -100 dBm. A measurement team has determined that the average path loss exponent in the system is n = 3. Determine the major radius of each cell if a sever-cell reuse pattern is used. What is the major radius if a four cell reuse pattern is used?

8.13. Show that if n = 4, a cell can be split into four smaller cells, each with half the radius and 1/6 of the transmitter power of the original cell. If extensive measurements show that the path loss exponent is three, how should the transmitter power be changed in order to split a cell into four smaller cells ? What impact will this have on the cellular geometry ? Explain your answer and provide drawings that show how the new cells would fit within the original macrocells. For simplicity use omnidirectional antennas.

REVIEW QUESTIONS

8.1. What do you understand by frequency reuse concept ? Define in brief.

8.2. What is the main reason to accept hexagonal shape for construction of a cell.

8.3. Define co-channel reuse ratio in detail.

8.4. What is the main difference between fixed channel assignment strategy and dynamic channel assignment strategy?

8.5. What is handoff ? Draw a step-by-step diagram of call handoff.

8.6. Illustrate the various advantages of handoff techniques.

8.7. What do you understand by inteference? List major sources of interference,

8.8. How system capacity gets affected by interference?

8.9. What do you understand by adjacent channel interference ? What are its major drawbacks?

8.10. What are the ways by which capacity of a cellular system can be increased.

8.11. What is the main difference between cell sectoring and microcell zone concept ?

8.12. Given a cellular system with hexagonal cells. The cell radius is R. The service area is partitioned into cell clusters. Frequency is reused from cell cluster to cell cluster. The geometric relation between adjacent co-channel cells can be described by the two non-negative integers i and j.

(*i*) With pictorial illustriation, explain what i and j represent?

(*ii*) Determine the distance D between the centers of two adjacent co-channel cells.

(*iii*) Derive an expression for the number of cells N in each cell cluster.

(*iv*) What should be considered in choosing a value for N?

8.13. Prove that the frequency reuse factor for a cellular system is given by K/J, where J is the average number of channels per cell and K is the total number of channels available to the cellular network provider without frequency reuse. The cell cluster can be replicated M times to provide a total capacity of C channels. Discuss the changes in the value of C when you increase or decrease the frequency reuse factor while keeping K constant.

8.14. Give a point to point brief note on channel planning for wireless system.

8.15. Explain the difference between improper and a proper handoff situation.

8.16. Discuss a method of locating co-channel cells in a cellular system.

8.17. What are the ways to improve coverage and capacity of cellular system?

9
Chapter

WIRELESS NETWORKING

9.1. *Introduction to Wireless Networks.* **9.2.** *Difference Between Wireless and Fixed Telephone Networks.* **9.3.** *The Public Switched Telephone Network* (*PSTN*). **9.4.** *Limitations in Wireless Networking.* **9.5.** *Merging Wireless Networks and the PSTN.* **9.6.** *Evolution and Development of Wireless Networks.* **9.7.** *Traffic Routing In Wireless Networks.* **9.8.** *Common Channel Signaling* (*CCS*). **9.9.** *Integrated Services Digital Network* (*ISDN*). **9.10.** *Broadband ISDN and ATM.* **9.11.** *Signaling System Number 7* (*SS7*). **9.12.** *Personal Communication Services/Networks* (*PCS/PCNS*). **9.13.** *Packet Vs. Circuit Switching for PCN* **9.14.** *Protocol For Network Access.* **9.15.** *Network Databases.*

9.1 INTRODUCTION TO WIRELESS NETWORKS

The demand for ubiquitous personal communications is driving the development of new networking techniques that correspond to mobile voice and data users who move throughout building, cities, or countries. Consider the cellular telephone system shown in Fig. 9.1 by block diagram.

The cellular telephone system is responsible for providing coverage throughout a particular territory, known as coverage region or market.

The interconnection of many such systems defines a wireless network capable of providing service to mobile users throughout a country continent or in any region.

To provide wireless communications within a particular geographic region (a city, for example), an integrated network of base stations must be deployed to provide sufficient radio coverage to all mobile users or subscribers.

The base stations, in turn, must be connected to a central hub known as mobile switching center (MSC). The MSC provides connectivity between the public switched telephone network (PSTN) and the numerous base stations, and ultimately between all of the wireless subscribers in a system. The PSTN develope the global telecommunications grid which connects conventional (landline) telephone switching centers (called central offices) with MSCs throughout the world.

Fig. 9.1 represents a typical cellular system of the early 1990 's, but there is currently a major thrust to develop new transport architectures for the wireless subscriber. For example, PCS may be distributed over the existing cable television

plant to neighbourhoods or city blocks, where microcells are used to provide local wireless coverage.

To connect mobile subscribers to the base stations, radio links are develop. Using a carefully defined communication protocol called common air interface (CAI) which in essence is a precisely defined handshake communication protocol. The common air interface defined exactly how mobile subscribers and base stations communicate over radio frequency and also defines the control channel signaling methods.

The CAI must allow a great deal of channel reliability to ensure that data is properly sent and received between the mobile and the base station, and as such specifies speech and channel coding.

At the base station, the air interface position *i. e.,* signaling and synchronization data of the mobile transmission is rejected and the remaining voice traffic is passed along, to the MSC on fixed networks. While each base station may handle on the order of 50 at the same time calls, a typical MSC is responsible for connecting as many as 100 base stations to the PSTN (as many as 5,000 calls at one time), so the connection between the MSC and the PSTN need substantial capacity at any instant of time for transmission.

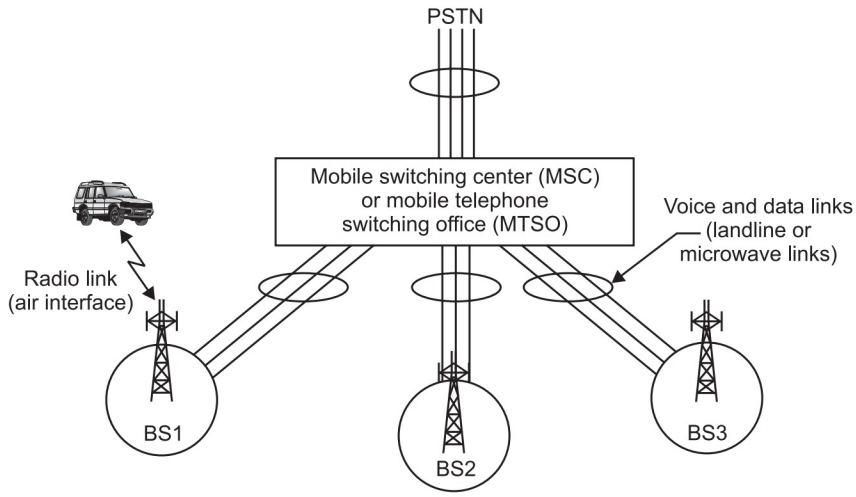


Fig. 9.1: Block diagram of a cellular system.

It becomes clear that networking strategies and standards may vary widely depending on whether a single voice circuit or an whole metropolitan population is served.

Unfortunately, the term network may be used to define a wide range of voice or data connections, from the case of a single mobile user to the base station, to the connection of a large MSC to the PSTN. This broad network definition shows a challenge in describing the large number of strategies and standards used in networking, and it is not feasible to cover all views of wireless networking in this chapter.

9.2 DIFFERENCE BETWEEN WIRELESS AND FIXED TELEPHONE NETWORKS

We can compare wireless networks and fixed telephone networks with the help of following points:

(*i*) Normally, in the public switched telephone network (PSTN), the information is sent over landline truncked lines (called trunks). These trunked lines include the fiber optic cables, copper cables, microwave links, and satellite links. The network configuration in the PSTN are virtually static, since the network connections may only be vary, when a subscriber vary the residence and requires reprogramming at the local central office (CO) of the subscriber. Wireless networks, on the other hand, are highly dynamic, by using network configuration being rearranged every time a subscriber moves into the coverage region of a different base station or a new market or a new region.

(*ii*) Also, networks are difficult to change whereas wireless networks must reconfigure themselves for users within small intervals of time (on the other of seconds) to arrange roaming and imperceptible handoffs between calls as a mobile move from one place to another place.

(*iii*) Further, the available channel bandwidth for fixed networks can be extended by installing high capacity cables (fiber-optic or coaxial cable), whereas, wireless networks are constrained by the meager RF cellular bandwidth provided for each user of mobile or mobile user.

9.3 THE PUBLIC SWITCHED TELEPHONE NETWORK (PSTN)

The PSTN is a highly integrated communications network that connects upto 70% of the world's inhabitants.

Before the 2001, the International Telecommunications Union estimated that there were 1 billion public landline telephone numbers, as compared to 600 million cellular telephone numbers, while landline telephones are being added at a 3% rate, wireless subscription are develop at greater than a 40% rate. Every telephone in the world is given calling access using the PSTN.

Each country is responsible for the regulation of the PSTN with in its borders. Overtime, some government telephone systems have become privatized by corporation which provide local and long distance service for profit or advantage.

In the PSTN, each city or a geographic grouping of towns is known as local access and transport area (LATA). Surrounding LATAs are connected by a company called a local exchange carrier (LEC). A LEC is a company that arrange intralata telephone service, and may be a local telephone company, or may be a telephone company that is regional in scope.

A long distance telephone company collects toll fees to arrange connections between different LATAs over its long distance network. These companies are

referred to an interexchange carriers (IXCs), and own and operate large optic and microwave radio networks which are connected to LECs throughout a country or continent (see Fig. 9.2), represent that it is allowed to provide long distance service over a large area.

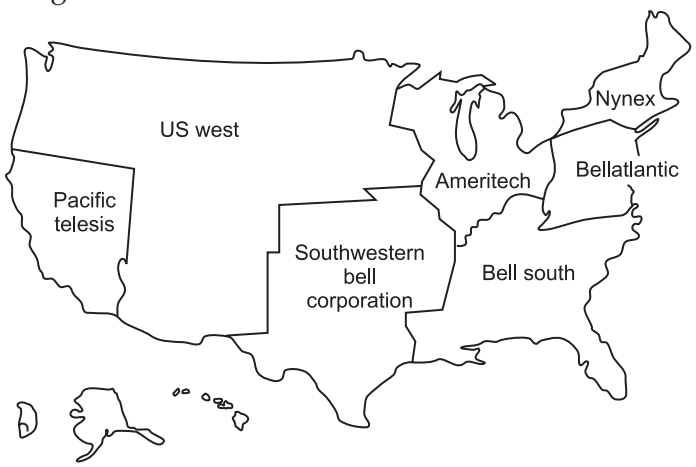


Fig. 9.2: Service area of US regional ball operating companies.

Fig. 9.3 is a simplified represent a local network called a local exchange. Each local exchange consists of a central office (CO) which provides PSTN connection to the customer premises equipment (CPE) which may be an personal phone at a residence or a private branch exchange (PBX) at a place of business.

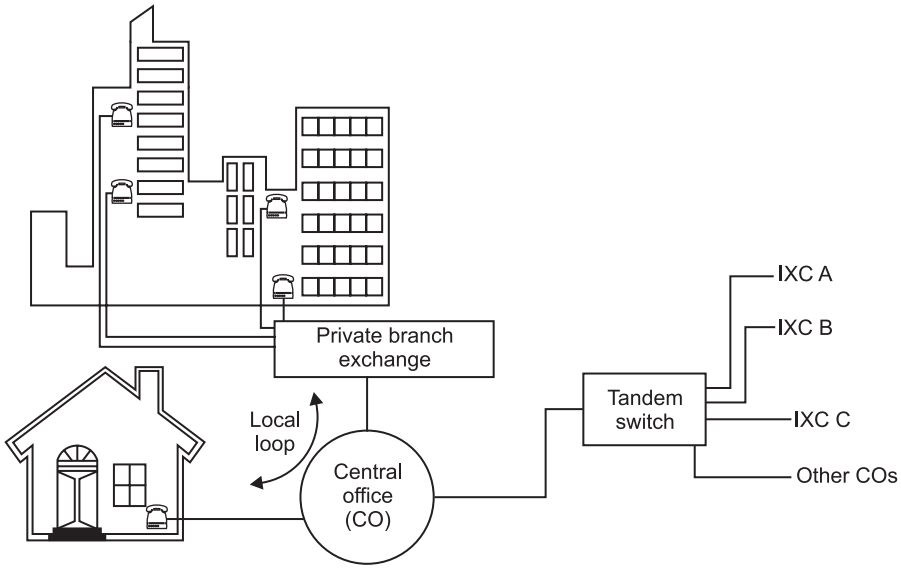


Fig. 9.3: Local landline telephone network.

The CO may handle large number of (million) telephone connections. The CO is connected to a tandem switch in turn connects the local exchange to the PSTN. The tandem switch physically connects the local telephone network to

the presence of point (POP) of trunked long distance lines provided by one or more IXCs. Sometimes EXCs connect directly to the CO switch to ignore local transport charges levied by the LEC.

Fig. 9.3 also shows how a PBX may be used to provide telephone connections throughout a building or campus.

A PBX support an organization or entity to provide internal calling and other in-building services which do not involve the EEC], as well as private networking between other organizational sites in addition to conventional local and long distance services which pass through the CO.

Telephone connections within a PBX are handled by the private owner, whereas connection of the PBX to the CO is provided and handeled by the LEC.

9.4 LIMITATIONS IN WIRELESS NETWORKING

(*i*) **Extremely complex system :** By compared with the local fixed telephone network, where all end-users are static, a wireless communications system is extremely complex.

First of all the wireless network requires an air interface between base stations and subscribers to provide telephone grade communications under a wide range of propagation conditions and for any possible user location. To a certain adequate area coverage, the deployment of many (sometimes hundreds) of base stations throughout a market is essential and each of these base stations must be connected to the MSC. Further, the MSC must eventually provide connections for each of the mobile users to the PSTN. This need simultaneous connections to the LEC, one or more DCCs and to other MSCs via a separate cellular signaling network.

(*ii*) **Handling capacity** : Generally, the demand for wireless communications has consistently exceeded the capacity of available technology, and this is most obvious in the design of MSCs. While a central office (CO) telephone switch may handle up to a million landline subscribers simultaneously, the most sophisticated MSCs in the middle of 1990s, are only able to handle 100,000 to 200,000 simultaneous cellular telephone subscribers or users.

(*iii*) **Hostile and random nature of radio channel**: A problem unique to wireless networks is the extremely hostile and varying nature of the radio channel, and since users may request service from any physical location while travelling over a wide range of velocities, the MSC is pressed to switch calls imperceptibly between base stations throughout the system. The radio spectrum available for this purpose is limited, so the wireless systems are constrained to operate in a fixed bandwidth to support an increasing number of users with time. Spectrally efficient modulation techniques, frequency reuse techniques and geographically distributed radio access points are vital components of wireless networks. As wireless systems grow, the essential addition of base stations increases the switching burden of the MSC. Because, the geographical location

of a mobile user changes constantly, extra overhead is required by all aspects of a wireless network, particularly at the MSC, to ensure seamless communications, regardless of the location of the user or subscriber.

9.5 MERGING WIRELESS NETWORKS AND THE PSTN

Throughout the world, first generation wireless systems (analog cellular and cordless telephones) were extended in the early and mid 1980s. As first generation wireless systems were being introduced, revolutionary advances were being made in the design of the PSTN by landline telephone companies.

Until the year 1980s, most analog landline telephone links throughout the world sent signaling informations along the same trunked lines as voice traffic. That is, a single physical connection was used to handle both signaling traffic and voice traffic for every user.

The advantage of a separate but parallel signaling channel allows the voice trunks to be used strickly for revenue generating voice traffic, and supports many more users on each trunked line so, during the mid 1980s.

The PSTN was transformed into two parallel networks- one dedicated to user traffic, and one dedicated to call signaling traffic. This technique is known as common channel signaling.

Common channel signaling is used in all modern telephone networks.

Must recently, dedicated signaling channels have been used by cellular MSCs to provide global signaling interconnection, thereby activating MSCs throughout the world to pass subscriber information.

In many of today's cellular telephone systems, voice traffice is forward on the PSTN while signaling information for each call is carried on a distinct (fixed) signaling channel.

In North America, the cellular telephone signaling network uses number 7 signaling system (SS7), and each MSC uses the IS-41 protocol to communicate with the other MSCs on the continent.

In first generation cellular systems, common signaling channels were not used, and signaling data was sent on the same trunked channel as the voice user. In second generation wireless systems, however, the air interfaces have been designed to provide parallel user and signaling channels for each mobile, so that each mobile receives the same features and services as fixed wireline telephones in the PSTN.

9.6 EVOLUTION AND DEVELOPMENT OF WIRELESS NETWORKS

Evolution and development of wireless networks may be explain in the form of following wireless networks

(*i*) First generation wireless networks.

(*ii*) Second generation wireless networks.

(*iii*) Third generation wireless networks.

9.6.1 First Generation Wireless Networks

First generation cellular and cordless telephone (phone) networks are based on analog technology. All first generation cellular systems use FM modulation and cordless telephones use a single base station to communicate with a single portable terminal. A typical example of a first generation cellular telephone system is the Advanced Mobile Phone Services (AMPS) system used in the United States. Basically, all first generation systems use the transport architecture shown in Fig. 9.4 with mobile user, and MSC.

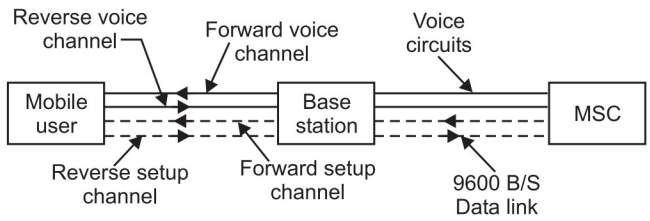


Fig. 9.4: Communication signaling between mobile, base station, and MSC in first generation wireless networks.

Fig. 9.4 shows a diagram of a first generation cellular radio network, which includes the mobile terminals, the base stations, and MSCs.

In first generation cellular networks, the system control for each market settle in the MSC, which maintains all mobile related information and controls each mobile handoff.

The MSC also performs all of the network management functions, such as call handling and processing, billing, and false detection within the market.

The MSC is interconnected with the PSTN using landline trunked lines (trunks) and a tandem switch. MSCs also are connected with other MSCs using dedicated signaling channels (see Fig. 9.5) for exchange of location, validation, and call signaling information.

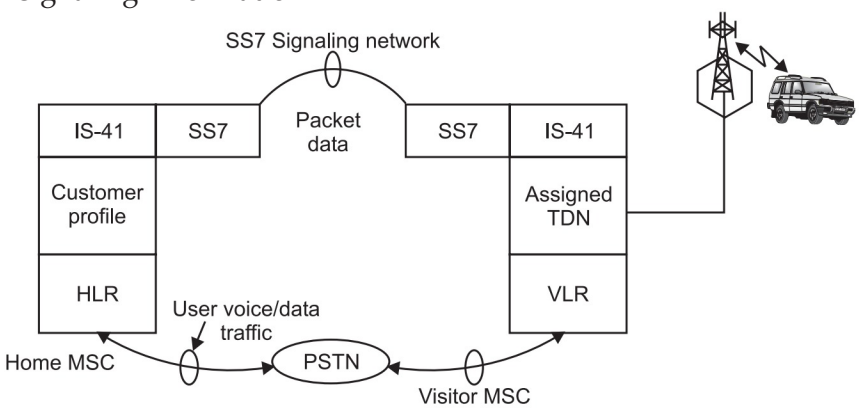


Fig. 9.5: The North American cellular network architecture.

In Fig. 9.5, the PSTN is a separate network from the SS7 signaling network for transmission.

In modern cellular telephone systems, long distance voice traffic is carried on the PSTN, but the signaling information used to provide call set-up and to inform MSCs about a particular user is forwarded on the SS7 network.

First generation wireless systems provide analog speech and inefficient, low-rate, data transmission between the base station and the mobile user. But the speech signals are usually digitized using a standard, time division multiplex format for transmission between the base station and the MSC are always digitized for distribution from the MSC to the PSTN.

Before the 1990s, US cellular carriers implemented the netwrok protocol standard IS-41 to allow different cellular systems to automatically accommodate subscribers who roam into their coverage region. This is known as interoperator roaming. IS-41 allows MSCs of different service providers to pass information about their subscribers to other MSCs on demand.

IS-41 relies on a feature of AMPS known as **autonomous registration.** Autonomous registration is a process by which a mobile notifies a serving MSC of its presence and location of it.

The mobile accomplishes this by periodically keying up and transmitting its identity information, which provide the MSC to constantly update its customer list.

The registration command is sent in the overhead message of each control channel at five to ten minute intervals, and includes a timer value which each mobile uses to determine the precise time at which it should respond to the serving base station with a registration transmission.

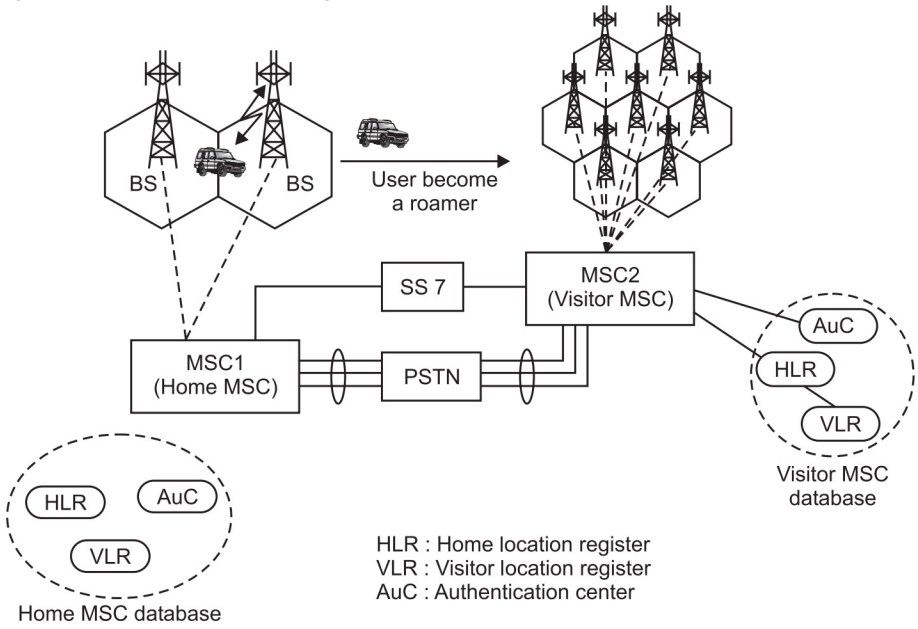


Fig. 9.6: Block diagram of a cellular radio network.

Every mobile reports its MIN and ESN during the brief registration transmission so that the MSC can validate and update the customer list within the market.

The MSC is able to differentiate home users from roaming users based on the MIN of each active user, and maintains a real-time user list in the home location register (HLR) and visitor location register (VLR) as shown in Fig. 9.6 with AuC.

IS-41 allows the MSCs of neighbouring systems to automatically handle the registration and location validation of roamers so that users no longer need to manually register as they travel. The visited system produce a VLR record for each new roamer and notifies the home system via IS-41 so it can update its own HLR.

9.6.2 Second Generation Wireless Networks

Second generation wireless systems described digital modulation and advanced call processing capabilities. Examples of second generation wireless systems include the Global System for Mobile (GSM), the TDMA and CDMA US digital standards (the telecommunications industry association IS-136 and IS-95 standards), second generation cordless telephone (CT2), the British standard for cordless telephony, the personal access communications system (PACS) local loop standard, and digital European cordless telephone (DECT), which is the European standard for cordless and office telephony system.

Second generation wireless networks have introduced new network architectures that have reduced the computational burden of the MSC (Mobile Switching Centre).

GSM introduced the concept of a base station controller (BSC), which is inserted between several base stations and the MSC.

In PACS/WACS, the BSC is known as radio port control unit. This architectural change has allowed the data interface between the base station controller and the MSC to be standardized, thereby support carriers to use different manufacturers for MSC and BSC components.

This trend in standardization and interoperability is new to second generation wireless networks. Eventually wireless network components, such as the MSC and BSC, will be available as off-the- shelf components, much like their wireline telephone counterparts.

All second generation systems use digital voice coding and digital modulation. The systems employ dedicated control channels within the air interface for same time exchanging voice and control information between the subscriber, the base station, and the MSC while a call is in progress.

Second generation systems also arrange dedicated voice and signaling trunks between MSCs and between each MSC and the PSTN.

In contrast to first generation systems, which were designed primarily for voice, second generation wireless networks have been specifically designed to

arrange paging and other data services such as facsimile and high data rate network access.

The network controlling structure is more distributed in second generation wireless systems, since mobile stations assume greater control functions.

In second generation wireless networks, the handoff process is mobile-controlled and is known as (called) Mobile Assisted Handoff (MAHO).

The mobile units in these networks perform several other functions not performed by first generation subscriber units, such as received power reporting, adjacent base station scanning, data encoding, and encryption.

DECT is an popular example of an second generation cordless telephone standard which allow each cordless phone to communicate with any of a number of base stations, by automatically selecting the base station with the greatest signal level. In DECT, the base stations consist greater control in terms of switching, signalling, and controlling handoffs.

In general, second generation systems have been designed to reduce the computational and switching burden at the base station or MSC, while providing more flexibility in the channel allocation scheme so that systems may be deployed rapidaly and in a less coordinated manner.

9.6.3 Third Generation Wireless Networks

The aim of third generation wireless networks is to arrange a single set of standards that can meet a wide range of wireless applications and provide universal access throughout the world.

In third generation wireless systems, the distinctions between cordless telephones and cellular telephones will disappear, and a universal personal communicator (a personal handset) will arrange access to a variety of voice, data, and video communication services.

Third generation systems will use Broadband Integrated Services Digital Network (B-ISDN) to arrange access to information networks, such as the internet and other public and private databases.

Third generation networks will carry many types of information (voice, data, and video), will operate in varied area (dense or sparsely populated regions), and will serve both stationary users and vehicular users traveling at high speeds. Packet radio communications will likely be used to distribute network control while providing a reliable information transfer.

The terms 3G personal communication system (PCS) and 3G Personal communication Network (PCN) are used to imply emerging third generation wireless systems for hand-held devices. Other names for PCS include future Public Land Mobile Telecommunication Systems (FPLMTS) for worldwide use which has more recently been called International Mobile Telecommunication (IMT-2000) and Universal Mobile Telecommunication System (UMTS) for advanced mobile personal services in Europe.

9.7 TRAFFIC ROUTING IN WIRELESS NETWORKS

Amount of traffic capacity required in a wireless network is highly dependent upon the type of traffic transfered. For example, a subscriber's telephone call (voice traffic) requires dedicated network access to arrange real-time communications, whereas control and signaling traffic may be bursty in nature and may be able to share network resources with other bursty users. Alternatively, some traffic may consist an urgent delivery schedule while some may consist no need to be sent in real-time. Type of traffic carried by a network determines the routing services, protocols, and call handling techniques which must be employed.

Two general routing services are arranged by networks. These are connection-oriented services (virtual circuit routing) and connectionless services (datagram services).

In connection oriented routing, the communications path between message source and destination is fixed for the whole duration of the message, and a call set-up procedure is required to dedicate network resources to both the called and calling parties.

Connection-oriented service relies heavily on error control coding to arrange data protection in case the network connection becomes noisy.

Connectionless routing, on the other hand, does not establish a firm connection for the traffic, and instead relies on packet-based transmissions.

Several packets from a message and each individual packet in a connectionless service is routed separately. Successive packets with in the same message might travel completely different routes and encounter widely varying delays throughout the network.

Packets sent using connectionless routing do not necessarily arrive in the order of transmission and must to be recorded at the receiver.

Connectionless routing often avoids having to retransmit an entire message but requires more overhead information for each packet. Typical packet overhead information includes the packet source address, destination address, routing information, and information needed to properly order packets at the .receiver.

In a connectionless service, a call set-up procedure is not required at the beginning of a call, and each message burst is treated independently by the network.

9.7.1 Circuit Switching

Circuit switching establishes a dedicated connection (A radio channel between the base and mobile, and a dedicated phone line between the MSC and the PSTN) for the whole duration of a call.

A mobile user may handoff to different base stations, there is always a dedicated radio channel to arrange service to the user, and the MSC dedicates a fixed, full duplex phone connection to the PSTN.

Wireless data networks are not well supported by circuit switching, due to their short, bursty transmissions which are often followed by periods of

inactivity. Often, the time required to establish a circuit exceeds the duration of the data transmission.

Circuit switching is best suited for dedicated voice only traffic, or for instance where data is continuously sent over long periods of time.

9.7.2 Packet Switching

Packet switching (also known as virtual switching) is the most common technique used to implement connectionless services and allows a large number of data users to remains virtually connected to the same physical channel in the network.

All users may access the network randomly and at will, call setup procedures are not needed to dedicate specific circuits when a particular user require to send data.

Packet switching breaks each message into smaller units for transmission and recovery of information.

When a message is broken into packets, a certain amount of control information is added to each packet to provide source and destination identification, as well as error recovery provisions from information.

Fig. 9.7 represents the sequential format of a packet transmission and packet consists of header information, the user data, and a trailer. The header specifies the beginning of a new packet and contains the source address, packet sequence number, and other routing and billing information. User data have the information which is generally protected with error control coding. Trailer contains a cyclic redundancy check sum which is used for error detection at the receiver.

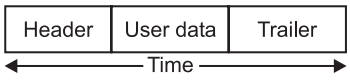


Fig. 9.7: Packet data format.

Fig. 9.8 describe the structure of a transmitted packet, which typically consists of five fields : flag bits, address field, control field, information field, and the frame check sequence field. Flag bits are specific (or reserved) bit sequence that represent the beginning and end of each packet. Address field contains the source and the destination address for transmitting message and for receiving acknowledgement. Control field defines functions such as transfer of acknowledgments, automatic repeat requests (ARQs), and packet sequencing. Information field contains the user data and may have variable length. Final field is the frame check sequence field or the CRC (cyclic redundancy check) that is used for error detection.

In contrast to circuit switching, packet switching provides excellent channel efficiency for bursty data transmissions of short length.

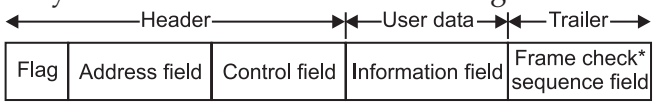


Fig. 9.8: Fields in a typical packet of data.

Advantage of packet-switched data is that the channel is utilized only when sending or receiving bursts of information. This benefit is valuable for the case of mobile services where the available bandwidth is limited.

Packet data approach supports intelligent protocols for data flow control and retransmission, which can provide highly reliable transfer in degraded channel conditions.

X.25 is a widely used packet radio protocol that defines a data interface for packet switching point of view.

9.7.3 The X.25 Protocol

X.25 was developed by CCITT (now ITU-T) to provide standard connectionless network access (packet switching) protocols for the three lowest layers (layers 1, 2 and 3) of the open systems interconnection (OSI) model (see Fig. 9.9 for the OSI layer hierarchy).

X.25 protocols provide a standard network interface between originating and terminating subscriber equipment (known as data terminal equipment or DTE), the base stations (called data circuit-terminating equipment or DCE), and the MSC (known as the data switching exchange or DSE).

Fig. 9.9 describe the hierarchy of X.25 protocols in the OSI model. The layer 1 protocol deals with the electrical, mechanical, procedural, and functional interface between the subscriber (DTE), and the base station (DCE). The layer 2 protocol describe the data link on the common air-interface between the subscriber and the base station. Layer 3 provides connection between the base station and the MSC, and is known as the packet layer protocol. Packet assemble disassemble (PAD) is used at layer 3 to connect networks using the X.25 interface with devices that are not equipped with a standard X.25 interface.

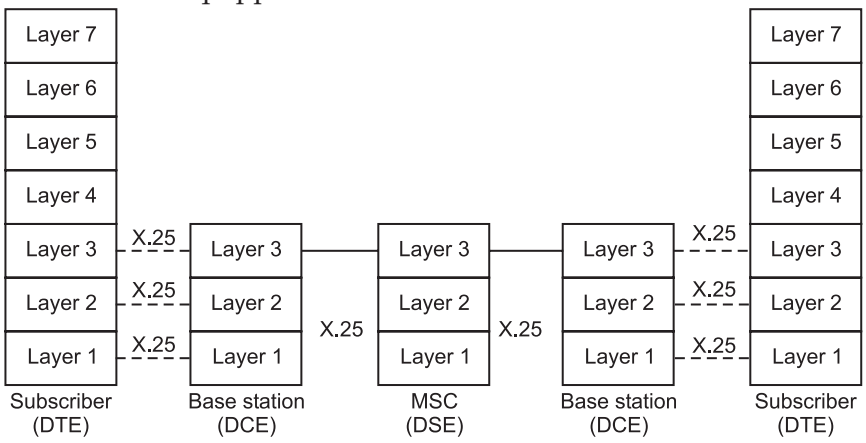


Fig. 9.9: Hierarchy of X.25 in OSI model.

X.25 protocol does not define the particular data rates or how packet-switched networks are implemented. Rather, X.25 arrange a series of standard functions and formats which give structure to the design of software that is used to arrange packet data on a generic connectionless network.

9.8 COMMON CHANNEL SIGNALING (CCS)

Common channel signaling (CCS) is a digital communications technique that provides same time transmission of user data, signaling data and other related traffic throughout a network. This is accomplished by using out-of-band signaling channels which logically distinct the network data from the user information (voice or data) on the similar channel.

For second generation wireless communications systems, CCS is used to pass user data and control/supervisory signals between the subscriber and the base station, between the base station and the MSC, and between MSCs.

Even by using the concept of CCS implies dedicated, parallel channels, it is implemented in a TDM format for serial data transmissions.

CCS is an out-of-band signaling technique which support much faster communications between two nodes within the PSTN. Instead of being constrained to signaling data rates which are on the order of audio frequencies, CCS provide signaling data rates from 56 Kbps to many megabits per second.

CCS provides a substantial increase in the number of users which are provide by truncked PSTN lines, but requires that a dedicated portion of the trunk time be used to provide a signaling channel used for network traffic.

CCS generally uses variable length packet sizes and a layered protocol structure.

Distributed Central Switching Office Forces

As more users subscribe to wireless services, backbone networks that link MSCs together will rely more heavily on network signaling to preserve message integrity, to support end-to-end connectivity for each mobile user, and to maintain a robust network that can recover from failures.

CCS define the foundations of network control and management functions in second and third generation networks.

Out-of-band signaling networks which connect MSCs throughout the world enable the entire wireless network to update and keep track of particular mobile users, wherever they happen to be.

Fig. 9.5 shows that how an MSC is connected to both the PSTN and the signaling networks.

As shown in Fig. 9.10, the CCS network architecture is composed of geographically distributed central switching offices, each with embeddei switching end points (SEPs), signaling transfer points (STPs), a service management system (SMS) and a database service management system (DBAS).

The MSC arrange the subscriber access to the PSTN via the SEP. The SEP implements a stored-program-control switching system known as the service control point (SCP) that uses CCS to set up calls and to access a network database. The SCP instructs the SEP to create billing records based on the call information recorded by the SCP.

The STP controls the switching of messages between nodes in the CCS network. For higher reliability of transmission (redundancy), SEPs are required to be connected to the SS7 network via at least two STPs. This combination of two STPs in parallel is known as a mated pair and provides connectivity to the network in the event one STP fails.

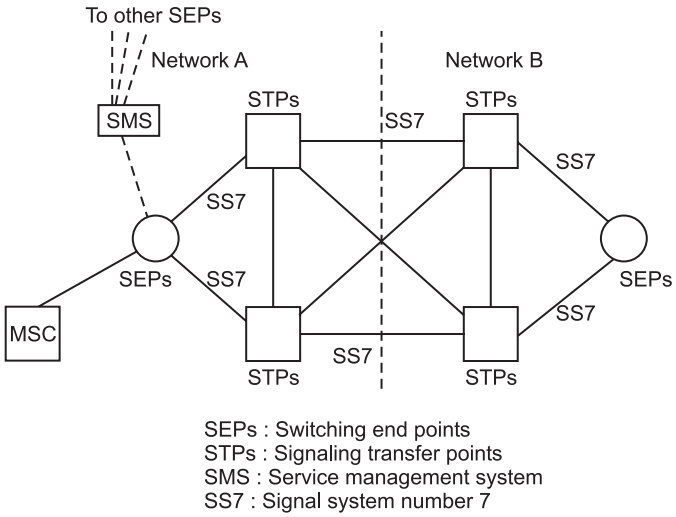


Fig. 9.10: Common channel signaling (CCS) network architecture showing STPs, SEPs and SMS embedded within a central switching office, based on SS7.

9.9 INTEGRATED SERVICES DIGITAL NETWORK (ISDN)

Integrated Services Digital Network (ISDN) is a complete network framework designed around the concept of common channel signaling. While telephone users throughout the world rely on the PSTN to transmit conventional voice traffic, new end-user data and signaling services can be provided with a parallel, dedicated signaling network. ISDN describe the dedicated signaling network that has been created to complement the PSTN for more flexible and efficient network access and signaling and may be thought of as a parallel world-wide network for signaling traffic that can be used to either route voice traffic on the PSTN or to support new data services between network nodes and the end-users.

ISDN support two distinct kinds of signaling components to end-users in a telecommunications networks.

(*i*) The first component supports traffic between the end user and the network, and is known as access signaling. Access signaling defines how end-users obtain access to the PSTN and the ISDN for communications or services and is determine by a suite of protocols known as the Digital Subscriber Signaling System Number (DSSN).

(*ii*) The second signaling component of ISDN is network signaling and is determine by the SS7 suite of protocol.

For wireless communications systems, the SS7 protocols within ISDN are critical to providing backbone network connectivity between MSCs throughout

the world, as they support network interfaces for common channel signaling traffic. ISDN provides a complete digital interface between end-users over twisted pair telephone lines.

ISDN interface is divided into three different types of channels. Information bearing channels called bearer channels (*B* channels) are used exclusively for end-user traffic (voice, data, video). Out-of-band signaling channels, called data channels (*D* channels), are used to send signaling and control information across the interface to end-users.

As shown in Fig. 9.11, ISDN provides integrated end-user access to both circuit-switched and packet-switched networks with digital end-to-end connectivity.

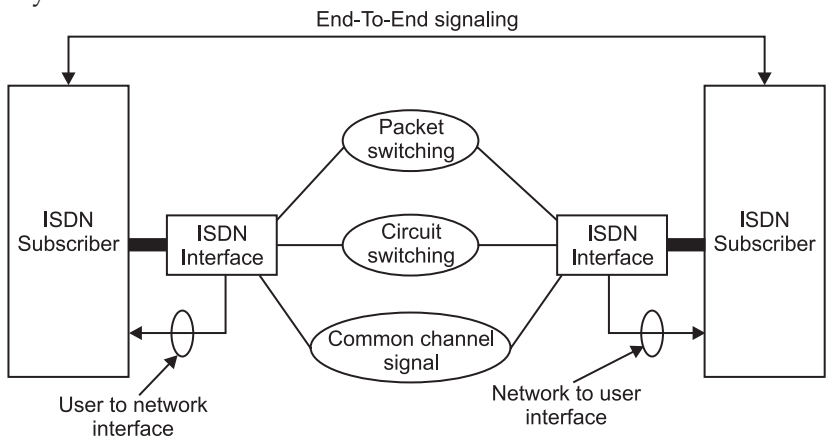


Fig. 9.11: Block diagram an ISDN.

ISDN end-users may select between two different interfaces, the basic rate interface (BRI) or the primary rate interface (PRI). The BRI is intended to serve small capacity terminals while the PRI is intended for large capacity terminals.

The *B* channels support 64 Kbps data for both the primary rate and the basic rate interface. The *D* channel support 64 Kbps for the primary rate and 16 Kbps for the basic rate.

BRI provides two 64 Kbps bearer channels and one 16 Kbps signaling channel (25 + *D*), whereas the PRI provide twenty-three 64 Kbps bearer channels and one 64 Kbps signaling channel (235 + *D*) for North America and Japan.

For wireless service subscribers, an ISDN basic rate interface is provided in exactly the same manner as for a fixed terminal. To differentiate between wireless and fixed subscribers, the mobile BRI defines signaling data (D channels in the fixed network) as control channels (C channels in the mobile network), so that a wireless subscriber has *2B* + *C* service.

9.10 BROADBAND ISDN AND ATM

With the proliferation of computer systems and video imaging, end-user applications are requiring the large bandwidths than the standard 64 Kbps *B* channel provided by ISDN. Recent work has defined ISDN interface standards

that increase the end-user transmission bandwidth to several Mb/s. This emerging networking technique is called the broadband ISDN (5-ISDN) and is based on Asynchronous Transfer Mode (ATM) technology which allows packet switching rates up to 2.4 Gbps and total switching capacities as high as 100 Gbps (gega bit per second).

ATM is a packet switching and multipexing technique which has been specifically designed to control both voice users and packet data users is a single physical channel.

ATM data rates vary from low traffic rates (64 Kbps) over twisted pair to over 100 Mbps over fiber optic cables for high traffic rates between network nodes.

ATM supports bidirectional transfer of data packets of fixed length between two end points, while preserving the order of transmission. ATM data units, known as cell, are routed based on header information in each unit (called a label) that identifies the cell as belonging to a specific ATM virtual connection. The label is determined upon virtual connection of a user and remains the same throughout the transmission for a particular connection.

The ATM header also includes data for congestion control, priority information for queuing of packets, and a priority which indicates which ATM packets can be dropped in case of congestion in the network.

Fig. 9.12, describe the cell format of ATM. ATM cells (packets) have a fixed length of 53 bytes, consisting of 49 bytes of data and 5 bytes of header information. Fixed length packets result in simple implementation of fast packet switches, since packets arrive synchronously at the switch. A compromise was made in selecting the length of ATM cells to accommodate both voice and data users.

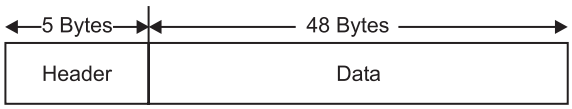


Fig. 9.12: Cell format of asynchronous transfer mode (ATM)

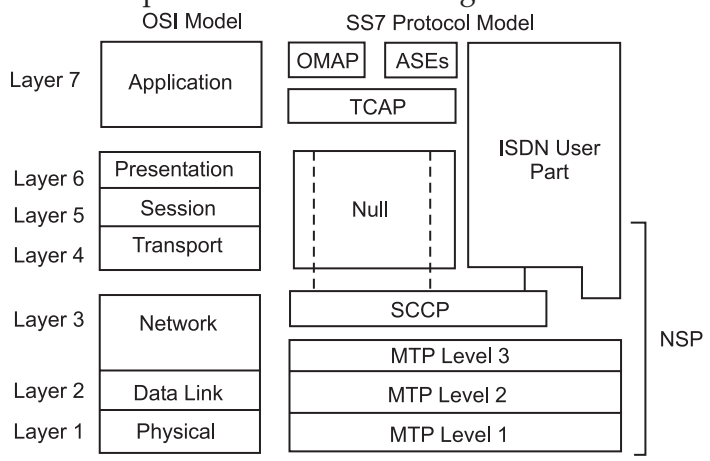
9.11 SIGNALING SYSTEM NUMBER 7 (SS7)

The SS7 signaling protocol is widely used for common channel signaling between interconnected networks.

SS7 is used to interconnect most of the cellular MSCs throughout the US, and is the key factor in empower autonomous registration and automated roaming in first generation cellular systems (1G system).

SS7 is an outgrowth of the out-of band signaling first developed by the CCITT under common channel signaling standard, CCS number 6. Further work caused SS7 to evolve along this lines of the ISO-OSI seven layer network definition, where a highly layered structure (transport from layer to layer) is used to provide network communication between two terminals.

A comparison of the OSI-7 network model and the SS7 protocol standard is given in Fig. 9.13. The lowest three layers of the OSI model are controlled in SS7 by the network service part (NSP) of the protocol, which in turn is made up of three Message transfer parts (MTPs) and the Signaling Connection Control Part (SCCP) of the SS7 protocol as shown in Fig. 9.13.



OMAP: OPERATIONS MAINTENANCE AND ADMINISTRATION PART
ASE: APPLICATION SERVICE ELEMENT
TCAP: TRANSACTION CAPABILITIES APPLICATION PART
SCCP: SIGNALING CONNECTION CONTROL PART
MTP : MESSAGE TRANSFER PART
NSP: NETWORK SERVICE PART

Fig. 9.13: SS7 protocol architecture.

The NSP provides ISDN nodes with a highly reliable and efficient means of exchanging signaling traffic using connectionless services. The SCCP in SS7 actually allows packet data network interconnections as well as connection-oriented networking to virtual circuit networks.

The NSP allows network nodes to communicate throughout the world without discuss for the application or context of the signaling traffic.

The function of the MTP is the ensure that signaling traffic can be transferred and delivered reliably between the end-users and the network.

MTP is provided at three levels. Fig. 9.14 shows the functionality of the various MTP levels that will be described.

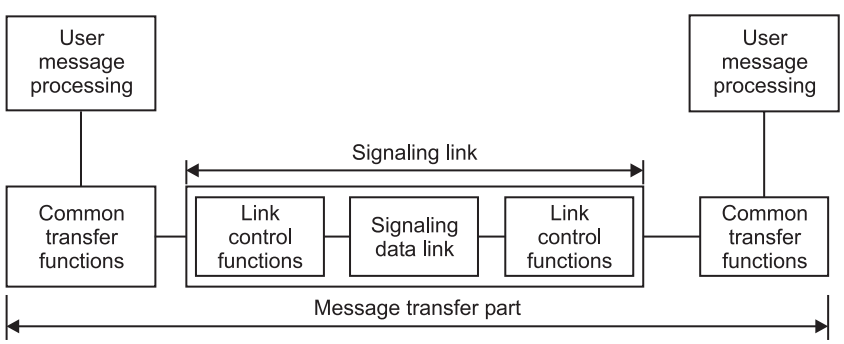


Fig. 9.14: Functional diagram of message transfer part.

Signaling data link functions (MTP level 1) provide an interface to the actual physical channel over which communication takes place.

Signaling link function (MTP level 2) correspond to the second layer in the OSI reference model and provide a reliable link for the transfer of traffic between two directly connected signaling points. Variable length packet messages, called message signal units (MSUs), are defined in MTP level 2. A wide range of error detection and correction features are provided in MTP level 2. MTP level 2 also provides flow control data between two signaling points as a means of sensing link failure.

Signaling network functions (MTP level 3) provide procedures that transfer messages between signaling nodes. As in ISDN, there are two types of MTP level 3 functions : signaling message handling and signaling network management. Signaling message handling is used to provide routing, distribution, and traffic discrimination.

The signaling connection control part (SCCP) provides enhancement to the addressing capabilities provide by the MTP. While the addressing capabilities of MTP are limited in nature, SCCP uses local addressing based on subsystem numbers (SSNS) to identify users at a signaling node. SCCP also provides the ability to address global tittle messages, such as 800 numbers or non-billed numbers. SCCP provides four classes of service : two are connectionless and two are connection-oriented, as shown in Table 9.1.

Table 9.1: Different Classes of Service Provided by SCCP

Class of Service	Type of Service
Class 0	Basic connection class
Class 1	Sequenced (MTP) connectionless class
Class 2	Basic connection-oriented class
Class 3	Flow control connection-oriented class

The SS7 user part provides call control and management functions and call set-up capabilities to the network. These are the higher layers in the SS7 reference model, and utilize the transport facilities provided by the MTP and the SCCP.

The SS7 user part includes the ISDN user part (ISUP), the transaction capabilities application part (TCAP) and the operations maintenance and administration part (OMAP). The telephone user part (TUP) and the data user part (DUP) are included in the ISUP.

The ISUP provides the signaling functions for carrier and supplementary services for voice, data and video in an ISDN environment. In the past, telephony requirements were lumped in the TUP. ISUP uses the MTP for transfer of messages between different exchanges. ISUP message includes a routing label that indicates the source and destination of the message, a circuit identification code (CIC) and a message code that serves to define the format and function of each message. They have variable lengths with a maximum of 272 octets.

The transaction capabilities application part in SS7 refers to the application layer which invokes the services of the SCCP and the MTP in a hierarchical format.

The OMAP functions include monitoring, coordination, and control functions to ensure that trouble free communications are possible. OMAP supports diagnostics are known throughout the global network to determine loading and specific subnetwork behaviours.

9.11.1 Signaling Traffic in SS7

Call set-ups, inter MSC handoffs, and location updates are the main activities that generate the maximum signaling traffic in a network, and which are all controlled under SS7. Table 9.2 shows the amount of signaling traffic that is generated for call set-up in GSM.

Table 9.2: Signaling Load for Call Setup and Handoffs in GSM

Call originating from a Mobile	Load
Information on the originating MSC and the terminating switch	120 bytes
Information on the originating MSC and the associated VLR	550 bytes
Call terminating at a Mobile	
Information on the switch and terminating MSC I	120 bytes
Information on the terminating MSC and associated VLR	612 bytes
Information on the originating switch and HLR	126 bytes
Inter MSC handoffs	
Information on the new MSC and associated	148 bytes
VLR Information on the new MSC and old MSC	383 bytes

An GSM location update records are updated in the network whenever a subscriber moves to a new location. The traffic required by the location update process as a subscriber moves within and between VLR areas is shown in Table 9.3 and in Table 9.4.

Table 9.3: Signaling Load for Location Updating in GSM

Location Updating	Load
Information on the current MSC and associated VLR	406 bytes
Information on the current VLR and HLR	55 bytes
Information on new VLR and old VLR	406 bytes
Information on the new VLR and old VLR	213 bytes
Information on the old VLR and HLR	95 bytes
Information on the new VLR and HLR	182 bytes

9.11.2 SS7 Services

The SS7 support three main type of services and which are-Alternate billing services, 800 services, and touchstar services and these services are explain briefly as the following:

(*i*) **Alternate billing services:** These services used by CCS network to enable the calling party to bill a call to a personal number (calling cord, third party number) from any number.

(*ii*) **800 services :** The Bell system introduce these services to support tall free access to the calling party to the services and data base which is offered by the private parties. The service subscriber paid the cost of processing of calls. There are two types of plan known as 800 database plan and 800 NXX plan offered by these services.

(*iii*) **Touchstar :** CLASS is the another kind of this services and it is a group of switch controlled services that provide its users with certain call management capabilities. These services include the following call return, call forwarding, repeat dialing, call block, call tracing and caller ID.

9.12 PERSONAL COMMUNICATION SERVICES/NETWORKS (PCS/PCNS)

The main function of personal communication network or personal communications services or system is to provide enabling user to access the telephone network, internet for different type of communication needs and ubiquitous wireless communication coverage. These services are provide by PCS/PCN to the user without regard for the location of user.

The PCS/PCN is basically based on ATN (Advance Intelligent Network).

The AIN will support to its user to consist a single telephone number to be used for both wireless and wireline services.

Murthy, Sheikh and Ashity introduce a architecture of PCS/PCN, which have three levels: access level, transport level and intelligent level.

(*i*) The intelligent level have the data base for the stored information about the network user.

(*ii*) The transport level control the transmission of information.

(*iii*) The access level allows ubiquitous access to every user in the network and also update the location of each user.

The PCS/PCN will be characterized by high densities user that will require improved network requirements.

The efficient signaling protocol and common channel signaling will play an effected role in PCS/PCN.

The Table 9.4 represent the data requirement for PCS/PCN.

Table 9.4: Potential Data Loads for Wireless Networks

Application	Average data rates (kbps)	Peak data rate (kbps)	Maximum delay (sec)	Maximum packet loss rate
e-mail, paging	0.01–0.1	1–10	< 10–100	$<10^{-9}$
Computer data	0.1–1	10–100	< 1–10	$<10^{-9}$
Telephony.	10–100	< 0.1–1	$< 10^{-4}$	
Digital audio	100–1000	100–1000	< 0.01–0.1	$<10^{-5}$
Video conference	100–1000	1000–10000	0.001–0.01	$<10^{-5}$

9.13 PACKET vs. CIRCUIT SWITCHING FOR PCN

The packet switching is better than circuit switching for the PCS/PCN and the packet switching is affected by the following factor:

(*i*) PCN will allow large population of infrequent users so the economic viability is depend on bandwidth sharing and infrastructural.

(*ii*) The relatively unreliable channel is more suited for packet switching than the circuit switching.

(*iii*) PCN will need a high capacity switching infrastructure for routing of traffic between cells.

(*iv*) PCN will need a wide range of services including voice, data e-mail and digital video.

9.14 PROTOCOL FOR NETWORK ACCESS

A new contention technique, knows as packet radio is used to transmit the data on a communication channel.

ALOHA protocols in an best example of contention technique in which users transmit whenever they have data and the users listen to the acknowledgment feedback to define that if a transmission has been successful or not. If the collision is occur than the user wait for some time and than retransmit the total data again.

The packet contention technique has a great advantage that it serve a large number of terminals with extremely low overhead.

The mobile transmitters can send the bursty traffic in the form of data packet to a common base station receiver can use random access.

The unslotted and slotted ALOHA are only valid for low traffic load transmission but for high traffic load transmission it is inefficient because for high traffic load transmission, more collisions problem is occur so to reduce this situation, a new technique known as carrier sense multiple access (CSMA) can be used.

In carrier sense multiple access (CSMA), firstly sense the channel that channel is free or busy for transmission, if the channel is busy then the sender wait for some time and then again sense and if the channel is free then sender transmit the data. By using this technique, collision problem will be reduce.

The packet reservation multiple access (PRMA) is the another transmission protocol used for packet voice transmission in a cellular system.

The PRMA is a TDM based multiple access protocol that support or provide a group of spatially dispersed terminals to transmit packet voice and low bit rate data over a common channel.

The important feature of this protocol is the utilization of user transmission to obtain the access to the radio resources.

The PRMA is a derivative of reservation ALOHA and reservation ALOHA is a combination of TDMA and slotted ALOHA.

Hence we can say PRMA has an important feature to access the network.

9.15 NETWORK DATABASES

In first generation wireless network, the MSCs of two system are not easily to communication with each other because the network control was limited to the MSCs.

The second and third generation system, reduce this problem by introduce the network access several database for authentication of mobiles, billing and location update etc.

There are various processing element are used in a network for the accessed like as visitor location database, home location database and authentication center.

For provide the interconnection between various MSCs of a wireless network, a distributed database has been introduce.

A distributed database hierarchical architecture is develope to provide the facilitate to the tracking and location update of the mobile subscriber.

Fig. 9.15 shows a hierarchical of distributed data base.

The database position of each access MAN node define subscriber in the related BSC control area and every BSCs is able to receive the MAN broadcast and updates its database if the mobile is in its region.

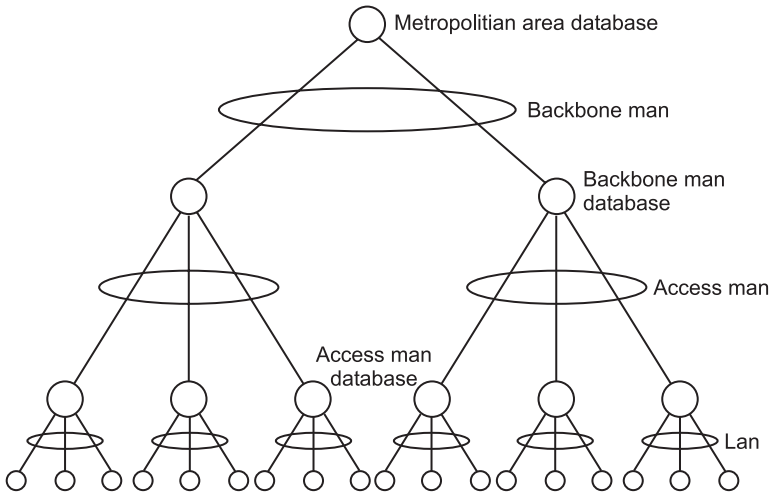


Fig. 9.15: Illustration of a hiearchical distributed database.

The access MAN database define the mobile's BSC area location and the backbone database (MAN) define the MAN access location.

This method of partition is a effected technique for minimize the time required to locate any mobile and also minimize the traffic congestion resulting from the heavy broadcast traffic required to locate a roaming mobile.

HIGHLIGHTS

1. To provide wireless communications with in a particular geographic region (a city for example) an integrated network of base stations must be deployed to provide sufficient radio coverage to all mobile users or subscribers.

2. Normally in the public switched telephone network (PSTN), the information is sent over landing truncked lines (called trunks). These trunked lines include the fiber optic cables, copper cables, microwave links and satellite links. The network configurations in the PSTN are virtually static, since the network connections may only be vary when a subscriber varies the resistance and requires reprogramming at the local central office (CO) of the subscriber, wireless network on the other hand, are highly dynamic by using network configuration being rearranged every time a subscriber moves into the coverage region of a different base station or a new market or a new region.
3. The PSTN is a highly integrated communications network that connects upto 70% of the world's inhabitants.
4. A long distance telephone company collects toll fees to arrange connections between different LATAS over its long distance network. These companies are referred to as interchange carriers (IXCs) and own and operate large optic and microwave radio networks which are connected to LECs throughout a country or continent.
5. Throughout the world, first generation wireless systems (analog cellular and cordless telephones) were extended in the early and mid 1980s. As first generation wireless systems were being introduced, revolutionary advances were being made in the design of the PSTN by land-line telephone companies.
6. Advantage of separate but parallel signaling channels allow the voice trunks used strickly for revenue generating voice traffic and support many more usei each trunked line so during the mid 1980s.
7. The PSTN was transformed into two parallel networks one dedicated to user traffic and another dedicated to call signaling traffic. This technique is known as common channel signalling.
8. Evolution and development of wireless network may be explain in the form of following wireless networks:
   (*i*) First generation wireless networks
   (*ii*) Second generation wireless networks.
   (*iii*) Third generation wireless networks.
9. First generation cellular and cordless telephone (phone) networks are based on analog technology. All first generation cellular systems use FM modulation and cordless telephones use a single base station to communicate with a single portable terminals. A typical example of a first generation cellular telephone system is the advanced mobile phone services (AMPS) system used in the united states. Basically all first generation systems use the transport architecture.
10. Second generation wireless system described digital modulation and advanced call processing capabilities. Examples of second generation wireless systems include the Global system for mobile (GSM), the TDMA and CDMA, US digital standards (the telecommunications industry association IS-136 and IS-95 standards), second generation cordless telephone (CT2). The British standard for cordless telephony. The personal access communications system (PACS) local loop standard and digital European cordless telephone (DECT), which is the European standard for cordless and office telephony system.
11. The aim of third generation wireless networks is to arrange a single set of standards that can meet a wide range of wireless applications and provides universal access throughout the world.

12. Two general routing services are arranged by networks. These are connection-oriented services (virtual circuit routing) and connectionless services (datagram services).

13. In connection oriented routing, the communication path between message source and destination is fixed for the whole duration of the message and a call set up procedure is required to dedicate network resources to both the called and calling parties.

14. Connection-oriented service relies heavily on error control coding to arrange data protection in case the network connection becomes noisy.

15. Circuit switching establishes a dedicated connection (a radio channel between the base and mobile, and a dedicated phone line between the MSC and the PSTN) for the whole duration of a call.

16. Packet switching (also known as virtual switching) is the most common technique used to implement connectionless services and allows a large number of data users to remain virtually connected to the same physical channel in the network.

17. X.25 protocols provide a standard network interface between originating and terminating subscriber equipment (known as data terminal equipment as DTE). The base station (called data circuit terminating equipment or DCE), and the MSC (known as the data switching exchange or DSE).

18. Common channel signaling (CCS) is a digital communication technique that provides same time transmission of user data, signaling data, and other related traffic throughout a network. This is accomplished by using out-of-band signaling channels which logically distinct. The network data from the user information (voice or data) on the similar channel.

19. Integrated services digital network (ISDN) is a complete network framework designed around the concept of common channel signaling. While telephone users throughout the world rely on the PSTN to transmit conventional voice.

18. Traffic, new end-user data and signaling services can be provided with a parallel, dedicated signaling network.

20. ATM is a packet switching and multiplexing technique which has been specifically designed to control both voice users and packet data users is a single physical channel.

21. SS7 is used to interconnect most of the cellular MSCs throughout the US and is the key factor in empower autonomous registration and automated roaming in first generation cellular system (1G system).

22. Call set-ups, inter MSC handoffs and location updates are the main activities that generate the maximum signaling traffic in a network and which are all controlled under SS7.

23. The SS7 support three main type of services and which are alternate billing services. 800 services and touchstar services, and these services are explain briefly as the following:

(*i*) Alternate billing services

(*ii*) 800 services

(*iii*) Touchstar services.

24. The main function of personal communication network or personal communication services or system is to provide assembling user to access the telephone network,

internet for different type of communication needs and ubiquitous wireless communication coverage. These services are provide by PCS/PCN to the user without regard for the location of user.

25. ALOHA protocols is a best example of contention technique in which users transmit whenever they have data and the user listen to the acknowledgment feedback to define that if a transmission has been successful or not. If the collision occurs then the user wait for some time and then retransmit the total data again.

26. In first generation wireless network, the MSCs of two systems are not easily to communication with each other because the network control was limited to the MSCs.

27. The second and third generation system, reduce this problem by introduce the network access several database for authentication of mobiles, billing and location update etc.

28. This method of partition is a effected technique for minimize the time required to locate any mobile and also minimize the traffic congestion resulting from the heavy broadcast traffic required to locate a roaming mobile.

REVIEW QUESTIONS

9.1. Explain the meaning of wireless network and also explain the application of wireless network.

9.2. How you can say that the fixed telephone network is differ from wireless telephone network ?

9.3. Explain the term public switched telephone network (PSTN).

9.4. What are the wireless networks limitations ?

9.5. Describe the development and evolution of wireless networks with 1G, 2G and 3G networks.

9.6. How you can say circuit switching help in traffic routing in wireless network ?

9.7. Explain the following:

(*i*) Circuit switching

(*ii*) Packet switching

(*iii*) X.25 protocol.

9.8. Explain in brief common channel signaling (CCS).

9.9. Explain in brief integrated service digital network (ISDN).

9.10. Describe the signaling system number 7 in brief.

9.11. What are the SS7 services ?

9.12. Define the signaling traffic in SS7.

9.13. What do you mean by personal communication services/networks (PCS/PCN) ?

9.14. Define the difference between packet switching and circuit switching.

9.15. Explain in brief ATM and broadband ISDN.

9.16. Explain the development of wireless multimedia.

9.17. What is common channel signaling?

9.18. Compare circuit switching and packet switching?

9.19. Write the performance of SS7.

10 Chapter

WIRELESS SYSTEMS AND STANDARDS

10.1. *Introduction*. **10.2.** *Global System for Mobile* (*GSM*). **10.3.** *CDMA Digital Cellular Standard* (*LS-95*). **10.4.** *IMT-2000 - A Third Generation Communication*. **10.5.** *UMTS- Universal Mobile Telecommunication System*. **10.6.** *PACS-Personal Access Communication System*. **10.7.** *AMPS and ETACS*. **10.8.** *United States Digital Cellular* (*IS-54 and IS-136*). **10.9.** *Wireless Local Loop* (*WLL*). **10.10.** *Bluetooth*.

10.1 INTRODUCTION

This chapter explain the wireless systems and standards used throughout the world. At the pioneering stage of wireless communication, analog systems were used which lie in the first generation of mobile systems. Recently, with more advancement digital systems were developed. Global system of mobile communication is an example of second generation digital system which was developed to work in the European wireless market. Some other second generation systems are the cordless telephone system (T-2), DECT, PDC etc.

The third generation systems tried to establish a common universal systems which support world wide communication.

In Europe the ETSI is defining UMTS and other important technologies parallel standardization activities on a worldwide level are undertaken within the International Telecommunication Union (ITU) under the banner of International Mobile Telecommunication in the year 2000 (EVIT 2000).

10.2 GLOBAL SYSTEM FOR MOBILE (GSM)

Global system for mobile (GSM) is a second generation cellular system standard. It was developed to solve the fragmentation problems of first cellular system in Europe.

GSM was the world's first cellular system with specify digital modulation and network level architectures and services and is the world's most popular 2G technology.

GSM was originally developed to provide as the Pan-European cellular service and promised a wide range of network services through the use of ISDN.

The setting of standards for GSM is provided by the European technical standards institute (ETSI).

GSM was first introduced into the European market in 1991. By the end of 1993, various non-European countries in South America, Asia and Australia had adopted GSM and the technically equivalent OFF short, DCS 1800, which supports personal communication services (PCS) in the 1.8 GHz to 2.0 GHz radio bands recently develope by governments throughout the world.

10.2.1 GSM Services and Features

(*a*) **GSM Services** : GSM services follow ISDN guidelines and are classified as either teleservices and data services.

User services may be classified into three major categories:

(*i*) *Telephone services :* These services include standard mobile.

Telephone and mobile-originated or base originated traffic.

Telephone services also support emergency calling and facsimile.

GSM also allow videotex and teletex, though they are not integral parts of the GSM standard.

(*ii*) *Data services or bearer :* These services are related to the layer 1, 2 and 3 of the open system interconnection (OSI) reference model within a network.

Data services also support the services include packet switched protocols and data rates from 300 bps to 9.6 kbps data rate.

Data may be transmitted using either a transparent mode (where GSM provides standards channel coding for the user data) or non-transparent mode (where GSM allow special coding efficiencies based on the particular data interface).

(*iii*) *Supplementary ISDN services :* These services are digital in nature and include call diversion, closed user groups and caller identification and are not available in analog mobile networks in the previous generation technology.

Supplementary services also include the short message services (SMS), which allows GSM subscribers and base station to transmit alpha-numeric pages of limited length while simultaneously carrying normal voice traffic.

SMS also support cell broadcast. SMS may be used for safety and advisory applications, like as the broadcast of highway or weather information to all GSM subscribers within reception range.

(*b*) **GSM Features** : One of the most popular features of GSM is the subcriber's identity module (SIM), which is a memory device that store information like as the subscriber's identification number, the networks and countries where the subscriber is entitled to service, privacy key and other user related information. A subscriber uses the SIM with a four digit personal ID number to activate service from any GSM phone. SIMs are available as smart cards (credit card sized cards and it can be inserted into any GSM phone) or plug-in-modules.

The second remarkable feature of GSM is the on-the-air-privacy which is provided by the system. The privacy is made possible by every digital bit stream,

transmit by GSM transmitter, according to a specific secret cryptographic key that is known only to the cellular carrier. This key changes with time for each user. Every carrier and GSM equipment manufacturer must sign the memorandum of understanding (MOU) before developing GSM equipment or extend a GSM system.

10.2.2 GSM System Architecture

The GSM system architecture consists of three major interconnected subsystem that interact between emphatic and with the user through certain network interface.

The subsystems are base station subsystem (BSS), network and switching subsystem (NSS), and the operation support subsystem (OSS), mobile station (MS) is also a subsystem but it is generally consider to be part of the BSS for architecture purpose.

The BSS, also known as the radio subsystem, support and manages radio transmission path between the mobile stations and the mobile switching center (MSC) of the system.

The BSS also manages the radio interface between the mobile station and all other subsystem of GSM system.

Each BSS consist of many base station controllers (BSCs) which connect the MS to the NSS by the MSC. The NSS manages the switching function of the system and allows the MSCs to communicate with other networks such as PSTN and ISDN networks.

The OSS supports the operation and maintenance of GSM and support system engineers to monitor, diagnose and troubleshoot all aspects of the GSM system. This subsystem interacts with the other GSM subsystem, and is provided solely for the staff of the GSM operating system or company which provides service facilities for the network.

Fig. 10.1 shows the block diagram of the GSM system architecture with all three subsystems BSS, NSS and PN.

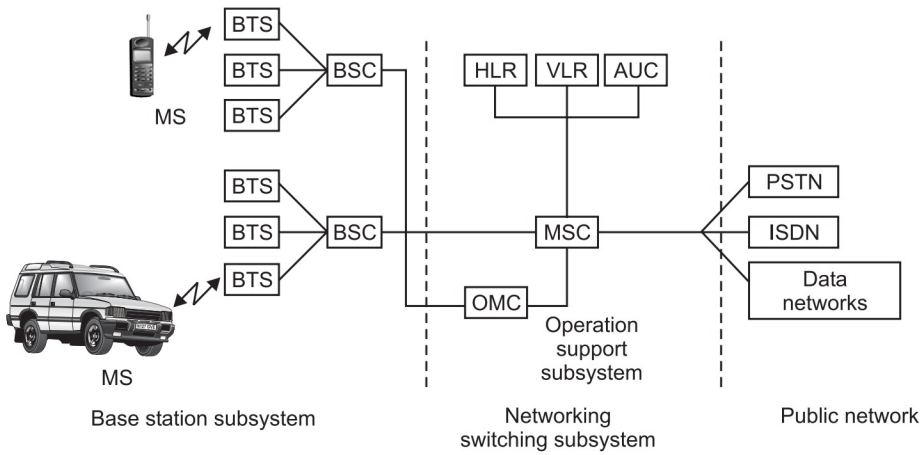


Fig. 10.1: GSM system architecture.

According to Fig. 10.1 the mobile stations (MSs) communicate with the base station subsystem (BSS) over the radio air interface. BSS support many BSCs, which connect to single MSC, and each BSC typically controls up to several hundred base transceiver stations (BTSs) in a system.

Some of the BTSs may be co-located at the BSC and other may be remotely distributed and physically connected to the BSC by microwave link or dedicated leased lines connection.

Mobile handoffs (called handover or HO in the GSM specification) between two BTS under the control of the same BSC are handled by the BSC and not the MSC. This simply and greatly reduces the switching burden of the MSC.

Fig. 10.2 show the interface which connects a BTS to a BSC is known as the Abis interface. The Abis interface carries traffic and maintenance data and is specified by GSM to be standardized for all manufacturers.

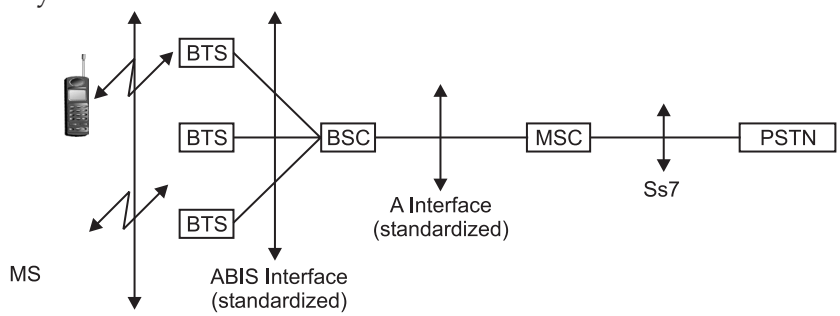


Fig. 10.2: The various interfaces used in GSM

The BSCs are physically connected using dedicated/leased lines or microwave link to the MSC (mobile switching centre).

The interface between the BSC and a MSC is called the *A* interface, which is standardized within GSM. The *A* interface uses an SS7 protocol known as the signaling correction control part (SCCP) which supports communication between MSC and the BSS, as well as network message between the individual subscribers and the MSC. The *A* interface support a service provider to use base stations and switching equipment.

The NSS handles the switching of the GSM calls between external network and the BSCs in the radio subsystem and is also allowed and responsible for the managing and providing external access to several customer data base. The MSC is the central unit in the NSS and controls the traffic among all the BSCs with in the system.

In the NSS, there are three different database called the home location register (HLR), visitor location register (VLR) and authentication centre (AUC). The HLR is a database which consist subscriber information and the location information for each user who stay in the same city as MSC. The VLR is a database which temporarily store the IMSI and customer information for each roaming subscriber who is visiting the coverage region of a particular MSC. Once a roaming mobile is logged in the VLR, the MSC send the necessary information

to visiting subscriber's HLR, so that calls to the roaming mobile can be properly routed over the PSTN by the roaming user's HLR. *The authentication center* is a strongly protected database which control the authentication and encryption key for every single subscriber in the HLR and VLR. The authentication center have a register known as equipment identity register (EIR), which identifies stolen or fraudulently altered phones that transmit data and that does not match with information contained in either the HLR or VLR in a system.

The OSS supports one or several operation maintenance centers (OMC) which are used to monitor and maintain the perfom once of each MS, BS, BSC and MSC within a GSM system.

The OSS has main three functions, which are described as following:

1. To maintain all telecommunication hardware and network operations with a particular market.
2. Manage all charging and billing procedures of an operation.
3. Manage all mobile equipment in the system.

10.2.3 GSM Radio Subsystem

GSM originally used two 25 MHz cellular bands set a side for all member countries, but now it is known as globally in many bands.

The 890-915 MHz band was for the subscriber to base transmission (reverse link) and the 935 to 960 MHz band was for base to subscriber transmission (forward link) for the GSM.

GSM uses FDD and a combination of TDMA and FDMA schemes to provide multiple access to mobile users in a system.

The available forward and reverse frequency bands are divided into 200 KHz wide channel called ARFCNs (absolute radio frequency channel numbers).

The ARFCN shows a forward and reverse channel pair and each channel is time shared between as many as eight subscribers using TDMA.

All the eight subscribers uses the same ARFCN and occupies a unique time slot (TS) per frame.

Radio transmissions on both the forward and reverse link are made at a channel data rate of 270.833 kbps. Thus the signaling bit time or duration is 3.692 u.s and the effective channel transmission rate per is 33.854 kbps (means kilobits per second).

Each TS has a time duration of 576.92 List as shown in Fig. 10.3 with central channel frame and ideal frame.

The total number of available channels within a 25 MHz bandwidth is 125. Since each radio channel consist of eight time slots so there is a total of 1000 traffic channels within GSM. In practical implementations, a guard band of 100 kHz is provided at the upper and lower end of GSM spectrum.

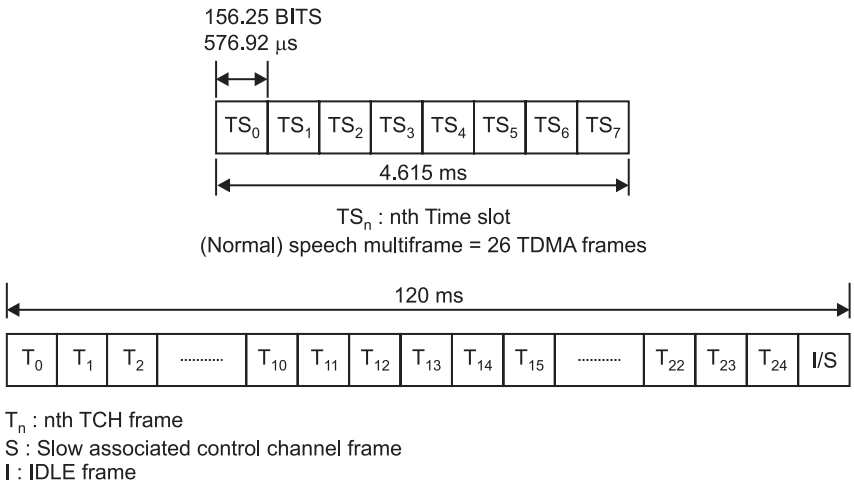


Fig. 10.3: The speech dedicated control channel frame and multiframe structure.

Table 10.1 summarizes the GSM air interface in which define the real value of parameter in terms of specifications.

Table 10.1: GSM Air Interface Specifications Summary

Parameter	Specifications
Reverse Channel Frequency	890-915 MHz
Forward Channel Frequency	935-960 MHz
ARFCN Number	0 to 124 and 975 to 1023
Tx/Rx Frequency Spacing	45 MHz
Tx/Rx Time Slot Spacing	3 Time slots
Modulation Data Rate	270.833333 kbps
Frame Period	4.615 ms
Users per Frame (Full Rate)	8
Time Slot Period	576.9 μs
Bit Period	3.692 μs
Modulation	0.3 GMSK
ARFCN Channel Spacing	200 kHz
Interleaving (maximum delay)	40 ms
Voice Coder Bit Rate	13.4 kbps

10.2.4 GSM Channels Types

A large amount of information is transmitted between the MS (or MSS) and the BS (or BSS). This information user information (voice or data) and control signaling data. Depending on the type of transmitted information, different logical channels are used. These are classified as the:

1. *Traffic channels* (*TCHs*)*:* Used to carry digitally encoded user information or user data and have identical function and formats on both the forward and reverse link in GSM.

2. *Control channels (CCHs) :* Used to carry signaling and synchronization commands between the base station and the mobile station with in a system.

Different logical channels structure in GSM is shown Fig. 10.4 with TCH and CCH types.

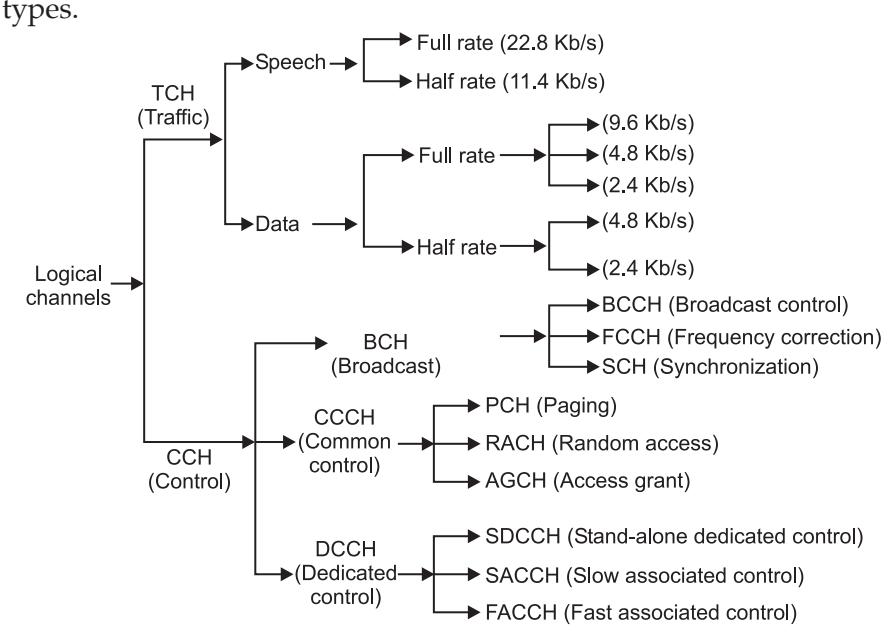


Fig. 10.4: GSM logical channels.

10.2.4.1 Traffic Channels (TCHs)

These are used to carry the encoded speech or user data both in the uplink and downlink directions for communication point of view.

It supports two information rates, full rate and half rate.

TCH are generally duplex in nature not the simplex.

When transmitted as full rate, user data is contained within one TS per frame, and when transmitted as half rate, user data is mapped onto the same time slot but is sent in alternate frames in GSM system.

TCH is divided into main two parts:

(a) Speech channel: It carries digitized user speech which is sampled at 22.8 kb/s for full rate and 11.4 kb/s for half rate for transmit speech information.

(b) Data channel: It carries user data which is sent at 9.8 kb/s, 4.8 kb/s and 2.4 kb/s for full rate and 4.8 kb/s and 2.4 kb/s for half rate for transmit data information.

10.2.4.2. Control Channels (CCHs)

These are used to carry the control and synchronization signals for easy trouble free communication with in a system.

These are main three types- broadcast channels (BCHs), common control channel (CCCH) and dedicated control channel (DCCH), which are divided into many channels, explain as below:

(*a*) *Broadcast channels* (*BCH*)*:* BCH is used to operates on the forward link of the call and transmit data only in the first time slot (TS) of assured GSM frame. It allow synchronization for all mobiles with in the cell.

It is again splited into three types described below:

(*i*) *Broadcast control channel* (*BCCH*) *:* BCCH represent a forward control (base to mobile) channel. This is used to broadcast information such as regarding the mobile's serving cell with the neighbouring cell and network identity.

(*ii*) *Frequency correction channel* (*FCCH*) *:* FCCH represent forward channel (base to mobile) or downlink channel. It carries a transfer information for frequency correction of the MS. This channel is required for the correct operation of the radio subsystem and permit an MS to accurately tune its frequency (local oscillator) to the exact frequency of the base station.

(*iii*) *Synchronization channel* (*SCH*)*:* It transfer the information for the frame synchronization of the MS. The SCH has a 64 bit binary sequence that is previously known to the MS. The MS can achieves the exact timing synchronization with respect to a GSM frame by correlating bits with the internally stored 64 bits. The SCH can carries base station identification code (BSIC) and reduced frame number (RFN) during the burst transmisstion of information.

(*b*) *Common control channel* (*CCCH*) *:* These are primarily used to carry signaling information necessary for accessing management related function. There are two forward link channels (BS -» MS) and one reverse link channel (MS-BS). The forward channels are PCH and AGCH and link channel is RACH. Now it can be again subdivided into three types:

(*i*) *Paging channel* (*PCH*)*:* PCH provides paging information from the base station to all mobiles in the cell and it is a either forward link or downlink channel.

(*ii*) *Random access channel* (*RACH*) *:* It is an uplink channel and is used by mobile to originate a call. It may also be used to show request assignment of a DCCH.

(*iii*) *Access grant channel* (*AGCH*)*:* It is a downlink channel and used by base station to carry data which instruct the MS to operate in a specific physical channel. Generally, AGCH is used by BS to respond to a RACH sent by a MS or MSC.

(*c*) *Dedicated control channel* (*DCCH*) *:* DCCH are bidirectional in nature and used for signaling and control after call establishment. It is again divided into three types and which are described as:

(*i*) *Stand-alone dedicated control channel* (*SDCCH*) *:* SDCCH is used for authentication of the MS, for location updates and for assignment to TCHs. It also secure that the MS and BS remain connected but the BS and the MSC verify the subscriber unit and allocated resources for the mobile.

(*ii*) *Slow associated control channel* (*SACCH*) *:* The SACCH is always associated with a TCH or SDCCH. It is used to forward general control information. It is deployed to transfer link quality and signal strength information in the uplink direction (MS to BS) and slow but continuously, the power control frame adjustment (time) information in the downlink direction (BS to MS).

(*iii*) *Fast associated control channel :* It is related only to a traffic channel (TCH) and used to transfer handoff information during an active call. It is based on frame stealing means it obtain access to the TCH to a time slot for which it is assigned.

10.2.5 Frame Structure for GSM

Each user transmits a burst of data during the time slot alloted to it. These data bursty may have one of five specific formats as defined in GSM and also shown in Fig. 10.5.

Fig. 10.5 illustrates the five type of data bursts used for various control and traffic burst *(i.e.* FCCH, SCH, RACH, dummy and normal burst).

Normal

3 Start BITS	58 BITS of encrypted data	26 Training bits	58 BITS of encrypted data	3 Stop BITS	8.25 BITS guard period

FCCH burst

3 Start BITS	142 Fixed BITS of all zeroes	3 Stop BITS	8.25 BITS guard period

SCH burst

3 Start BITS	39 BITS of encrypted data	64 Training BITS	39 BITS of encrypted data	3 Stop BITS	8.25 BITS guard period

RACH burst

8 Start BITS	41 BITS of synchronization	36 BITS of encrypted data	3 Stop BITS	68.25 BIT extended guard period

DUMMY burst

3 Start BITS	58 Mixed BITS	26 Training BITS	58 Mixed BITS	3 Stop BITS	8.25 BIT guard period

Fig. 10.5: Time slot data burst in GSM.

Normal bursts are used for TCH and DCCH transmission on both the forward and reverse link connection.

The FCCH and SCH bursts are used in the first time slot (TS) of specific frames to broadcast the frequency and time synchronization control message on the forward link connection.

The RACH burst is used by all mobile to access service from any base station.

The dummy burst is used as filter information for unused time slots on the forward link connection.

A simple GSM frame structure is shown in Fig. 10.6 with super frame, multi-frame and normal frame.

There are 8 time slots per TDMA frame and the frame period is 4.615 MS. Therefore a frame contain 8k 1.56.25 = 1250 bits/frame.

The rate of each frame is 270.833 kbps/1250 bits per frame or 216.66 frames per second.

The 13th or 26th frame are not used for traffic, but used for control point of view.

Each of the normal speech frames are grouped into larger structures called multiframes which in turn are grouped into super-frames and hyper-frames. All three are inter related.

One multi-frame contains 26 TDMA frames and one super-frames contains 51 multi-frames or 1326 TDMA frames, as shown in Fig. 10.6

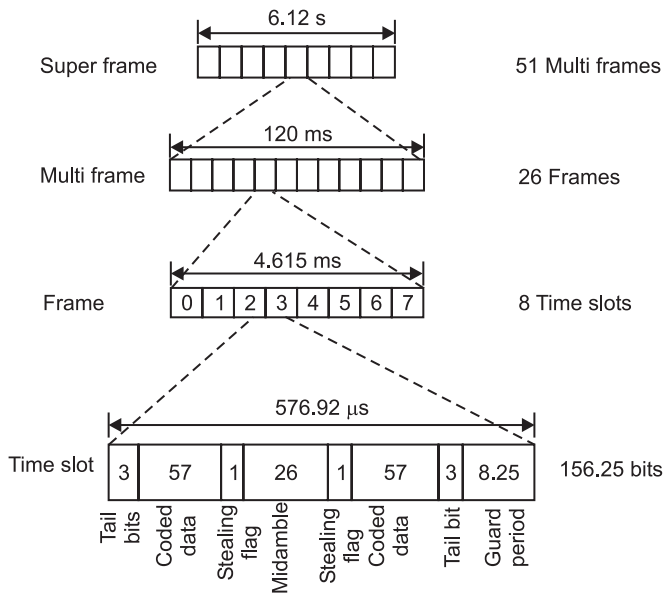


Fig. 10.6: GSM frame structure.

A hyperframe contains 2048 superframe or 2,715,648 TDMA frames. A complete hyper frame is sent about every 3 hours 28 minutes and 54 seconds and is important to GSM since the encryption algorithms rely on the specific frame number and sufficient security can only be obtained by using a large number of frames as provided by the hyper frame.

Each time slot have the tail bits, coded bits, stealing bits, midamble and guard period which are used to provide the information like as whether the time slot (TS) consist voice data or control data, to analyze the radio channel characteristic before decoding the user data, to measure the signal strength, to carry the information etc.

10.2.6 Signal Processing in GSM

Fig. 10.7 represent all of the GSM operations from transmitter to receiver.

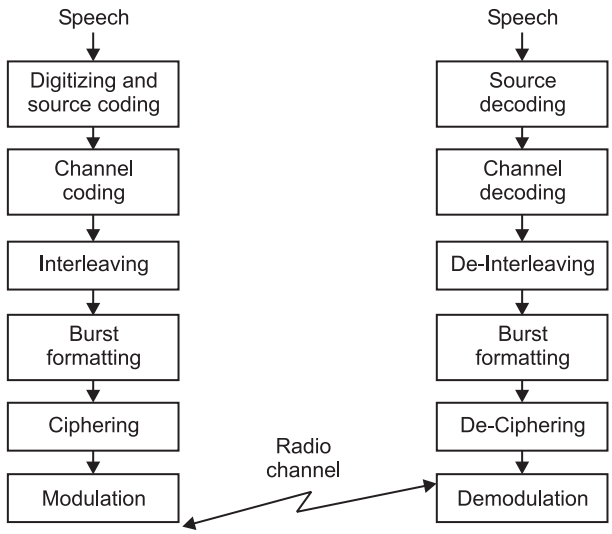


Fig. 10.7: GSM operations from speech input to speech output.

Speech Coding : The GSM speech coder is based on the residually excited linear predictive coder (RELPC), which is extended by including a long term predictor (LTP). The coder provides 260 bits for each 20 ms blocks of speech, which yields a bit rate of 13 kbps (kilobites per second).

TCH/FS, SACCH and FACCH Channel Coding : The output bits of speech coder are divided into two groups for error protection, based upon their significance in contributing to speech quality out of the total 260 bits in a frame, the important 50 bits, known as type la bits have three parity check (CRC) bits added to them. This facilitates the detection of non-correctable error at the receiver. The next 132 bits concern with the first 53 are recordered and appended by four trailing zero bits, thus providing a data block of 189 bits. The error protection coding scheme extend the gross data rate of the GSM speech signal with channel coding to 22.8 kbps data rate.

Channel Coding for Data Channels : The coding provided for GSM full rate data channels is based on handling 60 bits of user data at 5 ms intervals. According to the modified CCITTV 110 modem standard.

Channel Coding for Control Channels : GSM control channel messages are defined to be 184 bits long and are encoded by using a shortened cyclic fire code followed by a half rate convalution coder.

Interleaving : In order to reduce the effect of sudden fades on the received data, the total of 456 encoded bits within each 20 ms speech frame or control message frame are splited into eight 57 bit sub-blocks. These 8 sub-blocks which make up a single speech frame are spread over eight consecutive TCH time slot. If a burst is lost due to interference or fading, channel coding secure that through bits will still be received correctly to allow the error correction to work. Fig. 10.8 show the exactly how the speech frames are diagonally interleaved within the time slots. Note that TS contains 57 bits of data from the 0th sub-block of the wth speech coder frame and 57 bits of data from the 4th sub-block of the (n - l)st speech coder frame.

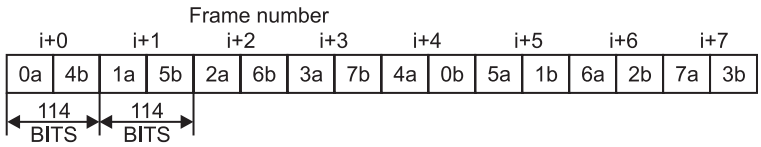


Fig. 10.8: Diagonal interleaving used for TCH/SACCH/FACCH data.

Ciphering : Ciphering modifies the contents of the eight interleaved blocks by using the encryption techniques known only to the particular mobile station and base transceiver station, security is for their enhanced by the fact that the encryption algorithm is varies from call to call. Two types of ciphering algorithms called *A*3 and *A*5 are used in GSM to prevent unauthorized network access and privacy for the radio transmission respectively. The *A3* algorithm is used to authenticate each mobile by verifying the users pass code within the SIM with the cryptographic key at the MSC. The *A5* algorithm provides the scrambling for the 114 coded data bits sent in each TS.

Burst formatting : Burst formatting adds binary data to the chip ciphered blocks, in order to help synchronization and equalization of the received signal.

Modulation : The modulation scheme used by GSM is 0.3 GMSK, where 0.3 describes the 3dB bandwidth of the gaussian pulse shaping filter with relation to the bit rate. The GMSK is a special type of digital FM modulation. The binary ones and zero are represented in GSM by shifting the RF carrier by +67.708 kHz. The channel data rate of GSM is 270.8333 kbps, which is exactly four times the RF frequency shift. This minimizes the bandwidth occupied by the modulation spectrum and hence improve channel capacity. The MSK modulated signal is passed through a gaussian filter to smooth the rapid frequency transitions which would otherwise spread energy into adjacent channels.

Frequency hopping : If user in a particular physical channel is transmitted cell have severe multipath problems, the cell may be defined as a hopping cell by the network operator, in which case slow frequency hopping may be implemented to combat the multipath or interference effects in the cell. Frequency hopping is completely specified by the service provider.

Equalization: Equalization is performed at the receiver with the help of the training sequences transmitted in the middamble of every time slot. The type of equalizer for GSM is not specified and is left up to the manufacturer.

Demodulation: The portion of the transmitted forward channel signal which is off interest to a particular user is determined by the assigned TS and ARFCN. The appropriate TS is demodulated with the aid of synchronization data provided by the burst formatting. After demodulation, the binary information is deciphered de-interleaved channel decoded and speech decoded.

10.3 CDMA DIGITAL CELLULAR STANDARD (LS-95)

The CDMA digital technique is titled as interim standard 95 (IS-95). It is a direct sequence spread spectrum CDMA system.

The main advantage of IS-95 over analog systems (AMPS) is to provide a twenty fold increases in the system capacity for information.

It was standardized by the US telecommunication industry association (TIA) in 1993 and was followed by an extended and revised version (IS-95A) in 1995.

IS-95, provide each user within a cell and users of neighbouring cells to use the same radio channel because it is based on direct sequence spread spectrum (DS-SS) CDMA technology. Hence it can reduce the need for frequency planning.

IS-95 uses a different modulation and spreading technique for the forward and reverse link connection.

On the forward link, the base station simultaneously transmit the user data for all mobile in the cell by using different spreading sequences.

On the reverse link, all mobiles have ideally a constant signal level from the power control (applied by base station to the mobile).

The speech coder used in IS-95 is qualcomm 9600 bps code excited linear predictive coder (QCELPC). This coder reduce the data rate during silent period of the transmission.

10.3.1 Frequency and Channel Specification

The range of frequency for forward and reverse link operation are describe by following equation:

For reverse link operation (MS $\rightarrow$ BS) = 824 - 849 MHz ... (*a*)

For forward link operation (BS $\rightarrow$ MS) = 869 - 894 MHz ... (*b*)

A forward and reverse channel pair is separated by 45 MHz for cellular band operation.

Some channel can be used by many users at a time having maximum user data rate 9.6 kbps (kilo byte per second).

The transmitted data gets spreaded by a spreading factor of 128 therefore the channel chip rate is 1.2288 Mchips/sec (9.6 Kbps x 128).

Every mobile in a given cell is assigned a different spreading sequence, providing perfect separation among the signals from several users.

RAKE receivers are used to resolve and combine the multipath component at both the base station and the subscriber's unit.

10.3.2 Forward CDMA channel

Forward channel is used to transmit the message from base station (BS) to mobile station (MS) (BS → MS).

The forward channel have a pilot channel, a synchronization channel, upto seven paging channel and 63 forward Traffic Channel (TCH).

Explaination of forward channel is given as follows:

(*i*) **Pilot channel**: The pilot channel allows a mobile station to use timing for the forward CDMA channel. Identify multipath component and provides signal strength comparison between each base station for identifying handoff conditions with in a system.

(*ii*) **Synchronization channel**: It provides synchronization message to the mobile stations and operates at 1200 bps. The synchronization channel consist system identification number (SID) and network identification number (NID) with it.

(*iii*) **Paging channel** : These channels are send paging message and controlling information from base station to mobile station and operates at 9600, 4800 and 2400 bits per second (bps). A forward channel can have maximum 7 paging channels.

(*iv*) **Forward traffic channel**: This is used to transmit voice or data to a mobile. There may be upto 63 forward traffic channels. These channels allow variable user data rates at 9600, 4800, 2400 or 1200 bps.

The modulation process for the forward traffic channel is shown in Fig. 10.9 with variable user data rates.

Data on the forward traffic channel is grouped into 20 ms frames.

The user data is first convalutionally coded and then formatted and interleaved to adjust the original user data rate then the signal is spread with a Walsh code and a long PN sequence at a rate of 1.2288 megachips per second.

The speech data rate applied to transmitter varies from 1200 bps to 9600 bps for transmit the data at the base station.

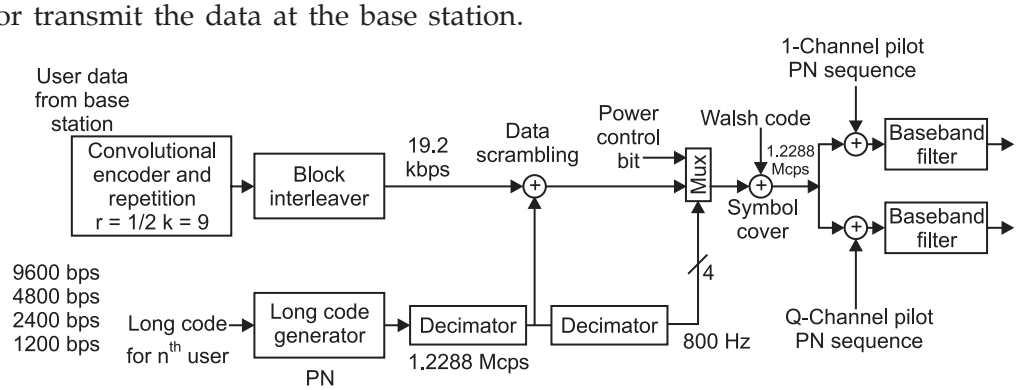


Fig. 10.9: Forward CDMA channel modulation process. 10.3.3. Reverse CDMA Channel :

Reverse channel is used to transmit the message from mobile station to base station (MS → BS).

The reverse channel structure allows upto 62 distinct traffic channels and 32 distinct access channels.

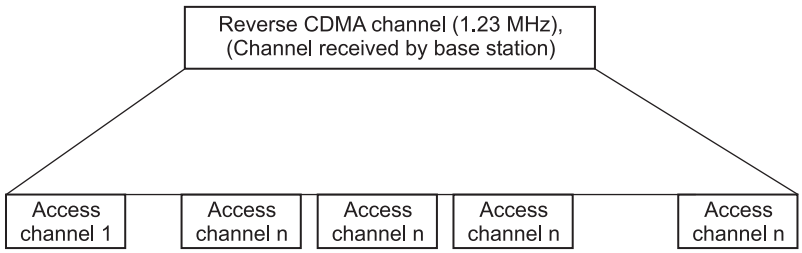


Fig. 10.10: Reverse channel in IS-95.

Details of different reverse CDMA channels is given below which are shown in Fig. 10.10.

(*i*) *Access channels :* An access channel is used when a mobile is required to initiate a call or to respond to paging channel, because many users may try to access at the same time. The data rate for access channel is 4800 bps (bit per second).

(*ii*) *Reverse traffic channel r:* These are used to transmit voice or data from MS to BS. Each reverse traffic channel is paired with a forward traffic channel with different data rate.

The reverse traffic channel modulation process is shown in Fig. 10.11.

User data on the reverse channel are grouped into 20 ms frame.

All the transmit data on the reverse channel are convalutionally encoded, block interleaved and modulated by a 64-ary orthogonal modulation and spread prior to its transmission.

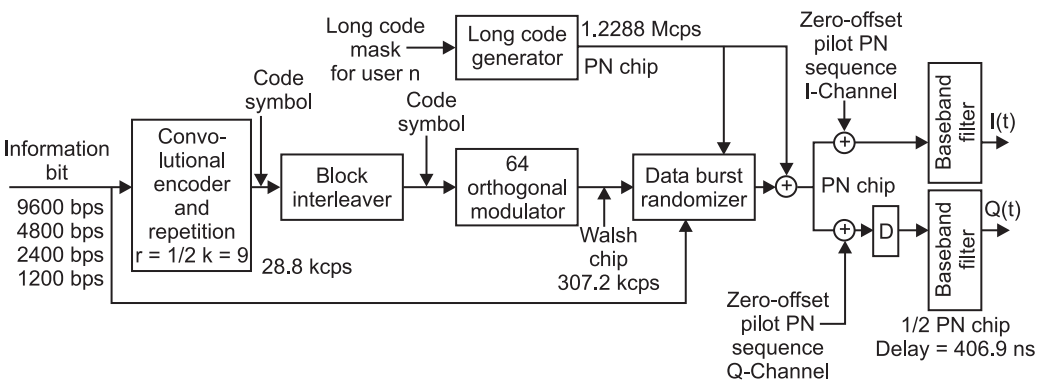


Fig. 10.11: Reverse IS-95 channel modulation process for a single users.

The burst data randomizer generates a masking pattern of 0 ‘s and 1 ‘s that randomly masks the redundant data.

A block of 14 bits taken from the long code and find the masking pattern and the last 14 bits of the long code are used in spreading.

Key features of IS-95 systems

(*i*) Provides extensive path diversity to the system.

(*ii*) Improved performance in difficult propagation environment.

(*iii*) Power control to the system.

(*iv*) Improvement in system capacity and coverage.

10.4 IMT-2000 - A THIRD GENERATION COMMUNICATION

The existing second generation (2G) systems are based on national or regional standards those provide mobility only within their radio environments. Hence they are unable to interoperate with each other, therefore the need of such system exists which will be able to provide world wide mobility. It is obtained by international mobile telecommunication system (IMT-2000).

IMT-2000 is a third generation (3G) system, tried to establish a common worldwide communication system which provide the idea of universal personal telecommunication (UPT).

IMT-2000 is a intended to provide telecommunication services to mobile and stationary users with a wireless link, covering wide range of user sectors (*e.g.* public, private, business, residential) and accomodating distinct types of user equipments (*e.g.* personal pocket terminals, vehicle, mounted terminals).

One of the service objectives of IMT-2000 is to provide multimedia access services.

Until 1997 it was known as future public land mobile telecommunication system (FPLMTS).

The standardization work on the network aspects of IMT-2000 is guide by study group - II (SG-II) of the ITU. This group is responsible for the specification of IMT-2000 signaling interfaces.

Working detail : The IMT-2000 systems family is a federation of IMT-2000 family members and each system within an DVIT-2000 family member will allow IMT- 2000 capabilities using capability sets CS-1, CS-2, CS-3,.....CS-*n*.

The DVIT-2000 family of systems concept is shown in Fig. 10.12.

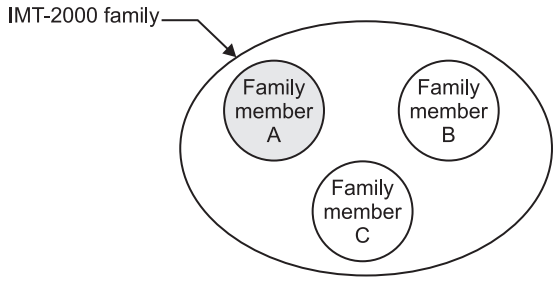


Fig. 10.12: The IMT-2000 family of system concept.

The overall functionality of IMT-2000 is depend on capability sets. These capability sets will reflect increasing level of services. It is shown by as they evolve from the initial capability set 1 (CS-1) to future set (CS-2, CS-3 etc.).

The first capability set perform a lot of function such as user authentication and ciphering user to network and network to network authentication provide multimedia services, provide roaming among the IMT-2000 family geographic position identification etc.

10.4.1 Frequency and channel specification

The range of frequency for reverse and forward link operation is given as following equations:

(*i*) For reverse link operation = 1885 - 2025 MHz ... (*a*)

(*ii*) For forward link operation = 2110- 2200 MHz ... (*b*)

In these frequency bands, 30 MHz frequency have been reserved for mobile satellite services (MSS), separately in each band (forward and reverse).

These frequency bands are designed by world radio conference(WRC).

In addition to the original frequency allocation the world radio conference (WRC) assigned new terrestrial IMT-2000 bands in the range of 800-1000 MHz, 1700-1900 MHz and 2500-2200 MHz.

This approach includes the reuse of 2G spectrum.

10.4.2 Key features of IMT-2000

It provide common frequency spectrum world wide (1.8-2.2 GHz).

It has wide range of telecommunication services (voice, data, multimedia, internet, confrences etc.).

It supports very high data rate upto 2 Mbps (megabit per second).

It can work in multiple radio environments (cellular, cordless LAN's, satellite).

It provide advanced security and performance features.

10.5 UMTS-UNIVERSAL MOBILE TELECOMMUNICATION SYSTEM

The European proposal for IMT-2000 prepared ETSI is known as universal mobile telecommunication system (UMTS).

UMTS will provide an integrated system in which users have access to desired service via uniform service access procedures.

UMTS services can be called indentical, to those offered to users by wireless terminals.

The integration of networks of UMTS system include the switching network and control network and some other important network, which are use in operation.

10.5.1 Channel specification

UMTS will provide highly flexible and efficient radio interface.

It uses a constant chipping rate of 3.84 microchip per second.

The information bits are multiplied with the chipping sequence to extend the signal.

To separate different users, the codes used for spreading should be (quasi) orthogonal *i.e.* their cross correction should be nearly zero or equal to zero.

10.5.2 Key features of UMTS

Distinct users data rate can be supported using in distinct spreading factors.

UMTS will provide a wide range of services to mobile and wireless users for communication.

UMTS can work in a variety of environments such as indoor, rural, outdoor and satellite based.

It provide global radio coverage for transmission.

10.5.3 UMTS network architecture

The network architecture of UMTS consist a large macrocell. There are a large number of microcell in a macrocell as shown in Fig. 10.13.

A microcell is made by combining many picocells as shown in Fig. 10.13.

In this architecture the whole traffic load is splited between macrocells, microcells and picocells.

The highly mobile traffic is operated on the macrocells which the local traffic operates on the micro and picocells. This traffic distribution minimize the number of handoffs required for the fast moving traffic.

As shown in Fig. 10.13, the macrocells cover the area not covered by other cells thus macrocells will also able to ignore to failures of the overlapped cells.

The UMTS architecture will provide radio coverage with a network of base stations interconnected to each other and to a fixed network exchange as shown in Fig. 10.13.

A metropolition area network (MAN) is one of the valid choices for the network interconnection.

Network reliability simply mean that a subcriber will always remain in the reach (approach) of the network even when it changes its location region.

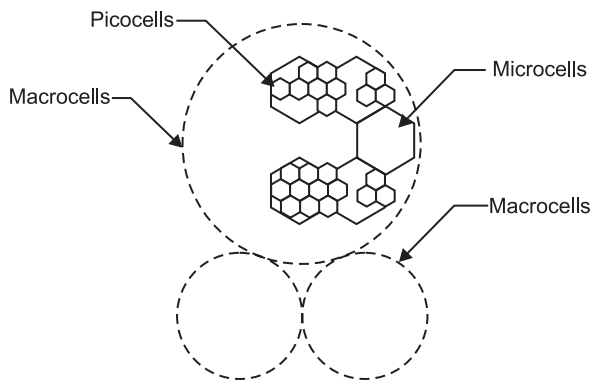


Fig. 10.13: UMTS network architecutre.

The network maintains a constant location information on each of the terminals. This will be updated whenever a mobile changes its location region.

Once the accurate location of the mobile has been accessed then the network will be able to route the calls through distributed network data base.

10.6 PACS- PERSONAL ACCESS COMMUNICATION SYSTEM

It is a third generation (3G) system and proposed by Bellcore in 1992.

It was developed as a universal wireless access system for communication.

The main objective of PACS is to integrate all forms of wireless local loop communication into one system with full telephone features in order to allow wireless connectivity for local exchange carriers (LECs).

It uses $\pi/4$ *A* QPSK for its modulation process.

10.6.1 PACS Architecture

It consist four main components showing in Fig. 10.14. With interfaced and interface.

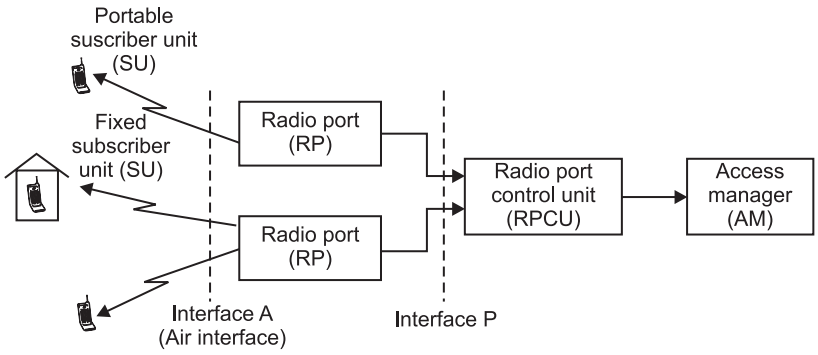


Fig. 10.14: PACS architecture.

1. Subscriber unit (SU) may be mobile handset or fixed PSTN terminal phone or any land line phone.
2. Radio ports (RPs) used to connect the subscriber unit with RPCU.
3. Radio point central unit (RPCU) used to control many radio parts.
4. Access manager (AM) provide the finale access of the subscriber unit over radio ports.

Subscribe units are connected to radio ports through air interface.

Interface *P* provides protocols to connects SU(through the RPS) to RPCU for processing.

10.6.2 Specification of PACS :

(*i*) It can work at both duplexing TDD or FDD technique.

(ii) It uses TDMA multiple access technology.

(*iii*) It uses 7t/4 DQPSK modulation technique.

(*iv*) Channel bit rate is 3.84 Kbps (kilo bit second).

(*v*) One frame duration is 2.5 ms (micro second).

(*vi*) Speech coding is 16 bit ADPCM. ADPCM provides low complexity and minimum cost.

10.6.3 Frame structure

Frame structure of PACS system for forward and reverse link is shown in Fig. 10.15 with TDM.

For forward link, each 2.5 ms frame is spliled into 8 ports. Each part of a frame having 312.5 μs time period.

In forward frame, synchronization bits, control bits, error correction bits and power control bits are transmitted combine with data bits.

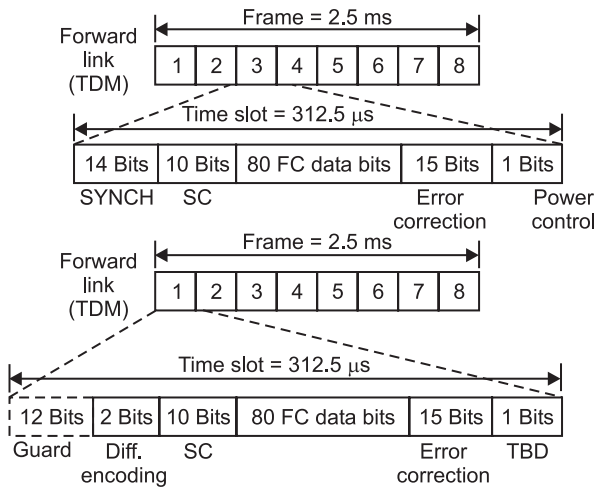


Fig. 10.15: PACS frame structure.

For reverse link, each parts out of 8 parts of a frame contain guard bit, encoding bit, system control bit and error correction bits.

10.6.4 Key Features

(*i*) PACS use four types of channels-system broadcasting channel, synchronization channel, fast channel and slow channel. All have different features.

(*ii*) PACS is a TDMA technology that supports either FDD or TDD, technology.

(*iii*) Reduced co-channel interference within a system.

(*iv*) PACS subscriber unit uses adaptive power control to minimize or, reduce battery consumption.

(*v*) PACS standards contains network intelligence.

10.7 AMPS AND ETACS

Advance mobile phone services (AMPS) was deployed in Chicago by Ameritch in urban and suburban area in late 1983.

AMPS use the 40 MHz frequency band out of 800 MHz band spectrum which are allocated by federal communication commission (FCC).

The AMPS system uses a 7 cell reuse pattern with cell-splitting and sectoring process to increase the capacity of the system.

For 30 kHz AMPS channel, there are 18 dB signal to interference (SIR) is require for satisfactory system performance.

European total access communication system (ETACS) is virtually identical to AMPS but it is scaled for 25 kHz and it was deployed in mid 1980.

AMPS and ETACS both, for radio transmission use the frequency modulation (FM) and frequency division duplexing (FDD).

AMPS and ETACS both are two frequency band one for forward link and second for reverse link.

For AMPS:

824 - 849 MHz = For reverse link (from mobile to base station).

869 - 894 MHz = For forward link (from base station to mobile).

For ETACS:

890 - 915 MHz = For reverse link (from mobile to base station).

935 - 960 MHz = For forward link (from base station to mobile).

The value of maximum deviation of FM modulator if differ for AMPS and ETACS. Means ± 12 kHz for AMPS and ± 10 kHz for ETACS.

The blank and burst data stream and control channel transmission use the 10 kbps AMPS and 8 kbps for ETACS for transmit.

An AMPS system use 21 control channels for each for the two service providers but the ETACS use 42 control channels for a single provider.

The AMPS and ETACS uses the different physical channels rate for the transmission of control and voice information.

A control channels also known as paging or setup channel.

A control channel is used by every base station in a system to simultaneously page subscriber units to alert them of incoming calls and to move connected calls to a voice channel.

The AMPS and ETACS radio interface specifications are shown in Table 10.2.

Table 10.2. AMPS and ETACS Radio Interface Specifications.

Parameter	AMPS Specification	ETACS Specification
Duplexing	FDD	FDD
Multiple Access	FDMA	FDMA
Duplexing	FDD	FDD
Channel Bandwidth	30 kHz	25 kHz
Traffic Channel per RF Channel	1	1
Reverse Channel Frequency	824-894 MHz	890-915 MHz
Forward Channel Frequency	869-894 MHz	935-960 MHz
Voice Modulation	FM	FM
Peak Deviation: Voice Channels	± 12 kHz	± 10 kHz
Control/Wideband Data	±8 kHz	± 6.4 kHz
Channel Coding for Data Transmission	BCH(40,28)onFC BCS(48,36)onRC	BCH(40,28)onFC BCH(48,36)onRC
Data Rate on Control/Wideband Channel	10 kbps	8 kbps
Spectral Efficiency	0.33 bps/Hz	0.33 bps/Hz
Number of Channels	832	1000

10.8 UNITED STATES DIGITAL CELLULAR (IS-54 AND IS-136)

The AMPS (advance mobile phase system) play an important role to provide the connectivity for communication with in a system but AMPS system has one main drawback that it does not support the demand of increased capacity in large cities.

To eliminate this problem, a new technology was develope known as united states digital cellular (IS-54 and IS-136) and this technique also support more users in a fixed spectrum.

The united states digital cellular also increased the number of the voice channels, hence it is also called digital AMPS.

The IS-54 use or support three 8 kbps encoded speech channels in every 30 kHz AMPS channel.

An extended version of IS-54 is known as IS-136, and it operate at 1.8 GHz.

The transmission data rate used by IS-54 is 48.6 kbps at it also use the same frequency spectrum used by AMPS.

It is based on multiple access technique like as TDMA.

The united state digital cellular (IS-54 and IS-136) provide some special services like as call transfer, call forwarding services, multiple party conference, call waiting, calling number identification.

10.9 WIRELESS LOCAL LOOP (WLL)

The fast growth of internet has produce a large demand of broadband internet and computer access for the home and business point of view throughout the whole world.

In the developing country, there is always a telecommunications backbone infrastructure is present and which provide a inexpensive, reliable, rapidly deployable broadband connectivity for exchange the information between each other.

Fixed wireless equipment is perfect suited for broadband internet connectivity in many services or in many instances and this approach is more popular for providing or support "last mile" broadband local loop access as well as point to point or point to multipoint communication.

The fixed wireless communication system has more advantage due to his well-defined, time invarient nature of the propagation channel between the fixed receiver and fixed transmitter.

The modern wireless system are allot the microwave and millimeter radio frequency are 28 GHz band and these frequency are ten time greater then the carrier frequency of 3G terrestrial cellular telephone network.

The carrier frequency of fixed wireless access terminal is very high so some time the radio channel behaves like an optical channel.

At very high frequency (microwave frequency), the fixed wireless network are only viable where there are no obstructions, such in the rural setting or in relatively that suburban.

The microwave wireless link can be used to produce a wireless local loop (WLL).

The local loop is from "last mile" of the telecommunication network that provide the connection between the central office (CO) and the individual homes and business as close as possible.

The WLL technology has an additional benefit that once the wireless equipment is paid for them then there are no additional costs for transport between the customer premises equipment (CPE) and CO.

The WLL system can be completed with copper wire based digital subscriber loop (DSL) technologies.

10.10 BLUETOOTH

The bluetooth specification is a joint or combined development by several companies-IBM, Intel, Nokia, Ericsson.

Bluetooth is a short range system, which support 10 cm to 10 m operating range and this range can be extended up to 100 m and this can be possible by using RF amplifiers for the transmitter.

The bluetooth device are more compact like as cellular and PCS phones, notebook computers and personal digital assistants (PDAs) and computer peripherals such as printers, modems and loads peakers.

By using the bluetooth standard, all these device can communicate with each other.

The bluetooth standard is a part of IEEE 802-11 wireless ethernet standard.

The bluetooth standard operate between 2-4 GHz ISM band and which is provided by IEEE 802.11 standard.

The bluetooth also support spread spectrum technique using frequency hopping.

In bluetooth, the transmitter use two level FSK, with a frequency deviation of between 140 kHz to l75 kHz.

The bluetooth technology use 1 Mb/s channel bit rate and in this system bit rate and band rate both are equal.

The bluetooth system, divide the whole data into packets and each packet is divided into one to five slots and each slot's length is 625 ms.

The simplest bluetooth networks also known as piconets which has 8 nodes.

One bluetooth device can be operated same time on two piconets.

HIGHLIGHTS

1. Global system for mobile communication (GSM) is a second generation cellular system standard and it was developed to solve the fragmentation problems of first generation cellular system in Europe.
2. GSM services follow ISDN guidelines and are classified as:

(*i*) Telephone services

(*ii*) Data services

(*iii*) Supplementary ISDN services.

3. The GSM support two features, one is "SIM" (subscriber's identity module), which is a memory device that store the subscribers related information and second feature is "on the air privacy" in which encrypting the digital bit stream transmit by the GSM transmitter.

4. A GSM system architecture consist three type of subsystems -

(*i*) Base station subsystem (BSS)

(*ii*) Network and switching subsystem (NSS)

(*iii*) Operation support subsystem (OSS). And the mobile station (MS) is also consider as a subsystem.

5. The GSM consist three different database- (*i*) HLR (Home location register) which consist subscriber's information and location information of every user who stay in same area or MSC (*ii*) VLR (Visitor location register), which store the IMSI and customer information for each roaming subscriber who is visiting in the coverage region (*iii*) AUC (Authentication center), which central the authentication and encryption key for every single subscriber in the HLR and VLR.

6. GSM use the logical channels for transmit the information, and these are classified into two parts:

(*i*) Traffic channels-used for carrying digitally encoded user information.

(*ii*) Control channels-used to carry signaling and synchronization commands between the BS and MS.

7. The CDMA digital technique is titled as interim standard 95 (IS-95) and it is a direct sequence spread spectrum CDMA system and the IS-95 provide a twenty fold increases in the system capacity for information.

8. A CDMA or (IS-95) system support two type of channels -

(*i*) Forward CDMA channels-which is used to transmit the information from base station (BS) to mobile station (MS).

(*ii*) Reverse CDMA channels-which is used to transmit the information from mobile station (MS) to base station (BS).

9. IMT-2000 is a third generation (3G) system, tried to establish a common worldwide communication system which provide the idea of universal personal telecommunication (UPT) and untill 1997, it was known as future public land mobile telecommunication system (FPLMTS).

10. The IMT-2000 use two links for transmission of information -

(*i*) Forward link-whose frequency range is 1885-2025 MHz

(*ii*) Reverse link-whose frequency range is 2110-2200 MHz.

11. The European proposal for IMT-2000 prepared by ETSI is known as universal mobile telecommunication system (UMTS) and it also provide an integrated system in which users have access to desired service via uniform service access procedures hence the UMTS can be called identical.

12. The UMTS can work in a variety of environments such as indoor, rural, outdoor and satellite based region and it also provide global radio coverage for transmission.

13. The personal access communication system (PACS) is a third (3G) generation system and proposed by Bellcore in 1992 and it was developed as a universal wireless access system for communication and it use the $\pi/4$ QPSK for its modulation process.
14. The PACS system can use both TDD and FDD duplexing technique for its operations. It use 3.84 kilobits per seconds channel bit rate and the frame duration is 2.5 microsecond.
15. The AMPS (Advance Mobile Phone Services) was deployed in late 1983 and by Ameritch and it use 40 MHz frequency band out of 800 MHz band spectrum which are allocated by federal communication commision (FCC).
16. The AMPS and ETACS both use the frequency modulation (FM) and frequency division duplexing (FDD) for radio transmission.
17. For AMPS, the forward frequency is 869 MHz - 894 MHz and reverse frequency is 824 MHz - 849 MHz but for ETACS, the forward frequency is 935 MHz - 960 MHz and reverse frequency range is 890-915 MHz.
18. The new develope technology united state digital cellular (IS-54 and IS-136) is used to eliminate the drawback of AMPS means the AMPS does not support the demand of increased capacity in large cities and IS-54 and IS-136 also support more users in a fixed spectrum.
19. The IS-54 support or use 48.6 kbps transmission rate and it also support TDMA multiple access technique.
20. The fixed wireless equipment is prefect suited for broadband internet connectivity in many services or in many instances and this approaches is more popular for providing or support "last mile" broadband local loop access as well as point to point to mulitpoint communication.
21. At very high frequency (microwave frequency), the fixed wireless network are only viable where there are no obstructions, such in the rural setting or in relatively flat suburban. So the microwave wireless link can be used to produce a Wireless Local Loop (WLL).
22. The bluetooth specification is a joint or combined development by several companies like IBM, intel, Nokia etc. and the bluetooth is a short range system, which supports 10 cm to 10 m operting range and this range can be extended upto 100 m and this can be possible by using RF amplifiers for the transmitter.
23. The bluetooth system, divide the whole data into packets and each packet is divided into one to five slot and each slot length is 625 μs and the simplest bluetooth network is also known as piconets.
24. During the signal processing in GSM, the transmitter, use the different technology or device for transmission and these are digitizing and source coding system, channel coding system, interleaving system, burst formatting system, ciphering system and modulation process.
25. During the signal processing in GSM, following things are used- source coding system, channel decoding, de-interleaving system, burst formatting system, de-ciphering system and demodulation process are used for receiving the signal from the transmitter.
26. The first generation system are based on analog system while the second generation system used the digital system and the most popular example is GSM and the third generation system tried to establish a common universal system which support worldwide communication.

REVIEW QUESTIONS

10.1. List out the various analog cellular system.

10.2. Write down the forward and reverse link channels in detail for IS-95.

10.3. Draw the reference architecture and signaling system for the GSM and explain each block.

10.4. List out various subsystem of GSM standard.

10.5. What is universal mobile telecommunication system ? Explain its working with reference of distributed data base and network reliability.

10.6. Explain the working of all traffic channels and control channels used in GSM.

10.7. Draw and explain GSM frame structure.

10.8. What is the concept of universal mobile telecommunication system and how it will be useful to the society ?

10.9. Discuss the whole structure of third generation cellular system.

10.10. Give an brief description of system architecture and interface, discuss GSM system for mobile communication.

10.11. Draw and explain the architecture of personal access communication system.

10.12. What are the key features of PACS ? Explain in detail.

10.13. Explain the following terms for GSM systems and their functions performed-

(*i*) MSC

(*ii*) VLR

(*iii*) BSS

(*iv*) HLR

10.14. Explain the processing in GSM.

10.15. Give an comparison between PACS and GSM.

10.16. Explain the following:

(*i*) Wireless local loop

(*ii*) Bluetooth

(*iii*) IS-136

(*iv*) AMPS

10.17. Giving a brief description of system architecture and interfaces, discuss GSM system for mobile communication.

11 Chapter

SATELLITE COMMUNICATION

11.1. *Introduction.* **11.2.** *Element of Satellite Communication.* **11.3.** *Reason for the Downlink Frequency to be Lower than the Uplink Frequency.* **11.4.** *Satellite Frequency Allocation.* **11.5.** *Satellite Orbit Mechanism.* **11.6.** *Orbital Parameters.* **11.7.** *Satellite Subsystem.* **11.8.** *Multiple Access Techniques.* **11.9.** *Satellite Applications.*

11.1 INTRODUCTION

The idea of communication through a satellite, in particular with a synchronous satellite was conceived by Arthur C Clarke, a famous British science fiction writer in 1945.

The latest growing and most recent field of communication involves the use various satellite relays, of which the first was launched in the year 1957. However, it may be noted that the synchronous orbit was not achieved until 1963.

The line of sight communication range was limited due to the Earth's curvature. Hence, for long distance communication we required to develop a different technique like satellite communication.

Satellite communication does not require any cables as repeated station for communication like long distance communication system.

Any point of globe can be communicated with the help of satellite communication. This has made information exchange between two countries easily.

With help of satellite we can cover a specific limited area for transmission the earth station within the coverage area can communicate directly with the satellite.

Satellite communication is a direct line of sight communication direct down to the earth. It solved many problems in the field of communication.

A satellite is launched will above the surface of earth. Satellite communication has a lot of applications.

In satellite communication, for transmission and reception separate frequencies are used so that they do not mix with each other.

Some satellites transmit information and receive control and operating signals from the earth stations. These satellites are used for air traffic control, radar, navigation, survey, weather forecasting etc.

Cameras and antennas are installed on the satellite so that it can take pictures of the earth's surface and other events.

11.2 ELEMENT OF SATELLITE COMMUNICATION

The basic elements of satellite communication system are earth stations, terrestrial system and users.

The general structure of a satellite communication system is shown in Figure 11.1.

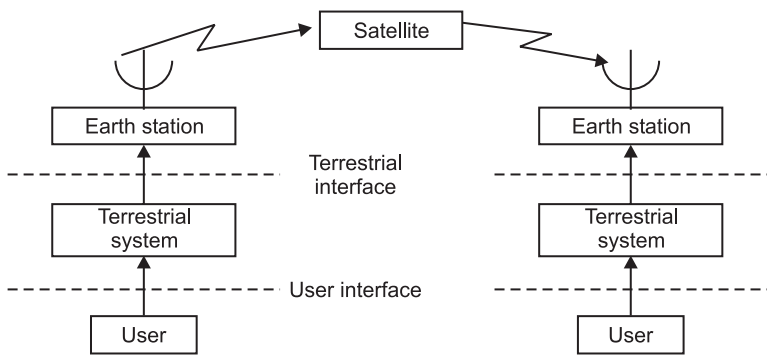


Fig. 11.1: Basic structure of a satellite communication system.

The user is connected to the earth station through a terrestrial network, may be a telephone switch as dedicated link to a earth station.

The Baseband signal generates by the user that is processed and transmitted to the satellite at the earth station. The satellite receives this signal, process it and transmits it back at a reduced frequency.

To avoid the interference downlink frequency spectrum should be different from uplink frequency spectrum.

The receiving earth stations will receive this signal using parabolic dish antennas pointed towards the satellite.

Modern communication satellites also employ frequency reuse to increase the number of transponders in the 500 MHz allocated to them.

The signal which is being transmitted upwards to the satellite is called as the 'up-link' and it is normally at a frequency of 6 GHz and the signal which is transmitted back to the receiving earth station is called as the 'down link' and it is normally at a frequency of 4 GHz.

The operation of satellite takes place at a very high signal frequencies in the microwave range.

The typical band of signal frequency allocations for the communication satellites are as follows:

1. C band : 6/4 GHz

(*i*) C band frequency of 6/4 GHz indicate that the downlink frequency is 4 GHz while the uplink frequency is 6 GHz.

(*ii*) 6/4 GHz band is most popular because of less propagation problem.

(*iii*) Sky attenuation and rain attenuation is low at GHz, so it is possible to build a receiving system.

2. Ku band : 14/12 GHz

14/12 GHz band is also used in commercial communication satellites because of over crowding of geostationary satellites at 6/4 GHz band.

3. Ka band: 30/20 GHz

(*i*) A third band is the 30/20 GHz band where 2.5 GHz band width has been allocated with a uplink 27.5 to 31 GHz and downlink of 18.1 to 21.2 GHz.

(*ii*) The power requirement of a satellite is satisfied by solar panels and a set of nickel cadmium batteries, carried by the satellite itself.

11.3 REASON FOR THE DOWNLINK FREQUENCY TO BE LOWER THAN THE UPLINK FREQUENCY

In satellite communication, frequency of downlink is smaller than that of the uplink, these are given as:

(a) Output power amplifier in transponder

The final power amplifier in the transponder generates more power at lower frequencies than at higher frequencies.

(b) Effective area of the receiving antenna :

(*i*) Effective area of an antenna is directly proportional to the square of the frequency.

(*ii*) The effective capture area of the downlink antenna should be more so as to receive more energy at the earth's receiving antenna. Thus, the frequency should be kept lower.

(c) Path loss

Path loss is less at lower frequencies than at higher frequencies.

(d) Beam width

Beam width should be wider as a satellite would send energy to a large number of earth stations. Lower the frequency wider is the beam width.

11.4 SATELLITE FREQUENCY ALLOCATION

Allocating frequencies to satellite services is a complicated process which requires international coordination and planning.

There are eight frequency bands that have been allocated for the use with satellites communication and these are given as:

Table 11.1: Frequency band for satellite communication.

Band	Uplink band (MHz)	Downlink band (MHz)	Bandwidth (GHz)
1. UHF—Military	292-312 (Appx.)	250-270 (Appx.)	0.02
2. S Band	3200-3700	1800-2300	0.5
3. C Band—Commercial	5925-6425	3700-4200	0.5
4. X Band—Military	7900-8400	7250-7750	0.5
5. Ku Band—Commercial	14000-14500	11700-12200	0.5
6. Ka Band—Commercial	27500-30000	17700-21200	3
7. Ka Band—Military	43500-45500	20200-21200	1
8. V Band—Nonmilitary	50000-51000	40000-41000	1

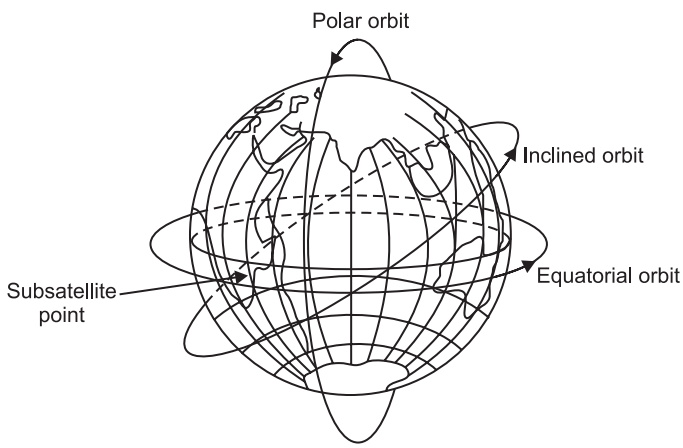


Fig. 11.2: Satellite orbits.

Depending on parameters like period for which satellite is visible to earth, delay time transmission path loss, basis coverage required etc., and the orbits of satellite is accordingly selected.

When the satellite rotates in an orbits above the equator it is called as on equatorial orbit or geostationary orbit. If the satellite rotating in this orbit has the angular velocity equal to the earth's angular velocity, then the satellite is said to be moving along with the earth.

The geostationary orbit the earth in the same direction as the earth's spin and at the same speed.

The rotational period of the earth around its axis is 23 hours and 56 minutes, so, the satellite orbiting in a geostationary orbit must complete its rotation around the earth's in the same period of 23 hours and 56 minutes.

Table 11.2. Frequency allocations for fixed satellite service and broadcasting satellite.

Frequency	Broadcasting satellite	Fixed satellite service
2500-2535 MHz	Down	Down regions II and III
2535-2655	Down	Down region II
2655-2690	Down	Up regions II and III
3400-3700		Down
3700-4200		Down
4500-4800		Down
5725-5850		Up region I
5850-5925		Up
5925-7075		Up
7250-7450		Down
7450-7550		Down
7900-8025		Up
8025-8400		Up
10.7-11.7 GHz	Up region I	Down
11.7-12.1	Down regions I and II	Down region II
12.1-12.2	Down	Down region II
12.2-12.3	Down regions I and II	Down region II
12.3-12.5	Down regions I and II	
12.5-12.75	Down region II	Up/Down
12.75-13.25		Up
14-14.5		Up
14.5-14.8	Up	
17.3-17.7	Up	
17.7-18	Up	
18.1-18.6		Down
18.6-18.8		Down
18:8-20.2		Down
20.2-21.2		Down
22.5-23.0	Down regions II and III	
27-27.5		Up regions II and HI
27.5-29.5		
29.5-31		Up
37.5-39.5		Down
39.5-40.5		Down
40.5-42.5	Down	
42.5–3.5	Down	Up

47.2-49.2	Up	
49.2-50.2		Up
50.4-51.4		Up
71.0-75.5		Up
81-84		Down
84-86	Down	
92-95		Up
102-105		Down
149-164		Down
202-217		Up
231-241		Down
265-275		Up

Table 11.3: Frequency allocation for general mobile satellite service.

Frequency	General mobile
806-890 MHz	Regions II and III
1530-1535	Limited use
1544-1545	Down
1545-1559	Down
1645.5-1646.5	Up
19.7-21.2 GHz	Down
29.5-31.0	Up
39.5-40.5	Down
43.5-47	Up
66-71	Up/Down
71-74	Up/Down
81-84	Up/Down
95-100	Up/Down
134-142	Up/Down
190-200	Up/Down
252-265	Up/Down

11.5 SATELLITE ORBIT MECHANISM

The satellite can be rotated around the earth different paths. These paths are called as orbits of the satellite.

When a satellite is moving in the orbit, it remains in the position as the centripetal force on if balances the gravitational force of the earth. This balance dependent on the distance from the earth, gravitational force on the earth, earth's radius and speed of the satellite.

The satellite orbits are used to cover the required application areas. Fig. 11.2 shows satellite orbits.

11.5.1 Equation of the orbit

A number of coordinate systems and reference planes can be used to describe the orbit of a satellite around a planet.

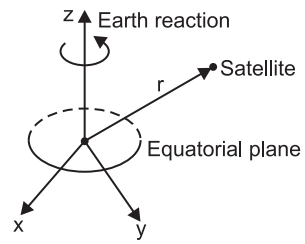


Fig. 11.3: Co-ordinate system used to describe the earth and a satellite.

It is assume that the center of mass of the earth satellite system coincides with the center of mass of the earth at the origin. With the coordinate system setup as in Fig. 11.3 and with the satellite mass m located at a vector distance r from the center of earth.

The gravitational force F on the satellite is given by

$$\overline{F} = \frac{-GM_E m\overline{r}}{r^2} \quad ...(11.1)$$

where

G = (Universal gravitational constant)

$G = 6.672 \times \text{Kr}^{\text{u}}\text{Nm/kg}^2$

m = Mass of earth

$\overline{r}$ = Unit vector in the r direction

We know that

$$\text{Force} = \text{Mass} \times \text{Acceleration} \quad ...(11.2)$$

Equation can be written as

$$\overline{F} = m\frac{d^2\overline{r}}{dt^2} \quad ...(11.3)$$

From equation 11.2 and 11.3, we get

$$\frac{\overline{r}}{r^3}\mu = \frac{d^2\overline{r}}{dt^2} \quad ...(11.4)$$

which yields

$$\frac{d^2\overline{r}}{dt^2} + \frac{\overline{r}}{r^3}\mu = 0 \quad ...(11.5)$$

Equation 11.5 is a second order vector linear differential equation and solution will involve six undetermine constants called the orbital elements, writing it as

$$r = r \cdot \overline{r} \quad ...(11.6)$$

Rearranging the equation 11.4, we get

$$\frac{1}{r}\frac{d^2r}{dt^2}+\frac{\mu r}{r^3} = 0 \qquad ...(11.7)$$

Now, motion defined by above equation is confined to a plane taking r x each term yields.

$$r\times\frac{d^2r}{dt^2} = 0 \qquad ...(11.8)$$

Invoking the rule for finding the derivative of a product and we get

$$\frac{d}{dt}\left(r\times\frac{dr}{dt}\right) = \frac{dr}{dt}\times\frac{dr}{dt}+r\times\frac{d^2r}{dt^2} \qquad ...(11.9)$$

We know that the cross product of any vector with itself is zero, so we can rewrite equation 11.9 as

$$\frac{d}{dt}\left(r\times\frac{dr}{dt}\right) = 0 \qquad ...(11.10)$$

This is equivalent to

$$r\times\frac{dr}{dt} = h \qquad ...(11.11)$$

where, h = constant

and the orbital angular momentum of the satellite.

If the orbit lies in a plane, then h can be a constant. So the problem of motion of satellite in three dimensions reduces the problem of motion in a plane.

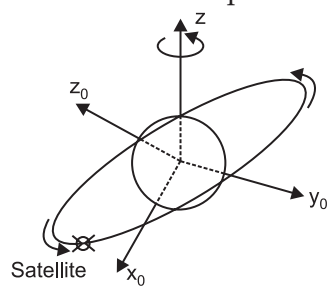


Fig. 11.4: Orbital plane co-ordinate system.

When the x_0 and y_0 axis lie in the orbital plane and z-axis is perpendicular to it.

The solution to equation 11.5 is difficult since the second derivative of r involves the second derivative of the unit vector r. The first two co-ordinate system lie in the plane and z_0 is normal to it.

Now, expressing equation 11.5 in terms of the new coordinate axis x_0, y_0, z_0 gives

$$\bar{x}_0\left(\frac{d^2x_0}{dt^2}\right)+\bar{y}_0\left(\frac{d^2y_0}{dt^2}\right)+\mu\frac{(x_0\bar{x}_0+y_0\bar{y}_0)}{(x_0^2+y_0^2)^{3/2}} = 0 \qquad ...(11.12)$$

Equation 11.12 can be solved easily if it is expressed in polar co-ordinates system ($r_{0'}$ ϕ_0)

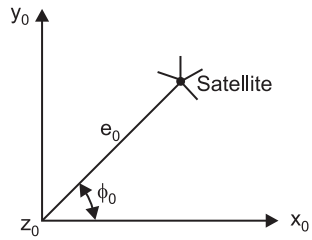


Fig. 11.5. Polar co-ordinate in the orbital plane.

With the polar coordinates system (Fig. 11.5) and using transformations, we get

$$x_0 = r_0 \cos \phi_0 \quad ...(11.13)$$

$$x_0 = r_0 \sin \phi_0 \quad ...(11.14)$$

$$y_0 = r_0 \sin \phi_0 \quad ...(11.15)$$

$$\overline{x}_0 = \overline{\phi}_0 \cos \overline{\phi}_0 - \overline{r}_0 \sin \phi_0 \quad ...(11.16)$$

Equating the vector components of r_0 and ϕ_0 in turn in equation, we get

$$\frac{d^2 r_0}{dt^2} - r_0 \left(\frac{d\phi_0}{dt} \right) = -\frac{\mu}{r_0^2} \quad ...(11.17)$$

and

$$r_0 \left(\frac{d^2 \phi_0}{dt^2} \right) + 2 \left(\frac{dr_0}{dt} \right) \left(\frac{d\phi_0}{dt} \right) = 0 \quad ...(11.18)$$

We can develop an equation for the radius of the satellite's orbit using standard mathematical procedures.

$$r_0 = \frac{P}{1 + e\cos(\phi_0 - \theta_0)} \quad ...(11.19)$$

Here, θ_0 is a constant and e is the eccentricity of an ellipse whose semilatus return P is given by

$$P = \left(\frac{h^2}{\mu} \right) \quad ...(11.20)$$

where h is magnitude of the orbital angular momentum of the satellite.

11.6 ORBITAL PARAMETERS

The satellite movement in an orbit follows the three Kepier's laws. According to first law, the satellite moves along a orbit on the orbit plane, so the equation, for which is:

$$r = \frac{P}{(1 + e\cos\phi_0)} \quad ...(11.21)$$

where

$$P = \text{Parameter}$$

e = Ecentricity

v = Central angle

Fig. 11.6 represent the above orbit of satellite. Infact the value of the ecentricity e determines the type of the orbit.

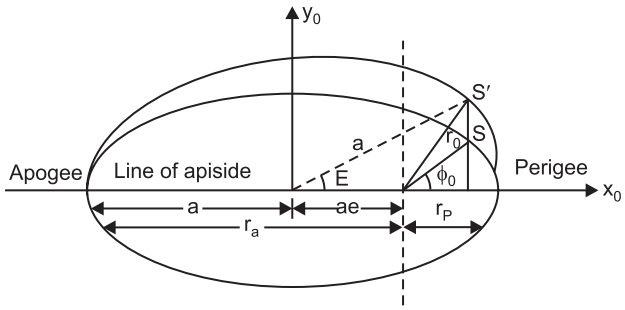


Fig. 11.6: Geometry of on elliptical orbit of communication satellite.

$$e = \begin{cases} 0 \rightarrow \text{Circle} \\ <1 \rightarrow \text{Ellipse} \\ 1 \rightarrow \text{Parabola} \\ >1 \rightarrow \text{Hyperbola} \end{cases} \quad ...(11.22)$$

In Fig. 11.6, a is the semi-major axis and b is semi minor axis and is linked to the orbital period by the expression.

$$T = 2\pi\sqrt{\frac{a^3}{\mu}} \quad ...(11.23)$$

where,

$$\mu = GM \quad ...(11.24)$$

G = Gravitation constant = 6.6×10^{-11} m^3/kgs^2

M = Earth mass = 5974×10^{14} kg

So, $\mu = 3.98 \times 10^4 m^3/s^2$

Satellite velocity for a circular orbit = $\sqrt{\frac{\mu}{a}}$ ...(11.25)

Infact the satellite velocity at on orbital point s (distance to the centre of the earth r) is given by

$$V_s^2 = \frac{2\mu}{r} - \frac{\mu}{a} \quad ...(11.26)$$

According to the Fig. 11.6 following definitions are made.

$$x = r_0 \cos \phi_0 = a (\cos E - e) \quad ...(11.27)$$

$$y = r_0 \cos \phi_0 = a \sin E(1 - e^2)^{1/2} \quad ...(11.28)$$

Ecentricity, $e = \frac{c}{a} = \frac{(r_\alpha - r_P)}{(r_a + r_P)}$ ...(11.29)

Apogee distance, $r_a = a + c = a(1 - c)$ ... (11.30)

Perigee distance, $r_P = a - c = a(1 - c)$ ... (11.31)

Semi-major axis, $$a = \left(\frac{r_a + r_P}{2}\right) = \frac{P}{(1-e^2)} \quad ...(11.32)$$

Locus parameter, $$P = a(1 - e^2) \quad ...(11.33)$$

$$P = \frac{2r_a r_P}{(r_a + r_P)}$$

Semi-minar axis, $$b = a(1 - e^2)^{1/2} \quad ...(11.34)$$

$$= (r_a r_P)^{1/2}$$

11.7 SATELLITE SUBSYSTEM

A complete satellite consists of several subsystems, but most important of them are :

1. Power supply subsystem
2. Attitude and orbit control system (AOCS)
3. Telemetry, tracking command and monitoring (TTC and M)
4. Communication subsystem.

11.7.1 Power Supply Subsystem

(*i*) The communication satellite got electrical power from cells. The solar cells are photo voltaic device that convert the sun's light energy into electrical energy.

(*ii*) The conversion efficiency of solar cell is around 12% to 15% and typical satellite solar cells generate 1 to 2 kW of power supply.

(*iii*) During the time of election of satellite into transfer orbit, the electro mechanical accumulators provide the electric power supply.

(*iv*) A spin stabilized satellite has a cylindrical body. The body is covered with solar cell. But as body is cylindrical only half of the solar cells are illuminated. Thus, the generation of less electrical power.

(*v*) A three axis stabilized has the cells arranged on flat panels. The cells can be rotated in order to maintain the incident of satellite.

(*vi*) Large communication satellites used for direct broadcast can generate power into 6 kW from the solar power.

11.7.2 Attitude and Orbit Control System (AOCS)

(*i*) Attitude and orbit control is used to control the orbit of satellite, besides helping to maintain stabilization and its position.

(*ii*) To move the satellite back to the correct orbit rocket motors, gas gets, internal devices are used. They control altitude of satellite.

(*iii*) The external forces acting on the satellite cause it to drift from its original position.

(*iv*) There are four operations needed for active attitude control.

I Detection of the satellite altitude

II Comparison with the reference axis

III Determination of the corrective torques

IV Correction of the attitude by actuators mounted **on** the satellite.

(*v*) It depends on weather the satellite is spin **stabilized or three-axis body stabilized.**

(*vi*) For orbit control, sensors in the satellite are used **to measure linear acceleration** changes in velocity, sensed by the velocity sensors, **are transmitting to earth's** controlling station.

(*vii*) In this control system, momentum wheels **cannot be used because they cause** rotational torques.

(*viii*) The changes are analyzed and appropriate commands are generated and sent to the satellite for correcting the velocity.

11.7.3 Telemetry, Tracking and Command (TT&C)

(*i*) TT&C subsystem is the most essential subsystem of the satellite because it kelps the satellite in working order and all sorts of manoeuvre are carried out through them.

(*ii*) These systems are partly on earth's station and partly on the satellite.

(*iii*) The telemetry system provides a data stream to the ground that reports on each of the satellite sub-systems independent of any other communications link.

(*iv*) The tracking system is located at the earth station and provides information on the range, elevation and azimuth angles of the satellite.

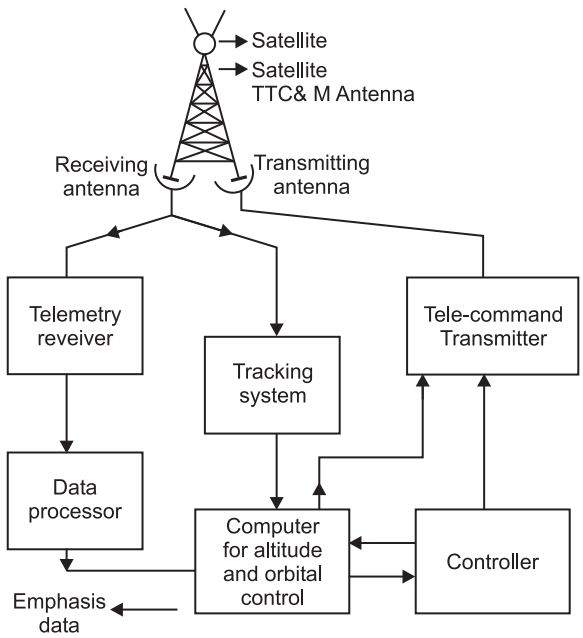


Fig. 11.7: Telemetry, tracking command and monitoring system.

(*v*) If there is any change in the orbital elements such as range, elevation and azimuth angles of the satellite then they are detected by the tracking system.

(*vi*) For command operation, the satellite contains a receiver which works only for command signals transmitted from TT&C earth station.

(*vii*) Using this receiver the command subsystem controls the satellite operation through all phases of the mission by receiving and decoding command from the TT&C earth station.

(*viii*) The satellite is facing the earth, it is supplied the entire electrical power so that control can be handed over to the TT&C and monitoring system.

11.7.4 Communication Sub-systems

(*i*) The communication subsystem is the major part of the operation of a space craft. It provides the receive and transmit coverage for the satellite.

(*ii*) The communication subsystem consist of a communication antenna and communication repeater.

(*iii*) The antenna may be used to provide a signal link for the telemetry, command and ranging system.

(*iv*) Repeaters provides the receive and transmit coverage for satellite.

11.8 MULTIPLE ACCESS TECHNIQUES

Communication satellites are used to carry telephone, video, and data signals can use both analog and digital modulation techniques.

The satellite transmission bandwidth in different frequency band varies in the range of 0.5 GHz to 3 GHz. This range bandwidth is much more than the bandwidth needed for a message to be transmitted. Thus, in order to utilize the frequency spectrum efficiency multiple access are introduced.

The multiple access provides wide geography coverage capability and the best utilization of broadcast nature of satellite channel.

Multiple access techniques allow interconnection among large number of earth station terminals simultaneously via satellite and multiple access techniques any one.

Earth station can communicate with all other earth stations using the same satellite.

The broadcast network is also on example of multiple access (multiuser) communication system. The down link transmission of satellite, TV transmission and radio are example of such system.

There are some multiple access techniques:

(*i*) Frequency division multiple access (FDMA)

(*ii*) Time division multiple access (TDMA)

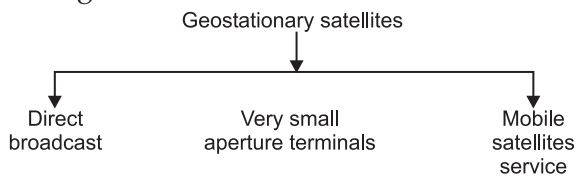
(*iii*) Space division multiple access (SDMA)

(*iv*) Code division multiple access (CDMA).

11.9 SATELLITE APPLICATIONS

In the field of satellite communication applications, telephone, data communication and satellite television are the major application areas.

The geostationary and other types of satellites can provide different types of services all over the globe.



The other application of geostationary satellites include 'Radarset' and Global Positioning Satellite System.

Satellites can also be used in a variety of scientific experimental studies including studying the characteristics of upper region of the atmosphere, probing others stars and plants.

Satellite can be used for wheather, TIROS-1 was launched by USA in 1960 and was used by a large number of countries to monitor the behaviour of atmosphere.

11.9.1 Direct Broadcast Satellite System

The direct broadcast satellite is used for home broadcasting of TV signals.

Planned broadcasting directly to home TV receivers takes place in the Ku-band (12 GHz). This service is known as direct broadcast satellite (DBS) service.

The major features of direct broadcast satellite is that it allows earth station to be mounted in domestic area.

The DBS TV transmission uplink frequency is about 14.5-14.8 GHz and the down link frequency is about 12 GHz. The dish size is of the order of 0.4 to 0.75 metre.

11.9.2 Very Small Aperture Terminals

The very aperture terminals (VSATs) satellite terminal network is used for long distance low density voice and data communication.

In VSAT the earth station antenna size is typically less than 2.4 urn in diameter.

They generally operate in the KU band, with earth stations diameters of 1 to 2 m. The transmitter power are 1 W or 2 W.

VSAT system operate in a star configuration, which means that the connection of one VSAT to another must be made through the hub.

The basic structure of a VSAT network consists of a hub station which provides a broadcast facility to all the VSATs in the network, VSATs themselves which access the satellite in some form of multiple access.

It is most economical transmission architecture for point to point communication. VSAT networks take the advantage of wide area broadcast capabilities of GEO satellites.

A geostationary satellite links a large number of VSATs with the main switching centre of region or a city. Every VSATs acts as a link to the local switching centre in village with final mile of telephony link being carried over wireless local loop.

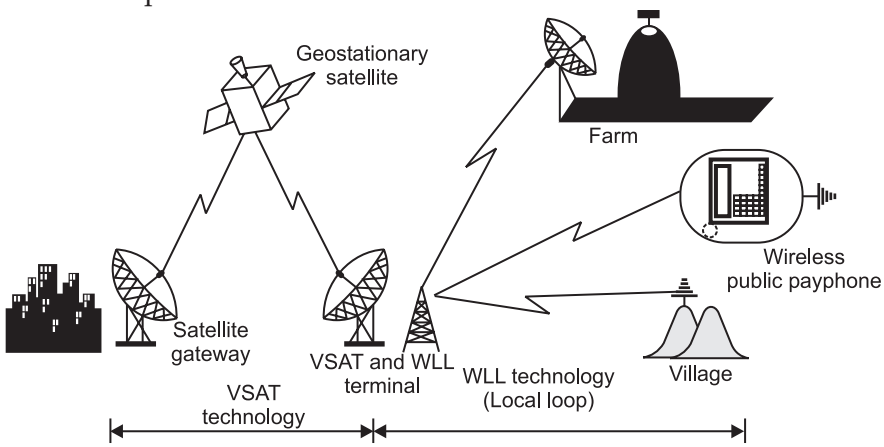


Fig. 11.8: VSAT/WLL communications network.

It uses wireless local loop (WLL) coupled with VSAT distribution architecture as shown in Fig. 11.8.

The basic requirements of such VSAT networks require the use of low power transmitter with small antenna but relatively high EIRP form the satellite.

11.9.3 Mobile Satellite Communication System (MSAT)

Radio communication in VHF and UHF bands between 30 MHz and 30000 MHz takes place in the transposphere.

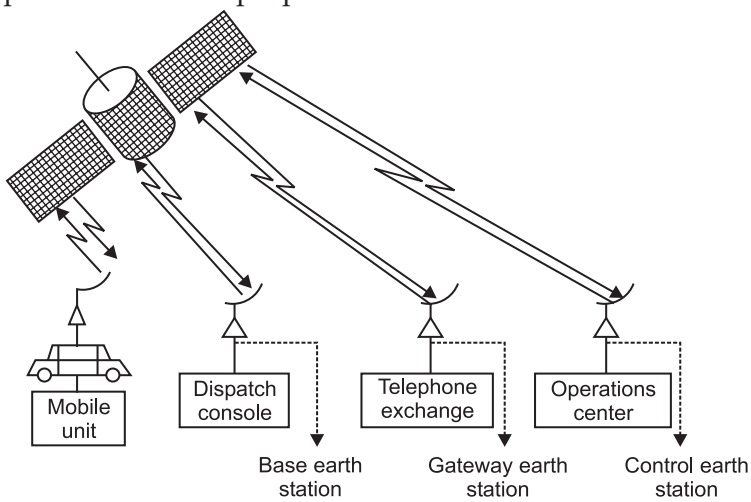


Fig. 11.9: Network structure of MSAT.

The two way communication system use in VHF/UHF bands is communication between a fixed base station and many mobile stations situated in vehicles, aircraft and ships in the frequency band of 30 MHz to 470 MHz.

MSATs are superior to cellular radio services which can be found in western nations on cars, buses, trucks, train and other specialized vehicles.

MSAT systems may be easily said to have a marriage between satellite and mobile phones. For example, Canadian MSAT and the MSS of United States, it provide direct satellite service to mobile phones and thus are bound to provide greater improvement in rural communications.

In such mobile satellite systems, the direct service by satellite to mobile phones requires one or more satellite geosynchronous orbit, with high effective radiated power and operating in the 800 MHz range.

Direct communication links can be established between a satellite and a variety of stations: mobile units, portable units, base earth station, gateway earth stations, and a central earth station.

The MSAT network structure is capable of accomodating for kinds of communications, these are:

(*i*) Mobile to mobile

(*ii*) Mobile to dispatcher

(*iii*) Mobile to public switched telephone network

(*iv*) Satellite and network control.

These network is capable of serving many mobile stations, a some what smaller number of base station and gateways and the operations center.

MSAT terminals give the communication services to the moving vehicles. MSATs terminals can be called as VSAT-on-wheels, have several features common with VSATs.

Mobile station use two classes of mobile antennas namely the steered and non-steered. Steered antennas (with more than 8 dB gain) is directional with respect to azimuth and tracks the satellite being use as the mobile units and the non-steered (4 to 8 dB gain) is omnidirectional with regard to azimuth.

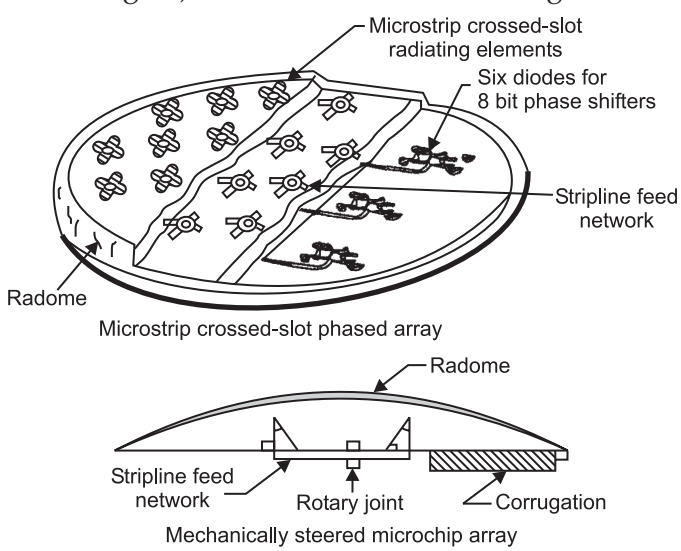


Fig. 11.10: Steered antennas

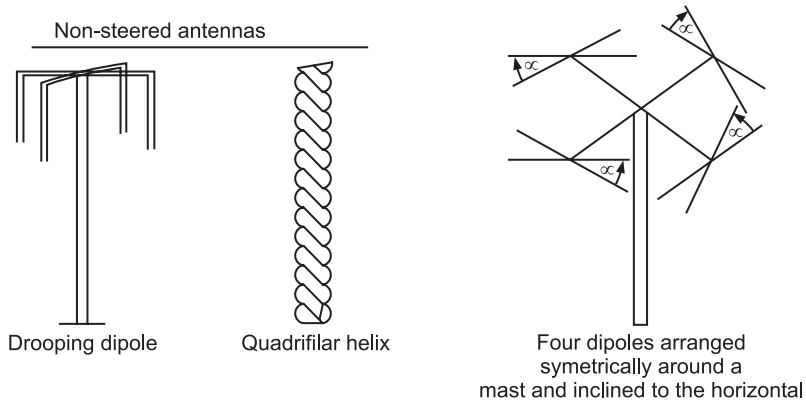


Fig. 11.11.

HIGHLIGHTS

1. The idea of communication through a satellite in particular with a synchronous satellite was conceived by Arthur C Clarke a famous British science fiction writer in 1945.
2. Satellite communication does not require any cables or repeater stations for communication like long distance communication system.
3. With help of satellite we can cover a specific limited area for transmission. The earth station within the coverage area can communicate directly with the satellite.
4. Some satellites transmit information and receive control and operating signals from the earth stations. These satellites are used for air traffic control, radar, navigation, survey, weather forecasting etc.
5. The basic elements of satellite communication system are earth stations, terrestrial system and users.
6. The receiving earth stations will receive this signal using parabolic dish antennas pointed towards the satellite.
7. Modern communication satellites also employ frequency reuse to increase the number of transponders in the 500 MHz.
8. The power requirement of a satellite is satisfied by solar panels and a set of nickel cadmium batteries, carried by the satellite itself.
9. In satellite communication, frequency of downlink is small than that of the uplink.
10. The final power amplifier in the transponder generates more power at lower frequencies than at higher frequencies.
11. The effective capture area of the downlink antenna should be more so as to receive more energy at the earth's receiving antenna. Thus, the frequency should be kept lower.
12. Beam width should be wider as a satellite would set. Energy to a large number of earth stations. Lower the frequency wider is the beam width.
13. If the satellite rotating in this orbit, has the angular velocity equal to the earth's angular velocity, then the satellite is said to be moving along with the earth.

14. The geostationary orbit and the earth in the same direction as the earth's spin and at the same speed.

15. When a satellite is moving in the orbit, it remains a position as the centripetal force on it balances the gravitational force of the earth. This balance dependent on the distance from the earth gravitational force on the earth, earth's radius and speed of the satellite.

16. A number of coordinate systems and reference planes can be used to describe the orbit of a satellite around a planet.

17. A complete Satellite consists of several subsystems, but most important of these are:

(*a*) Power supply system

(*b*) Attitude and orbit control system (AOCS)

(*c*) Telemetry, tracking command and monitoring (TTC&M)

(*d*) Communication subsystem.

18. The communication satellites set electrical power from cells. The solar cells are photo voltaic device that convert the sun's light energy into electrical energy.

19. TT&C subsystem is the most essential subsystem of the satellite because it keeps the satellite in working order and all sorts of manoeuver are carried out through them.

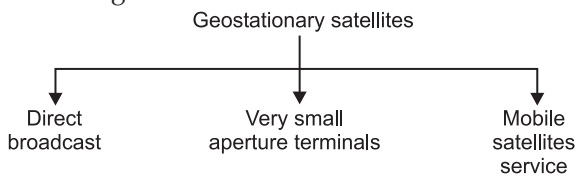
20. The communication subsystem is the major part of the operation of a space craft. It provides the receive and transmit coverage for the satellite.

21. The antenna may be used to provide a signal linker link for the telemetry, command and ranging system.

22. Communication satellites are used to carry telephone, video, and data signals can use both analog and digital modulation techniques.

23. Multiple access techniques allow interconnection among large number of earth station terminals. Simultaneously via satellite and multiple access techniques any one.

24. The geostationary and other types of satellites can provide different types of services all over the glob.



25. Satellite can also be used in a variety of scientic experimental studies including, studying the characteristics of upper resion of the atmosphere, probing other stars and planets.

26. The very aperture terminals (VSAT) satellite terminal network is used for long distance low density voice and data communications.

27. A geostationary satellite links a larger number of VSATs with the main switching centre of region or a city. Every VSATs acts as a link to the local switching centre in village with final mile of telephony link being carried over wireless local loop.

28. The two way communication chief use in VHF/UHF bands is communications

25. between a fixed base station and many mobile stations situated in vehicles, aircraft and ships in the frequency band of 30 MHz to 470 MHz.

29. The MSAT network structure is capable of accomodating for kinds of communications. These are:

(*a*) Mobile to mobile

(*b*) Mobile to dispatcher

(*c*) Mobile to public switched telephone network

(*d*) Satellite and network control

30. Steered antennas (with more than 8 dB gain) is directional with respect to azimuth and tracks the satellite being use as the mobile units.

REVIEW QUESTIONS

11.1. What do you mean by satellite communication ? Explain in brief.

11.2. What are the elements of a satellite communication ? Explain each of them with a suitable block diagram.

11.3. Define the following terms:

(*i*) Uplink frequency

(*ii*) Downlink frequency

11.4. Why the downlink frequency lower then the uplink frequency ?

11.5. Define the following bands:

(*i*) C-band

(*ii*) Ku-band

(*iii*) Ka-band.

116. Explain in brief satellite frequency allocation.

11.7. What do you means by orbital mechanism of satellite and also explain the equation of orbit?

11.8. Draw the block diagram of a satellite showing various subsystems and explain the function of each block in brief.

11.9. Discuss the various application area of VSAT and data broadcast satellite.

11.10. Describe the telemetry, tracking and command subsystem with reference to an earth station

11.11. What do you means by multiplexing and also explain the multiple access techniques?

11.12. Define the orbital parameters?

11.13. Define the following subsystems:

(*i*) Communications subsystems

(*ii*) Power supply subsystems.

11.14. Describe the field of satellite communication application.

11.15. Discuss the direct broadcast satellite system.

11.16. Discuss the mobile satellite communication system ?

11.17. Write down the advantages and disadvantages of satellite communication.

11.18. Give block diagram and description of earth station and transponder used for satellite communication.

12 Chapter

MODULATION TECHNIQUES FOR MOBILE RADIO

12.1. *Introduction.* **12.2.** *Frequency Modulation vs. Amplitude Modulation.* **12.3.** *Amplitude Modulation.* **12.4.** *Angle Modulation.* **12.5.** *Digital Modulation-an Overview.* **12.6.** *Linear Modulation Techniques.* **12.7.** *Differential Phase Shift Keying (DPSK).* **12.8.** *Quadrature Phase Shift Keying (QPSK).* **12.9.** *QPSK Transmission and Detection Techniques.* **12.10.** *π/4 QPSK.* **12.11.** *π/4 QPSK Transmission Techniques.* **12.12.** *π/4 QPSK Detection Techniques.* **12.13.** *Constant Envelope Modulation.* **12.14.** *Combined Linear and Constant Envelope Modulation Techniques.*

12.1 INTRODUCTION

Modulation is the process of encoding information from a message source in a manner suitable for transmission. It generally involves translating a baseband message signal (called the source) to a bandpass signal at frequencies that are very high when compared to the baseband frequency.

The bandpass signal is called the modulated signal and the base band message signal is called the modulating signal. Modulation may be done by varying the amplitude, phase, or frequency of a high frequency carrier in accordance with the amplitude of the message signal.

Demodulation is the process of extracting the baseband message from the carrier so that it may be processed and interpreted by the intended receiver (also called the sink).

This chapter describes various modulation techniques that are used in mobile communication system. Analog modulation schemes that are employed in first generation mobile radio system, as well as digital modulation scheme proposed for use in present and future systems, are covered.

Since digital modulation offers numerous benefits and is already being used to replace conventional analog system, the primary emphasis of this chapter is on digital modulation schemes. However, since analog systems are in widespread use, and will continue to exist, they are treated first.

The utility goal of a modulation technique is to transport the message signal through a radio channel with the best possible quality while occupying the least amount of radio spectrum, new advances in digital signal processing continue to bring about new forms of modulation and demodulation.

This chapter describes many practical modulation schemes, receiver architectures, design tradeoffs, and their performance under various types of channel impairments.

12.2 FREQUENCY MODULATION vs. AMPLITUDE MODULATION

Frequency modulation (FM) is the most popular analog modulation technique used in mobile radio systems. In FM, the amplitude of the modulated carrier signal is kept constant while its frequency is varied by the modulating message signal. Thus, FM signals have all their information in the phase on frequency of the carrier.

In amplitude modulation (AM) schemes, there is a linear relationship between the quality of the received signal and the power of the received signal since AM signals superimpose the exact relative amplitudes of the carrier.

FM offers many advantages over amplitude modulation (AM), which makes it a better choice for many mobile radio applications.

Frequency modulation has better noise immunity when compared to amplitude modulation. Since signals are represented as frequency variations rather than amplitude variations, FM signals are less susceptible to atmospheric and impulse noise, which tend to cause rapid fluctuations in the amplitude of the received radio signal.

Message amplitude variation do not carry information in FM, so burst noise does not affect FM system performance as much as AM systems, provided that the FM received signal is above the FM threshold.

In an FM system, it is possible to tradeoff bandwidth occupancy for improved noise performance. Unlike AM, in an FM system, the modulation index, and hence bandwidth occupancy, can be varied to obtain greater signal-to-noise performance.

AM signals are able to occupy less bandwidth as compared to FM signals, since the transmission system is linear.

An FM signal is a constant envelope signal, due to the fact that the envelope of the carrier does not change with changes in the modulating signal. Hence the transmitted power of an FM signal is constant regardless of the amplitude of the message signal. The constant envelope of the transmitted signal allows efficient class *C* power amplifiers to be used for RF power amplification of FM. In AM, however, it is critical to maintain linearity between the applied message and the amplitude of the transmitted signal, thus linear class *A* or *AB* amplifiers.

FM systems require a wider frequency band in the transmitting media (generally several times as large as that needed for AM) in order to obtain the advantages of reduced noise and capture effect.

Both AM and FM may be demodulated using inexpensive non-coherent detectors.

AM is easily demodulated using an envelope detector whereas FM is demodulated using a discriminator or slope detector.

AM may be detected coherently with a product detector, and in such cases AM can out perform FM in weak signal conditions since FM must be received above threshold.

12.3 AMPLITUDE MODULATION

In amplitude modulation, the amplitude of a high frequency carrier signal is varied in accordance to the instantaneous amplitude of the modulating message signal.

If $A_c \cos(2\pi f_c t)$ is the carrier signal and $m(t)$ is the modulating message signal, the AM signal can be represented as

$$S_{AM}(t) = A_C\,[1 + m(t)]\cos(2nf_c t) \qquad ...(12.1)$$

The modulation index K of an AM signal is defined as the ratio of the peak message signal amplitude to the peak carrier amplitude. For a sinusoidal modulating signal

$$m(t) = \left(\frac{A_m}{A_C}\right)\cos(2\pi f_m t),$$ the modulation index is given by

$$K = \frac{A_m}{A_C} \qquad ...(12.2)$$

The modulation index is often expressed as a percentage, and is called percentage modulation.

Fig. 12.1. shows a sinusoidal modulating signal and the corresponding AM signal. For the case shown in Fig. 12.1, $A_m = 0.5\,A_C$, and the signal is said to be 50% modulated.

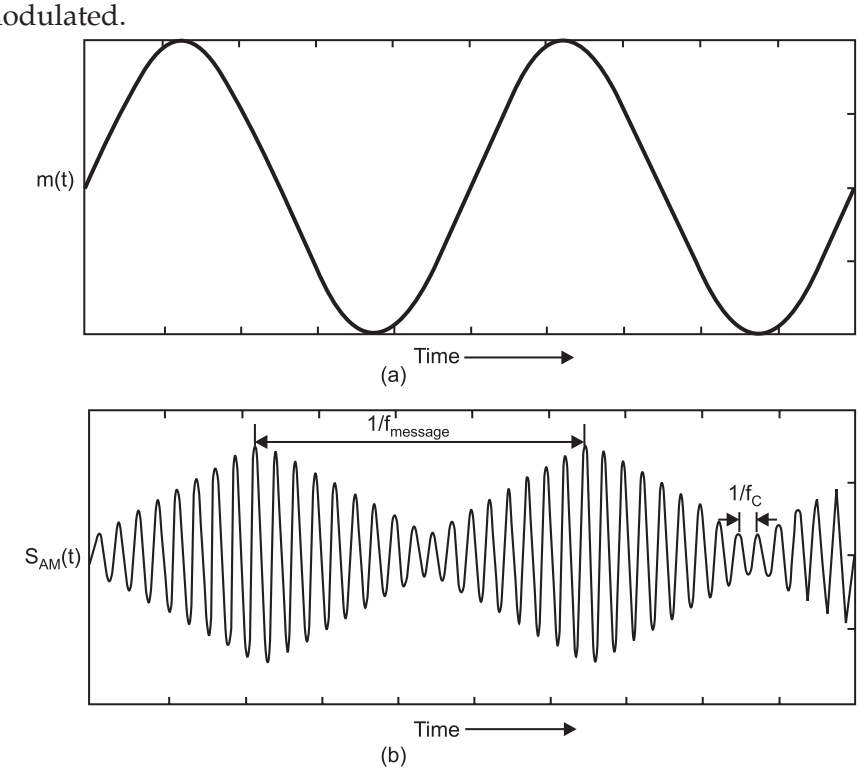


Fig. 12.1: (a) A sinusoidal modulating signal and (b) the corresponding AM signal with modulation index 0.5

A percentage of modulation greater than 100% will distort the message signal if detected by an envelope detector. Equation 12.1 may be written as

$$S_{AM}(t) = Re\ [g(t)\ \exp\ (j2\pi f_c t)] \quad ...(12.3)$$

where $g(t)$ is the complex envelope of the AM signal given by

$$g(t) = A_c\ [1 + m(t)] \quad ...(12.4)$$

The spectrum of an AM can be written as

$$S_{AM}(f) = \frac{1}{2}A_c[\delta(f-f_c)+M(f-f_c)+\delta(f+f_c)+M(f+f_c)] \quad ...(12.5)$$

where 5(*) is the unit impulse function and $M(f)$ is the message signal spectrum. Fig. 12.2 represent an AM spectrum for a message signal whose magnitude spectrum is a triangular function. From Fig. (12.2), the AM spectrum consists of an impulse at the carrier frequency and two sidebands, which replicate the message spectrum. The sidebands above and below the carrier frequency are called the upper and lower sidebands respectively.

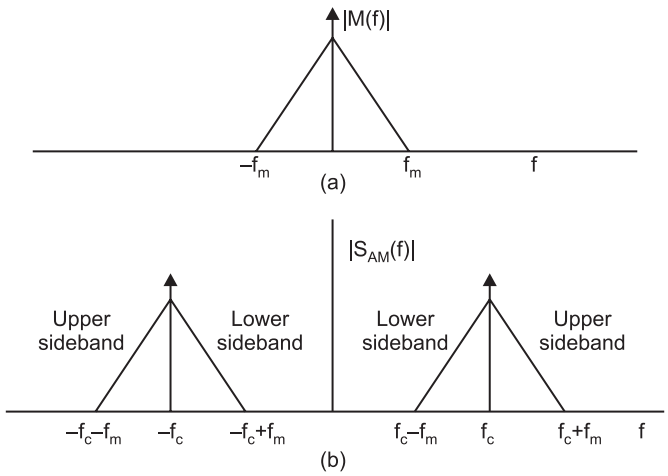


Fig. 12.2: (a) Spectrum of a message signal; (b) spectrum of the corresponding AM signal.

The RF bandwidth of an AM signal is equal to

$$B_{AM} = 2f_m \quad ...(12.6)$$

where f_m is the maximum frequency contained in the modulating message signal. The total power in an AM signal can be written as

$$P_{AM} = \frac{1}{2}A_C^2[1+2[m(t)]+[m^2(t)]] \quad ...(12.7)$$

where [*] represent the average value. If the modulating signal is $M(t) = K \cos (2nf_m t)$, so equation (12.7) can be written as

$$P_{AM} = \frac{1}{2}A_C^2[1+P_m] = P_C\left[1+\frac{K^2}{2}\right] \quad ...(12.8)$$

where $P_c = \frac{A_C^2}{2}$ is the power in the carrier signal.

$P_m = [m^2(t)]$ is the power in the modulating signal $m(t)$ and K is the modulation index.

12.3.1 Single Sideband AM

Since both the sidebands of an AM signal carry the same information, it is possible to remove one of them without losing any information.

Single sideband (SSB) AM systems transmit only one of the sidebands (either upper or lower) about the carrier, and hence occupy only half the bandwidth of conventional

AM systems.

An SSB signal can be mathematically expressed as

$$S_{SSB}(t) = A_C[m(t)\cos(2\pi f_c t) \mp \hat{m}(t)\sin(2\pi f_c t)] \quad ...(12.9)$$

where the negative sign in equation (12.9) is used for upper sideband SSB and the positive sign is used for lower sideband SSB. The term $\hat{m}(t)$ denotes the Hilbert transform of $m(t)$ which is given by

$$\hat{m}(t) = m(t) \otimes h_{HT}(t) = m(t) \otimes \frac{1}{\pi t} \quad ...(12.10)$$

and $H_{HT}(f)$, the Fourier transform of $h_{HT}(t)$, corresponds to $\alpha - 90°$ phase shift network.

$$H(f) = \begin{cases} -j & f > 0 \\ j & f < 0 \end{cases} \quad ...(12.11)$$

The two common techniques used for generating an SSB signal are the filter method and the balanced modulator method.

In the filter method, SSB signals are generated by passing a double sideband AM signal through a bandpass filter which removes one of the sidebands. A block diagram of such a modulator is shown in Fig. 12.3 (*a*). Excellent sideband suppression can be obtained using crystal filters at an intermediate frequency (IF).

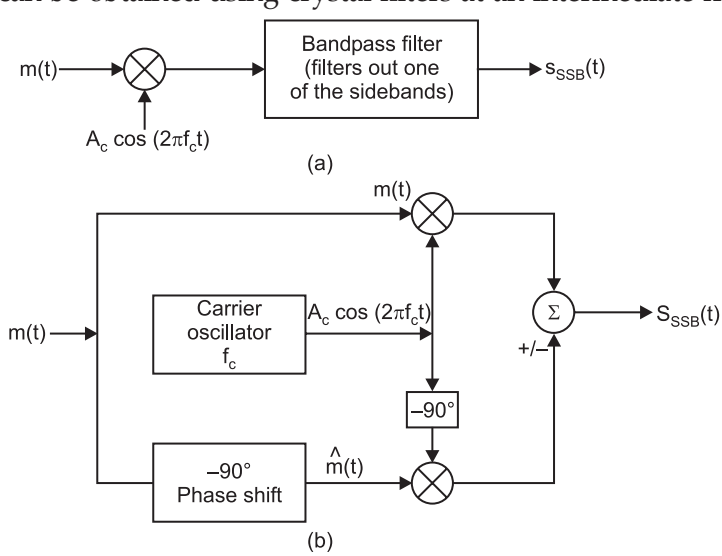


Fig. 12.3: Generation of SSB using (a) a sideband filter and (b) a balanced modulator.

Fig. 12.3 (*b*) shows a block diagram of a balanced modulator which is a direct implementation of equation 12.9. The modulating signal is split into two identical signals, one which modulates the in-phase carrier and the other which is passed through **a** - 90° phase shifter before modulating a quadrature carrier. The sign used for the quadrature component determines whether US SB or LSSB is transmitted.

12.3.2 Demodulation of AM Signals

AM demodulation techniques may be broadly divided into two categories: Coherent and non-coherent demodulation. Coherent demodulation requires knowledge of the transmitted carrier frequency and phase at the receiver, whereas non-coherent detection requires no phase information.

In practical AM receivers, the received signal is filtered and amplified at the carrier frequency and then converted to an intermediate frequency (IF) using a superheterodyne receiver. The IF signal retains the exact spectral shape as the RF signal.

Fig. 12.4 shows a block diagram of a product detector which forms a coherent demodulator for AM signals.

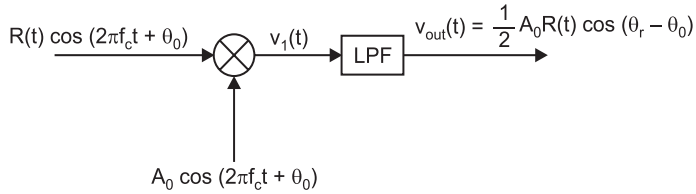


Fig. 12.4: Block diagram of a product detector.

A product detector (also called a phase detector) is a down converter circuit which convert the input bandpass signal to a baseband signal. If the input to the product detection is an AM signal of the form $R(t)$ cos $(27f_c t + \theta_r)$, the output of the multiplier can be expressed as

$$v_1(t) = R(t) \cos (2\pi f_c t + \theta_r) A_0 \cos (2\pi f_c t + \theta_0) \quad ...(12.12)$$

where f_c is the oscillator carrier frequency, and θ_r and θ_0 are the received signal phase and oscillator phases, respectively.

Equation 12.12 may be rewritten as

$$v_1(t) = \frac{1}{2} A_0 R(t) \cos (\theta_r - \theta_0) + \frac{1}{2} A_0 R(t) \cos [\pi 2 f_c t + \theta_r + \theta_0] \quad ...(12.13)$$

Since the low pass filter following the product detector removes the double carrier frequency term, the output is

$$v_{\text{out}}(t) = \frac{1}{2} A_0 R(t) \cos[\theta_r - \theta_r] = K\, R(t) \quad ...(12.14)$$

where K is a gain constant.

Equation 12.14 shows that the output of the low pass filter is the demodulated AM signal.

AM singnals are often demodulated using non-coherent envelope detectors which are easy and cheap to build.

An ideal envelope detector is a circuit that has an output proportional to the real envelope of the input signal. If the input to the envelope detector is represented as $R(t)$ cos $(2\pi f_c t + \theta_r)$, then the output is given by

$$v_{\text{out}}(t) = K\,|R(t)| \qquad ...(12.15)$$

where K is a gain constant.

Envelope detectors are useful when the input signal power is at least 10 dB greater than the noise power, whereas product detectors are able to process AM signals with input signal-to-noise ratios well below 0 dB.

12.4 ANGLE MODULATION

FM is part of a more general class of modulation known as angle modulation. Angle modulation varies a sinusoidal carrier signal in such a way that the angle of the carrier is varied according to the amplitude of the modulating baseband signal.

In this method, the amplitude of the carrier wave is kept constant (this is why FM is called constant envelope).

There are a number of ways in which the phase $\theta(t)$ of a carrier signal may be varied in accordance with the baseband signal; the two most important classes of angle modulation being frequency modulation and phase modulation.

Frequency modulation (FM) is a form of angle modulation in which the instantaneous frequency of the carrier signal is varied linearly with the baseband message signal $m(t)$, as shown in equation 12.16.

$$S_{\text{FM}}(t) = A_c \cos\,[2\pi f_c t + \theta(t)] = A_C \cos\left[2\pi f_c t + 2\pi k_f \int_{-\infty}^{t} m(\eta)d\eta\right] \qquad ...(12.16)$$

where A_c is the amplitude of the carrier frequency, and Kf is the frequency deviation constant (measured in units of Hz/v). If the modulating signal is a sinusoid of amplitude A_m, and frequency f_m, then the FM signal may be expressed as

$$S_{FM}(t) = A_c \cos\left[\pi f_c t + \frac{k_f A_m}{f_m} \sin(2\pi f_m t)\right] \qquad ...(12.17)$$

Phase modulation (PM) is a form of angle modulation in which the angle 0(0 of the carrier signal is varied linearly with the baseband message signal $m(t)$, as shown in equation 12.18.

$$S_{\text{PM}}(t) = A_c \cos\,[2nf_c t + k_\theta m(t)] \qquad ...(12.18)$$

In equation 12.18, k_θ is the phase deviation constant (measured in units of radians/volt).

The frequency modulation index β_f defines the relationship between the message amplitude and the bandwidth of the transmitted signal, and is given by

$$\beta_f = \frac{k_f A_m}{w} = \frac{\Delta_f}{w} \quad \text{... (12.19)}$$

where A_m is the peak value of the modulating signal, Δ_f is the peak frequency deviation of the transmitter, and w is the maximum bandwidth of the modulating signal.

The phase modulation index β_P is given by

$$\beta_P = k_\theta A_m = \Delta\theta \quad \text{... (12.20)}$$

where Δ_θ is the peak phase deviation of the transmitter.

12.4.1 Spectra and Bandwidth of FM Signals

When a sinusoidal test tone is used such that $m(t) = A_m \cos 2\pi f_m t$, the spectrum of $5_{FM}(t)$ contains a carrier component and an infinite number of sidebands located on either side of the carrier frequency, spaced at integer multiples of the modulation frequency f_m. Since $S_{FM}(t)$ is a nonlinear function of $m(t)$, the spectrum of an FM signal must be evaluated on a case-by-case basis for a particular modulating wave shape of interest. It can be shown that for a sinusoidal message, amplitudes of the spectral components are given by Bessel functions of the modulation index βf

An FM signal has 98% of the total transmitted power in a RF bandwidth B_T, given by

$$B_T = 2(\beta_f + 1)f_m \quad \text{(upper band)} \quad \text{...(12.21)}$$

$$B_T = 2\Delta_f \quad \text{(lower band)} \quad \text{...(12.22)}$$

The above approximation of FM bandwidth is called Carson' rule.

Carson's rule states that for small values of modulation index ($\beta f < 1$), the spectrum of an FM wave is effectively limited to the carrier frequency f_c, and one pair of sideband frequencies *at* $f_c \pm f_m$, and that for large values of modulation index, the bandwidth approaches, and is only slightly greater than, $2\Delta_f$.

12.4.2 FM Modulation Methods

There are basically two methded of generating an FM signal:

(*i*) The direct method,

(*ii*) The indirect method.

In the direct method, the carrier frequency is directly varied in accordance with the input modulating signal. In the indirect method, a narrowband FM signal is generated using a balanced modulator, and frequency multiplication is used to increase both the frequency deviation and the carrier frequency to the required level.

12.4.2.1. Direct Method

In this method, voltage-controlled oscillators (VCOs) are used to vary the frequency of the carrier signal in accordance with the baseband signal amplitude variations. These oscillators use devices with the reactance that can be varied by the application of a voltage, where the reactance causes the instantaneous

frequency of the VCO to change proportionally. The most commonly used variable reactance device is the voltage-variable capacitor called a varactor. The voltage-variable capacitor may be obtained, for example, by using a reverse biased *p-n* junction diode. The larger the reverse voltage applied to such a diode, the smaller the transition capacitance will be of the diode. By incorporating such a device into a standard Hartley or Colpitts oscillator, FM signals can be generated. Fig. 12.5 shows a simple reactance modulator. While VCOs offer a simple way to generate narrowband FM signals, the stability of the center frequency (carrier) of the VCO becomes a major value when it is used for wideband FM generation. The stability of the VCO can be improved by incorporating a phase locked loop (PLL) which locks the centre frequency to a stable crystal reference frequency.

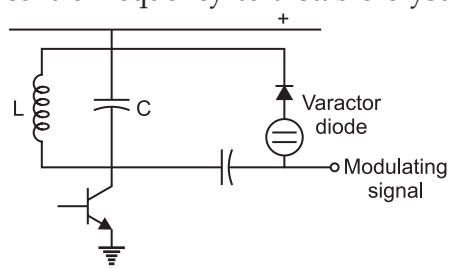


Fig. 12.5: A simple reactance modulator in which the capacitance of varactor diode is changed to vary the frequency of a simple oscillator. This circuit serves as a VCO.

12.4.2.2 Indirect Method

The indirect method of generating FM is based on aproximating a narrowband FM signal as the sum of a carrier signal and a single sideband (SSB) signal where the sideband is 90° out of phase with the carrier.

Using a Taylor series for small values of $\theta(t)$, equation 12.16 can be expressed as

$$S_{FM}(t) \cong A_c \cos 2\pi f_c t - A_c\theta(t) \sin 2\pi f_c t \qquad ...(12.23)$$

where the first term represents the carrier and the second term represents the sideband.

A simple block diagram of the indirect FM transmitted is shown in Fig. 12.6. A narrowband FM signal is generated using a balanced modulator which modulates a crystal controlled oscillator.

The maximum frequency deviation is kept constant and small in order to maintain the validity of equation 12.23, and hence the output is a narrowband FM signal.

A wideband FM signal is then produced by multiplying in frequency the narrowband FM signal using frequency multipliers.

A disadvantage of using the indirect method for wideband FM generation is that the phase noise in the system increases with the frequency multiplying factor *N*.

12.4.3 FM Detection Techniques

There are many ways to recover the original information from an FM signal. The objective of all FM demodulators is to produce a transfer characteristic that is the inverse of that of the frequency modulator. That is, a frequency demodulator should produce an output voltage with an instantaneous amplitude that is directly proportional to the instantaneous frequency of the input FM signal. Thus, a frequency-to-amplitude converter circuit is a frequency demodulator.

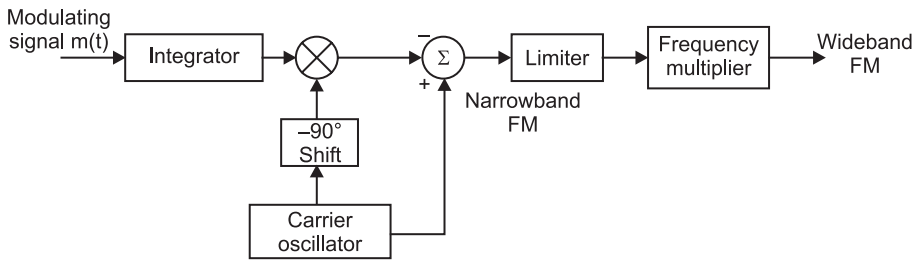


Fig. 12.6: Indirect method for generating a wideband FM signal. A narrowband FM signal is generated using a balanced modulator and then frequency multiplied to generate a wideband FM signal.

Various techniques such as slope detection, zero-crossing detection, phase locked discrimination, and quadrature detection are used to demodulate FM. Devices which perform FM demodulation are often called frequency discriminators.

12.4.3.1 Slope Detector

It can be easily shown that FM demodulation can be performed by taking the time derivative (often called slope detection) of the FM signal, followed by envelope detection.

A block diagram of such an FM demodulator is shown in Fig. 12.7. The FM signal is first passed through an amplitude limiter which removes any amplitude perturbation which the signal might have undergone due to fading in the channel, and produce a constant envelope signal.

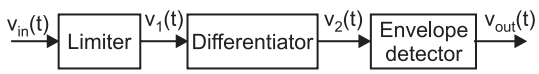


Fig. 12.7: Block diagram of a slope detector type FM demodulator.

Using equation 12.16 the signal at the output of the limiter can be represented as

$$v_1(t) = v_1 \cos[\pi f_c t + \theta(t)] = v_1 \cos\left[2\pi f_c t + 2\pi k_f \int_{-\infty}^{t} (\eta) d\eta\right] \quad ...(12.24)$$

Equation 12.24 can be differentiated in practice by passing the signal through a filter with a transfer function that has gain that increases linearly with frequency. Such a filter is called a slope filter (which is where the term slope detector derives its name). The output of the differentiator then becomes

$$v_2(t) = -v_1\left[2\pi f_c t + \frac{d\theta}{dt}\right]\sin[2\pi f_c t + \theta(t)] \quad ...(12.25)$$

The output of the envelope detector becomes

$$v_{\text{out}}(t) = v_1\left[2\pi f_c + \frac{d}{dt}\theta(t)\right]$$

$$= v_1 2\pi f_c + v_1 2\pi K_f\, m(t) \qquad ...(12.26)$$

The above equation shows that the output of the envelope detector contains a dc term proportional to the carrier frequency and a time-varying term proportional to the original message signal $m(t)$. The dc term can be filtered out using a capacitor to obtain the desired demodulated signal.

12.4.3.2 PLL for FM Detection

The phase locked loop (PLL) method is another popular technique to demodulate an FM signal.

The PLL is a closed loop control system which can track the variations in the received signal phase and frequency.

A block diagram of a PLL circuit is shown in Fig. 12.8. It consists of a voltage controlled oscillator $H(s)$ with an output frequency which is varied in accordance with the demodulated output voltage level. The output of the voltage controlled oscillator is compared with the input signal using a phase comparator, which produces an output voltage proportional to the phase difference. The phase difference signal is then feedback to the VCO to control the output frequency. The feedback loop functions in a manner that facilitates locking of the VCO frequency to the ouput frequency. Once the VCO frequency is locked to the input frequency, the VCO continues to track the variations in the input frequency. Once this tracking is achieved, the control voltage to the VCO is simply the demodulated FM signal.

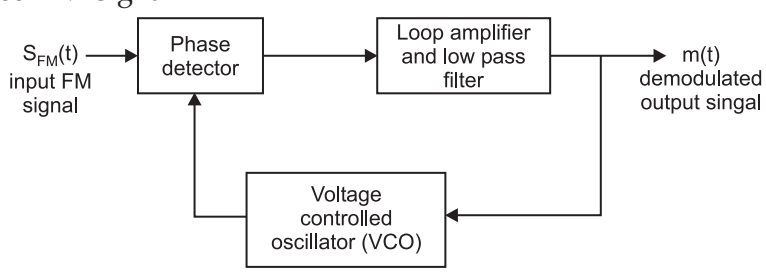


Fig. 12.8: Block diagram of a PLL used as a frequency demodulator.

12.5 DIGITAL MODULATION–AN OVERVIEW

Modern mobile communication systems use digital modulation techniques. Advancements in very large-scale integration (VLSI) and digital signal processing (DSP) technology have made digital modulation more cost effective than analog transmission systems.

Digital modulation offers many advantages over analog modulation. Some advantages include greater noise immunity and robustness to channel impairments, easier

multiplexing of various forms of information (*e.g.*, voice, data, and video), and greater security.

Digital transmissions accommodate digital error-control codes which detect and/or correct transmission errors, and support complex signal conditioning and processing techniques such as source coding, encryption, and equalization to improve the performance of the overall communication link.

New multipurpose programmable digital signal processors have made it possible to implement digital modulators and demodulators completely in software.

In digital wireless communication systems, the modulating signal (*e.g.*, the message) may be represented as a time sequence of symbols or pulses, where each symbol has m finite states. Each symbol represents n bits of information, where $n = \log_2 m$ bits/symbol.

Many digital modulation schemes are used in modern wireless communication systems, and many more are sure to be introduced. Some of these techniques have subtitle differences between one another, and each technique belongs to a family of related modulation methods. For example, phase shift keying (PSK) may be either coherently or differentially detected; and may have two, four, eight or more possible levels (e.g., n = 1, 2, 3 or more bits) per symbol, depending on the manner in which information is transmitted within a single symbol.

12.5.1 Factors that Influence the Choice of Digital Modulation

Several factors influence the choice of a digital modulation scheme. A desirable modulation scheme provides low bit error rates at low received signal-to-noise ratios, performs well in multipath and fading conditions, occupies a minimum of bandwidth, and is easy and cost-effective to implement.

The performance of a modulation scheme is often measured in terms of its power efficiency and bandwidth efficiency.

Power efficiency describes the ability of a modulation technique to preserve the fidelity of the digital message at low power levels. In a digital communication system, in order to increase noise immunity, it is necessary to increase the signal power.

The power efficiency, η_P (sometimes called energy efficiency) of a digital modulation scheme is a measure of how favourably this tradeoff between fidelity and signal power is made, and is often expressed as the ratio of the signal energy per bit to noise power spectral density $\left(\frac{E_b}{N_0}\right)$ required at the receiver input for a certain probability of error (say 10^{-5}).

Bandwidth efficiency describes the ability of a modulation scheme to accommodate data with in a limited bandwidth.

Bandwidth efficiency reflects how efficiently the allocated bandwidth is utilized and is defined as the ratio of the throughput data rate per Hertz in a given bandwidth.

If R is the data rate in bits per second, and B is the bandwidth occupied by the modulated RF signal, then bandwidth efficiency n_B is expressed as

$$\eta_B = \frac{R}{B} \text{bps/Hz} \qquad ...(12.27)$$

The system capacity of a digital mobile communication system is directly related to the bandwidth efficiency of the modulation scheme, since a modulation with a greater value of η_B will transmit more data in a given spectrum allocation.

There is a fundamental upper bound on achievable bandwidth efficiency. Shannon's channel coding theorem states that for an arbitrarily small probability of error, the maximum possible bandwidth efficiency is limited by the noise in the channel, and is given by the channel capacity formula. Note that Shannon's bound applies for AWGN non-fading channels.

$$\eta_{B\ max} = \frac{C}{B} = \log_2\left(1 + \frac{S}{N}\right) \qquad ...(12.28)$$

where C is the channel capacity (in bps), B is the RF bandwidth, and $\frac{S}{N}$ is the signal-to-N noise ratio.

In the design of a digital communication system, very often there is a tradeoff between bandwidth efficiency and power efficiency, for example, adding error control coding to a message increases the bandwidth occupancy, but at the same time reduces the required received power for a particular bit error rate, and hence trades bandwidth efficiency for power efficiency.

12.6 LINEAR MODULATION TECHNIQUES

Digital modulation techniques may be broadly classified as linear and non linear. In linear modulation techniques, the amplitude of the transmitted signal, s(t), varies linearly with the modulating digital signal, $m(t)$.

Linear modulation techniques are bandwidth efficient and hence are very attractive for use in wireless communication systems where there is an increasing demand to accommodate more and more users within a limited spectrum.

In a linear modulation scheme, the transmitted signal $S(i)$ can be expressed as

$$s(t) = Re[Am(t)\ \exp(j2\pi f_c t)] \qquad ...(12.29)$$

$$s(t) = A[m_R(t)\ \cos(2\pi f_c t) - m_1(t)\ sin(2\pi\ f_c t)] \qquad ...(12.30)$$

where A is the amplitude, f_c is the carrier frequency, and $m(t) = m_R(t) + jm_1(t)$ is a complex envelope representation of the modulated signal which is in general complex form.

From equation 12.30, it is clear that the amplitude of the carrier varies linearly with the modulating signal. While linear modulation schemes have very good spectral efficiency, they must be transmitted using linear RF amplifiers which have poor power efficiency.

The most popular linear modulation techniques include pulse-shaped QPSK, OQPSK, and $\frac{\pi}{4}$ QPSK, which are discussed subsequently.

12.6.1 Binary Phase Shift Keying (BPSK)

In binary phase shift keying (BPSK), the phase of a constant amplitude carrier signal is switched between two values according to the two possible signals m_1 and m_2 corresponding to binary 1 and 0, respectively. Normally, the two phases are separated by 180°.

The sinusoidal carrier has an amplitude A_c and energy per bit $E_b = \frac{1}{2}A_C^2\, T_b$,, then the transmitted BPSK signal is either

$$S_{BPSK}(t) = \sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_c t + \theta_c) \quad 0 \le t \le \text{(binary 1)} \quad ...(12.31)$$

or

$$S_{BPSK}(t) = \sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_c t + \pi + \theta_c) \quad ...(12.32)$$

$$= -\sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_c t + \theta_c) \;\; 0 \text{ £ } t \text{ £ } T_b\text{(binary 0)} \quad ...(12.33)$$

It is often convenient to generalize m_x and m_2 as a binary data signal $m(t)$, which takes on one of two possible pulse shapes. Then the transmitted signal may be represented as

$$S_{BPSK}(t) = m(t)\sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_c t + \theta_c) \quad ...(12.34)$$

The BPSK signal is equivalent to a double sideband suppressed carrier amplitude modulated waveform, where $\cos(2\pi f_c t)$ is applied as the carrier, and the data signal $m(t)$ is applied as the modulating waveform. Hence a BPSK signal can be generated using a balanced modulator.

12.6.1.1 Spectrum and Bandw idth of BPSK

The BPSK signal using a polar base band data waveform $m(t)$ can be expressed in complex envelope form as

$$S_{BPSK}(t) = Re\{g_{BPSK}(t)\,\exp(j2\pi f_c t)\} \quad ...(12.35)$$

where $g_{BPSK}(t)$is the complex envelope of is signal given by

$$g_{BPSK}(t) = \sqrt{\frac{2E_b}{T_b}}m(t)e^{j\theta} \quad ...(12.36)$$

The power spectral density (PSD) of the complex envelope can be shown to be

$$Pg_{BSPSK}(f) = 2E_b\left(\frac{\sin \pi f T_b}{\pi f T_b}\right)^2 \quad ...(12.37)$$

Hence, the PSD of a BPSK signal at RF is given by

$$P_{BPSK}(f) = \frac{E_b}{2}\left[\left(\frac{\sin \pi(f-f_c)T_b}{\pi(f-f_c)T_b}\right)^2 + \left(\frac{\sin \pi(-f-f_c)T_b}{\pi(-f-f_c)T_b}\right)^2\right] \quad ...(12.38)$$

The PSD of the BPSK signal for both rectangular and raised cosine rolloff pulse shapes is plotted in Fig. 12.9.

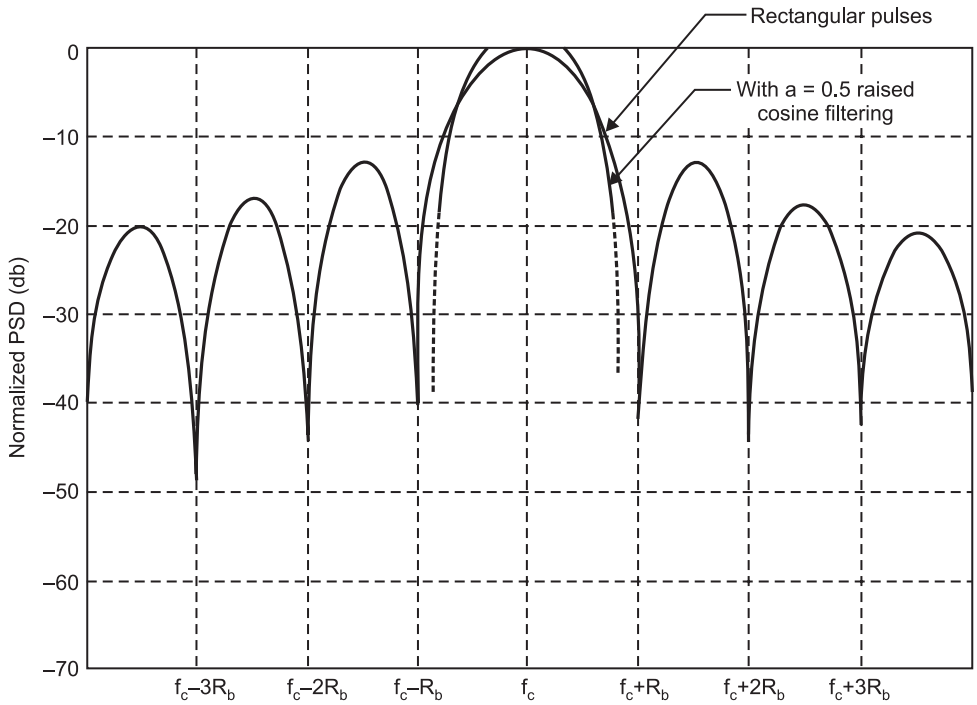


Fig. 12.9: Power spectral density (PSD) of a BPSK signal.

12.6.1.2 BPSK Receiver

If no multipath inpairments are induced by the channel, the received BPSK signal can be expressed as

$$S_{BPSK}(t) = m(t)\sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_c t + \theta_c + \theta_{ch}) \quad ...(12.39)$$

$$= m(t)\sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_c t + \theta) \quad ...(12.40)$$

where θ_{ch} is the phase shift corresponding to the time delay in the channel.

BPSK uses coherent or synchronous demodulation, which requires that information about the phase and frequency of the carrier be available at the receiver.

If a low level pilot carrier signal is transmitted along with the BPSK signal, then the carrier phase and frequency may be recovered at the receiver using a phase locked loop (PLL). If no pilot carrier is transmitted, a costas loop or squaring loop may be used to synthesize the carrier phase and frequency from the received BPSK signal.

Fig. 12.10 shows the block diagram of a BPSK receiver along with the carrier recovery circuits.

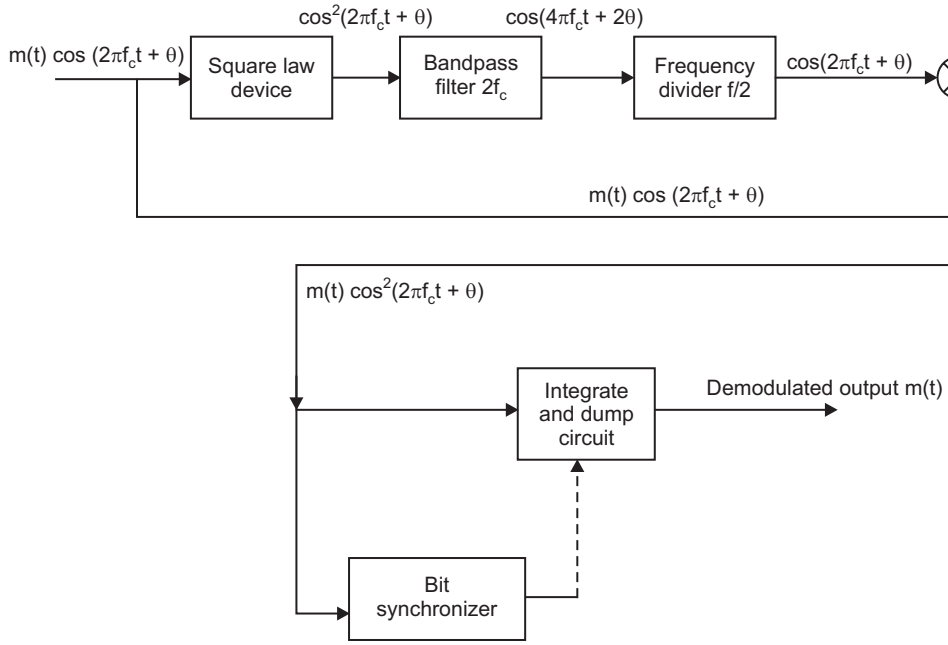


Fig. 12.10: BPSK receiver with carrier recovery circuits.

The received signal $(2\pi f_c t + \theta)$ is squared to generate a dc signal and an amplitude varying sinusoid at twice the carrier frequency. The dc signal is filtered out using a bandpass filter with centre frequency tuned to $2f_c$. A frequency divider is then used to recreate the waveform for $(2\pi f_c t + \theta)$. The output of the multiplier after the frequency divider is given by

$$m(t)\sqrt{\frac{2E_b}{T_b}}\cos^2(2\pi f_c t + \theta) = m(t)\sqrt{\frac{2E_b}{T_b}}\left[\frac{1}{2} + \frac{1}{2}\cos 2\,(2\pi f_c t + \theta)\right] \quad ...(12.41)$$

This signal is applied to an integrate and dump circuit which forms the low pass filter segment of a BPSK detector. If the transmitter and receiver pulse shapes are matched, then the detection will be optimum.

A bit synchronizer is used to facilitate sampling of the integrator output precisely at the end of each bit period. At the end of each bit period, the switch at the output of the integrator closes to dump the output signal to the decision circuits depending on wheather the integrater output is above or below a certain threshold, the decision circuit decides that the received signal corresponds to a binary 1 or 0.

The threshold is set at an optimum level such that the probability of error is minimized. If it is equally likely that a binary 1 or 0 is transmitted then the voltage level corresponding to the midpoint between the detector output voltage levels of binary 1 and 0 is used as the optimum threshold.

12.7 DIFFERENTIAL PHASE SHIFT KEYING (DPSK)

Differential PSK is a non-coherent form of phase shift keying which avoids the need for a coherent reference signal at the receiver. Non-coherent receivers are easy and cheap to build, and hence are widely used in wireless communications.

In DPSK systems, the input binary sequence is first differentially encoded and then modulated using a BPSK modulator.

The differentially encoded sequence $\{d_k\}$ is generated from the input binary sequence $\{m_k\}$ by complementing modulo-2 sum of m_k and d_{k-1}. The effect is to leave the symbol d_k unchanged from the previous symbol if the incoming binary symbol m_k is 1, and to toggle d_k if m_k is 0.

Table 12.1 illustrates the generation of a DPSK signal for a sample sequence m_k which follows the relationship $d_k = \overline{m_k \oplus d_{k-1}}$.

Table 12.1: Illustration of the Differential Encoding Process.

(mk)		1	0	0	1	0	1	1	0
$(d_k\text{-}x)$		1	1	0	1	1	0	0	0
(d_k)	1	1	0	1	1	0	0	0	1

A block diagram of a DPSK transmitter is shown in Fig. 12.11. It consists of a one bit delay element and a logic circuit interconnected so as to generate the differentially encoded sequence from the input binary sequence. The output is passed through a product modulator to obtain the DPSK signal. At the receiver, the original sequence is recovered from the demodulated differentially encoded signal through a complementary process, as shown in Fig. 12.11.

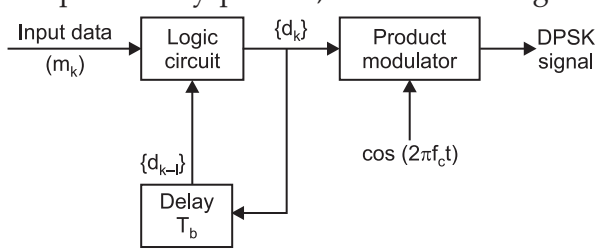


Fig. 12.11: Block diagram of a DPSK transmitter.

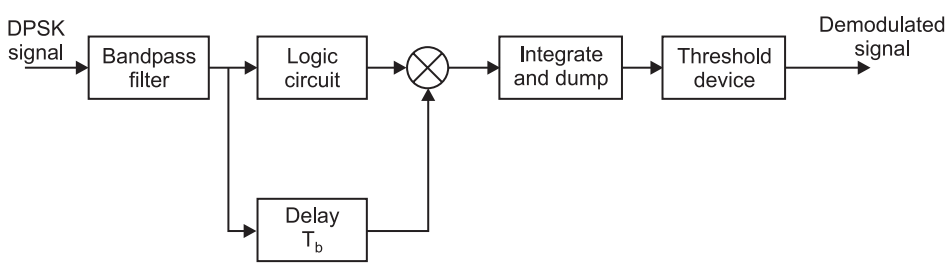


Fig. 12.12: Block diagram of DPSK receiver.

12.8 QUADRATURE PHASE SHIFT KEYING (QPSK)

Quadrature phase shift keying (QPSK) has twice the bandwidth efficiency of BPSK, since two bits are transmitted in a single modulation symbol.

The phase of the carrier takes on one of four equally spaced values, such as $0, \frac{\pi}{2}, \pi$ and $\frac{3\pi}{2}$, where each value of phase corresponds to a unique pair of message bits.

The QPSK signal for this set of symbol may be defined as

$$S_{\text{QPSK}}(t) = \sqrt{\frac{2E_s}{T_s}}\cos\left[2\pi f_c t + (i-1)\frac{\pi}{2}\right] \quad 0 \le t \le T_s \; i = 1, 2, 3, 4 \qquad ...(12.42)$$

where T_s is the symbol duration and is equal to twice the bit period.

Using trigonometric identities, the above equations can be rewritten for the interval $0 < \le t \le T_s$ as

$$S_{\text{QPSK}}(t) = \sqrt{\frac{2E_s}{T_s}}\cos\left[(i-1)\frac{\pi}{2}\right]\cos(2\pi f_c t) - \sqrt{\frac{2E_s}{T_s}}\sin\left[(i-1)\frac{\pi}{2}\right]\sin(2\pi f_c) \qquad ...(12.43)$$

If basis functions $\phi_1(t) = \sqrt{\frac{2}{T_s}}\cos(2\pi f_c t), \phi_2(t) = \sqrt{\frac{2}{T_s}}\sin(2\pi f_c t)$ are defined over the interval $0 \le t \le T_s$ for the QPSK signal set, then the four signals in the set can be expressed in terms of the basis signals as

$$S_{QPSK}(t) = \left\{\sqrt{E_s}\cos\left[(i-1)\frac{\pi}{2}\right]\theta_1(t) - \sqrt{E_s}\sin\left[(i-1)\frac{\pi}{2}\right]\theta_2(t)\right\} i = 1,2,3,4 \qquad ...(12.44)$$

Based on this representation, a QPSK signal can be depicted using a two-dimensional constellation diagram with four points as shown in Fig. 12.13 (*a*). Fig. 12.13 (*b*) shows another QPSK signal set where the phase values are $\frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}$, and $\frac{7\pi}{4}$.

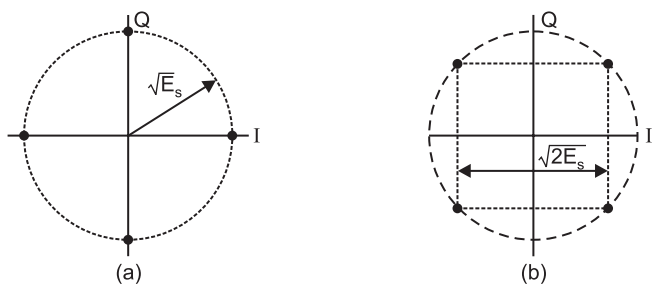


Fig. 12.13. (a) QPSK constellation where the carrier phases are 0, π/2,π, 3π/2; (b) QPSK constellation where the carrier phases are π/4, 3π/4, 5π/4, 7π/4.

From the constellation diagram of a QPSK signal, it can be seen that the distance between adjacent points in the constellation is $\sqrt{2E_s}$. Since each symbol corresponds to two bits, then $E_s = 2E_b$, thus the distance between two neighboring points in the QPSK constellation is equal to $2\sqrt{E_b}$. Substituting this in equation $P_s \le \sum_{j=1}\left(\frac{d_{ij}}{\sqrt{2N_0}}\right)$, the average probability of bit error in the additive white Gaussian noise (AWGN) channel is obtained as

$$P_{e,\,QPSK} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \qquad ...(12.45)$$

When compared to BPSK, QPSK provides twice the spectral efficiency with exactly the same energy efficiency.

Spectrum and Bandwidth of QPSK signals

The power spectral density of a QPSK signal can be obtained in a manner similar to that used for BPSK, with the bit periods T_b replaced by symbol periods T_r Hence, the PSD of a QPSK signal using rectangular pulses can be expressed as

$$P_{QPSK}(f) = \frac{E_s}{2}\left[\left(\frac{\sin \pi(f-f_c)T_s}{\pi(f-f_c)T_s}\right)^2 + \left(\frac{\sin \pi(-f-f_c)T_s}{\pi(-f-f_c)T_s}\right)^2\right] \quad ...(12.46)$$

$$= \frac{E_s}{2}\left[\left(\frac{\sin 2\pi(f-f_c)T_b}{2\pi(f-f_c)T_b}\right)^2 + \left(\frac{\sin 2\pi(-f-f_c)T_b}{2\pi(-f-f_c)T_b}\right)^2\right] \quad ...(12.47)$$

The PSD of a QPSK signal for rectangular and raised cosine filtered pulses is plotted in Fig. 12.14. The null-to-null RF bandwidth is equal to the bit rate R_b, which is half that of a BPSK signal.

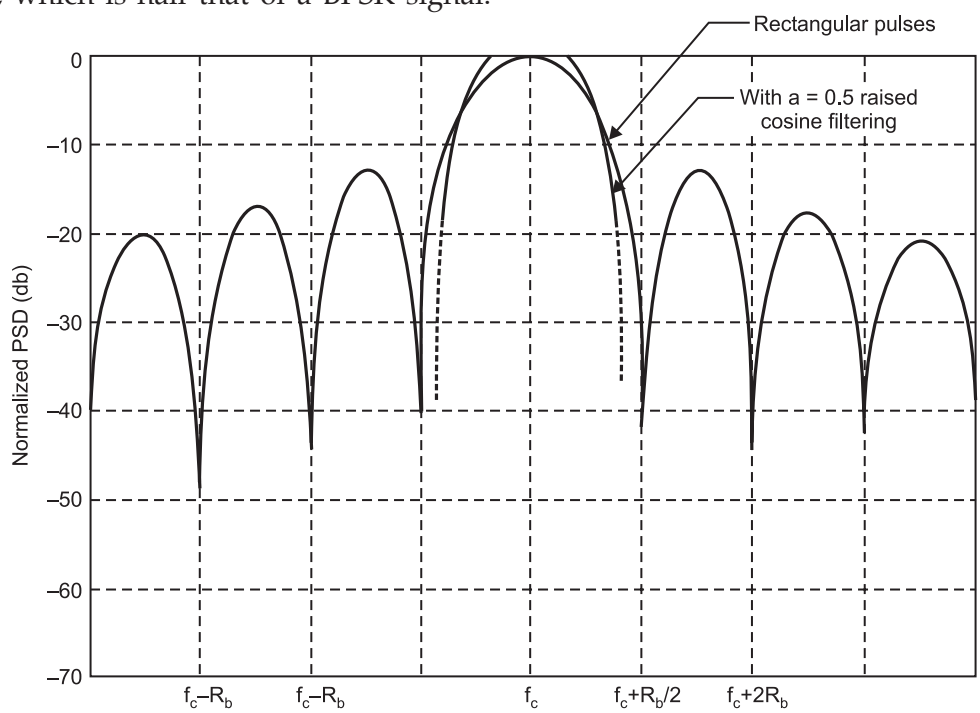


Fig. 12.14: Power spectral density of a QPSK signal.

12.9 QPSK TRANSMISSION AND DETECTION TECHNIQUES

Fig. 12.15 shows a block diagram of a typical QPSK transmitter.

The unipolar binary message stream has bit rate R_b and is first converted into a bipolar non-return-to zero (NRZ) sequence using a unipolar to bipolar converter. The bit stream $m(t)$ is then split into two bit streams $m_1(t)$ and $m_Q(t)$

(in-phase and quadrature streams), each having a bit rate of $R_s = \frac{R_b}{2}$. The bit stream $m_1(t)$ is called the "even" stream and $m_Q(t)$ is called the "odd" stream. The two binary sequences are separately modulated by two carriers $f_1(t)$ and $f_2(t)$, which are in quadrature. The two modulated signals, each of which can be considered to be a BPSK signal, are summed to produce a QPSK signal. The filter at the output of the modulator confines the power spectrum of the QPSK signal within the allocated band.

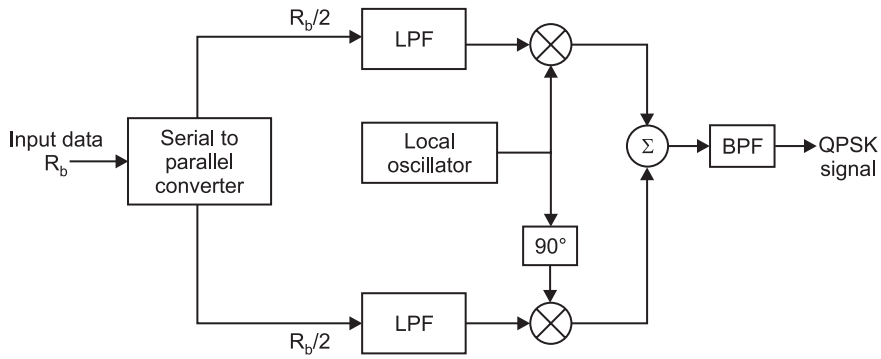


Fig. 12.15: Block diagram of a QPSK transmitter.

Fig. 12.16 shows a block diagram of a coherent QPSK receiver. The frontend bandpass filter removes the out-of-band noise and adjacent channel interference. The filtered output is split into two parts, and each part is coherently demodulated using the in-phase and quadrature carriers. The coherent carriers used for demodulation are recovered from the received signal using carrier recovery circuits of the type described in Fig. 12.10. The output of the demodulators are passed through decision circuits which generate the in-phase and quadrature binary streams. The two components are then multiplexed to reproduce the original binary sequence.

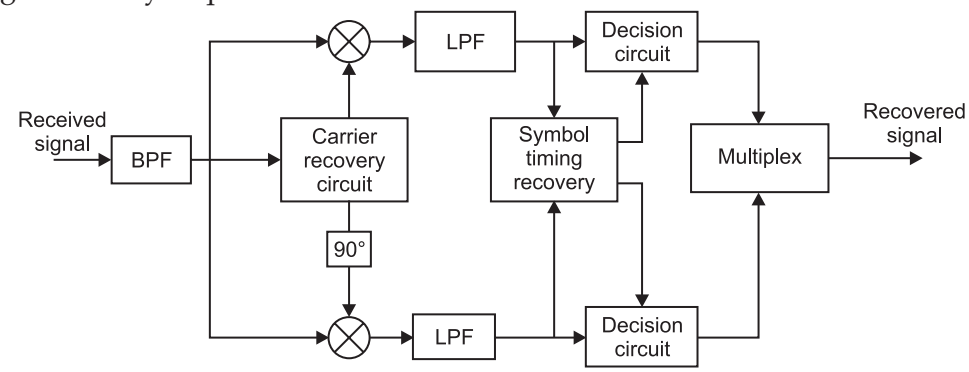


Fig. 12.16: Block diagram of a QPSK receiver.

12.10 π/4 QPSK

The π/4 shifted QPSK modulation is a quadrature phase shift keying technique which offers a compromise between OQPSK and QPSK in terms of

the allowed maximum phase transitions. It may be demodulated in a coherent or non-coherent fashion.

In $\pi/4$ QPSK, the maximum phase change in limited to $\pm$ 135°, as compared to 180° for QPSK and 90° for OQPSK. Hence the band limited $\pi/4$ QPSK signal preserves the constant envelope property better than band limited QPSK, but is more susceptible to envelope variations than OQPSK.

$\pi/4$ QPSK signals are differentially encoded to facilitate easier implementation of differential detection or coherent demodulation with phase ambiguity in the recovered carrier. When differentially encoded, $\pi/4$ QPSK is called $\pi/4$ DQPSK.

In a $\pi/4$ QPSK modulator, signaling points of the modulated signal are selected from two QPSK constellations which are shifted by $\pi/4$ with respect to each other.

Fig. 12.17: shows the two constellations along with the combined constellation where the links between two signal points indicate the possible phase transitions.

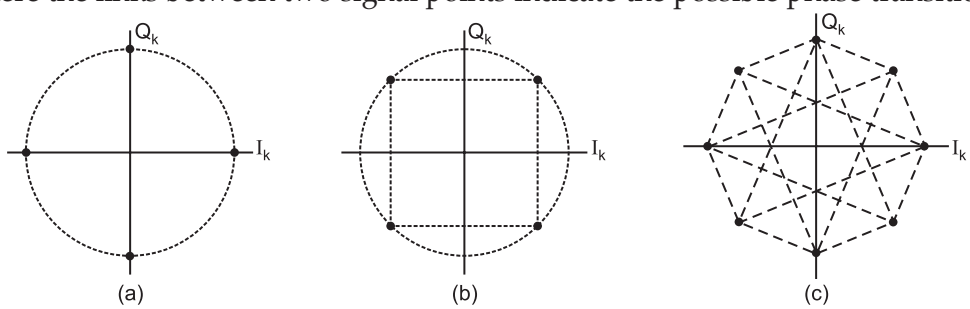


Fig. 12.17: Constellation diagram of a $\pi/4$ QPSK signal: (a) possible states for θ_k when $\theta_{k-1} = n\pi/4$; (b) possible states when $\pi_{k-1} = n\pi t/2$; (c) all possible states.

Switching between two constellations, every successive bit ensures that there is at least a phase shift which is an integer multiple of $\pi/4$ radians between successive symbols. This ensures that there is a phase transition for every symbol, which enable a receiver to perform timing recovery and synchronization.

12.11 $\pi/4$ QPSK TRANSMISSION TECHNIQUES

A block diagram of a generic $\pi/4$ QPSK transmitter is shown in Fig. 12.18.

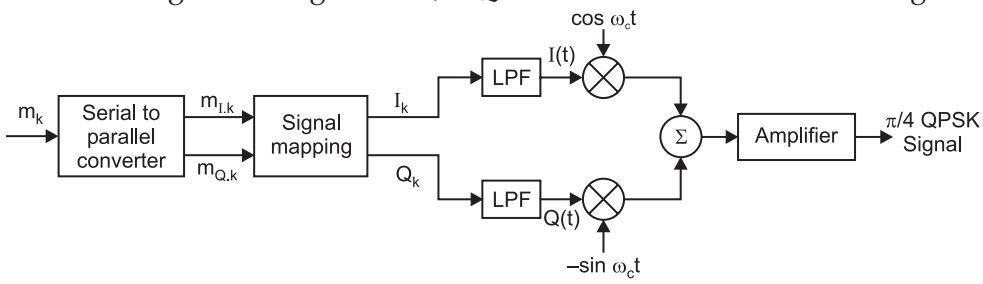


Fig. 12.18: Generic $\pi/4$ QPSK transmitter.

The input bit stream is partitioned by a serial-to-parallel (S/P) converter into two parallel data streams $m_{I'k}$ and $m_{Q'k'}$ each with a symbol rate equal to half that of the incoming bit rate. The k_{th} in-phase and quadrature pulses, I_k and

Q_h are produced at the output of the signal mapping circuit over time $kT \le t \le (k + 1)T$ and are determined by their previous values, I_{k-1} and θ_{k-1}, as well as θ_k, which itself is a function of ϕ_k, which is a function of the current input symbol m_{Ik} and m_{0k}. I_k and ϕ_k, represent rectangular pulses over one symbol duration having amplitudes given by

$$I_k = \cos\theta_k = I_{k-1}\cos\theta_k - Q_{k-1}\sin\phi_k \quad ...(12.48)$$

$$Q_k = \sin\theta_k = I_{k-1}\sin\phi_k + Q_{k-1}\cos\phi_k \quad ...(12.49)$$

where,

$$\theta_k = \theta_{k-1} + \theta_k$$

θ_k and θ_{k-1} are phases of the k_{th} and k–1st symbols. The phase shift θ_k is related to the input symbols m_{Ik} and m_{Qk} according to Table 12.2.

Table 12.2: Carrier Phase Shifts Corresponding to Various Input Bit Pairs

Information bits m_{Ik}, m_{Qk}	Phase shift ϕ_k
11	$\pi/4$
01	$3\pi/4$
00	$-3\pi/4$
10	$-\pi/4$

Just as in a QPSK modulator, the in-phase and quadrature bit stream I_k and Q_k are then separately modulated by two carriers which are in quadrature with one another, to produce the $\pi/4$ QPSK waveform given by

$$S_{\pi/4\text{QPSK}}(t) = I(t)\cos\omega_c t - Q(t)\sin\omega_c t \quad ...(12.50)$$

$$I(t) = \sum_{k=0}^{N-1} I_k P\left(t - kT_s - \frac{T_s}{2}\right) = \sum_{k=0}^{N-1}\cos\theta_k P\left(t - kT_s - \frac{T_s}{2}\right) \quad ...(12.51)$$

$$Q(t) = \sum_{k=0}^{N-1} Q_k P\left(t - kT_s - \frac{T_s}{2}\right) = \sum_{k=0}^{N-1}\sin\theta_k P\left(t - kT_s - \frac{T_s}{2}\right) \quad ...(12.52)$$

Both I_k and Q_k are usually passed through raised cosine roll off pulse shaping filters before modulation, in order to reduce the bandwidth occupancy.

The function $P(t)$ in equations 12.51 and 12.52 corresponds to the pulse shape, and T_s in the symbol period. Pulse shaping also reduces the spectral restoration problem which may be significant in fully saturated nonlinear amplified systems.

It should be noted that the values of I_k and Q_k and the peak amplitude of the waveform $I(t)$ and $Q(t)$ can take one of the five possible values, 0, +1, –1, $+1/\sqrt{2}$, $-1/\sqrt{2}$.

12.12 π/4 QPSK DETECTION TECHNIQUES

Due to ease of hardware implementation, differential detection is often employed to demodulate $\pi/4$ QPSK signals. There are various types of detection techniques that are used for the detection of $\pi/4$ QPSK signals. They include baseband differential detection, IF differential detection, and FM discriminator detection.

While both the baseband and IF differential detector determine the cosine and sine functions of the phase difference, and then decide on the phase difference accordingly, the FM discriminator detects the phase difference directly in a non coherent manner.

12.12.1 Baseband Differential Detection

Fig. 12.19 shows a block diagram of a baseband differential detector.

Fig. 12.19. Block diagram of a baseband differential detector.

The incoming $\pi/4$ QPSK signal is quadrature demodulated using two local oscillator signals that have the same frequency as the unmodulated carrier at the transmitter, but not necessarily the same phase. If $\phi_k = \tan^{-1}\left(\frac{Q_k}{I_k}\right)$ is the phase of the carrier due to the k_{th} data bit, the output w_k and z_k from the two low pass filters in the in-phase and quadrature arms of the demodulator can be expressed as

$$w_k = \cos(\phi_k - \gamma) \quad ...(12.53)$$

$$z_k = \sin(\phi_k - \gamma) \quad ...(12.54)$$

where γ is a phase shift due to noise, propagation, and interference. The phase y is assumed to change much slows than ϕ_k, so it is essentially a constant.

The two sequence w_k and z_k are passed through a differential decoder which operates on the following rule.

$$x_k = w_k w_{k-1} + z_k z_{k-1} \quad ...(12.55)$$

$$y_k = z_k w_{k-1} - w_k z_{k-1} \quad ...(12.56)$$

The output of the differential decoder can be expressed as

$$x_k = \cos(\phi_k - \gamma)\cos(\phi_{k-1} - \gamma) + \sin(\phi_k - \gamma)\sin(\phi_{k-1} - \gamma)$$

$$= \cos(\phi_k - \phi_{k-1}) \quad ...(12.57)$$

$$y_k = \sin(\phi_k - \gamma)\cos(\phi_{k-1} - \gamma) + \cos(\phi_k - \gamma)\sin(\phi_{k-1} - \gamma)$$

$$= \sin(\phi_k - \phi_{k-1}) \quad ...(12.58)$$

The output at the differential decoder is applied to the decision circuit, which uses Table 12.2 to determine

$$S_I = 1, \text{ if } x_k > 0 \quad \text{or } S_r = 0, \text{ if } x_k < 0 \quad ...(12.59)$$

$$S_Q = 1, \text{ if } y_k > 0 \quad \text{or } S_Q = 0, \text{ if } y_k < 0 \quad ...(12.60)$$

where, IF and Sg are the detected bits in the in-phase and quadrature arms, respectively.

It is important to ensure the local receiver oscillator frequency is the same as the transmitter carrier frequency, and that it does not drift. Any drift in the carrier frequency will cause a drift in the output phase which will lead to BER degradation.

12.12.2 IF Differential Detector

The IF differential detector shown in Figure 12.20 avoids the need for a local oscillator by using a delay line and two phase detectors. The received signal

in converted to IF and is bandpass filtered. The bandpass filter is designed to match the transmitted pulse shape, so that the carrier phase is preserved and noise power is minimized.

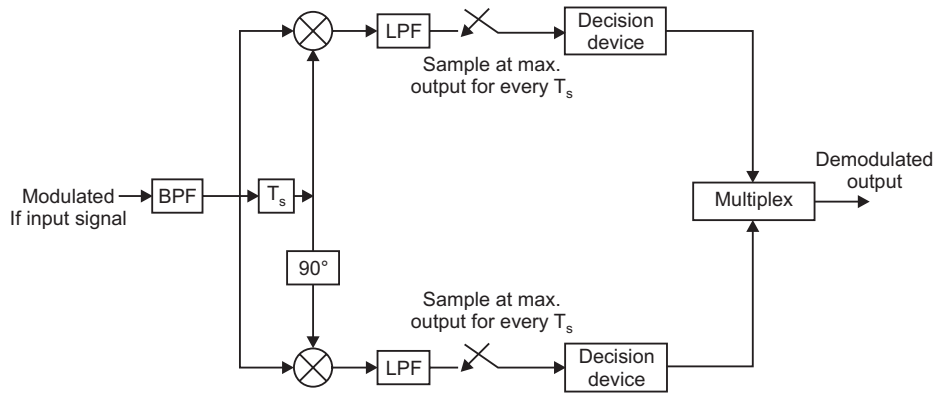


Fig. 12.20: Block diagram of an IF differential detector for π/4 QPSK.

To minimize the effect of ISI and noise, the bandwidth at the filters are chosen to be $0.57/T_s$. The received IF signal is differential decoded using a delay line and two mixers. The bandwidth as the signal at the output of the differential detector is twice that as the baseband signal at the transmitter end.

12.12.3 FM Discriminator

Figure 12.21 shows a block diagram as an FM discriminator detector for $\frac{\pi}{4}$ QPSK.' The input signal is first filtered using a bandpass filter that is matched to the transmitted signal. The filtered signal is then hardlimited to remove any envelope fluctuations. Hardlimiting preserves the phase changes in the input signal and hence no information is lost. The FM discriminator extracts the instantaneous frequency deviation of the received signal which, when integrated over each symbol period gives the phase difference between two sampling instants. The phase difference is then detected by a four level threshold comparator to obtain the original signal. The phase difference can also be detected using a modulo - 2π phase detector. The modulo - 2π phase detector improves BER performance and reduces the effect as click noise.

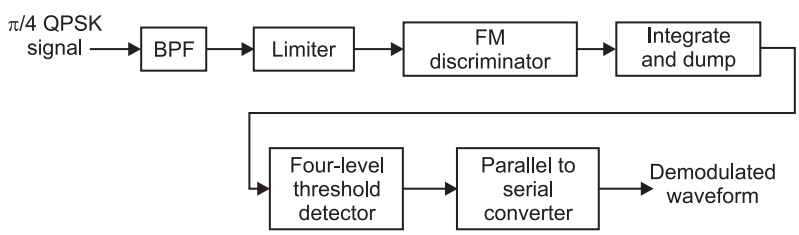


Fig. 12.21: FM discriminator detector for π/4 DQPSK demodulation.

12.13 CONSTANT ENVELOPE MODULATION

Many practical mobile radio communications, systems use non-linear modulation methods, where the amplitude of the carrier is constant, regardless as the variation in the modulating signal.

The constant envelope family of modulations has the advantage of satisfying a number of conditions, some of which are:

(*i*) Power efficient class C amplifiers can be used without introducing degradation in the spectrum occupancy of the transmitted signal.

(*ii*) Low out-of-band radiation of the order of - 60 dB to - 70 dB can be achieved.

(*iii*) Limiter-discriminator detector can be used, which simplifies receiver design and provides high immunity against random FM noise and signal fluctuation due to Rayleigh fading.

While constant envelope modulation have many advantages, they occupy a larger bandwidth than linear modulation schemes. In situations where bandwidth efficiency is more important than power efficiency, constant envelope modulation is not well-suited.

12.13.1 Binary Frequency Shift Keying

In binary frequency shift keying (BFSK), the frequency of a constant amplitude carrier signal is switched between two values according to the two possible message states (called high and low tones), corresponding to a binary 1 or 0.

Depending on how the frequency variations are imparted into the transmitted waveform, the FSK signal will have either a discontinues phase or continuous phase between bits. In general, an FSK signal may be represented as

$$S_{FSK}(t) = V_H(t) = \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c + 2\pi\Delta f)t \qquad 0 \le \Sigma T_b \text{ (binary 1)} \qquad ...(12.61)$$

$$S_{FSK}(t) = V_L(t) = \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c - 2\pi\Delta f)t \qquad 0 \le t \le T_b \text{ (binary 0)} \qquad ...(12.62)$$

where, $2\pi\ \Delta f$ is a constant offset from the nominal carrier frequency.

One obvious way to generate an FSK signal is to switch between two independent oscillators according to whether the data bit is 0 or a 1. Normally, this form of FSK generation results in a waveform that is discontinuous at the switching times, and for this reason this type at FSK is called discontinuous FSK. A discontinuous FSK signal is represented as

$$S_{\text{FSK}}(t) = V_H(t) = \sqrt{\frac{2E_b}{T_b}} \cos(2\pi\ f_H t + \theta_1) \qquad 0 \le t \le T_b \text{ (binary 1)} \qquad ...(12.63)$$

$$S_{FSK}(t) = V_L(t) = \sqrt{\frac{2E_b}{T_b}} \cos(2\pi\ f_L t + \theta_2) \qquad 0 \le t \le T_b \text{ (binary 0)} \qquad ...(12.64)$$

The more common method for generating an FSK signal is to frequency modulate a single carrier oscillator using the message waveform. This type of modulation is similar to analog FM generation, except that the modulating signal $m(t)$ is a binary waveform. Therefore, FSK may be represented as

$$S_{FSK}(t) = \sqrt{\frac{2E_b}{T_b}}\cos[2\pi f_c t + \theta(t)] \quad ...(12.65)$$

$$= \sqrt{\frac{2E_b}{T_b}}\cos\left[2\pi f_c t + 2\pi k_f \int_{-\infty}^{t} m(\eta)d\eta\right] \quad ...(12.66)$$

12.13.1.1 Spectrum and Bandwidth of BFSK Signals

The power spectral density of a binary FSK signal consists of discrete frequency components at f_c, $f_c + n\Delta f$, $f_c - n\,\Delta f$, where n is an integer.

The transmission bandwidth B_T of an FSK signal is given by Carson's rule as

$$B_T = 2\,\Delta f + 2B \quad ...(12.67)$$

where B is the bandwidth of the digital baseband signal.

Assuming that first null bandwidth is used the bandwidth of rectangular pulses is $B = R$. Hence, the FSK transmission bandwidth becomes

$$B_T = 2\,(Df + R) \quad ...(12.68)$$

If a raised cosine pulse-shaping filter is used, then the transmission bandwidth reduces to

$$B_T = 2A/+\,(1 + a)\,R \quad ...(12.69)$$

where a is the rolloff factor of the filter.

12.13.1.2. Coherent Detection of Binary FSK

A block diagram of a coherent detection scheme for demodulation of binary FSK signals is shown in figure 12.22. The receiver shown is the optimum detector for coherent binary FSK in the presence of additive white Caussian noise. It consists of two correlators which are supplied with locally generated coherent reference signals. The difference of the correlator outputs is then compared with a threshold comparator. If the difference signal has a value greater than the threshold, the-receiver decides in favour of a 1, otherwise, it decides in favour of a 0. It can be shown that the probability of error for a coherent FSK receiver is given by

$$P_{e,\,FSK} = Q\left(\sqrt{\frac{E_b}{N_B}}\right) \quad ...(12.70)$$

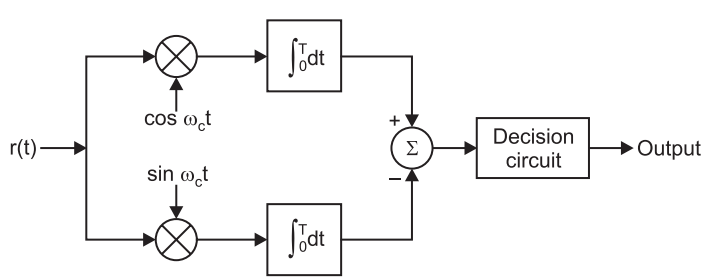


Fig. 12.22: Coherent detection of FSK signals.

12.13.1.3 Non-coherent Detection of Binary FSK

Unlike phase-shift keying, it is possible to detect FSK signals in the presence of noise without using a coherent carrier reference. A block diagram of a non-coherent FSK receiver is shown in figure 12.23. the receiver consists of a pair of matched filters followed by envelope detectors. The filter in the upper path is matched to the FSK signal of frequencyf_H and the filter in the lower path is matched to the signal of frequency $f_{L'}$ These matched filters function as bandpass filters centered at f_H *and* $f_{L'}$ respectively. The outputs of the envelope detectors are sampled at every $t = kT_b$, where k is an integer, and their values compared. Depending on the magnitude of the envelope detector output, the comparator decides whether the data bit was a 1 or 0.

The average probability of error of an FSK system employing non-coherent detection is given by

$$P_{e\text{FSK},\text{NC}} = \frac{1}{2}\exp\left(-\frac{E_b}{2N_0}\right) \qquad ...(12.71)$$

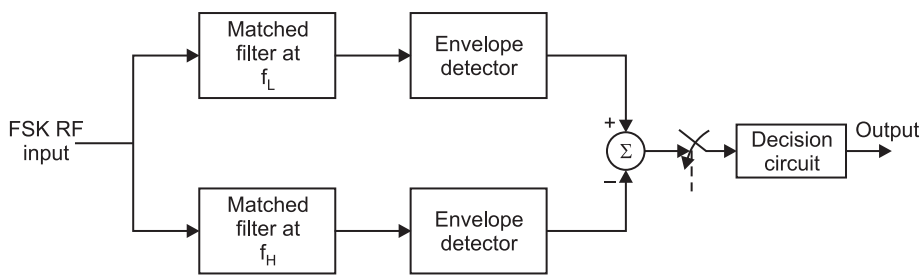


Fig. 12.23: Block diagram of non-coherent FSK receiver.

12.13.2 Minimum Shift Keying (MSK)

Minimum shift keying (MSK) is a special type of continuous phase-frequency shift keying (CPFSK) where in the peak frequency deviation is equal to $\frac{1}{4}$ the bit rate. In other words, MSK is continuous phase FSK with a modulation index of 0.5. The modulation index of an FSK signal is similar to the FM modulation index, and is defined as $K_{FSK} = \frac{(2\Delta F)}{R_b}$, where ΔF is the peak RF frequency deviation and R_b is the bit rate.

A modulation index of 0.5 corresponds to the minimum frequency spacing that allows two FSK signals to be coherently orthogonal, and the name minimum shift keying implies the minimum frequency separation (*i.e.*, bandwidth) that allows orthogonal detection. Two FSK signals $V_H(t)$ and $V_L(t)$ are said to be orthogonal if

$$\int_0^T V_H(t)V_L(t)dt = 0 \qquad ...(12.72)$$

MSK is sometimes referred to as fast FSK, as the frequency spacing used is only half as much as that used in conventional non-coherent FSK.

MSK is a spectrally efficient modulation scheme and is particularly attractive for use in mobile radio communication systems. It possesses properties such as constant envelope, spectral efficiency, good BER performance, and self-synchronizing capability.

An MSK signal can be thought of as a special form of OQPSK where the baseband rectangular pulses are replaced with half-sinusoidal pulses. These pulses have shapes like the St. Louis arch during a period of 2 T_b.

There are a number of variations of MSK that exist, in the literature. For example, while one version of MSK uses only positive half-sinusoids as the basic pulse shape, another version uses alternating positive and negative half-sinusoids as the basic pulse shape. However, all variations of MSK are continuous phase FSK employing different techniques to achieve spectral efficiency.

Figure 12.24 shows the power spectral density of an MSK signal. The spectral density of QPSK and OQPSK are also drawn for comparison. From 12.24 it is seen that the MSK spectrum has lower side lobes than QPSK and OQPSK.

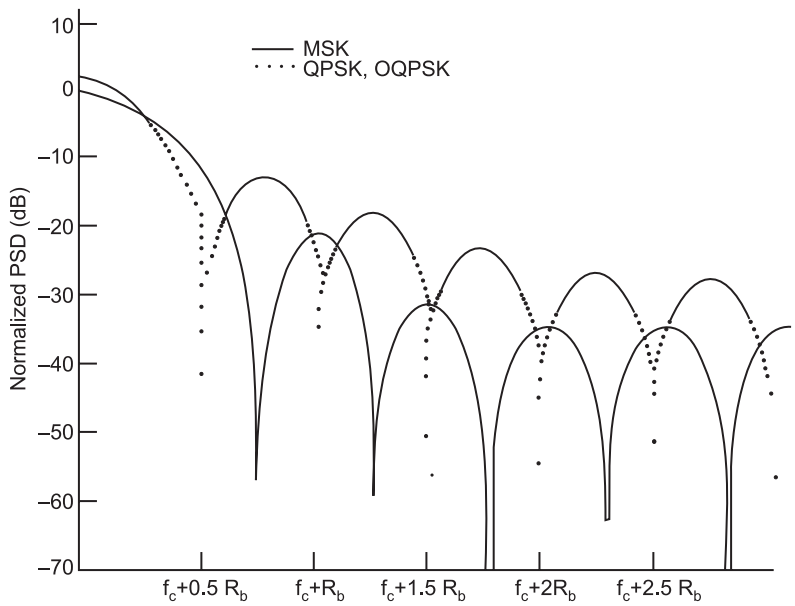


Fig. 12.24: Power spectral density of MSK signals as compared to QPSK and OQPSK signals.

12.13.2.1. MSK Transmitter and Receiver

Figure 12.25 shows a typical MSK modulator. Multiplying a carrier signal with $\cos\left[\frac{\pi t}{2T}\right]$ produces two phase-coherent signal at $f_c + \frac{1}{4T}$ and $f_c - \frac{1}{4T}$. These two FSK signals are separated using two narrow bandpass filters and appropriately combined to form the in-phase and quadrature carrier components $x(t)$ and $y(t)$, respectively. These carriers are multiplied with the odd and even bit streams, $m_1(t)$ and $m_Q(t)$, to produce the MSK modulated signal $S_{MSK}(t)$.

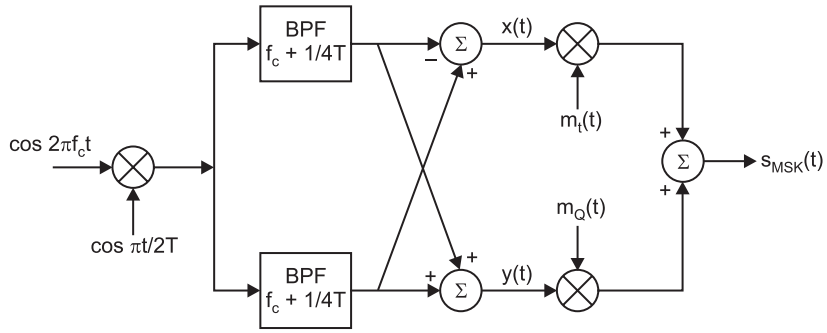


Fig.12.25: Block diagram of an MSK transmitter. Note that $m_Q(t)$ and $m_0(t)$ are offset by T_b.

The block diagram of an MSK receiver is shown in Figure 12.26. The received signal $S_{MSK}(t)$ (in the absence of noise and interference) is multiplied by the respective in-phase and quadrature carrier $x(t)$ and $y(t)$. The output of the multipliers are integrated over two bit periods and dumped to a decision circuit at the end of each two bit periods. Based on the level of the signal at the output of the integrator, the threshold detector decider whether the signal is a 0 or *a* 1. The output data streams correspond to $m_I(t)$ and $m_Q(t)$, which are offset combined to obtain the demodulated signal.

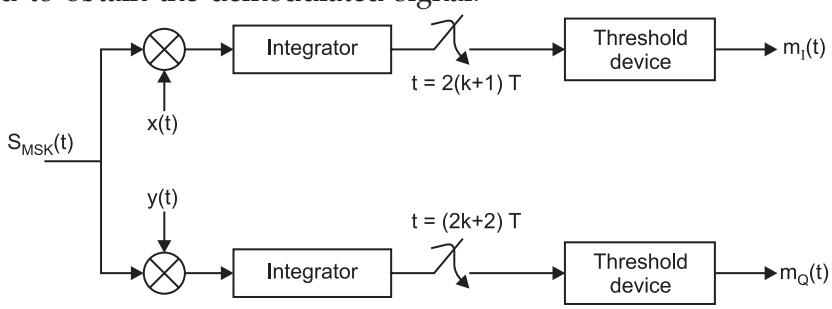


Fig. 12.26: Block diagram of an MSK receiver.

12.14 COMBINED LINEAR AND CONSTANT ENVELOPE MODULATION TECHNIQUES

Modern modulation techniques exploit the fact that digital baseband data may be sent by varying both the envelope and phase (or frequency) of an RF carrier. Because the envelope and phase offer two degrees of freedom, such modulation techniques map baseband data into four or more possible RF carrier signals. Such modulation techniques are called M-ary modulation, since they can represent more signals than if just the amplitude or phase were varied alone.

In an M-ary signaling scheme , two or more bits are grouped together to form symbols and are of Mpassible signals, $s_1(t)$ $s_2(t)$,...., s_m (t) is transmitted during each symbol period of duration T_s usually, the number of possible signals is M= 2^n, where n is an integer.

Depending on whether the amplitude, phase, or frequency of the carrier is varied, the modulation scheme is called *M*-ary ASK, *M*-ary PSK, or *M*-ary FSK.

Modulations which alter both the amplitude and phase of the carrier are the subject of active research.

M-ary modulation schemes achieve better bandwidth efficiency at the expense of power efficiency.

12.14.1 M-ary Phase Shift Keying (MPSK)

In M-ary PSK, the carrier phase takes on one of *M* possible values, namely, $\theta_i = 2\ (i - 1)\ \pi/M$, where $i = 1,2$, M. The modulated waveform can be expressed as

$$S_i(t) = \sqrt{\frac{2E_s}{T_s}}\cos\left(2\pi f_c t + \frac{2\pi}{M}(i-1)\right),\ 0 \le t \le T_s\ i = 1, 2 \ldots, M \qquad \ldots(12.73)$$

where $E_s = [\log_2 M]E_b$ is the energy per symbol and $T_s = [\log_2 M]\ T_b$ is the symbol period. The above equation can be rewritten *n* quadrature form as

$$S_i(t) = \sqrt{\frac{2E_s}{T_s}}\cos\left[(i-1)\frac{2\pi}{M}\right]\cos(\pi f_c t)\ i = 1, 2, \ldots, M \quad \ldots(12.74)$$

$$= \sqrt{\frac{2E_s}{T_s}}\sin\left[(i-1)\frac{2\pi}{M}(2\pi f_c t)\right] \qquad \ldots(12.75)$$

By choosing orthogonal basis signals

$$\phi_1(t) = \sqrt{\frac{2}{T_s}}\cos(2\pi f_c t), \quad \text{and} \quad \phi_2(t) = \sqrt{\frac{2}{T_s}}\sin(2\pi f_c t)$$

defined over the internal $0 \le t \le T_s$, the *M*-ary PSk signal set can be expressed as

$$S_{M-PSK}(t) = \left\{\sqrt{E_s}\cos\left[(i-1)\frac{\pi}{2}\right]\phi_1(t), -\sqrt{E_s}\sin\left[(i-1)\frac{\pi}{2}\right]\phi_2(t)\right\}\ i = 1, 2, \ldots, M \ \ldots(12.76)$$

Since there are only two basis signals, the constellation of *M-ary* PSK is two dimensional. The M-ary message points are equally spaced on a circle of radius $\sqrt{E_s}$ centered at the origin. The constellation diagram of an 8-ary PSK signal set is illustrated in figure 12.27. It is clear from figure 12.27 that MPSK is a constant envelope signal when no pulse shaping is used.

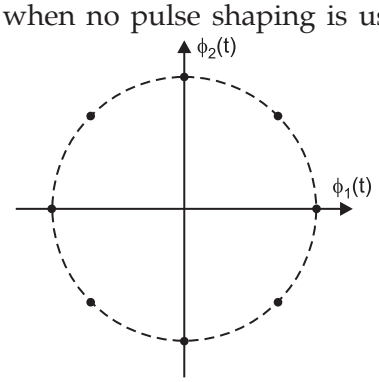


Fig. 12.27: Constellation diagram of an /W-ary PSK system (M = 8).

From the geometry of figure 12.27, it is easily seen that the distance between adjacent symbols is equal to $2\sqrt{E_s}\sin\left(\frac{\pi}{M}\right)$. Hence, the average symbol error probability of an M-ary PSK system is given by

$$Pe \leq 2Q\left(\sqrt{\frac{2E_b \log_2 M}{N_0}}\sin\left(\frac{\pi}{M}\right)\right) \quad ...(12.77)$$

Just as in BPSK and QPSK modulation, M-ary PSK modulation is either coherently detected or differentially encoded for non-coherent differential detection. The symbol error probability of a differential M-ary PSK system in AWGN channel for M ≥ 4 is approximated by

$$\text{Pe} \approx 2Q\left[\sqrt{\frac{4E_s}{N_0}}\sin\left(\frac{\pi}{2M}\right)\right] \quad ...(12.78)$$

The power spectral density (PSD) of an M-ary PSK signal can be obtained in a manner similar to that described for BPSK and QPSK signals. The symbol duration T_s of an M-ary PSK signal is related to the bit duration T_b by

$$T_S = T_b \log_2 M \quad ...(12.79)$$

The PSD of the M-ary PSK signal with rectangular pulses is given by

$$P_{\text{MPSK}} = \frac{E_s}{2}\left[\left(\frac{\sin\pi(f-f_c)T_s}{\pi(f-f_c)T_s}\right)^2 + \left(\frac{\sin\pi(f-f_c)T_s}{\pi(f-f_c)T_s}\right)^2\right] \quad ...(12.80)$$

$$P_{\text{MPSK}} = \frac{E_b \log_2 M}{2}\left[\left(\frac{\sin\pi(f-f_c)T_b\log_2 M}{\pi(f-f_c)T_b\log_2 M}\right)^2 + \left(\frac{\sin\pi(-f-f_c)T_b\log_2 M}{\pi(f-f_c)T_b\log_2 M}\right)^2\right] \quad ...(12.81)$$

The bandwidth and power efficiency of M-PSK systems using ideal Nyquist pulse shaping in AWGN for various values by M are listed in Table 12.3.

Table 12.3: Bandwidth and Power Efficiency of M-ary PSK Signals

M	2	4	8	16	32	64
$\eta_B = R_b/B^*$	0.5	1	1.5	2	2.5	3
E_b/N_0 for BER = 10^{-6}	10.5	10.5	14	18.5	23.4	28.5

*B: First null bandwidth of M-ary PSK signals.

12.14.2 M-ary Quadrature Amplitude Modulation (QAM)

In M-ary PSK modulation, the amplitude of the transmitted signal was constrained to remain constant, thereby yielding a circular constellation. By allowing the amplitude to also vary with the phase, a new modulation scheme called quadrature amplitude modulation (QAM) is obtained.

Figure 12.28 shows the constellation diagram of 16-ary QAM. The constellation consists of a square lattice of signal points. The general form of an M-ary QAM signal can be defined as

$$S_i(t) = \sqrt{\frac{2E_{\min}}{T_s}} a_i \cos(2\pi f_c t) + \sqrt{\frac{2E_{\min}}{T_s}} b_i \sin(2\pi f_c t) \quad ...(12.82)$$

$$o \le t \le T \qquad i = 1, 2, ...M$$

where, $E_{\min}$ is the energy of the signal with the lowest amplitude, and a_t and b_i are a pair of independent integers chosen according to the location of the particular signal point.

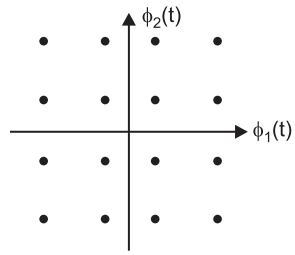


Fig. 12.28: Constellation diagram of an M-ary QAM (M = 16) signal set.

If rectangular pulse shapes are assumed, the signal $S_i(t)$ may be expanded in terms of a pair of basis functions defined as

$$\phi_1(t) = \sqrt{\frac{2}{T_s}} \cos(2\pi f_c t) \quad 0 \le t \le T_s \quad ...(12.83)$$

$$\phi_2(t) = \sqrt{\frac{2}{T_s}} \sin(2\pi f_c t) \quad 0 \le t \le T_s \quad ...(12.84)$$

The coordinates of the ith message point are

$$a_i\sqrt{E_{\min}} \text{ and } b_i\sqrt{E_{\min}}$$

where (a_t, b_i) is an element of the L by L matrix given by

$$\{a_i, b_i\} = \begin{bmatrix} (-L+1),(L-1) & (-L+3,L-1) & .... & (L-1,L-1) \\ (-L+1,L-3) & (-L+3,L-3) & .... & (L-3,L-3) \\ \vdots & & & \\ (-L+1,-L+1) & (-L+3,-L+1) & .... & (L-1,-L+1) \end{bmatrix}$$

where, $L = \sqrt{M}$

It can be shown that the average probability of error in a AWGN channel for M-ary QAM, using coherent detection, can be approximated by

$$P_e \cong 4\left(1 - \frac{1}{\sqrt{M}}\right) Q\left(\sqrt{\frac{2E_{\min}}{N_0}}\right) \quad ...(12.86)$$

In terms of the average signal energy E_{av}, this can be expressed as

$$P_e \cong 4\left(1 - \frac{1}{\sqrt{M}}\right) Q\left(\sqrt{\frac{3E_{av}}{(M-1)N_0}}\right) \quad ...(12.87)$$

The power spectrum and bandwidth efficiency of QAM modulation is identical to M-ary PSK modulation. In terms of power efficiency, QAM is superior to M-ary PSK.

Table 12.4 lists the bandwidth and power efficiencies of a QAM signal for various values of M, assuming optimum raised casine rolloff filtering in AWGN.

Table 12.4: Bandwidth and Power Efficiency of QAM

M	4	16	64	256	1024	4096
η_B	1	2	3	4	5	6
E_b/N_0 for BER = 10^{-6}	10.5	15	18.5	24	28	33.5

12.14.3 M-ary frequency shift keying (MFSK) and OFDM

In M-ary FSK modulation, the transmitted signals are defined by

$$S_t(t) = \sqrt{\frac{2E_s}{T_s}}\cos\left[\frac{\pi}{T_s}(n_c+i)t\right] \quad 0 \le t \le T_s, \;\; i = 1, 2, ..., M \quad ...(12.88)$$

where, $f_c = \dfrac{n_c}{2T_s}$ for some fixed integer n_c.

The M transmitted signals are of equal energy and equal duration, and the signal frequencies are separated by $\frac{1}{2}T_s$ Hz, making the signals orthogonal to one another.

For coherent M-ary FSK, the optimum receiver consists of a bank of M correlators, or matched filters, which are turned to the M distinct carriers. The average probability of error based on the union bound is given by

$$P_e \le (M-1)Q\left(\sqrt{\frac{E_b \log_2 M}{N_0}}\right) \quad ...(12.89)$$

For non-coherent detection using matched filter followed by envelope detectors, the average probability of error is given by

$$P_e = \sum_{k=1}^{M-1}\left(\frac{(-1)^{k+1}}{k+1}\right)\left(\frac{M-1}{k}\right)\exp\left(\frac{-kE_s}{(k+1)N_0}\right) \quad ...(12.90)$$

Using only the leading terms of the binomial expansion, the probability of error can be bounded as

$$P_e \le \frac{M-1}{2}\exp\left(\frac{E_s}{2N_0}\right) \quad ...(12.91)$$

The channel bandwidth of a coherent as M-ary FSK signal may be defined as

$$B = \frac{R_b(M+3)}{2\frac{\log_2}{2}M} \quad ... (12.92)$$

and that of a non-coherent MFSK may be defined as

$$B = \frac{R_b M}{2\log_2 M} \quad ... (12.93)$$

This implies that the bandwidth efficiency of an *M-ary* FSK signal decrease with increasing M. Therefore, unlike M-PSK signals, M-FSK signals are bandwidth

inefficient. However, since all the *M* signals are orthogonal, there is no crowding in the signal space, and hence the power efficiency increases with *M*.

Table 12.5 provides a listing of bandwidth and power efficiency of M-FSK signals for various values of *M*.

Table 12.5: Bandwidth and Power Efficiency of Coherent M-ary FSK

M	2	4	8	16	32	64
η_B	0.4	0.57	0.55	0.42	0.29	0.18
E_b/N_0 for BER = 10^{-6}	13.5	10.8	9.3	8.2	7.5	6.9

The orthogonality characteristic of MFSK has led researchers to explore orthogonal frequency division multiplexing (OFDM) as a means of providing power efficient signaling for a large number of users on the same channel. Each frequency in equation is modulated with binary data (on/off) to provide a number of parallel carriers each containing a portion of user data.

MFSK and OFDM modulation methods are being explored for high speed data connections as part of the IEEE 802.11 a standards activities to provide 54 Mbps WLAN connections, as well as for high speed line-of-sight and non-line-of-sight microwave connections for MMDS operations.

HIGHLIGHTS

1. Modulation is the process of encoding information from a message source in a manner suitable for transmission. It generally involves translating a baseband message signal (called the source) to a bandpass signal at frequencies that are very high when compared to the baseband frequency.
2. Frequency Modulation (FM) is the most popular modulation technique used in mobile radio systems. In FM, the amplitude of the modulated carrier signal is kept constant while its frequency is varied by the modulating message signal. Thus, FM signals have all their information in the phase or frequency of the carrier.
3. FM systems require a wider frequency band is the transmitting media (generally several times as large as that needed for AM) in order to obtain the advantages of reduced noise and capture effect.
4. In amplitude modulation, the amplitude of a high frequency carrier signal is varied in accordance to the instantaneous amplitude of the modulating message signal.
5. Since both, the sidebands of an AM signal carry the same information, it is possible to remove one of them without losing any information.
6. The two common techniques used for generating an SSB signal are the filter method and the balanced modulator method.
7. AM demodulation techniques may be broadly divided into two categories: coherent and non-coherent demodulation. Coherent demodulation requires knowledge of the transmitted carrier frequency and phase at the receiver, whereas non-coherent detection requires no phase information.
8. FM is part of a more general class of modulation known as angle modulation. Angle modulation varies a sinusoidal carrier signal in such a way that the angle

of the carrier is varied according to the amplitude of the modulating baseband signal.

9. There are basically two methods of generating an FM signal:
   (*i*) The direct method
   (*ii*) The indirect method
10. The indirect method of generating FM. It is based on approximating a narrowband FM signal as the sum of a carrier signal and a single side band (SSB) signal where the sideband is 90° out of phase with the carrier.
11. It can be easily shown that FM demodulation can be performed by taking the time derivative (often called slope detection) of the FM signal, followed by envelope detection.
12. The Phase Locked Loop (PLL) method is another popular technique to demodulate an FM signal.
13. The PLL is a closed loop control system which can track the variations in the received signal phase and frequency.
14. Modern mobile communication systems use digtal modulation techniques. Advancements in Very Large Scale Integration (VLSI) and digital signal processing (DSP) technology have made digital modulation more cost-effective than analog transmission systems.
15. Several factors influence the choice of a digital modulation scheme. A desirable modulation scheme provides low bit error rates at low received signal-to-noise ratios, performs well in multipath and fading conditions, occupies a minimum of bandwidth, and is easy and cost-effective to implement.
16. Digital modulation techniques may be broadly classified as linear and nonlinear. In linear modulation techniques, the amplitude of the transmitted signal, $S(t)$, varies linearly with the modulating digital signal, $m(t)$.
17. In binary phase shift keying (BPSK), the phase of a constant amplitude carrier signal is switched between two values according to the two possible signals W_1 and m_2 corresponding to binary 1 and 0, respectively. Normally, the two phase are repeated by 180°.
18. BPSK uses coherent or synchronous demodulation, which requires that information about the phase and frequency of the carrier be available at the receiver.
19. Differential PSK is a non-coherent form of phase shift keying which avoids the need for a coherent reference signal at the receiver. Non-coherent receivers are easy and cheap to build, and hence are widely used in wireless communications.
20. Quadrature phase shift keying (QPSK) has twice the bandwidth efficiency of BPSK, since two bits are transmitted in a single modulation symbol.
21. The $\pi/4$ shifted QPSK modulation is a quadrature phase shift keying technique which offers a compromise between OQPSK and QPSK in terms of the allowed maximum phase transitions. It may be demodulated in a coherent or non-coherent fashion.
22. Switching between two constellations, every successive bit ensures that there is at least a phase shift which is an integer multiple of $\pi/4$ radions between successive symbols. This ensures that there is a phase transition for every symbol, which enables a receiver to perform timing recovery and synchronization.

23. While both the baseband and IF differential detector determine the cosine and sine function of the phase difference, and then decide on the phase difference, and then decide on the phase difference accordingly, the FM discriminator detects the phase difference directly in a non-coherent manner.

24. Many practical mobile radio communications system use nonlinear modulation methods, where the amplitude of the carrier is constant, regardless of the variation in the modulating signal.

25. In binary frequency shift keying (BFSK), the frequency of a constant amplitude carrier signal is switched between two values according to the two possible message.

24. states (called high and low tones), corresponding to a binary 1 or 0.

26. The power spectral density of a binary FSK signal consists of discrete frequency components at f_c, $f_c + n\Delta f$, $f_c - n\Delta f$, where n is an integer.

27. MSK is a spectrally efficient modulation scheme and is particularly attractive for use in mobile radio communication systems. It possesses properties such as constant envelope, spectral efficiency, good BER performance, and self-synchronizing capability.

28. In an M-ary signaling scheme, two or more bits are grouped together to form symbols and one of M possible signals, $S_1(t)$, $S_2(t)$, ... $S_M(t)$ is transmitted during each symbol period of duration T_S. Usually, the number of possible signals is $M = 2^n$, where n is an integer.

29. In M-ary PSK modulation, the amplitude of the transmitted signal was constrained to remain constant, thereby yielding a circular constellation. By allowing the amplitude to also vary with the phase, a new modulation scheme called quadrature amplitude modulation (QAM) is obtained.

30. The M transmitted signals are of equal energy and equal duration, and the signal frequencies are separated by $1/2T_S$ Hz, making the signals orthogonal to one another.

31. The orthogonality characteristic of MFSK has led researchers to explore Orthogonal Frequency Division Multiplexing (OFDM) as a means of providing power efficient signaling for a large number of users on the same channel. Each frequency in equation is modulated with binary data (on/off) to provide a number of parallel carriers each containing a portion of user data.

REVIEW QUESTIONS

12.1. How does Gaussian MSK differ from MSK? Discuss GMSK in brief.

12.2. Explain with the help of a block diagram the working of quadrature modulation based MSK modulator.

12.3. What are various modulation techniques used for mobile networks ?

12.4. Derive the probability of error expressions for DPSK and non-coherent orthogonal binary FSK.

12.5. Verify that the BPSK receiver shown in Fig. 12.10 is able to recover a digital message (t). 12.6. Define the difference between amplitude modulation and frequency modulation.

12.7. Explain in brief angle modulation and also explain the spectra and bandwidth of FM signal.

12.8. What do you means by FM detection techniques ?

12.9. How many types of linear modulation techniques are used ? Explain.

12.10. Explain the difference between DPSK and QPSK.

12.11. Define the $\pi/4$ QPSK.

12.12. Explain with the help of block diagram, the working of binary phase shift keying(BPSK).

12.13. How many types of constant envelope modulation techniques are used ?

12.14. Explain:

(*i*) QAM

(*ii*) OFDM

(*iii*) BFSK

(*iv*) MFSK

12.15. Plot the BER *vs.* E_b/N_0 performance for BPSK, DPSK, QPSK and non-coherent FSK in additive Goussian white noise. List advantage and disadvantage of each modulation method from the mobile communication standpoint.

13 Chapter

ARCHITECTURE AND APPLICATIONS OF WIRELESS NETWORKS

13.1. *Wireless USB.* **13.2.** *Wireless Network Design Analysis.* **13.3.** *GPRS* (*General Packet Radio Service*). **13.4.** *Wireless Application Protocol* (*WAP*). **13.5.** *Bluetooth.* **13.6.** *Cellular Digital Packet Data* (*CDPD*). **13.8.** *Connectionless Protocol.* **13.9.** *Connection-Oriented Protocol.* **13.10.** *Internet Protocol.* **13.11.** *IP Address.* **13.12.** *OSI Model.* **13.13.** *Transmission Control Protocol.* **13.14.** *User Datagram Protocol* (*UDP*). **13.15.** *Wireless Location Service Architectures.*

13.1 WIRELESS USB

Wireless USB is a short-range, high-bandwidth wireless radio communication protocol created by the Wireless USB Promoter Group. Wireless USB is sometimes abbreviated as "WUSB", although the USB Implementers Forum discourages this practice and instead prefers to call the technology "Certified Wireless USB" to differentiate it from competitors. Wireless USB is based on the WiMedia Alliance's Ultra-WideBand (UWB) common radio platform, which is capable of sending 480 Mbit/s at distances up to 3 metres and 110 Mbit/s at up to 10 metres. It was designed to operate in the 3.1 to 10.6 GHz frequency range, although local regulatory policies may restrict the legal operating range for any given country.

An upcoming 1.1 specification will increase speed to 1 Gbit/s and working frequencies up to 6 GHz.

"Wireless USB" is also the name of a separate wireless protocol created by Cypress Semiconductor for human interface devices.

13.1.1 Uses

Wireless USB is used in game controllers, printers, scanners, digital cameras, MP3 players, hard disks and flash drives. It is also suitable for transferring parallel video streams. Kingston released a Wireless USB universal docking station in August, 2008.

13.1.2 Development

The Wireless USB Promoter Group was formed in February 2004 to define the Wireless USB specification. The group consists of Agere Systems (now merged with LSI Corporation), Hewlett-Packard, Intel, Microsoft, NEC Corporation, Philips and Samsung.

In May 2005, the Wireless USB Promoter Group announced the completion of the Wireless USB specification.

In June 2006, five companies showed the first multi-vendor interoperability demonstration of Wireless USB. A laptop with an Intel host adapter using an Alereon PHY was used to transfer high definition video from a Philips wireless semiconductor solution with a Realtek PHY, all using Microsoft Windows XP drivers developed for Wireless USB.

In October 2006 the FCC approved the first complete Host Wire Adapter (HWA) and Device Wire Adapter (DWA) wireless USB solution from WiQuest Communications for both outdoor and indoor use. The first retail product was shipped by IOGEAR using Alereon, Intel and NEC silicon in mid-2007. Around the same time, Belkin, Dell, Lenovo and D-Link began shipping products that incorporated WiQuest technology. These products included embedded cards in the notebook PCs or Hub/Adapter solutions for those PCs that do not currently include Wireless USB. In 2008, a new Wireless USB Docking Station from Kingston was made available through Dell. This product was unique as it was the first product on the market to support video and graphics over a USB connection, by using Display Link USB graphics technology. Kingston's Docking Station enables wireless connectivity between a notebook PC and an external monitor, speakers, and existing wired USB peripherals. Imation announced Q408 availability of a new external Wireless HDD. Both of these products are based on WiQuest technology.

13.2 WIRELESS NETWORK DESIGN ANALYSIS

One of the imperatives of the modern combat leader is enabling his forces to exchange information easily between all services, at all echelons in the chain of command, in all forms of information to include voice, data, and video. The DoD (Designing of Digital Communication) master plan for information systems stresses affordability, interoperability, COTS (Commercial off the Shelf), easy connectivity between services, and support for mobile forces over a wide zone of battle. Part of this requirement means that communications systems must be able to link squad-level units and sometimes even individuals into a mobile communication environment with hardware resembling cellular telephones, laptop computer, and video cameras. Currently there are two existing wireless/ mobile networking technologies that will support the DoD master plan; there are cellular telephone technology and wireless local area network IEEE 802.1lx. This research focuses on the IEEE 802.1lb and the 802.11a standards.

13.2.1 IEEE 802.11b

The IEEE 802.11b is now the most common wireless technology standard used in offices, airports, and coffee shops. The IEEE established the 802.1lb standard in 1997; it can provide data rates of 1 -11 Mbps from 600 metres to 25 miles. 802. 11b operates in the 2.4-2.83 GHz range. 802.11b supports multi-hop wireless network devices meaning that each radio supports more than one

peer-to-peer link and acts as a repeater or bridge between the established links. This results in a dynamic, infrastructure less network capable of supporting up to 200 users per gateway. In times of heavy congestion the frequency spectrum can be divided into 11 different channels (in the US).

13.2.2 IEEE 802.11a

The IEEE ratified the 802.11a standard in 1999. The 802.11a standard can provide data rates of 1 -54 (72 in the Turbo Mode) Mbps from 300 metres to 25 miles; it operates in the 5.15-5.35 GHz range. 802.11b and 802.11a use identical framing; however, 802.11a uses an entirely different encoding scheme called Orthogonal Frequency Division Multiplexing or OFDM. The 802.11a also supports multi-hop wireless devices and during heavy congestion it can divide the frequency spectrum into 10 different channels. Although segmented, the total bandwidth available for IEEE 802.11a applications is almost 4 times that of 802.11b because the 802.11b spectrum is plagued by saturation from wireless phones, microwave ovens and other emerging wireless technologies, such as Bluetooth.

13.2.3 Technology Analysis

In real world applications, vendors recommend using both the 802. 11b and 802.11a standards in designing new wireless networks. Because 802.11a and 802.11b operate in different frequencies they can coexist within the same network. The coexistence of both technologies enables the system administrator to better manage the network by spreading the network over two different frequency domains and up to 21 different channels within the frequency spectrum. Several vendors, such as Cisco and Netgear, produce several suites of devices that can support a combined 802.11b and 802.11a wireless network.

The Wireless Network design for information systems must support a maneuver Brigade's existing ground and signal platforms so that connectivity in tactical situations could be considered. One solution for a Brigade wireless network design is the Cisco AIRONET 1200 Series. The Cisco AIRONET 1200 Series products are compliant with IEEE 802.11a and 802.11b standards. The modular design of the Cisco AIRONET 1200 Series allows for both single- and dual-band configurations. The 802.11a radio supports data rates up to 54 Mbps and ten non-overlapping channels that offer high performance and scalability. The 802.11b radio provides data rates up to 11 Mbps and eleven non-overlapping channels, which supports widely deployed 802.11b clients. With a variety of directional and omni directional antenna the clients can be out to 25 Miles from an Access Point. The Omni directional antenna provides optimal coverage patterns and maximum range.

The Cisco AIRONET 1200 Series simplifies wireless LAN management because many of the same management tools and capabilities available in wired networks are used on the wireless network. The 1200 Series supports network management through Simple Network Management Protocol (SNMP), Telnet, and a Web browser to aid in troubleshooting, monitoring, software download,

and event logging. The CiscoWorks Wireless LAN Solution Engine is also available as a management tool.

13.3 GPRS (GENERAL PACKET RADIO SERVICE)

GPRS (General Packet Radio Service) is a step between GSM and 3G cellular networks. GPRS offers faster data transmission via a GSM network within a range 9.6Kbits to 115Kbits. This new technology makes it possible for users to make telephone calls and transmit data at the same time. (For example, if you have a mobile phone using GPRS, you will be able to simultaneously make calls and receive e-mail massages.). This topic will entail an intermediate language and serves to introduce this technology to both the developer and end-user alike and discuss the underlying architecture and basic functionality of GPRS followed by its many applications.

GPRS gives subscribers access to data communication applications such as e-mail, corporate networks, and the Internet using their mobile phones. The GPRS service uses the existing GSM network and adds new packet-switching network equipment. GPRS employs packet switching, which means that the GPRS mobile phone has no dedicated circuit assigned to it. Only when data is transferred is a physical channel created. After the data has been sent, it can be assigned to other users. This allows for the most efficient use of the network.

When packet-switched data leaves the GPRS/GSM network, it is transferred to TCP-IP networks such as the Internet or X.25. Thus, GPRS includes new transmission and signaling procedures as well as new protocols for inter working with the IP world and other standard packet networks. Mobile phones currently available do not work with the new GPRS technology. The industry's mobile phone vendors are working on both existing and new phones that will support both GSM and packet switching. It is also becoming more common for laptops and PDA's (Personal Digital Assistants) to have a GPRS phone integrated in them. Perfectly dubbed a 2.5G system the General Packet Radio Service has bridged us to the 3G (UMTS) of today.

13.3.2 GPRS Telephones

Due to the fact that more than one channel is used for downlink, the GPRS mobile phones make possible greater data transmission speeds. There are several types of phones with regard to the number of channels they use for data transmission.

(*i*) Type 2 + 1 – two downlink channels and one uplink data transmission channel.

(*ii*) Type 3 + 1 – three downlink channels and one uplink data transmission channel.

(*iii*) Type 4 + 1 – four downlink channels and one uplink data transmission channel.

(*iv*) The GPRS mobile phones can be classified into the following three classes in terms of the possibility of simultaneous calls (via GSM) and data transmission (via GPRS).

- Class A – Simultaneous calls (via GSM) and data transmission (via GPRS)
- Class B – Automatic switching between the GSM and the GPRS mode is possible according to telephone settings.
- Class C – Hand operated switching between the GSM and the GPRS mode

13.3.3 Data Transmission Speeds

The supported data transmission speed per channel is 13.4Kbits. Depending on the type of phone, the following data transmission speeds are theoretically possible:

(i) Type 2 + 1: Receive 26.8 Kbits and send 13.4 Kbits.
(*ii*) Type 3 + 1: Receive 40.2 Kbits and send 13.4 Kbits.
(*iii*) Type 4 + 1: Receive 53.6 Kbits and send 13.4 Kbits.

13.3.4 The Network

In the core network, the existing MSCs (Mobile Services Switching Centers) are based upon circuit-switched technology, and they cannot handle the GPRS style packet traffic. Thus two new and very important components, called GPRS Support Nodes, are added:

(*i*) Serving GPRS Support Node (SGSN)
(*ii*) Gateway GPRS Support Node (GGSN)

The SGSN can be viewed as a "packet-switched MSC;" it delivers packets to mobile stations (MSs) within its service area. SGSNs send queries to home location registers (HLRs) to obtain profile data of GPRS subscribers. SGSNs detect new GPRS MSs in a given service area, process registration of new mobile subscribers, and keep a record of their location inside a given area. Therefore, the SGSN performs mobility management functions such as mobile subscriber attach/detach and location management.

GGSNs are used as interfaces to external IP networks such as the public Internet, other mobile service provider's GPRS services. GGSNs maintain routing information that is necessary to "tunnel" the protocol data units (PDUs) to the SGSNs that service particular MSs. Other functions include network and subscriber screening and address mapping, as well as authentication and charging functions. One (or more) GGSNs may be provided to support multiple SGSNs. More detailed technical descriptions of the SGSN and GGSN are provided in a later section.

13.3.5 Roots of GPRS

As first stated GPRS will complement rather than replace the current data services available through today's GSM digital cellular networks, such as Circuit Switched Data and Short Message Service. It also provides the type of data capabilities planned for "third generation" cellular networks, but years ahead of them. Table 13.1 below is a timeframe of GSM data services and their availability.

Table 13.1: Roadmap of Data Services for GSM

	Timeframe	Capabilities	Notes
9.6 kbps service	Available today	Circuit-switched data and fax	Service available from most GSM operators today.
14.4 kbps service	Available today	Higher speed circuit-switched data and fax	Works identically to 9.6 kbps service only at higher speed
Direct IP Access	Available through some carriers today	Circuit-switched connection directly to Internet	Reduces call set-up time and provides a stepping-stone to packet data.
High-speed circuit-switched data service (HSCSD)	Available today	High speed rates to 56 kbps	A software-only upgrade for carriers not requiring expensive infrastructure.
GPRS	Available today	High speed packet data with transmission speeds over 100 kbps, with most user devices offering about 56 kbps	Extremely capable and flexible mobile communications.
EDGE	Available within three years	High speed packet data which will triple the rates available with GPRS	Final high-speed data technology for existing networks.
Third generation cellular	Available within three to five years	High speed packet data to 2 Mbps	Completely new airlink.

According to the specifications provided by the European Telecommunications Standards Institute (ETSI), the highest speed for a single user session (or time slot) is the coding scheme CS4, which allows 21.4 kbps per time slot. Thus, theoretically, a GPRS connection can provide a data transmission speed of up to 171.2 Kbps (approximately three times that of a fixed-line 56 K dial-up) if all eight slots are used. GPRS's rival, HSCSD, can achieve up to 57.6 kbps. However, it is unlikely that network operators will let a single user use up all the time slots. Even Nokia admitted that realistically GPRS could achieve only about 43 Kbps while Ericsson thinks 56 Kbps is achievable. Currently, GSM systems are running at 9.6 kilobits. A comparison of Data Transfer Speeds (in kbps) follows in Table 13.2.

Table 13.2: A Comparison of Data Transfer Speeds (in Kbps)

56 KDial-Up	GSM	HSCSD (maximum speed)	GPRS (maximum speed)	GPRS (realistic speed)
56	9.6	57.6	171.2	43 to 56

GPRS could possibly be the technology that will allow consumers to really begin to pursue the mobile Internet. GPRS is considered one step ahead of HSCSD (High Speed Circuit Switched Data) and a step towards 3G (Third-

generation) networks. It is the step to 2.5G for GSM and TDMA (Time Division Multiple Access) service providers.

GPRS is ideal for Wireless Application Protocol (WAP) services because of the cost saving WAP over GPRS bring to mobile operators and cellular consumers. Costs are reduced because GPRS radio resources are only needed while the message is being transferred. For the end user, that means you only pay for the time it takes to download the data and information that you need. For the GSM operator, that means that you will be able to provide high speed Internet access to consumers at a reasonable cost, because you will bill mobile phone users for only the amount of data that they transfer rather than billing them for the length of them that they are connected to the network.

With GPRS-enabled mobile phones, services are received faster than with traditional GSM phones. GPRS offers an increase in data throughput rates, so information retrieval and database access is faster, more usable and more convenient. At its best, GPRS is transparent, allowing the user to concentrate on the task in hand rather than on the technology.

13.3.6 GPRS History

Table 13.3

Date	Milestone
Throughout 1999-2000	Network operators place trial and commercial contracts for GPRS 1999-2000 infrastructure.
	Incorporation of GPRS infrastructure into GSM networks.
	First trial GPRS services become available.
	Typical single user throughput is likely to be 28 kbps.
	For example, T-Mobil is planning a GPRS trial at Expo2000 in Hanover in the Summer of 2000.
Start of 2001	Basic GPRS capable terminals begin to be available in commercial quantities.
Throughout 2001	Network operators launch GPRS services commercially an roll out GPRS.
	Vertical market and executive GPRS early adopters begin using it regularly for nonvoice mobile communications.
2001/2002	Typical single user throughput is likely to be 56 kbps.
	New GPRS specific applications, higher bitrates, greater network capacity solutions, more capable terminals become available, fueling GPRS usage.
2002	Typical single user throughput is likely to be 112 kbps.
	GPRS Phase 2/EDGE begins to emerge in practice.
2002	GPRS is routinely incorporated into GSM mobile phones and has reached critical mass in terms of usage. (This is the equivalent to the status of SMS in 1999)
2002/2003	3GSM arrives commercially.

13.3.7 GPRS Architecture

Network Protocols Used

There are several protocols used in the network equipment. These protocols operate in both the data and signalling planes. The following is a brief description of each protocol layer shown in Figure 13.1 on the next page.

- *Sub-Network Dependent Convergence Protocol* (*SNDCP*): the protocol that maps a network-level protocol, such as IP or X.25, to the underlying logical link control. SNDCP also provides other functions such as compression, segmentation and multiplexing of network-layer messages to a single virtual connection.
- *Logical Link Control* (*LLC*)*:* a data link layer protocol for GPRS which functions similar to Link Access Protocol - D (LAPD). This layer assures the reliable transfer of user data across a wireless network.
- *Base Station System GPRS Protocol* (*BSSGP*): BSSGP processes routing and quality of service (QoS) information for the BSS. BSSGP uses the Frame Relay Q.922 core protocol as its transport mechanism.
- *GPRS Tunnel Protocol* (*GTP*): protocol that tunnels the protocol data units through the IP backbone by adding routing information. GTP operates on top of TCP/UDP over IP.
- *GPRS Mobility Management* (*GMM*): protocol that operates in the signalling plane of GPRS and handles mobility issues such as roaming, authentication, and selection of encryption algorithms.
- *Network Service:* protocol that manages the convergence sub-layer that operates between BSSGP and the Frame Relay Q.922 Core by mapping BSSGP's service requests to the appropriate Frame Relay services.
- *BSSAP+:* protocol that manages paging for voice and data connections and optimizes paging for mobile subscribers. BSSAP+ is also responsible for location and routing updates as well as mobile station alerting.

Figure 13.1 shows the GPRS protocol stack, and Figures 13.2 show GPRS architecture.

These support nodes are necessary to integrate GPRS into the previously Established architecture of GSM. The support nodes (GSN) take on the task of delivering and routing data packets between the mobile stations and the PDNs (packet data network). Figure 13.2 shows the interfaces between these new nodes and the GSM network. Notice how the signaling data is manipulated

The blue Gb interface connects the dinosaur component (BSC) with the SGSN. Using two other, unique interfaces - Gn and Gp - user and signaling data are both transmitted between the GSNs. The Gp interface is only used if the two GSNs are in different PLMNs. Otherwise, the Gn interface is used (GSNs are in the same PLMN). A backbone IP-based GPRS network is used to connect all GSNs. Here the GSNs package the PDN packets and tunnel (transmit) them using the GTP (GPRS tunneling protocol). The type of backbone

depends upon whether or not the GSNs exist within the same PLMN. Figure 13.2 shows two "intra-PLMN" backbones within an "inter-PLMN" network. Intra-PLMN backbones have GSNs within the same PLMN, inter-PLMN networks have GSNs in differing PLMN.

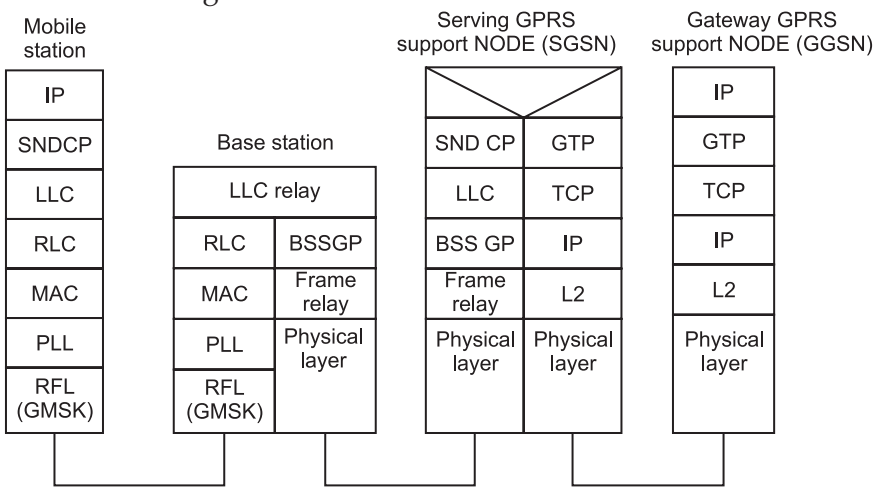


Fig. 13.1: GPRS protocol stack.

In order to fully understand the building blocks of our topic, GPRS, it suffices to first look at the make-up of the GMS system.

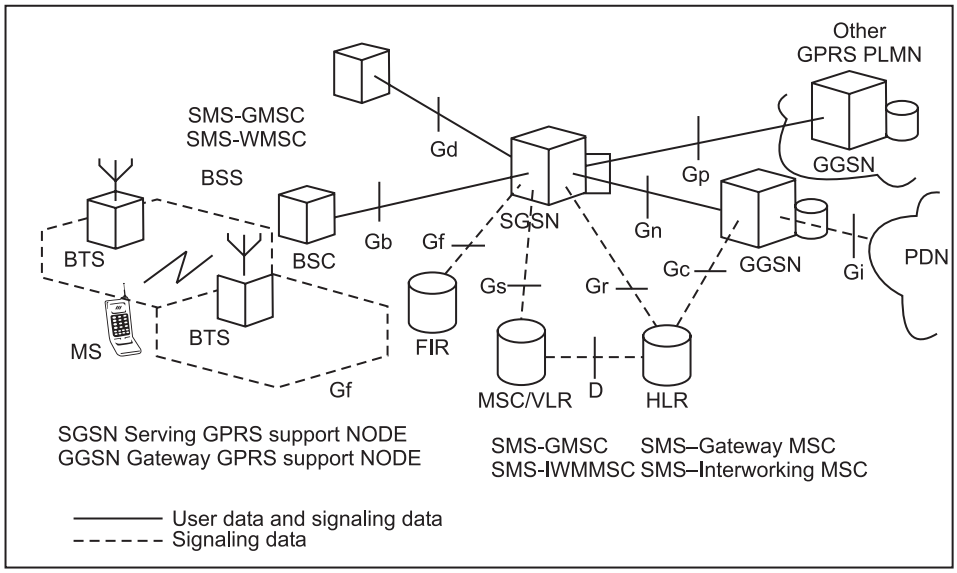


Fig.13.2: GPRS system architecture.

13.4 WIRELESS APPLICATION PROTOCOL (WAP)

13.4.1 The Basics Behind WAP

The Wireless Application Protocol has become the standard for communication between server applications and its clients. That is exactly what the Wireless

Application Protocol is; for example, WAP is used as a standardized method so that a cellular phone can talk to a server among the cellular network that it belongs to. Because WAP has become so global, it no longer is bounded by the means of the cellular market. WAP has become the link of the Internet to the Mobile World, bridging a gap between two of the top industries of the world.

WAP also follows a model similar to the Internet. The Internet itself has a layered protocol stack. The portable device using WAP has browser software that connects to a WAP Gateway and sends requests to receive data from web servers. Data could be a web page or an email. The content is then sent back to the portable device, and depending on the capability of the portable device to receive and view data, the data is received and viewable.

WAP founders include Ericsson, Nokia, Motorola, and Phone.Com (formerly Unwired Planet). In December 1997, these three large companies, all with strong influence on the Mobile market, formed the WAP Forum, an organization with open membership and now with over 300 members worldwide. The purpose of this forum is to make sure that the specifications of WAP do not go astray.

Basic specifications of WAP include: micro browsing, scripting, wireless telephone applications, and a layered protocol stack to be discussed in the next section.

13.4.2 Some Quick Technical Issues

Before beginning a description about the WAP Stack, the user should be briefed on a slight technical introduction. WAP is the connection between the Internet and the Wireless World, and there are two things that need to be addressed:

(*i*) *HTML vs. HDML vs. WML:* Hyper Text Markup Language (HTML) is the language that is used to translate data to a web browser to be viewed by a user. Handheld Device Markup Language (HDML) is used to incorporate handheld or portable device applications such as a micro browser. Wireless Markup Language (WML) is a language derived from HTML that is used with WAP to incorporate WAP features.

(*ii*) *How to incorporate WAP:* Here is; an example: We have a user that uses a WAP-compliant mobile phone. That is, his phone has WAP features on it such as a small email application and a mini browser. He decided to view something on his mini browser, thus requesting to see data. The request is passed to the WAP Gateway of his mobile service network, where their servers retrieve the information in standard HTML format or perhaps even in WML if the server can make the translation through a filter in the WAP Gateway. One the request is sent through the WAP Gateway, the WAP Gateway in turn sends the data to the WAP client to view on his micro browser. Diagram goes here.

WAP Architecture is divided into several layers. This is often called the 'WAP Stack'.Below are technical descriptions of each layer.

1 -Wireless Application Environment (WAE)
2-Wireless Session Protocol (WSP)
3-Wireless Transaction Protocol (WTP)
4-Wireless Transport Layer Security (WTLS)
5-Wireless Data Protocol (WDP)
6-Bearers
WAP LAYERED STAK

1. Wireless Application Environment (WAE)

The goal of WAE is to establish an environment that allows users to build applications that can be used over a wide variety of wireless systems that can be implemented efficiently. WAE is able to be used with Phone, the World Wide Web's HTML, and should work with other technologies of the web, such as Uniform Resource Locator (URL) and Hyper Text Transfer Protocols (HTTP), and other modern technologies that are found within a mobile network.

In the WAP model, WAE would be the application environments that sends and receives requests from the Clients to the Gateway to the Origin Server. WAE relies heavily on URL and HTTP ideals by assuming the existence of a gateway server and support from the WAP gateway in the network.

WAE is composed of user agents such as browsers, text editors, date book or phonebook. WAE is also composed of scripting, higher-lever programming languages and image formats. WAE uses languages such as WML Script (similar to Java Script) and WML (similar to HTML).

2. Wireless Session Protocol (WSP)

The Wireless Session Protocol is the layer that deals with communication between client and proxy or server. The WSP provides dialog between client and server. WSP provides the following services:

(*i*) Opens a session of communication between client and server.

(*ii*) Establishes a protocol and negotiation between client and server.

(*iii*) Exchanges encoded data between client and server.

(*iv*) Exchanges request and replies between client and server.

(*v*) Supports several asynchronous transmissions of data.

3. Wireless Transaction Protocol (WTP)

The Wireless Transaction Protocol, in a simple definition, deals with the transaction of data. WTP handles transactions, re-transmission of data, and with the separation and concatenation of data.

This particular protocol has a separate interface that manages and referees the WTP layer and the settings of the handheld device. This management application is known as the WTP Management Entity. For WTP to work, the following factors are important:

(*i*) The handheld device is within coverage area of base agent.

(*ii*) The handheld device is turned on and has a power and is reliable.

(*iii*) Resources are adequate especially with the CPU and memory.

(*iv*) WTP settings are correctly inputted.

The WTP Management Entity makes sure the above factors are valid during the transaction session. Of course, it is assumed that the user wants to receive and transmit data.

4. Wireless Transport Layer Security (WTLS)

The Wireless Transport Layer Security is the layer that handles security of data and validity of data between two communicating to manage, start, and finish security issues between two portable devices.

To transport data, WTLS needs: the source address and port number to identify the message creator, and from where the message is being sent, the destination address and port number to which data is being sent, and of course, the data itself. WTLS has a connection interface which provides a connection protocol between client and server.

5. Wireless Data Protocol (WDP)

The Wireless Data Protocol acts as the communication layer between the upper level protocols (WTLS, WTP, and WSP), and the bearer services. WDP allows the upper layers to function independently from the wireless network at hand, as long as the WTP layer is specifically set to the settings of the bearer settings.

The function of the WDP is to provide a stable environment so that any of the underlying bearers can operate using WAP. WDP can be adapted to different bearers with different services however the services offered by WDP remains constant thus providing a continuous interface to the upper layers of the WAP stack.

6. Bearers

The bearers of WAP are the products or other types of medium that implements WAP in their network and in their technology, such as CDPD, Mobitex, and GSM.

Of course every great protocol or specification out there for wireless technology will definitely have its drawbacks. With WAP, there are several problems that may or may never get addressed during its development:

(*i*) Since WAP is designed for portable devices, important details matter: such as dealing with small screen sizes (a Palm Pilot or mobile phone). A small screen size does not allow high resolution or high pixel images to be seen clearly.

(*ii*) Because WAP is designed for portable devices, there is limited device memory so that not a lot of data can be stored.

(*iii*) WAP allows two devices to communicate only, and data transfers are usually done with Infrared, therefore the devices have limited range of data transfer.

(*iv*) WAP also works on devices that have limited bandwidth.

(*v*) There are no "cookies" available to hold the session together.

13.5 BLUETOOTH

13.5.1 What is Bluetooth?

Those who are not so savvy in the technology and wireless worlds may or may not know what is Bluetooth. The word 'Bluetooth' tends to be placed in the same sentence with the words 'WAP' and 'Wireless'. According to the Bluetooth website, a frequently asked question is: is Bluetooth a protocol, a standard, a technology, or an actual product?

A quick answer to this question is: Bluetooth is not an actual product. Bluetooth is a standard for wireless communication between short-ranged radio links to PCs to mobile phones.

The Special Interest Group, also known as its acronym SIG, includes companies such as: 3Com, Ericsson, IBM, Intel, Lucent, Microsoft, Motorola, Nokia and Toshiba, among the few big companies. SIG contains many members of smaller companies as well.

13.5.2 The Technology behind Bluetooth:

(*i*) **Global Band** : For Bluetooth to work globally, the frequency band must be established and available universally. Bluetooth only uses one frequency band that is available worldwide, and this is the ISM (Industrial-Scientific-Medical) because it runs at the appropriate frequency required for Bluetooth and in fact, ISM is the only band on the market that runs to this necessary frequency.

(*ii*) **Frequency Hopping** : When working with Bluetooth, one must realize that the connections of Bluetooth are primarily dependant on frequency since its short-ranged communication take into consideration interferring signals: such as garage door openers, walkie-talkies, microwaves, toaster ovens, and so on. Bluetooth takes care of this by hopping to different channels depending on where the Bluetooth device is located.

Table 13.4: Packet data in op-code format. Taken from "Bluetooth - The Universal Radio Interface for ad hoc, wireless connectivity" by Jaap Haartsen (March 1998).

ACCESS CODE	PACKET HEADER	PAYLOAD
72 BIT	54 BIT	0-2745 BIT

Fixed Packet Format

The standard of frequency hopping is easy to use and understand and thus implemented by Bluetooth. Frequency hopping divides the frequency band into different channels that are "hopable" - meaning that the frequency will hop from

one channel to the other, depending where the Bluetooth portable product is located. When there is some sort of interference - for example, maybe a garage opener - then the channel "hops" to another channel so that transmission and receiving of data is still easy to do. However interference handling is not always guaranteed and there are times when an interference will jam the entire Bluetooth device's connection. This is where error handling comes in and takes care of such problems by restoring some error bits.

13.6 CELLULAR DIGITAL PACKET DATA (CDPD)

Cellular Digital Packet Data, otherwise CDPD, is a mobile technology that transmits data over the system offered by Advanced Mobile Phone Service (AMPS). CDPD was primarily developed by IBM but its development has now moved to a communal effort among several big companies including Ameritech, AirTouch, and GTE - now Verizon. In the early stages of CDPD development, CDPD offered packet-switching data transmission over analog cellular networks and voice transmission over cellular voice networks.

CDPD devices now take voice channels over an analog cellular network such as AMPS described above. These channels have the ability to transmit relatively short data packets and use these data packets to start a packet-switching session. Like Bluetooth, CDPD uses a channel hopping service among the cellular frequency it uses.

The Wireless Data Forum is an organization that works with the specifications of CDPD and maintains its technology. Unlike other national organizations such as WAP, SIG, or MOA, the Wireless Data Forum not only maintains the specifications of CDPD, but also works on the development of its marketing, provides equipment, and offers technical support. The Wireless Data Forum is aiming to implement CDPD wherever possible.

13.6.1 CDPD Architecture

CDPD works primarily on systems - end systems (ES's) and intermediate systems (IS's). CDPD Architecture is similar to the model of the World Wide Web. The mobile end system (M-ES) acts as a client who uses the CDPD network over the wireless environment. The fixed end system (F-ES) acts as a server, a host, or a gateway depending on the environment. The fixed end system responds to the mobile end system to give whatever data it has requested.

For the M-ES to be able to communicate with the F-ES, it usually will have a radio modem that attaches to it and manages all the data, data traffic, and other types of protocols that are implemented with the CDPD Architecture.

There are also the intermediate systems (IS's) which works within the network. There are two types of IS's: the simple IS that acts as an IP hub or router and has no interest in the network it resides in. The other type of intermediate system is the mobile data intermediate system (MD-IS) which works based on the information it has about the environmental network that it lives in. MD-IS will take care of switching, logins and logouts, status, encryption, and mobility for the device using CDPD.

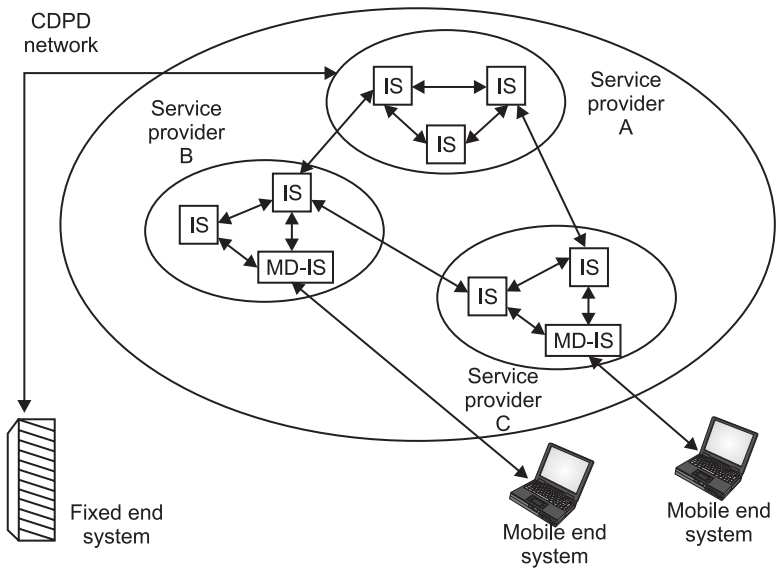


Fig. 13.3: A CDPD Network

I-Mode

I-Mode is the wireless technology produced by the Japanese-based company NTT DoCoMo. I-Mode simply allows cellular phone users to use the Internet on their cellular phones. This technology has just expanded to the European market and is looking to expand and implement WAP in the near future.

I-Mode uses CHTML (Compact Hyper Text Markup Language) and is a twenty-four hour connection service. There are no dial-up services or external phone number to dial too. Once the cellular phone is turned on, the user is connected to the Internet. Email is instant and large files will be easy to store. The user is billed by how much data they transmit and receive, not how many minutes they use the Internet. The cost is added to the bill. I-Mode has not been launched in the United States yet, but has been introduced to some parts in Europe. NTT DoCoMo is currently developing software in English while WAP continues to develop with I-Mode some I-Mode specifications as the two will probably collaborate in the near future.

Table 13.5: Protocols discussed in this survey paper.

Protocol	Organization	Mostly Founding	Developers	Features
WAP	WAP Forum	PDAs	Motorola Nokia Ericsson Phone.com	Microbrowsing scripting electronic notification telephony applications
Bluetooth	SIG	PDAs	Ericsson Intel Nokia Motorola	Frequency hopping global band channel hopping

vlOBITEX	MOA		Telia Mobitel (Sweeden)	Radio frequency
CDPD	WDF	Mobile devices	IBM Ameritech AirTouch Verizon	Voice data transmillion
I-Mode	None	Cellular phone	NTTDoCoMo	Internet Services on the cellular phone large capacity constant connection

137. IEEE 802.11

IEEE 802.11 is a set of standards for wireless local area network (WLAN) computer communication, developed by the IEEE LAN/MAN Standards Committee (IEEE 802) in the 5 GHz and 2.4 GHz public spectrum bands.

The 802.11 family includes over-the-air modulation techniques that use the same basic protocol. The most popular are those defined by the 802.11b and 802.11g protocols, and are amendments to the original standard. 802.11-1997 was the first wireless networking standard, but 802.11b was the first widely accepted one, followed by 802.11g and 802.11n. Security was originally purposefully weak due to export requirements of some governments, and was later enhanced via the 802.11i amendment after governmental and legislative changes. 802.11n is a new multi-streaming modulation technique that is still under draft development, but products based on its proprietary pre-draft versions are being sold. Other standards in the family (c-f, h, j) are service amendments and extensions or corrections to previous specifications.

802.11b and 802.11g use the 2.4 GHz ISM band, operating in the United States under Part 15 of the US Federal Communications Commission Rules and Regulations.

Because of this choice of frequency band, 802.11b and g equipment may occasionally suffer interference from microwave ovens and cordless telephones. Bluetooth devices, while operating in the same band, in theory do not interfere with 802.11b/g because they use a frequency hopping spread spectrum signaling method (FHSS) while 802.11b/g uses a direct sequence spread spectrum signaling method (DSSS). 802.11 a uses the 5 GHz U-NH band, which offers 8 non-overlapping channels rather than the 3 offered in the 2.4 GHz ISM frequency band.

The segment of the radio frequency spectrum used varies between countries. In the US, 802.11a and 802.11g devices may be operated without a license, as allowed in Part 15 of the FCC Rules and Regulations. Frequencies used by channels one through six (802.11b) fall within the 2.4 GHz amateur radio band. Licensed amateur radio operators may operate 802.11b/g devices under Part 97 of the FCC Rules and Regulations, allowing increased power output but not commercial content or encryption.

13.7.1 Protocols

The original version of the standard IEEE 802.11 was released in 1997 and clarified in 1999, but is today obsolete. It specified two net bit rates of 1 or 2 megabits per second (Mbit/s), plus forward error correction code. It specified three alternative physical layer technologies: diffuse infrared operating at 1 Mbit/s; frequency-hopping spread spectrum operating at 1 Mbit/s or 2 Mbit/s; and direct-sequence spread spectrum operating at 1 Mbit/s or 2 Mbit/s. The latter two radio technologies used microwave transmission over the Industrial Scientific Medical frequency band at 2.4 GHz. Previous WLAN technologies used lower frequencies, such as the U.S. 900 MHz ISM band.

13.7.2. IEEE 802.11a

The 802.11a standard uses the same data link layer protocol and frame format as the original standard, but an OFDM based air interface (physical layer). It operates in the 5 GHz band with a maximum net data rate of 54 Mbit/s, plus error correction code, which yields realistic net achievable throughput in the mid-20 Mbit/s[citation needed].

Since the 2.4 GHz band is heavily used to the point of being crowded, using the relatively un-used 5 GHz band gives 802.11 a a significant advantage. However, this high carrier frequency also brings a disadvantage: The effective overall range of 802.11a is less than that of 802.11b/g; 802.11a signals cannot penetrate as far as those for 802.11b because they are absorbed more readily by walls and other solid objects in their path due to their smaller wavelength.

13.7.3 IEEE 802.11b

802.11b has a maximum raw data rate of 11 Mbit/s and uses the same media access method defined in the original standard. 802.11b products appeared on the market in early 2000, since 802.11b is a direct extension of the modulation technique defined in the original standard. The dramatic increase in throughput of 802.11b (compared to the original standard) along with simultaneous substantial price reductions led to the rapid acceptance of 802.11b as the definitive wireless LAN technology.

802.11b devices suffer interference from other products operating in the 2.4 GHz and. Devices operating in the 2.4 GHz range include: microwave ovens, Bluetooth svices, baby monitors and cordless telephones.

13.7.4 IEEE 802.11g

In June 2003, a third modulation standard was ratified: 802.11g. This works in the 4 GHz band (like 802.11b), but uses the same OFDM based transmission scheme as 802.11a. It operates at a maximum physical layer bit rate of 54 Mbit/s exclusive of forward error correction codes, or about 19 Mbit/s average throughput [citation needed]. 502.11g hardware is fully backwards compatible with 802.11b hardware and therefore is encumbered with legacy issues that reduce throughput when compared to 802.11a by -21%.

The then-proposed 802.11g standard was rapidly adopted by consumers starting in January 2003, well before ratification, due to the desire for higher data rates, and reductions in manufacturing costs. By summer 2003, most dual-band 802.11a/b products became dual-band/tri-mode, supporting a b/g in a single mobile adapter card or access point. Details of making b and g work well together occupied much of the lingering technical process; in an 802.11g network, however, activity by an 802.11b participant will reduce the data rate of the overall 802.11g network.

Like 802.11b, 802.11g devices suffer interference from other products operating in the 2.4 GHz band. Devices operating in the 2.4 GHz range include: microwave ovens, Bluetooth devices, baby monitors and cordless telephones.

13.7.5. IEEE 802.11n

802.11n is a proposed amendment which improves upon the previous 802.11 standards by adding multiple-input multiple-output (MIMO) and many other newer features. The TGN workgroup is not expected to finalize the amendment until December 2009 Enterprises, however, have already begun migrating to 802.11n networks based on Draft 2 of the 802.11n proposal. A common strategy for many businesses is to set up 802.11n networks to support existing 802.11b and 802.11g client devices and while gradually moving to 802.11n clients as part of new equipment purchases.

13.7.6 Channels and international compatibility

802.11 divides each of the above-described bands into channels, analogously to how radio and TV broadcast bands are carved up but with greater channel width and overlap. For example the 2.4000-2.4835 GHz band is divided into 13 channels each of width 22 MHz but spaced only 5 MHz apart, with channel 1 centered on 2412 MHz and 13 on 2472, to which Japan adds a 14th channel 12 MHz above channel 13.

Availability of channels is regulated by country, constrained in part by how each country allocates radio spectrum to various services. At one extreme Japan permits the use of all 14 channels (with the exclusion of 802.11 g/n from channel 14), while at the other Spain allowed only channels 10 and 11 (later all of the 14 channels have been allowed), and France that allowed only 10, 11, 12 and 13 (now channels 1 to 13 are allowed). Most other European countries are almost as liberal as Japan, disallowing only channel 14, while North America and some Central and South American countries further disallow 12 and 13.

Besides specifying the centre frequency of each channel, 802.11 also specifies (in Clause 17) a spectral mask defining the permitted distribution of power across each channel. The mask requires that the signal be attenuated by at least 30 dB from its peak energy at ± 11 MHz from the centre frequency, the sense in which channels are effectively 22 MHz wide. One consequence is that stations can only use every fourth or fifth channel without overlap, typically 1,6 and 11 in the Americas, 1,5,9 and 13 in Europe, etc. Another is that channels 1-13 effectively

require the band 2401-2483 MHz, the actual allocations being for example 2400-2483.5 in the UK, 2402-2483.5 in the US, etc.

Since the spectral mask only defines power output restrictions up to ± 22 MHz from the center frequency to be attenuated by 50 dB; it is often assumed that the energy of the channel extends no further than these limits. It is more correct to say that, given the separation between channels 1, 6, and 11, the signal on any channel should be sufficiently attenuated to minimally interfere with a transmitter on any other channel. Due to the near-far problem a transmitter can impact a receiver on a "non-overlapping" channel, but only if it is close to the victim receiver (within a meter) or operating above allowed power levels.

Although the statement that channels 1,6, and 11 are "non-overlapping" is limited to spacing or product density, the 1-6-11 guideline has merit. If transmitters are closer together than channels 1, 6, and 11 (for example, 1, 4, 7, and 10), overlap between the channels may cause unacceptable degradation of signal quality and throughput. However, overlapping channels may be used under certain circumstances. This way, more channels are available

13.7.7 Frames

Current 802.11 × standards define "frame" types for use in transmission of data as well as management and control of wireless links. Frames are divided into very specific and standardized sections. Each frame has a 2 byte frame control field that provides detailed information on the wireless link. This field is segmented 11 ways and will be presented in order, with the first two bits reserved for identification of the protocol being used (*e.g.* 802.11g, 802.11b, etc.). These respectively two and four bit fields are used for identification of which frame type is used. The next two segments are reserved for type and subtype. The next two bits are the To DS and From DS fields. They indicate whether a frame is headed for a distributed system. All frames will have one of these bits set. The More Fragmentation bit is set most notable when higher level packets have been partitioned and will be set for all non-final sections. Some management frames may require partitioning as well. Sometimes frames require retransmission, and for this there is a Retry bit which is set to one when a frame is resent. This aids in the elimination of duplicate frames station side. The Power Management bit indicates the power management state of the sender after the completion of a frame exchange. Access points are required to manage the connection and will never set the power saver bit. The More Data bit is used to buffer frames received in a distributed system. The access point uses this bit to facilitate stations in power saver mode. It indicates that at least one frame is available and addresses all stations connected. The WEP bit is modified after processing a frame. It is toggled to one after a frame has been decrypted or if no encryption is set it will have already been one. The last bit is the Order bit and is only set when the "strict ordering" delivery method is employed. Frames and fragments are not always sent in order as it causes a transmission performance

penalty. The next two bytes are reserved for the Duration ID field. This field takes on one of three forms, Duration, contention-free period (CFP), and PS-Poll. An 802.11 frame can contain up to four address fields. Six bytes are reserved for each address field. Each field is numbered are used for different purposes. Address 1 is the receiver, Address 2 is the transmitter, and Address 3 is used for filtering purposes by the receiver. As addresses are only 46 bits long and there are 48 bits reserved for each address, the first bit has a special function. A 0 indicates a single stations address (unicast), while a 1 represent a group of stations (multicast). If all the bits are 1 's then the frame is broadcast to all station connected to an access point. The Sequence Control field is a two byte section used for identifying message order as well as eliminating duplicate frames. The first 4 bits are used for the fragmentation number and the last 12 bits are the sequence number. The Frame Body field is variably size, from 0-2132 bytes, and contains information from higher layers. The Frame Check Sequence (FCS) is the last four bytes in the standard 802.11 frame. Often referred to as the Cyclic Redundancy Check (CRC), it allows for integrity check of retrieved frames. As frames are about to be sent the FCS is calculated and appending. When a station receives a frame it can calculate the FCS of the frame and compare it to the one received. If they match it is assumed that the frame was not distorted during transmission.

(*a*) **Management Frames** allow for the maintenance of communication. Some common 802.11 subtypes include:

(*i*) **Authentication frame** : 802.11 authentications begin with the WNIC sending an authentication frame to the access point containing its identity. With an open system authentication the WNIC only sends a single authentication frame and the access point responds with an authentication frame of its own indicating acceptance or rejection. With shared key authentication, after the WNIC sends its initial authentication request it will receive an authentication frame from the access point containing challenge text. The WNIC sends an authentication frame containing the encrypted version of the challenge text to the access point. The access point ensures the text was encrypted with the correct key by decrypting it with its own key. The result of this process determines the WNIC's authentication status.

(*ii*) **Association request frame** : Sent from a station it enables the access point to allocate resources and synchronize. The frame carries information about the WNIC including supported data rates and the SSID of the network the station wishes to associate with. If the request is accepted the access point reserve memory and establishes and association ID for the WNIC.

(*iii*) **Association response frame:** Sent from an access point to a station containing the acceptance or rejection to an association request. If it is

an acceptance the frame will contain information such an association ID and supported data rates.

(*iv*) **Beacon frame:** Sent periodically from an access point to announce its presence and provide the SSED, and other parameters for WNICs within range.

(*v*) **Deauthentication frame:** Sent from a station wishing to terminate connection from another station.

(*vi*) **Disassociation frame:** Sent from a station wishing to terminate connection. It's an elegant way to allow the access point to relinquish memory allocation and remove the WNIC from the association table.

(*vii*) **Probe request frame:** Sent from a station when it requires information from another station.

(v*iii*) **Probe response frame:** Sent from a station containing capability information, supported data rates, etc., after receiving a probe request frame.

(*ix*) **Reassociation request frame:** A WNIC sends a resuscitation request when it drops from range of the currently associated access point and finds another access point with a stronger signal. The new access point coordinates the forwarding of any information that may still be contained in the buffer of the previous access point.

(*x*) **Reassociation response frame:** Sent from an access point containing the acceptance or rejection to a WNIC

(*xi*) **Reassociation request frame:** The frame includes information required for association such as the association ID and supported data rates.

(*b*) **Control frames** facility in the exchange of data frames between stations. Some common 802.11 control frames include:

(*i*) **Acknowledgement (ACK) frame:** After receiving a data frame the receiving station will send an ACK frame to the sending station if no errors are found. If the sending station doesn't receive an ACK frame within a predetermined period of time the sending station will resend the frame.

(*ii*) **Request to Send (RTS) frame:** The RTS and CTS frames provide an optional collision reduction scheme for access point with hidden stations. A station sends a RTS frame to as the first step in a two-way handshake required before sending data frames.

(*iii*) **Clear to Send (CTS) frame:** A station responds to an RTS frame with a CTS frame. It provides clearance for the requesting station to send a data frame. The CTS provides collision control management by including a time value for which all other stations are to hold off transmission while the requesting stations transmits.

Data frames carry packets from web pages, files, etc. within the body.

13.7.8 Community networks

With the proliferation of cable modems and DSL, there is an ever-increasing market of people who wish to establish small networks in their homes to share their broadband Internet connection.

Many hotspot or free networks frequently allow anyone within range, including passersby outside, to connect to the Internet. There are also efforts by volunteer groups to establish wireless community networks to provide free wireless connectivity to the public.

13.7.9 Security

In 2001, a group from the University of California, Berkeley presented a paper describing weaknesses in the 802.11 Wired Equivalent Privacy (WEP) security mechanism defined in the original standard; they were followed by Fluhrer, Mantin, and Shamir's paper entitled "Weaknesses in the Key Scheduling Algorithm of RC4". Not long after, Adam Stubblefield and AT&T publicly announced the first verification of the attack. In the attack they were able to intercept transmissions and gain unauthorized access to wireless networks.

The IEEE set up a dedicated task group to create a replacement security solution, 802.11i (previously this work was handled as part of a broader 802. 11e effort to enhance the MAC layer). The Wi-Fi Alliance announced an interim specification called Wi-Fi Protected Access (WPA) based on a subset of the then current IEEE 802.11i draft. These started to appear in products in mid-2003. IEEE 802.11i itself was ratified in June 2004, and uses government strength encryption in the Advanced Encryption Standard AES, instead of RC4, which was used in WEP. The modern recommended encryption for the home/consumer space is WPA2 (AES PreShared Key) and for the Enterprise space is WPA2 along with a radius server the strongest is EAP-TLS.

In January 2005, IEEE set up yet another task group, TGw, to protect management and broadcast frames, which previously were sent unsecured.

13.8 CONNECTIONLESS PROTOCOL

In telecommunications, connectionless describes communication between two network end points in which a message can be sent from one end point to another without prior arrangement. The device at one end of the communication transmits data to the other, without first ensuring that the recipient is available and ready to receive the data. The device sending a message simply sends it addressed to the intended recipient. As such there are more frequent problems with transmission than with connection-oriented protocols and it may be necessary to resend the data several times. Connectionless protocols are often disfavoured by network administrators because it is much harder to filter malicious packets from a connectionless protocol using a firewall. The Internet Protocol (IP) and User Datagram Protocol (UDP) are connectionless protocols, but TCP/IP (the most common use of IP) is connection-oriented.

Connectionless protocols are usually described as Stateless because the endpoints have no protocol-defined way to remember where they are in a "conversation" of message exchanges. The alternative to the connectionless approach uses connection-oriented protocols, which are sometimes described as stateful because they can keep track of a conversation.

13.8.1 List of connectionless protocols

(*i*) IP (internet layer, can also be used for connections)
(*ii*) UDP
(*iii*) ICMP
(*iv*) IPX
(*v*) TIPC
(*vi*) NetBEUI

13.9 CONNECTION-ORIENTED PROTOCOL

A connection-oriented networking protocol is one that delivers a stream of data in the same order as it was sent, after first establishing a communication session. It may be a circuit switched connection, or a virtual circuit connection in a packet switched network. In the latter case, it identifies traffic flows by some connection identifier rather than by explicitly listing source and destination addresses. Typically, this connection identifier is a small integer (10 bits for Frame Relay, 24 for ATM, for example). This makes network switches substantially faster (as routing tables are just simple look-up tables, and are trivial to implement in hardware). The impact is so great, in fact, that even characteristically connectionless protocols, such as IP traffic, are being tagged with connection-oriented header prefixes (*e.g.,* as with MPLS, or IPv6's built-in Flow ID field). Example of a connection-oriented protocol at the transport layer is the TCP protocol.

Note that connection-oriented protocols are not necessarily reliable protocols. ATM and Frame Relay, for example, are both examples of a connection-oriented, unreliable protocol. There are also reliable connectionless protocols as well, such as x.25 when it passes data in I-frames. But this combination is rare, and reliable-connectionless is uncommon in commercial and academic networks.

Note that connection-oriented protocols handle real-time traffic substantially more efficiently than connectionless protocols, which is why ATM has yet to be replaced by Ethernet for carrying real-time, isochronous traffic streams, especially in heavily aggregated networks like backbones, where the motto "bandwidth is cheap" fails to deliver on its promise. Experience has also shown that over-provisioning bandwidth does not resolve all quality of service issues. Hence, gigabit Ethernet is not expected to replace ATM at this time.

Some connection-oriented protocols have been designed or altered to accommodate both connection-oriented and connectionless data.

13.9.1 List of Connection-oriented protocols

(*i*) TCP
(*ii*) DCCP
(*iii*) Connection-oriented Ethernet
(*iv*) Phone Call — user must dial the telephone, get an answer before transmitting data
(*v*) ATM
(*vi*) Frame Relay
(*vii*) TIPC
(*viii*) SCTP
(*ix*) IPX/SPX

13.10 INTERNET PROTOCOL

The Internet Protocol (IP) is a protocol used for communicating data across a packet-switched internetwork using the Internet Protocol Suite, also referred to as TCP/ IP.

IP is the primary protocol in the Internet Layer of the Internet Protocol Suite and has the task of delivering datagrams (packets) from the source host to the destination host solely based on their addresses. For this purpose the Internet Protocol defines addressing methods and structures for datagram encapsulation. The first major version of addressing structure, now referred to as Internet Protocol Version 4 (IPv4) is still the dominant protocol of the Internet, although the successor, Internet Protocol Version 6 (IPv6) is actively deployed worldwide.

The TCP/IP model

Application Layer

BGP • DHCP • DNS • FTP • Gopher • GTP • HTTP • EVIAP • IRC • NNTP • NTP • POP • RIP • RPC • RTCP • RTP • RTSP • SDP • SIP • SMTP • SNMP
• SOAP • SSH • STUN • Telnet • TIME • TLS/SSL • XMPP •

Transport Layer

TCP • UDP • DCCP • SCTP • RSVP • ECN •

Internet Layer

IP (IPv4, IPv6) • ICMP • ICMPv6 • IGMP • IPsec •

Link Layer

ARP • RARP • NDP • OSPF • Tunnels (L2TP) • Media Access Control (Ethernet, DSL, ISDN, FDDI) • Device Drivers •

13.10.1 IP Encapsulation

Data from an upper layer protocol is encapsulated as packets/datagrams (the terms are basically synonymous in IP). Circuit setup is not needed before a host may send packets to another host that it has previously not communicated with (a characteristic of packet-switched networks), thus IP is a connectionless protocol. This is in contrast to Public Switched Telephone Networks that require

the setup of a circuit before a phone call may go through (connection-oriented protocol).

13.10.2 Services Provided by IP

Because of the abstraction provided by encapsulation, DP can be used over a heterogeneous network, i.e., a network connecting computers may consist of a combination of Ethernet, ATM, FDDI, Wi-Fi, token ring, or others. Each link layer implementation may have its own method of addressing (or possibly the complete lack of it), with a corresponding need to resolve IP addresses to data link addresses. This address resolution is handled by the Address Resolution Protocol (ARP) for IPv4 and Neighbor Discovery Protocol (NDP) for IPv6.

13.10.3 Reliability

The design principles of the Internet protocols assume that the network infrastructure is inherently unreliable at any single network element or transmission medium and that it is dynamic in terms of availability of links and nodes. No central monitoring or performance measurement facility exists that tracks or maintains the state of the network. For the benefit of reducing network complexity, the intelligence in the network is purposely mostly located in the end nodes of each data transmission, end-to-end principle. Routers in the transmission path simply forward packets to next known local gateway matching the routing prefix for the destination address.

As a consequence of this design, the Internet Protocol only provides best effort delivery and its service can also be characterized as unreliable. In network architectural language it is a connection-less protocol, in contrast to so-called connection-oriented modes of transmission. The lack of reliability allows any of the following fault events to occur:

(*i*) data corruption

(*ii*) lost data packets

(*iii*) duplicate arrival

(*iv*) out-of-order packet delivery; meaning, if packet 'A' is sent before packet 'B', packet 'B' may arrive before packet 'A'. Since routing is dynamic and there is no memory in the network about the path of prior packets, it is possible that the first packet sent takes a longer path to its destination.

The only assistance that the Internet Protocol provides in Version 4 (IPv4) is to ensure that the IP packet header is error-free through computation of a checksum at the routing nodes. This has the side-effect of discarding packets with bad headers on the spot. In this case no notification is required to be sent to either end node, although a facility exists in the Internet Control Message Protocol (ICMP) to do so.

IPv6, on the other hand, has abandoned the use of IP header checksums for the benefit of rapid forwarding through routing elements in the network.

The resolution or correction of any of these reliability issues is the responsibility of an upper layer protocol. For example, to ensure in-order delivery the upper layer may have to cache data until it can be passed to the application.

In addition to issues of reliability, this dynamic nature and the diversity of the Internet and its components provide no guarantee that any particular path is actually capable of, or suitable for performing the data transmission requested, even if the path is available and reliable. One of the technical constraints is the size of data packets allowed on a given link. An application must assure that it uses proper transmission characteristics. Some of this responsibility lies also in the upper layer protocols between application and IP. Facilities exist to examine the maximum transmission unit (MTU) size of the local link, as well as for the entire projected path to the destination when using IPv6. The IPv4 internetworking layer has the capability to automatically fragment the original datagram into smaller units for transmission. In this case, IP does provide re-ordering of fragments delivered out-of-order.

Transmission Control Protocol (TCP) is an example of a protocol that will adjust its segment size to be smaller than the MTU. User Datagram Protocol (UDP) and Internet Control Message Protocol (ICMP) disregard MTU size thereby forcing IP to fragment oversized datagrams.

13.10.4 IP addressing and routing

Perhaps the most complex aspects of IP are IP addressing and routing. Addressing refers to how end hosts become assigned IP addresses and how subnetworks of IP host addresses are divided and grouped together. IP routing is performed by all hosts, but most importantly by internetwork routers, which typically use either interior gateway protocols (IGPs) or external gateway protocols (EGPs) to help make IP datagram forwarding decisions across IP connected networks

13.10.5 Version history

In May, 1974, the Institute of Electrical and Electronic Engineers (IEEE) published a paper entitled "A Protocol for Packet Network Interconnection." The paper's authors, Vint Cerf and Bob Kahn, described an internetworking protocol for sharing resources using packet-switching among the nodes. A central control component of this model was the "Transmission Control Program" (TCP) that incorporated both connection-oriented links and datagram services between hosts. The monolithic Transmission Control Program was later divided into a modular architecture consisting of the Transmission Control Protocol at the connection-oriented layer and the Internet Protocol at the internetworking (datagram) layer. The model became known informally as TCP/IP, although formally it was henceforth referenced as the Internet Protocol Suite.

The Internet Protocol is one of the determining elements that define the Internet. The dominant internetworking protocol (Internet Layer) in use today is IPv4; with number 4 assigned as the formal protocol version number carried in every IP datagram. IPv4 is described in RFC-791.

The successor to IPv4 is IPv6. Its most prominent modification from Version 4 is the addressing system. IPv4 uses 32-bit addresses (~4 billion, or ~4.3* 109,

addresses) while IPv6 uses 128-bit addresses (-340 undecillion, or-3.4* 1038 addresses). Although adoption of IPv6 has been slow, as of June 2008, all United States government systems have demonstrated basic infrastructure support for IPv6 (if only at the backbone level).

Version numbers 0 through 3 were development versions of IPv4 used between 1977 and 1979. Version number 5 was used by the Internet Stream Protocol (IST), an experimental stream protocol. Version numbers 6 through 9 were proposed for various protocol models designed to replace IPv4: SIPP (Simple Internet Protocol Plus, known now as IPv6), TP/IX (RFC 1475), PIP (RFC 1621) and TUBA (TCP and UDP with Bigger Addresses, RFC 1347). Version number 6 was eventually chosen as the official assignment for the successor Internet protocol, subsequently standardized as IPv6.

In 2004, a Chinese project called IPv9 was briefly mentioned in the press as a possible competitor to IPv6. The proposal had no affiliation with or support by any international standards body, and appears to have gained no traction even within China.

13.11 IP ADDRESS

An Internet Protocol (IP) address is a numerical identification (logical address) that is assigned to devices participating in a computer network utilizing the Internet Protocol for communication between its nodes. Although IP addresses are stored as binary numbers, they are usually displayed in human-readable notations, such as 208.77.188.166 (for IPv4), and 2001 : db8 : 0 : 1234 : 0 : 567 : 1 : 1 (for IPv6). The role of the IP address has been characterized as follows: "A name indicates what we seek. An address indicates where it is. A route indicates how to get there."

The original designers of TCP/IP defined an IP address as a 32-bit number and this system, now named Internet Protocol Version 4 (IPv4), is still in use today. However, due to the enormous growth of the Internet and the resulting depletion of the address space, a new addressing system (IPv6), using 128 bits for the address, was developed (RFC 1883).

The Internet Protocol also has the task of routing data packets between networks, and IP addresses specify the locations of the source and destination nodes in the topology of the routing system. For this purpose, some of the bits in an IP address are used to designate a subnetwork. (In CEDR notation, the number of bits used for the subnet follows the IP address. *e.g.* 208.77.188.166/24). An IP address can be private, for use on a LAN, or public, for use on the Internet or other WAN.

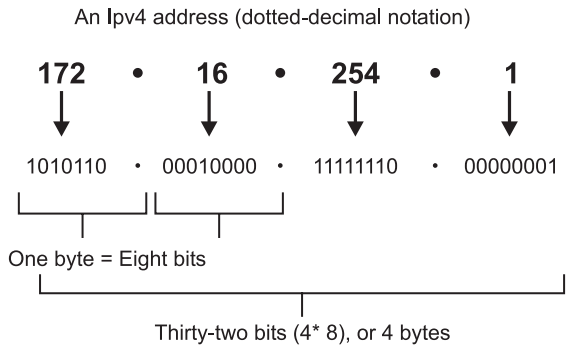
Early specifications intended IP addresses to each be uniquely assigned to a particular computer or device. However, it was found that this was not always necessary as private networks developed and address space needed to be conserved (IPv4 address exhaustion). RFC 1918 specifies private address spaces (also known as non-routable addresses) that may be reused by anyone; today, such private networks typically connect to the Internet through Network Address Translation (NAT). In addition, technologies such as anycast addressing

have been developed to allow multiple hosts at the same IP address but in different portions of the Internet to service requests by network clients.

The Internet Assigned Numbers Authority (IANA) manages the global IP address space. IANA works in cooperation with five Regional Internet Registries (RIRs) to allocate IP address blocks to Local Internet Registries (Internet service providers) and other entities.

13.11.1. IP Versions

The Internet Protocol (IP) has two versions currently in use. Each version has its own definition of an IP address. Because of its prevalence, "IP address" typically refers to those defined by IPv4.



An illustration of an IP address (version 4), in both dot-decimal notation and binary.

13.11.2 IP Version 4 Addresses

IPv4 uses 32-bit (4-byte) addresses, which limits the address space to 4,294,967,296 (232) possible unique addresses. However, IPv4 reserves some addresses for special purposes such as private networks (~ 18 million addresses) or multicast addresses (-270 million addresses). This reduces the number of addresses that can be allocated as public Internet addresses, and as the number of addresses available is consumed, an IPv4 address shortage appears to be inevitable in the long run. This limitation has helped stimulate the push towards IPv6, which is currently in the early stages of deployment and is currently the only offering to replace IPv4.

IPv4 addresses are usually represented in dot-decimal notation (four numbers, each ranging from 0 to 255, separated by dots, e.g. 208.77.188.166). Each part represents 8 bits of the address, and is therefore called an octet. It is possible, although less common, to write IPv4 addresses in binary or hexadecimal. When converting, each octet is treated as a separate number. (So 255.255.0.0 in dot-decimal would be FF.FF.00.00 in hexadecimal.)

13.11.3. IPv4 networks

In the early stages of development of the Internet protocol, network administrators interpreted an IP address as a structure of network number and host number. The highest order octet (most significant eight bits) were

designating the network number, and the rest of the bits were called the rest field or host identifier and used for host numbering within a network. This method soon proved inadequate as local area networks developed that were not part of the larger networks already designated by a network number. In 1981, the Internet addressing specification was revised with the introduction of classful network architecture.

Classful network design allowed for a larger number of individual assignments. The first three bits of the most significant octet of an IP address was defined as the "class" of the address, instead of just the network number and, depending on the class derived, the network designation was based on octet boundary segments of the entire address. The following table gives an overview of this system.

Table 13.6

Class	First octet in binary	Range of first octet	Network ID	Host ID	Possible number of networks	Possible number of hosts
A	oxxxxxxx	0-127	a	b.c.d	128 = (2^7)	16,777,214 = (2^{24}–2)
B	10XXXXXX	128-191	a.b	c.d	16,384 = (2^{14})	65,534 = (2^{16}–2)
C	110XXXXX	192-223	a.b.c	d	2,097,152 = (2^{21})	254 = (2^8–2)

Although classful network design was a successful developmental stage, it proved unscalable in the rapid expansion of the Internet and was abandoned in 1993 when Classless Inter-Domain Routing (CIDR) was introduced (RFC 1517, RFC 1518, RFC 1519) to define a new concept for the allocation of IP address blocks and new rules of routing protocol packets using IPv4 addresses. CIDR is based on variable-length subnet masking (VLSM) to allow allocation and routing on arbitrary-length prefixes.

13.11.4 IPv4 Private Addresses

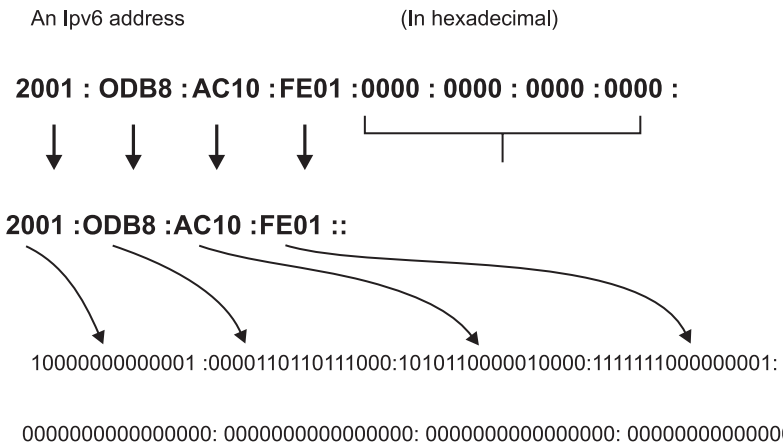
Computers not connected to the Internet, such as factory machines that communicate only with each other via TCP/IP, need not have globally-unique IP addresses. Three ranges of IPv4 addresses for private networks, one range for each class (A, B, C), were reserved in RFC 1918. These addresses are not routed on the Internet, and thus need not be coordinated with an IP address registry.

Table 13.7

IANA Reserved Private Network Ranges	Start of range	End of range	Total addresses
24-bit Block (/8 prefix, 1 × A)	10.0.0.0	10.255.255.255	16,777,216
20-bit Block (/12prefix, 16 × B)	172.16.0.0	172.31.255.255	1,048,576
16-bit Block (/16 prefix, 256 × Q	192.168.0.0	192.168.255.255	65,536

An user may use block. Typical, a network adminstrator will divide a blcok' into subbets; for example, many home routers automatically use a default address range of 192.168.0.0-192.168.0.255(192.168.0.0/24)

13.11.5 IP Version 6 Address



An illustration of an IP address (version 6), in hexadecimal and binary.

The rapid exhaustion of IPv4 address space, despite conservation techniques, prompted the Internet Engineering Task Force (IETF) to explore new technologies to expand the Internet's addressing capability. The permanent solution was deemed to be a redesign of the Internet Protocol itself. This next generation of the Internet Protocol, limed to replace IPv4 on the Internet, was eventually named Internet Protocol Version 6 (IPv6) The address size was increased from 32 to 128 bits (16 bytes), which, even with a generous assignment of network blocks, is deemed sufficient for the foreseeable future. Mathematically, the new address space provides the potential for a maximum of 2128, or about 3.403 x 1038 unique addresses.

The new design is not based on the goal to provide a sufficient quantity of addresses done, but rather to allow efficient aggregation of subnet routing prefixes to occur at outing nodes. As a result, routing table sizes are smaller, and the smallest possible individual allocation is a subnet for 264 hosts, which is the size of the square of the size of the entire IPv4 Internet. At these levels, actual address utilization rates will be small an any IPv6 network segment. The new design also provides the opportunity to separate the addressing infrastructure of a network segment-that is the local administration of the segment's available space from the addressing prefix used to route external traffic for a network. IPv6 has facilities that automatically change the routing prefix of entire networks should the global connectivity or the routing policy change without requiring internal redesign or renumbering.

The large number of IPv6 addresses allows large blocks to be assigned for specific purposes and, where appropriate, to be aggregated for efficient routing. With a large address space, there is not the need to have complex address conservation methods as used in classless inter-domain routing (CIDR).

Windows Vista, Apple Computer's Mac OS X, all modern Linux distributions, and an increasing range of other operating systems include native support for the protocol, but it is not yet widely deployed in other devices.

Example of an IPv6 address:

2001:0db8:85a3:08d3:1319:8a2e:0370:7334

13.11.6 IPv6 Private Addresses

Just as IPv4 reserves addresses for private or internal networks, there are blocks of addresses set aside in IPv6 for private addresses. In IPv6, these are referred to as unique local addresses (ULA). RFC 4193 sets aside the routing prefix fc00::/7 for this block which is divided into two /8 blocks with different implied policies (cf. IPv6). The addresses include a 40-bit pseudorandom number that minimizes the risk of address collisions if sites merge or packets are misrouted.

Early designs (RFC 3513) used a different block for this purpose (fecO::), dubbed site-local addresses. However, the definition of what constituted sites remained unclear and the poorly defined addressing policy created ambiguities for routing. The address range specification was abandoned and must no longer be used in new systems.

Addresses starting with fe80: — called link-local addresses are assigned only in the local link area. The addresses are generated usually automatically by the operating system's IP layer for each network interface. This provides instant automatic network connectivity for any IPv6 host and means that if several hosts connect to a common hub or switch, they have an instant communication path via their link-local IPv6 address. This feature is used extensively, and invisibly to most users, in the lower layers of IPv6 network administration. None of the private address prefixes may be routed in the public Internet.

13.11.7 IP Subnetworks

The technique of subnetting can operate in both IPv4 and IPv6 networks. The IP address is divided into two parts: the network address and the host identifier. The subnet mask (in IPv4 only) or the CIDR prefix determine how the IP address is divided into network and host parts.

The term subnet mask is only used within IPv4. Both IP versions however use the Classless Inter-Domain Routing (CIDR) concept and notation. In this, the IP address is followed by a slash and the number (in decimal) of bits used for the network part, also called the routing prefix. For example, an IPv4 address and its subnet mask may be 192.0.2.1 and 255.255.255.0, respectively. The CIDR notation for the same IP address and subnet is 192.0.2.1/24, because the first 24 bits of the IP address indicate the network and subnet.

13.11.8 Static and Dynamic IP Addresses

When a computer is configured to use the same IP address each time it powers up, this is known as a Static IP address. In contrast, in situations when the computer's IP address is assigned automatically, it is known as a Dynamic IP address.

Static IP addresses are manually assigned to a computer by an administrator. The exact procedure varies according to platform. This contrasts with dynamic IP addresses, which are assigned either randomly or assigned by a server

using Dynamic Host Configuration Protocol (DHCP). Even though IP addresses assigned using DHCP may stay the same for long periods of time, they can generally change. In some cases, a network administrator may implement dynamically assigned static IP addresses. In this case, a DHCP server is used, but it is specifically configured to always assign the same IP address to a particular computer, and never to assign that IP address to another computer. This allows static IP addresses to be configured in one place, without having to specifically configure each computer on the network in a different way.

In the absence of both an administrator (to assign a static IP address) and a DHCP server, the operating system may assign itself an IP address using stateless autoconfiguration methods, such as Zeroconf. These IP addresses, known as link-local addresses, default to the 169.254.0.0/16 address range in IPv4.

In IPv6, every interface, whether using static or dynamic address assignments, also receives a local-link address automatically in the fe80: :/64 subnet.

(*i*) **Uses of dynamic addressing:** Dynamic IP addresses are most frequently assigned on LANs and broadband networks by Dynamic Host Configuration Protocol (DHCP) servers. They are used because it avoids the administrative burden of assigning specific static addresses to each device on a network. It also allows many devices to share limited address space on a network if only some of them will be online at a particular time. In most current desktop operating systems, dynamic IP configuration is enabled by default so that a user does not need to manually enter any settings to connect to a network with a DHCP server. DHCP is not the only technology used to assigning dynamic IP addresses. Dialup and some broadband networks use dynamic address features of the Point-to-Point Protocol.

(*ii*) **Sticky dynamic IP address:** A sticky dynamic IP address or sticky IP is a term created by cable and DSL users to describe a dynamically assigned IP address that does not change often. This is however an informal term, as a sticky EP does not differ in any way from other dynamic IP address.

Even though IP addresses may not change often for cable or DSL users, the addresses are still controlled by the standard DHCP process. Since the modems are often online for extended periods of time, the leases on the IP addresses are commonly renewed, and therefore may not change.

Should the modem be turned off, a new IP address may be assigned when the modem is turned back on, as a different host on the network may have been assigned the old IP address. IP address changes may also be triggered by resetting the DHCP server configuration; therefore causing the modem to receive a new IP address.

Some users take advantage of stickiness to create websites on the cheap, since a static IP address is usually more expensive. And some ISPs counter this by forcibly re-assigning IP addresses every 24 hours or so.

(*iii*) **Uses of static addressing:** Some infrastructure situations have to use static addressing, such as when finding the Domain Name System host that will translate domain names to IP addresses. Static addresses are also convenient, but not absolutely necessary, to locate servers inside an enterprise. An address obtained from a DNS server comes with a time to live, or caching time, after which it should be looked up to confirm that it has not changed. Even static IP addresses do change as a result of network administration (RFC 2072).

13.11.9 Modifications to IP Addressing

13.11.9.1 *IP Blocking and Firewalls*

Firewalls are common on today's Internet. For increased network security, they control access to private networks based on the public IP of the client. Whether using a blacklist or a whitelist, the IP address that is blocked is the perceived public IP address of the client, meaning that if the client is using a proxy server or NAT, blocking one IP address might block many individual people.

13.11.9.2 *IP Address Translation*

Multiple client devices can appear to share IP addresses: either because they are part of a shared hosting web server environment or because an IPv4 network address translator (NAT) or proxy server acts as an intermediary agent on behalf of its customers, in which case the real originating IP addresses might be hidden from the server receiving a request. A common practice is to have a NAT hide a large number of IP addresses in a private network. Only the "outside" interface of the NAT need to have Internet-routable addresses.

Most commonly, the NAT device maps TCP or UDP port numbers on the outside to individual private addresses on the inside. Just as a telephone number may have site-specific extensions, the port numbers are site-specific extensions to an IP address.

In small home networks, NAT functions usually take place in a residential gateway device, typically one marketed as a "router". In this scenario, the computers connected to the router would have 'private' IP addresses and the router would have a 'public' address to communicate with the Internet. This type of router allows several computers to share one public IP address.

13.12 OSI MODEL

13.12.1 Introduction

The Open Systems Interconnection Basic Reference Model (OSI Reference Model or OSI Model) is an abstract description for layered communications and computer network protocol design. It was developed as part of the Open Systems Interconnection (OSI) initiative. In its most basic form, it divides network architecture into seven layers which, from top to bottom, are the Application, Presentation, Session, Transport, Network, Data-Link, and Physical Layers. It is therefore often referred to as the OSI Seven Layer Model.

A layer is a collection of conceptually similar functions that provide services to the layer above it and receives service from the layer below it. For example, a layer that provides error-free communications across a network provides the path needed by applications above it, while it calls the next lower layer to send and receive packets that make up the contents of the path.

Even though it has been largely superseded by newer IETF, IEEE, and indeed OSI protocol developments (subsequent to the publication of the original architectural standards), the basic OSI model is considered an excellent place to begin the study of network architecture. Not understanding that the pure seven-layer model is more historic than current, many beginners make the mistake of trying to fit every protocol under study into one of the seven basic layers. Especially the attempts of cross-layer optimization break the boundaries of the original layer scheme. Describing the actual layer concept with implemented systems is not always easy to do as most of the protocols in use on the Internet were designed as part of the TCP/IP model, and may not fit cleanly into the OSI Model.

13.12.2 History

In 1977, work on a layered model of network architecture was started, and the International Organization for Standardization (ISO) began to develop its OSI framework architecture (The ISO is a worldwide federation of national standards bodies from some 130 countries, one from each country). OSI has two major components: an abstract model of networking, called the Basic Reference Model or seven-layer model, and a set of specific protocols.

Note: The standard documents that describe the OSI model can be freely downloaded from the ITU-T as the X.200-series of recommendations. [2] A number of the protocol specifications are also available as part of the ITU-T X series. The equivalent ISO and ISO/IEC standards for the OSI model are available from the ISO, but only some of the ISO/IEC standards are available as cost-free downloads.

All aspects of OSI design evolved from experiences with the CYCLADES network, which also influenced Internet design. The new design was documented in ISO 7498 and its various addenda. In this model, a networking system is divided into layers. Within each layer, one or more entities implement its functionality. Each entity interacts directly only with the layer immediately beneath it, and provides facilities for use by the layer above it.

Protocols enable an entity in one host to interact with a corresponding entity at the same layer in another host. Service definitions abstractly describe the functionality provided to an (N)-layer by an (N-l) layer, where N is one of the seven layers of protocols operating in the local host.

13.12.3 Description of OSI Layers

13.12.3.1 *Layer 7: Application Layer*

The application layer is the OSI layer closest to the end user, which means that both the OSI application layer and the user interact directly with the software

application. This layer interacts with software applications that implement a communicating component. Such application programs fall outside the scope of the OSI model. Application layer functions typically include identifying communication partners, determining resource availability, and synchronizing communication. When identifying communication partners, the application layer determines the identity and availability of communication partners for an application with data to transmit. When determining resource availability, the application layer must decide whether sufficient network resources for the requested communication exist. In synchronizing communication, all communication between applications requires cooperation that is managed by the application layer. Some examples of application layer implementations include Telnet, File Transfer Protocol (FTP), and Simple Mail Transfer Protocol (SMTP).

13.12.3.2 *Layer 6: Presentation Layer*

The Presentation Layer establishes a context between Application Layer entities, in which the higher-layer entities can use different syntax and semantics, as long as the Presentation Service understands both and the mapping between them. The presentation service data units are then encapsulated into Session Protocol Data Units, and moved down the stack.

This layer provides independence from differences in data representation (*e.g.*, encryption) by translating from application to network format, and vice versa. The presentation layer works to transform data into the form that the application layer can accept. This layer formats and encrypts data to be sent across a network, providing freedom from compatibility problems. It is sometimes called the syntax layer.

The original presentation structure used the Basic Encoding Rules of Abstract Syntax Notation One (ASN. 1), with capabilities such as converting an EBCDIC-coded text file to an ASCII-coded file, or serializing objects and other data structures into and out of XML. ASN. 1 has a set of cryptographic encoding rules that allows end-to-end encryption between application entities.

13.12.3.3 *Layer* 5: *Session Layer*

The Session Layer controls the dialogues/connections (sessions) between computers. It establishes, manages and terminates the connections between the local and remote application. It provides for full-duplex, half-duplex, or simplex operation, and establishes checkpointing, adjournment, termination, and restart procedures. The OSI model made this layer responsible for "graceful close" of sessions, which is a property of TCP, and also for session checkpointing and recovery, which is not usually used in the Internet Protocol Suite. The Session Layer is commonly implemented explicitly in application environments that use remote procedure calls (RPCs).

13.12.3.4 *Layer* 4: *Transport Layer*

The Transport Layer provides transparent transfer of data between end users, providing reliable data transfer Services to the upper layers. The Transport Layer controls the reliability of a given link through flow control, segmentation/

desegmentation, and error control. Some protocols are state and connection oriented. This means that the Transport Layer can keep track of the segments and retransmit those that fail.

Although not developed under the OSI Reference Model and not strictly conforming to the OSI definition of the Transport Layer, the best known examples of a Layer 4 protocol are the Transmission Control Protocol (TCP) and User Datagram Protocol (UDP).

Of the actual OSI protocols, there are five classes of transport protocols ranging from class 0 (which is also known as TPO and provides the least error recovery) to class 4 (which is also known as TP4 and is designed for less reliable networks, similar to the Internet). Class 0 contains no error recovery, and was designed for use on network layers that provide error-free connections. Class 4 is closest to TCP, although TCP contains functions, such as the graceful close, which OSI assigns to the Session Layer.

Perhaps an easy way to visualize the Transport Layer is to compare it with a Post Office, which deals with the dispatch and classification of mail and parcels sent. Do remember, however, that a post office manages the outer envelope of mail. Higher layers may have the equivalent of double envelopes, such as cryptographic presentation services that can be read by the addressee only. Roughly speaking, tunneling protocols operate at the Transport Layer, such as carrying non-IP protocols such as IBM's SNA or Novell's IPX over an IP network, or end-to-end encryption with IPsec. While Generic Routing Encapsulation (GRE) might seem to be a Network Layer protocol, if the encapsulation of the payload takes place only at endpoint, GRE becomes closer to a transport protocol that uses IP headers but contains complete frames or packets to deliver to an endpoint. L2TP carries PPP frames inside transport packet.

13.12.3.5 *Layer* 3: *Network Layer*

The Network Layer provides the functional and procedural means of transferring variable length data sequences from a source to a destination via one or more networks, while maintaining the quality of service requested by the Transport Layer. The Network Layer performs network routing functions, and might also perform fragmentation and reassembly, and report delivery errors. Routers operate at this layer and sending data throughout the extended network and making the Internet possible. This is a logical addressing scheme in which values are chosen by the network engineer. The addressing scheme is hierarchical.

The best-known example of a Layer 3 protocol is the Internet Protocol (IP). It manages the connectionless transfer of data one hop at a time, from end system to ingress router, router to router, and from egress router to destination end system. It is not responsible for reliable delivery to a next hop, but only for the detection of errored packets so they may be discarded. When the medium of the next hop cannot accept a packet in its current length, IP is responsible for fragmenting into sufficiently small packets that the medium can accept it.

A number of layer management protocols, a function defined in the Management Annex, ISO 7498/4, belong to the Network Layer. These include routing protocols, multicast group management, Network Layer information and error, and Network Layer address assignment. It is the function of the payload that makes these belong to the Network Layer, not the protocol that carries them.

13.12.3.6 ***Layer* 2: *Data Link Layer***

The Data Link Layer provides the functional and procedural means to transfer data between network entities and to detect and possibly correct errors that may occur in the Physical Layer. Originally, this layer was intended for point-to-point and point-to-multipoint media, characteristic of wide area media in the telephone system. Local area network architecture, which included broadcast-capable multiaccess media, was developed independently of the ISO work, in IEEE Project 802. IEEE work assumed sublayering and management functions not required for WAN use. In modern practice, only error detection, not flow control using sliding window, is present in modern data link protocols such as Point-to-Point Protocol (PPP), and, on local area networks, the IEEE 802.2 LLC layer is not used for most protocols on Ethernet, and, on other local area networks, its flow control and acknowledgment mechanisms are rarely used. Sliding window flow control and acknowledgment is used at the Transport Layer by protocols such as TCP, but is still used in niches where X.25 offers performance advantages.

Both WAN and LAN services arrange bits, from the Physical Layer, into logical sequences called frames. Not all Physical Layer bits necessarily go into frames, as some of these bits are purely intended for Physical Layer functions. For example, every fifth bit of the FDDI bit stream is not used by the Data Link Layer.

(*i*) **WAN Protocol architecture:** Connection-oriented WAN data link protocols, in addition to framing, detect and may correct errors. They also are capable of controlling the rate of transmission. A WAN Data Link Layer might implement a sliding window flow control and acknowledgment mechanism to provide reliable delivery of frames; that is the case for SDLC and HDLC, and derivatives of HDLC such as LAPB and LAPD.

(*ii*) **IEEE 802 LAN architecture:** Practical, connectionless LANs began with the pre-IEEE Ethernet specification, which is the ancestor of IEEE 802.3. This layer manages the interaction of devices with a shared medium, which is the function of a Media Access Control sublayer. Above this MAC sublayer is the media-independent IEEE 802.2 Logical Link Control (LLC) sublayer, which deals with addressing and multiplexing on multiaccess media.

While IEEE 802.3 is the dominant wired LAN protocol and IEEE 802.11 the wireless LAN protocol, obsolescent MAC layers include Token Ring and FDDI. The MAC sublayer detects but does not correct errors.

13.12.3.7 Layer 1: Physical Layer

The Physical Layer defines the electrical and physical specifications for devices. In particular, it defines the relationship between a device and a physical medium. This includes the layout of pins, voltages, cable specifications, Hubs, repeaters, network adapters, Host Bus Adapters (HBAs used in Storage Area Networks) and more.

To understand the function of the Physical Layer in contrast to the functions of the Data Link Layer, think of the Physical Layer as concerned primarily with the interaction of a single device with a medium, where the Data Link Layer is concerned more with the interactions of multiple devices (i.e., at least two) with a shared medium. The Physical Layer will tell one device how to transmit to the medium, and another device how to receive from it (in most cases it does not tell the device how to connect to the medium). Obsolescent Physical Layer standards such as RS-232 do use physical wires to control access to the medium.

The major functions and services performed by the Physical Layer are:

(*i*) Establishment and termination of a connection to a communications medium.

(*ii*) Participation in the process whereby the communication resources are effectively shared among multiple users. For example, contention resolution and flow control.

(*iii*) Modulation, or conversion between the representation of digital data in user equipment and the corresponding signals transmitted over a communications channel. These are signals operating over the physical cabling (such as copper and optical fiber) or over a radio link.

Parallel SCSI buses operate in this layer, although it must be remembered that the logical SCSI protocol is a Transport Layer protocol that runs over this bus. Various Physical Layer Ethernet standards are also in this layer; Ethernet incorporates both this layer and the Data Link Layer. The same applies to other local-area networks, such as Token ring, FDDI, and IEEE 802.11, as well as personal area networks such as Bluetooth and IEEE 802.15.4.

13.12.4 Interfaces

Neither the OSI Reference Model nor OSI protocols specify any programming interfaces, other than as deliberately abstract service specifications. Protocol specifications precisely define the interfaces between different computers, but the software interfaces inside computers are implementation-specific.

For example, Microsoft Windows' Winsock, and Unix's Berkeley sockets and System V Transport Layer Interface, are interfaces between applications (Layer 5 and above) and the transport (Layer 4). NDIS and ODI are interfaces between the media (Layer 2) and the network protocol (Layer 3).

Interface standards, except for the Physical Layer to media, are approximate implementations of OSI Service Specifications

13.12.5 Comparison with TCP/IP

In the TCP/IP model of the Internet, protocols are deliberately not as rigidly designed into strict layers as the OSI model. RFC 3439 contains a section entitled "Layering considered harmful." However, TCP/IP does recognize four broad layers of functionality which are derived from the operating scope of their contained protocols, namely the scope of the software application, the end-to-end transport connection, the internetworking range, and lastly the scope of the direct links to other nodes on the local network.

Even though the concept is different than in OSI, these layers are nevertheless often compared with the OSI layering scheme in the following way: The Internet Application Layer includes the OSI Application Layer, Presentation Layer, and most of the Session Layer. Its end-to-end Transport Layer includes the graceful close function of the OSI Session Layer as well as the OSI Transport Layer. The internetworking layer (Internet Layer) is a subset of the OSI Network Layer, while the Link Layer includes the OSI Data Link and Physical Layers, as well as parts of OSI's Network Layer. These comparisons are based on the original seven-layer protocol model as defined in ISO 7498, rather than refinements in such things as the internal organization of the Network Layer document.

The presumably strict consumer/producer layering of OSI as it is usually described does not present contradictions in TCP/IP, as it is permissible that protocol usage does not follow the hierarchy implied in a layered model. Such examples exist in some routing protocols (e.g., OSPF), or in the description of tunneling protocols, which provide a Link Layer for an application, although the tunnel host protocol may well be a Transport or even an Application Layer protocol in its own right.

The TCP/IP design generally favours decisions based on simplicity, efficiency and ease of implementation.

13.13 TRANSMISSION CONTROL PROTOCOL

13.13.1 Introduction

The Transmission Control Protocol (TCP) is one of the core protocols of the Internet Protocol Suite. TCP is so central that the entire suite is often referred to as "TCP/IP." Whereas IP handles lower-level transmissions from computer to computer as a message makes its way across the Internet, TCP operates at a higher level, concerned only with the two end systems, for example a Web browser and a Web server. In particular, TCP provides reliable, ordered delivery of a stream of bytes from one program on one computer to another program on another computer. Besides the Web, other common applications of TCP include e-mail and file transfer. Among its management tasks, TCP controls message size, the rate at which messages are exchanged, and network traffic congestion.

13.13.2 Significance of TCP

TCP provides a communication service at an intermediate level between an application program and the Internet Protocol (IP). That is, when an application

program desires to send a large chunk of data across the Internet using IP, instead of breaking the data into IP-sized pieces and issuing a series of IP requests, the software can issue a single request to TCP and let TCP handle the IP details.

IP works by exchanging pieces of information called packets. A packet is a sequence of bytes and consists of a header followed by a body. The header describes the packet's destination and, optionally, the routers to use for forwarding and generally in the right direction until it arrives at its final destination. The body contains the data which IP is transmitting. When IP is transmitting data on behalf of TCP, the content of the IP packet body is TCP payload.

Due to network congestion, traffic load balancing, or other unpredictable network behaviour, IP packets can be lost or delivered out of order. TCP detects these problems, requests retransmission of lost packets, rearranges out-of-order packets, and even helps minimize network congestion to reduce the occurrence of the other problems. Once the TCP receiver has finally reassembled a perfect copy of the data originally transmitted, it passes that datagram to the application program. Thus, TCP abstracts the application's communication from the underlying networking details.

13.13.3 Applicability

TCP is used extensively by many of the Internet's most popular application protocols and resulting applications, including the World Wide Web, E-mail, File Transfer Protocol, Secure Shell, and some streaming media applications.

However, because TCP is optimized for accurate delivery rather than timely delivery, TCP sometimes incurs relatively long delays (in the order of seconds) while waiting for out-of-order messages or retransmissions of lost messages, and it is not particularly suitable for real-time applications such as Voice over IP. For such applications, protocols like the Real-time Transport Protocol (RTP) running over the User Datagram Protocol (UDP) are usually recommended instead.

TCP header									
Bit offset	Bits 4–7 0–3		8–15						16–31
0	Source port								Destination port
32	Sequence number								
64	Acknowledge number								
96	Data offset	Reserved	CWR ECE URC	ACK	PSH	RST	SYN	FIN	Window size
128	Checksum								
160	Optional (Optional urgent pointer)								
160/192+	Data								

Fig. 13.4

TCP is a reliable stream delivery service that guarantees delivery of a data stream sent from one host to another without duplication or losing data. Since packet transfer is not reliable, a technique known as positive acknowledgment with retransmission is used to guarantee reliability of packet transfers. This fundamental technique requires the receiver to respond with an acknowledgment message as it receives the data. The sender keeps a record of each packet it sends, and waits for acknowledgment before sending the next packet. The sender also keeps a timer from when the packet was sent, and retransmits a packet if the timer expires. The timer is needed in case a packet gets lost or corrupt.

TCP (Transmission Control Protocol) consists of a set of rules: for the protocol, that are used with the Internet Protocol, and for the IP, to send data "in a form of message units" between computers over the Internet. At the same time that the IP takes care of handling the actual delivery of the data, the TCP takes care of keeping track of the individual units of data "packets" (or more accurately, "segments") that a message is divided into for efficient routing through the net. For example, when a HTML file is sent to you from a Web server, the TCP program layer of that server takes the file as a stream of bytes and divides it into segments, numbers the segments, and then forwards them individually to the IP program layer. The IP program layer then turns each TCP segment into an IP packet by adding a header which includes (among other things) the destination IP address. Even though every packet has the same destination IP address, they can get routed differently through the network. When the client program in your computer gets them, the TCP stack (implementation) reassembles the individual segments and ensures they are correctly ordered as it streams them to an application.

13.13.4 TCP segment structure

A TCP segment consists of two sections:

(*i*) header

(*ii*) data

The TCP header consists of 11 fields, of which only 10 are required. The eleventh field is optional (pink background in table) and aptly named "options".

(*i*) Source port (16 bits) – identifies the sending port

(*ii*) Destination port (16 bits) – identifies the receiving port

(*iii*) Sequence number (32 bits) – has a dual role

- If the SYN flag is set, then this is the initial sequence number and the sequence number of the first data byte is this sequence number plus 1.
- If the SYN flag is not set, then this is the sequence number of the first data byte.

(*iv*) Acknowledgement number (32 bits) - if the ACK flag is set then the value of this field is the next expected byte that the receiver is expecting.

(*v*) Data offset (4 bits) - specifies the size of the TCP header in 32-bit words. The minimum size header is 5 words and the maximum is 15 words thus giving the minimum size of 20 bytes and maximum of 60 bytes. This field gets its name from the fact that it is also the offset from the start of the TCP packet to the data.

(*vi*) Reserved (4 bits) - for future use and should be set to zero

(*vii*) Flags (8 bits) (aka Control bits) - contains 81-bit flags

- CWR (1 bit) - Congestion Window Reduced (CWR) flag is set by the sending host to indicate that it received a TCP segment with the ECE flag set (added to header by RFC 3168).
- ECE (ECN-Echo) (1 bit) - indicate that the TCP peer is ECN capable during 3-way handshake (added to header by RFC 3168).
- URG (1 bit) - indicates that the Urgent pointer field is significant
- ACK (1 bit) - indicates that the Acknowledgment field is significant
- PSH (1 bit) - Push function
- RST (1 bit) - Reset the connection
- SYN (1 bit) - Synchronize sequence numbers
- FIN (1 bit) - No more data from sender

(*viii*) Window (16 bits) - the size of the receive window, which specifies the number of bytes (beyond the sequence number in the acknowledgment field) that the receiver is currently willing to receive.

(*ix*) Checksum (16 bits) - The 16-bit checksum field is used for error-checking of the header and data.

(*x*) Urgent pointer (16 bits) - if the URG flag is set, then this 16-bit field is an offset from the sequence number indicating the last urgent data byte

(*xi*) Options (Variable bits) - the total length of the option field must be a multiple of a 32-bit word and the data offset field adjusted appropriately

- 0 - End of options list
- 1 - No operation (NOP, Padding)
- 2 - Maximum segment size
- 3 - Window scale
- 4 - Selective Acknowledgement ok
- 5 - NIL
- 6 - NIL
- 7 - NIL
- 8 - Timestamp

The last flield is not a part of the header. The contents of this field are whatever the upper layer protocol wants but this protocol is not set in the header and is presumed based on the port selection.

- Data (Variable bits): As you might expect, this is the payload, or data portion of a TCP packet. The payload may be any number of

application layer protocols. The most common are HTTP, Telnet, SSH, FTP, but other popular protocols also use TCP. Other side this protocol is popular for the other protocols.

(a) Connection establishment

To establish a connection, TCP uses a three-way handshake. Before a client attempts to connect with a server, the server must first bind to a port to open it up for connections: this is called a passive open. Once the passive open is established, a client may initiate an active open. To establish a connection, the three-way (or 3-step) handshake occurs:

(*i*) The active open is performed by the client sending a SYN to the server.

(*ii*) In response, the server replies with a SYN-ACK.

(*ii*) Finally the client sends an ACK back to the server.

At this point, both the client and server have received an acknowledgment of the connection.

Example:

(*i*) The initiating host (client) sends a synchronization packet (SYN flag set to 1) to initiate a connection. It sets the packet's sequence number to a random value x.

(*ii*) The other host receives the packet, records the sequence number x from the client, and replies with an acknowledgment and synchronization (SYN-ACK). The Acknowledgment is a 32-bit field in TCP segment header. It contains the next sequence number that this host is expecting to receive (x + 1). The host also initiates a return session. This includes a TCP segment with its own initial Sequence Number of value y.

(*iii*) The initiating host responds with the next Sequence Number (x + 1) and a simple Acknowledgment Number value of y + 1, which is the Sequence Number value of the other host + **1**.

(b) Data Transfer

There are a few key features that set TCP apart from User Datagram Protocol:

(*i*) Ordered data transfer - the destination host rearranges according to sequence number

(*ii*) Retransmission of lost packets - any cumulative stream not acknowledged will be retransmitted

(*iii*) Discarding duplicate packets

(*iv*) Error-free data transfer

(*v*) Flow control - limits the rate a sender transfers data to guarantee reliable delivery. When the receiving host's buffer fills, then next acknowledgement contains a 0 in the window size, to stop transfer and allow the data in the buffer to be processed.

(*vi*) Congestion control - sliding window

(c) Ordered data transfer, retransmission of lost packets and discarding duplicate packets

TCP uses a sequence number to identify each byte of data. The sequence number identifies the order of the bytes sent from each computer so that the data can be transferred reliably and in order, regardless of any fragmentation, disordering, or packet loss that occurs during transmission. For every byte transmitted the sequence number must be incremented. In the first two steps of the 3-way handshake, both computers exchange an initial sequence number (ISN). This number can be arbitrary, and should in fact be unpredictable, in order to avoid a TCP Sequence Prediction Attack.

TCP primarily uses a cumulative acknowledgment scheme, where the receiver sends an acknowledgment signifying that the receiver has received all data preceding the acknowledged sequence number. Essentially, the first data byte in a segment is assigned a sequence number, which is inserted in the sequence number field, and the receiver sends an acknowledgment specifying the sequence number of the next byte they expect to receive. For example, if computer A sends 4 bytes with a sequence number of 100 (conceptually, the four bytes would have a sequence number of 100,101, 102, & 103 assigned) then the receiver would send back an acknowledgment of 104 since that is the next byte it expects to receive in the next packet. By sending an acknowledgment of 104, the receiver is signalling that it received bytes 100, 101, 102, & 103 correctly. If, by some chance, the last two bytes were corrupted then an acknowledgment value of 102 would be sent since 100 & 101 were received successfully.

In addition to cumulative acknowledgments, TCP receivers can also send selective acknowledgments to provide further information.

If the sender infers that data has been lost in the network, it retransmits the data.

(d) Error-free Data Transfer

Sequence numbers and acknowledgments cover discarding duplicate packets, retransmission of lost packets, and ordered-data transfer. To assure correctness a checksum field is included.

The TCP checksum is a quite weak check by modern standards. Data Link Layers with high bit error rates may require additional link error correction/ detection capabilities. If TCP were to be redesigned today, it would most probably have a 32-bit cyclic redundancy check specified as an error check instead of the current checksum. The weak checksum is partially compensated for by the common use of a CRC or better integrity check at layer 2, below both TCP and IP, such as is used in PPP or the Ethernet frame. However, this does not mean that the 16-bit TCP checksum is redundant: remarkably, introduction of errors in packets between CRC-protected hops is common, but the end-to-end 16-bit TCP checksum catches most of these simple errors. This is the end-to-end principle at work.

(e) Flow Control

TCP uses an end-to-end flow control protocol to avoid having the sender send data too fast for the TCP receiver to reliably receive and process it. Having a mechanism for flow control is essential in an environment where machines of diverse network speeds communicate. For example, when a fast PC sends data to a slow hand-held PDA, the PDA needs to regulate the influx of data, or protocol software would be overrun quickly. Similarly, flow control is essential if the application that is receiving the data is reading it more slowly than the sending application is sending it.

TCP uses a sliding window flow control protocol. In each TCP segment, the receiver specifies in the receive window field the amount of additional received data (in bytes) that it is willing to buffer for the connection. The sending host can send only up to that amount of data before it must wait for an acknowledgment and window update from the receiving host.

TCP sequence numbers and receive windows behave very much like a clock. The receive window shifts each time the receiver receives and acknowledges a new segment of data. Once it runs out of sequence numbers, the sequence number loops back to 0.

When a receiver advertises a window size of 0, the sender stops sending data and starts the persist timer. The persist timer is used to protect TCP from a deadlock situation that could arise if the window size update from the receiver is lost and the sender has no more data to send while the receiver is waiting for the new window size update. When the persist timer expires, the TCP sender sends a small packet so that the receiver sends an acknowledgement with the new window size.

If a receiver is processing incoming data in small increments, it may repeatedly advertise a small receive window. This is referred to as the silly window syndrome, since it is inefficient to send only a few bytes of data in a TCP segment, given the relatively large overhead of the TCP header. TCP senders and receivers typically employ flow control logic to specifically avoid repeatedly sending small segments. The sender-side silly window syndrome avoidance logic is referred to as Nagle's algorithm.

(f) Congestion Control

The final main aspect of TCP is congestion control. TCP uses a number of mechanisms to achieve high performance and avoid 'congestion collapse', where network performance can fall by several orders of magnitude. These mechanisms control the rate of data entering the network, keeping the data flow below a rate that would trigger collapse.

Acknowledgments for data sent, or lack of acknowledgments, are used by senders to infer network conditions between the TCP sender and receiver. Coupled with timers, TCP senders and receivers can alter the behaviour of the flow of data. This is more generally referred to as congestion control and/or network congestion avoidance.

Modern implementations of TCP contain four intertwined algorithms: Slow-start, congestion avoidance, fast retransmit, and fast recovery (RFC2581).

In addition, senders employ a retransmission timer that is based on the estimated round-trip time (or RTT) between the sender and receiver, as well as the variance in this round trip time. The behaviour of this timer is specified in RFC 2988. There are subtleties in the estimation of RTT. For example, senders must be careful when calculating RTT samples for retransmitted packets; typically they use Karn's Algorithm or TCP timestamps (see RFC 1323). These individual RTT samples are then averaged over time to create a Smoothed Round Trip Time (SRTT) using Jacobson's algorithm. This SRTT value is what is finally used as the round-trip time estimate.

Enhancing TCP to reliably handle loss, minimize errors, manage congestion and go fast in very high-speed environments are ongoing areas of research and standards development. As a result, there are a number of TCP congestion avoidance algorithm variations.

(g) Maximum Segment Size

The Maximum segment size (MSS) is the largest amount of data, specified in bytes, that TCP is willing to send in a single segment. For best performance, the MSS should be set small enough to avoid IP fragmentation, which can lead to excessive retransmissions if there is packet loss. To try to accomplish this, typically the MSS is negotiated using the MSS option when the TCP connection is established, in which case it is determined by the maximum transmission unit (MTU) size of the data link layer of the networks to which the sender and receiver are directly attached. Furthermore, TCP senders can use Path MTU discovery to infer the minimum MTU along the network path between the sender and receiver, and use this to dynamically adjust the MSS in order to avoid IP fragmentation within the network.

(h) Selective Acknowledgments

Relying purely on the cumulative acknowledgment scheme employed by the original TCP protocol can lead to inefficiencies when packets are lost. For example, suppose 10,000 bytes are sent in 10 different TCP packets, and the first packet is lost during transmission. In a pure cumulative acknowledgment protocol, the receiver cannot say that it received bytes 1,000 to 9,999 successfully, but failed to receive the first packet, containing bytes 0 to 999. Thus the sender would then have to resend all 10,000 bytes.

In order to solve this problem TCP employs the selective acknowledgment (SACK) option, defined in RFC 2018, which allows the receiver to acknowledge discontinuous blocks of packets that were received correctly, in addition to the sequence number of the last contiguous byte received successively, as in the basic TCP acknowledgment.

The acknowledgement can specify a number of SACK blocks, where each SACK block is conveyed by the starting and ending sequence numbers of a

contiguous range that the receiver correctly received. In the example above, the receiver would send SACK with sequence numbers 1,000 and 10,000. The sender will thus retransmit only the first packet, bytes 0 to 999.

The SACK option is not mandatory and it is used only if both parties support it. This is negotiated when connection is established. SACK uses the optional part of the TCP header (see TCP segment structure for details). The use of SACK is widespread, the all popular TCP stacks support it. Selective acknowledgment is also used in SCTP.

(i) Window Scaling

For more efficient use of high bandwidth networks, a larger TCP window size may be used. The TCP window size field controls the flow of data and is limited to between 2 and 65,535 bytes.

Since the size field cannot be expanded, a scaling factor is used. The TCP window scale option, as defined in RFC 1323, is an option used to increase the maximum window size from 65,535 bytes to 1 Gigabyte. Scaling up to larger window sizes is a part of what is necessary for TCP Tuning.

The window scale option is used only during the TCP 3-way handshake. The window scale value represents the number of bits to left-shift the 16-bit window size field. The window scale value can be set from 0 (no shift) to 14.

Many routers and packet firewalls rewrite the window scaling factor during a transmission. This causes sending and receiving sides to assume different TCP window sizes. The result is non-stable traffic that is very slow. The problem is visible on some sending and receiving sites which are behind the path of broken routers.

TCP window scaling can be a particular problem on Linux and Windows Vista systems.

(j) TCP Timestamps

TCP timestamps, defined in RFC 1323, help TCP compute the round-trip time between the sender and receiver. Timestamp options include a 4-byte timestamp value, where the sender inserts its current value of its timestamp clock, and a 4-byte echo reply timestamp value, where the receiver generally inserts the most recent timestamp value that it has received. The sender uses the echo reply timestamp in an acknowledgment to compute the total elapsed time since the acknowledged segment was sent.

TCP timestamps are also used to help in the case where TCP sequence numbers encounter their 232 bound and "wrap around" the sequence number space. This scheme is known as Protect Against Wrapped Sequence numbers, or PAWS.

(k) Out of Band Data

One is able to interrupt or abort the queued stream instead of waiting for the stream to finish. This is done by specifying the data as urgent. This will tell the receiving program to process it immediately, along with the rest of the

urgent data. When finished, TCP informs the application and resumes back to the stream queue. An example is when TCP is used for a remote login session, the user can send a keyboard sequence that interrupts or aborts the program at the other end. These signals are most often needed when a program on the remote machine fails to operate correctly. The signals must be sent without waiting for the program to finish its current transfer.

Unfortunately, TCP OOB data was not designed for the modern Internet. The urgent pointer only alters the processing on the remote host and doesn't expedite any processing on the network itself. When it gets to the remote host there are two slightly different interpretations of the protocol which means only single bytes of OOB data are reliable. This is assuming it's reliable at all as it's one of the least commonly used protocol elements and tends to be poorly implemented.

(l) Forcing Data Delivery

Normally, TCP waits for the buffer to exceed the maximum segment size before sending any data. This creates serious delays when the two sides of the connection are exchanging short messages and need to receive the response before continuing. For example, the login sequence at the beginning of a session begins with the short message "Login," and the session cannot make any progress until these five characters have been transmitted and the response has been received. This process can be seriously delayed by TCP's normal behaviour.

However, an application can force delivery of octets to the output stream using a push operation provided by TCP to the application layer. This operation also causes TCP to set the PSH flag or control bit to ensure that data will be delivered immediately to the application layer by the receiving transport layer.

In the most extreme cases, for example when a user expects each keystroke to be echoed by the receiving application, the push operation can be used each time a keystroke occurs. More generally, application programs use this function to force output to be sent after writing a character or line of characters. By forcing the data to be sent immediately, delays and wait time are reduced.

(m) Connection Termination

The connection termination phase uses, at most, a four-way handshake, with each side of the connection terminating independently. When an endpoint wishes to stop its half of the connection, it transmits a FIN packet, which the other end acknowledges with an ACK. Therefore, a typical tear-down requires a pair of FIN and ACK segments from each TCP endpoint.

A connection can be "half-open", in which case one side has terminated its end, but the other has not. The side that has terminated can no longer send any data into or receive any data from the connection, but the other side can (but generally if it tries, this should result in no acknowledgment and therefore a timeout, or else result in a positive RST, and either way thereby the destruction of the half-open socket).

It is also possible to terminate the connection by a 3-way handshake, when host A sends a FIN and host B replies with a FIN & ACK (merely combines 2 steps into one) and host A replies with an ACK. [3] This is perhaps the most common method.

It is possible for both hosts to send FINs simultaneously then both just have to ACK. This could possibly be considered a 2-way handshake since the FIN/ACK sequence is done in parallel for both directions.

Some host TCP stacks may implement a "half-duplex" close sequence, as Linux or HP-UX do. If such a host actively closes a connection but still has not read all the incoming data the stack already received from the link, this host will send a RST instead of a FIN. This allows a TCP application to be sure that the remote application has read all the data the former sent - waiting the FIN from the remote side when it will actively close the connection. Unfortunately, the remote TCP stack cannot distinguish between a Connection Aborting RST and this Data Loss RST - both will cause the remote stack to throw away all the data it received, but the application still didn't read.

Some application protocols may violate the OSI model layers, using the TCP open/close handshaking for the application protocol open/close handshaking - these may find the RST problem on active close.

For a usual program flow like above, a TCP/IP stack like that described above does not guarantee that all the data will arrive to the other application unless the programmer is sure that the remote side will not send anything.

13.13.5 TCP Ports

TCP uses the notion of port numbers to identify sending and receiving application end-points on a host, or Internet sockets. Each side of a TCP connection has an associated 16-bit unsigned port number (0-65535) reserved by the sending or receiving application. Arriving TCP data packets are identified as belonging to a specific TCP connection by its sockets, that is, the combination of source host address, source port, destination host address, and destination port. This means that a server computer can provide several clients with several services simultaneously, as long as a client takes care of initiating any simultaneous connections to one destination port from different source ports.

Port numbers are categorized into three basic categories': well-known, registered, and dynamic/private. The well-known ports are assigned by the Internet Assigned Numbers Authority (IANA) and are typically used by system-level or root processes. Well-known applications running as servers and passively listening for connections typically use these ports. Some examples include: FTP, SSH, TELNET, SMTP and HTTP. Registered ports are typically used by end user applications as ephemeral source ports when contacting servers, but they can also identify named services that have been registered by a third party.

Dynamic/private ports can also be used by end user applications, but are less commonly so. Dynamic/private ports do not contain any meaning outside of any particular TCP connection.

TCP is a complex protocol. However, while significant enhancements have been made and proposed over the years, its most basic operation has not changed significantly since its first specification RFC 675 in 1974, and the v4 specification RFC 793, published in September 1981. RFC 1122, Host Requirements for Internet Hosts, clarified a number of TCP protocol implementation requirements. RFC 2581, TCP Congestion Control, one of the most important TCP-related RFCs in recent years, describes updated algorithms to be used in order to avoid undue congestion. In 2001, RFC 3168 was written to describe explicit congestion notification (ECN), a congestion avoidance signalling mechanism.

The original TCP congestion avoidance algorithm was known as "TCP Tahoe", but many alternative algorithms have since been proposed (including TCP Reno, TCP Vegas, FAST TCP, TCP New Reno, and TCP Hybla).

TCP Interactive (iTCP) is a research effort into TCP extensions that allows applications to subscribe to TCP events and register handler components that can launch applications for various purposes, including application assisted congestion control.

13.13.7 TCP Over Wireless

TCP has been optimized for wired networks. Any packet loss is considered to be the result of congestion and the congestion window size is reduced dramatically as a precaution. However, wireless links are known to experience sporadic and usually temporary losses due to fading, shadowing, hand off, and other radio effects, that cannot be considered congestion. After the (erroneous) back-off of the congestion window size, due to wireless packet loss, there can be a congestion avoidance phase with a conservative decrease in window size. This causes the radio link to be underutilized. Extensive research has been done on the subject of how to combat these harmful effects. Suggested solutions can be categorized as end-to-end solutions (which require modifications at the client and/or server), link layer solutions (such as RLP in CDMA2000), or proxy based solutions (which require some changes in the network without modifying end nodes).

13.13.8 Hardware Implementations

One way to overcome the processing power requirements of TCP is to build hardware implementations of it, widely known as TCP Offload Engines (TOE). The main problem of TOEs is that they are hard to integrate into computing systems, requiring extensive changes in the operating system of the computer or device. The first company to develop such a device was Alacritech.

13.13.9 Debugging

A packet sniffer, which intercepts TCP traffic on a network link, can be useful in debugging networks, network stacks and applications which use TCP

by showing the user what packets are passing through a link. Some networking stacks support the SO DEBUG socket option, which can be enabled on the socket using setsockopt. That option dumps all the packets, TCP states and events on that socket which will be helpful in debugging, netstat is another utility that can be used for debugging.

13.13.10 Alternatives

For many applications TCP is not appropriate. One big problem (at least with normal implementations) is that the application cannot get at the packets coming after a lost packet until the retransmitted copy of the lost packet is received. This causes problems for real-time applications such as streaming multimedia (such as Internet radio), real-time multiplayer games and voice over IP (VoIP) where it is sometimes more useful to get most of the data in a timely fashion than it is to get all of the data in order.

For both historical and performance reasons, most storage area networks (SANs) prefer to use Fibre Channel protocol (FCP) instead of TCP/IP.

Also for embedded systems, network booting and servers that serve simple requests from huge numbers of clients (*e.g.* DNS servers) the complexity of TCP can be a problem. Finally some tricks such as transmitting data between two hosts that are both behind NAT (using STUN or similar systems) are far simpler without a relatively complex protocol like TCP in the way.

Generally where TCP is unsuitable the User Datagram Protocol (UDP) is used. This provides the application multiplexing and checksums that TCP does, but does not handle building streams or retransmission giving the application developer the ability to code those in a way suitable for the situation and/or to replace them with other methods like forward error correction or interpolation.

SCTP is another IP protocol that provides reliable stream oriented services not so dissimilar from TCP. It is newer and considerably more complex than TCP so has not yet seen widespread deployment. However, it is especially designed to be used in situations where reliability and near-real-time considerations are important.

Venturi Transport Protocol (VTP) is a patented proprietary protocol that is designed to replace TCP transparently in order to overcome perceived inefficiencies related to wireless data transport.

TCP also has some issues in high bandwidth utilization environments. The TCP congestion avoidance algorithm works very well for ad-hoc environments where it is not known who will be sending data, but if the environment is predictable, a timing based protocol such as ATM can avoid the overhead of the retransmits that TCP needs.

13.14 USER DATAGRAM PROTOCOL (UDP)

The User Datagram Protocol (UDP) is one of the core members of the Internet Protocol Suite, the set of network protocols used for the Internet. With UDP,

computer applications can send messages, sometimes known as datagrams, to other hosts on an Internet Protocol (IP) network without requiring prior communications to set up special transmission channels or data paths. UDP is sometimes called the Universal Datagram Protocol. The protocol was designed by David P. Reed in 1980 and formally defined in RFC 768.

UDP uses a simple transmission model without implicit hand-shaking dialogues for guaranteeing reliability, ordering, or data integrity. Thus, UDP provides an unreliable service and datagrams may arrive out of order, appear duplicated, or go missing without notice. UDP assumes that error checking and correction is either not necessary or performed in the application, avoiding the overhead of such processing at the network interface level. Time-sensitive applications often use UDP because dropping packets is preferable to using delayed packets. If error correction facilities are needed at the network interface level, an application may use the Transmission Control Protocol (TCP) or Stream Control Transmission Protocol (SCTP) which are designed for this purpose.

UDP's stateless nature is also useful for servers that answer small queries from huge numbers of clients. Unlike TCP, UDP is compatible with packet broadcast (sending to all on local network) and multicasting (send to all subscribers).

Common network applications that use UDP include: the Domain Name System (DNS), streaming media applications such as IPTV, Voice over IP (VoIP), Trivial File Transfer Protocol (TFTP) and many online games.

13.14.1 Ports

UDP applications use datagram sockets to establish host-to-host communications. Sockets bind the application to service ports that function as the endpoints of data transmission. A port is a software structure that is identified by the port number, a 16 bit integer value, allowing for port numbers between 0 and 65,535. Port 0 is reserved, but is a permissible source port value if the sending process does not expect messages in response.

Ports 1 through 1023 (hexadecimal 0 × 3FF) are named "well-known" ports and on Unix-like operating systems, binding to one of these ports requires superuser (root) access.

Ports 1024 through 49,151 (0 × BFFF) are registered ports.

Ports 49,152 through 65,535 (0 × FFFF) are used as temporary ports primarily by clients when communicating to servers.

13.14.2 Packet structure

UDP is a minimal message-oriented Transport Layer protocol that is documented in IETF RFC 768.

In the Internet Protocol Suite, UDP provides a very simple interface between the Internet Layer below (e.g., IPv4) and the Application Layer above.

UDP provides no guarantees to the upper layer protocol for message delivery and a UDP sender retains no state on UDP messages once sent (for this reason UDP is sometimes called the Unreliable Datagram Protocol). UDP adds only application multiplexing and checksumming of the header and payload. If any kind of reliability for the information transmitted is needed, it must be implemented in upper layers.

Table 13.8

+	Bits 0-15	16-31
0	Source Port	Destination Port
32	Length	Checksum
64	Data	

The UDP header consists of only 4 fields. The use of two of those is optional (pink background in table 13.8).

(*i*) **Source port**: This field identifies the sending port when meaningful and should be assumed to be the port to reply too if needed. If not used, then it should be zero.

(*ii*) **Destination port**: This field identifies the destination port and is required.

(*iii*) **Length**: A 16-bit field that specifies the length in bytes of the entire datagram: header and data. The minimum length is 8 bytes since that's the length of the header. The field size sets a theoretical limit of 65,535 bytes (8 byte header + 65, 527 bytes of data) for a UDP datagram. The length includes the UDP header, so the minimum size for a UDP datagram is 8 (8 byte header with no data). The practical limit for the data length which is imposed by the underlying IPv4 protocol is 65,507 bytes.

(*iv*) **Checksum** : The 16-bit checksum field is used for error-checking of the header and data.

With IPv4 when UDP runs over IPv4, the method used to compute the checksum is defined within RFC 768:

Checksum is the 16-bit one's complement of the one's complement sum of a pseudo header of information from the IP header, the UDP header, and the data, padded with zero octets at the end (if necessary) to make a multiple of two octets.

In other words, all 16-bit words are summed together using one's complement (with the checksum field set to zero). The sum is then one's complemented. This final value is then inserted as the checksum field. Algorithmically speaking, this is the same as for IPv6.

The difference is in the data used to compute the checksum. Included is a pseudo-header that contains information from the IPv4 header:

Table 13.9

+	Bits 0 - 7	8 –15	16-23	24-31
0	Source address			
32	Destination address			
64	Zeros	Protocol	UDP length	
96	Source Port		Destination Port	
128	Length		Checksum	
160	Data			

The source and destination addresses are those in the IPv4 header. The protocol is that for UDP. The UDP length field is the length of the UDP header and data.

If the checksum is calculated to be zero (all 0s) it should be sent as negative zero (all 1 ‘s). If a checksum is not used it should be sent as zero (all 0s) as zero indicates an unused checksum. With IPv6 when UDP runs over IPv6, the checksum is no longer considered optional, and the method used to compute the checksum is changed, as per RFC2460:

Any transport or other upper-layer protocol that includes the addresses from the IP header in its checksum computation must be modified for use over IPv6, to include the 128-bit IPv6 addresses instead of 32-bit IPv4 addresses.

When computing the checksum, a pseudo-header that mimics the IPv6 header is included:

Table 13.10

+	Bits 0 - 7	8-15	16-23	24-31
0	Source address			
32				
64				
96				
128		Destination address		
160				
192				
224				
24				
256	UDP length			
288	Zeros			Next Header
320	Source Port	Destination Port		
352	Length	Checksum		
384	Data			

The source address is the one in the IPv6 header. The destination address is the final destination; if the IPv6 packet doesn't contain a Routing header, that will be the destination address in the IPv6 header; otherwise, at the originating node, it will be the address in the last element of the Routing header, and, at the receiving node, it will be the destination address in the IPv6 header. The Next Header value is the protocol value for UDP: 17. The UDP length field is the length of the UDP header and data.

If the checksum is calculated to be zero (all 0s) it should be sent as negative zero (all 1's).

Lacking reliability, UDP applications must generally be willing to accept some loss, errors or duplication. Some applications such as TFTP may add rudimentary reliability mechanisms into the application layer as needed. Most often, UDP applications do not require reliability mechanisms and may even be hindered by them. Streaming media, real-time multiplayer games and voice over IP (VoIP) are examples of applications that often use UDP. If an application requires a high degree of reliability, a protocol such as the Transmission Control Protocol or erasure codes may be used instead.

Lacking any congestion avoidance and control mechanisms, network-based mechanisms are required to minimize potential congestion collapse effects of uncontrolled, high rate UDP traffic loads. In other words, since UDP senders cannot detect congestion, network-based elements such as routers using packet queuing and dropping techniques will often be the only tool available to slow down excessive UDP traffic. The Datagram Congestion Control Protocol (DCCP) is being designed as a partial solution to this potential problem by adding end host TCP-friendly congestion control behaviour to high-rate UDP streams such as streaming media.

While the total amount of UDP traffic found on a typical network is often in the order of only a few per cent, numerous key applications use UDP, including: the Domain Name System (DNS) (since most DNS queries only consist of a single request followed by a single reply), the simple network management protocol (SNMP), the Dynamic Host Configuration Protocol (DHCP) and the Routing Information Protocol (RIP).

13.14.3 Difference between TCP and UDP

TCP ("Transmission Control Protocol") is a connection-oriented protocol, which means that upon communication it requires handshaking to set up end-to-end connection. A connection can be made from client to server, and from then on any data can be sent along that connection.

(*i*) **Reliable** - TCP manages message acknowledgment, retransmission and timeout. Many attempts to reliably deliver the message are made. If it gets lost along the way, the server will re-request the lost part. In TCP, there's either no missing data, or, in case of multiple timeouts, the connection is dropped.

(*ii*) **Ordered** - If two messages are sent along a connection, one after the other, the first message will reach the receiving application first. When data packets arrive in the wrong order, the TCP layer holds the later data until the earlier data can be rearranged and delivered to the application.

(*iii*) **Heavyweight** - TCP requires three packets just to set up a socket, before any actual data can be sent. It handles connections, reliability and congestion control. It is a large transport protocol designed on top of IP.

(*iv*) **Streaming** - Data is read as a "stream," with nothing distinguishing where one packet ends and another begins. Packets may be split or merged into bigger or smaller data streams arbitrarily.

UDP is a simpler message-based connectionless protocol. In connectionless protocols, there is no effort made to set up a dedicated end-to-end connection. Communication is achieved by transmitting information in one direction, from source to destination without checking to see if the destination is still there, or if it is prepared to receive the information.

(*i*) **Unreliable** - When a message is sent, it cannot be known if it will reach its destination; it could get lost along the way. There is no concept of acknowledgment, retransmission and timeout.

(*ii*) **Not ordered** - If two messages are sent to the same recipient, the order in which they arrive cannot be predicted.

(*iii*) **Lightweight** - There is no ordering of messages, no tracking connections, etc. It is a small transport layer designed on top of IP.

(*iv*) **Datagrams** - Packets are sent individually and are guaranteed to be whole if they arrive. Packets have definite bounds and no split or merge into data streams may exist.

13.15 WIRELESS LOCATION SERVICE ARCHITECTURES

13.15.1 Introduction

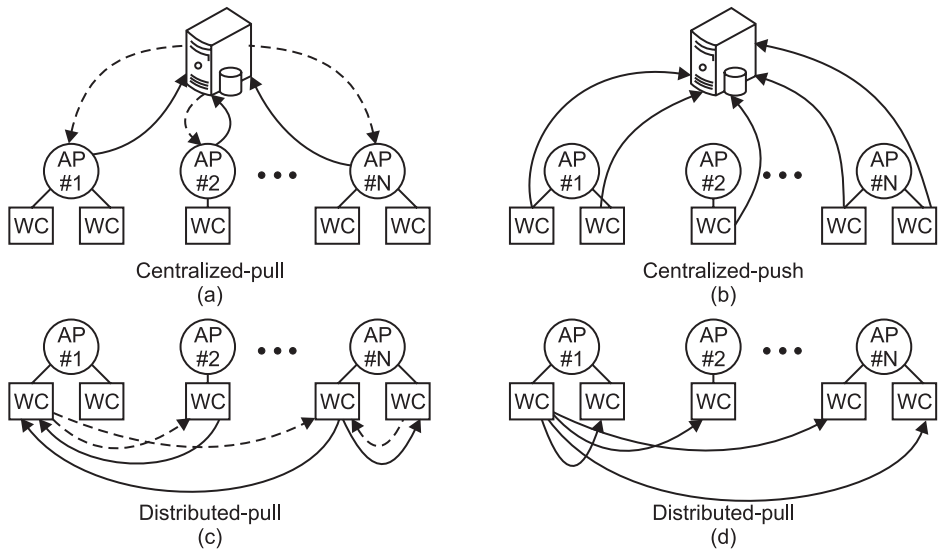
This topic describes the basic architectures for implementing wireless location services. The four architectures presented vary according to the amount of centralization and the methodology of how (push/pull) the location information is obtained from the client. These architectures are analyzed using both qualitative and quantitative methods. Qualitative comparisons utilize the amount of messages and size of data flow for various numbers of clients. Qualitative descriptions are provided for areas such as privacy, security, and ease of implementation. An understanding of the architecture options and their relative strengths and weaknesses is necessary in understanding Locator@CMU and how it differs from other implementations of wireless location services.

13.15.2 Taxonomy

Wireless location service architectures can be classified according to at least two main properties. The first property is whether a central server is used to store

the location information and respond to client queries. Architectures that have this property are "centralized", while architectures in which there is no server to store location data and these clients directly communicate with each other is said to be "distributed". The second property used to classify architectures is if the location updates are "pushed" or must be "pulled" from the client. The cross-product of these properties is the four basic architectures types for use in location services (Figure 13.5).

The first architecture *(a)* shown in Figure 13.5 is a centralized-pull architecture using polling of Access Points to retrieve location information. The second architecture (*b*) in Figure 13.5 is a centralized push architecture where the clients must send their information to the server. The third architecture Figure 13.5 shows a distributed-pull model in which a client responds to at and sends the requestor back the information. The last architecture (*d*) in Figure 13.5 is a distributed-push model where the client periodically sends information to subscribing requestors. Each of these architectures is explored in more depth in Sections 13.15.2.1 and 13.15.2.2.



Dotted lines represent a request for location information. Solid lines represent location information being sent. AP stands for Access Point and WC stands for Wireless Client.

Fig. 13.5: Common architectures for wireless location services.

13.15.2.1 *Centralized Architectures*

In the centralized architectures, a core server computer stores all location data and moderates the distribution of this data to requestors. Any requestor that desires the location of a user or group of users sends the request to the server instead of to the individual users. The server computer processes the request and uses privacy filters to return the appropriate information back to the requestor.

Context-aware applications that need a user's location can communicate directly with the server computer to obtain the desired information. The server

computer additionally provides associations between a wireless MAC Address and a user. This association is possible since users are required to initially register with the server and provide basic information and rules for their account. The privacy rules and constraints can be entered directly on the server computer. The user must trust the server to correctly implement and apply these rules. The server based architecture uses a database to keep track of the current state of all users. Additionally, central servers have the option of storing the history of movements.

The centralized architectures require a powerful database in order to organize and store all of the present and past information to be recorded. This information is often sub-divided into different databases and tables for maximum efficiency and performance. Modern database techniques are used to allow for near instant retrieval of current location information while historical information can incur slight delays to fully retrieve figure 13.6 shows a matrix that outlines the basic properties of the two centralized architectures and those of the distributed ones. More specific details of the centralized-pull and centralized push architecture types are discussed in Sections 13.15.2.1.1 and 13.15.2.1.2.

Method Update Location	Architecture: Centralized	Architecture: Distributed
Poll	- Core server - Passive Approach Access Points used for location information - No software required on wireless devices - Polling updates locations at set intervals	- No Server - Each location request in sent from requester to requestee - Privacy implematented on requestee - Software required on devices
Push	- Core server - Active Approach - Software required on wireless devices - Location updated sent based on time interval or events	- No Server - Requires subscription type service on requestees - Requestors send Subscription information to requestees - Software required on devices

Fig. 13.6: Matrix of common architectures.

13.15.2.1.1 *Centralized-Pull Architecture*

This centralized-pull architecture relies on the network infrastructure to obtain the clients' location information. No software or messages are sent directly from the client. This is a passive wireless location service that utilizes the client's normal communication with the access points. Locator@CMU is implemented as centralized-pull architecture. A server computer directly polls the necessary wireless access points. Each access point returns a list of the connected wireless devices in terms of their unique MAC Addresses and signal strengths. This information is used to determine the wireless device's location. On an 802.11 network this approach yields a limited location granularity. In a typical indoor environment,. a user location is defined to be within a 100 foot radius of the access point.

It should also be noted that besides polling the network infrastructure centralized-pull architecture could poll the clients directly if they had appropriate software installed and running. The common implementation of this approach is centralized pull-on-demand. Instead of polling at a set time interval the clients are polled only when an application is requesting information from the server. This request causes the server to check its cache and if the location information is not 'recent' the server will pull the information from the client and then relay it back to the requesting application. This approach is very similar to the distributed-pull described in Section 13.15.2.2.1 except that a server is used as an intermediary between the wireless clients. Due to this similarity and the more common implementation of the access point pulling for centralized-pull the centralized-pull on demand will not be analyzed separately.

13.15.2.1.2 ***Centralized-Push Architecture***

The centralized-push architecture like that of the centralized-pull utilizes a server computer to record and distribute wireless clients' location information. Unlike the pull model, in the push model it is the responsibility of each client to send information on its location to the server computer. The location data that is pushed to the server could range from precise coordinates generated by an advanced technique such as signal strength or less precise information such as the access point to which the device is connected.

The client push can be implemented as a time or event based mechanism. In a time based push, each client independently sends its location information at a set interval (*i.e.* 2 minutes, 10 minutes, etc.). The event based push is based on a user action or state change. User movement is a common event chosen (i.e. anytime a user changes location by more than 10 feet, update the server). Other possible events include turning on or off the client program.

One significant difference between the centralized-push and centralized-pull model is the necessity for a program to be running on the client for sending messages to the server. In a centralized-pull model, no software needs to run on the client while in the push model each client must install and maintain software to send the necessary information to the server. This can pose a significant burden, but also allows for greater flexibility and granularity of location information.

13.15.2.2 ***Distributed Architectures***

In a completely distributed architecture clients must communicate directly with each other to obtain desired location information. No server is necessary to store or poll location data. Therefore, each client can implement its own desired privacy rules and moderate all requests locally.

One common problem in a distributed system is the communication between the clients. Without a central server it can be difficult for one client to know how to send a message to another client. There are a number of possible solutions ranging from complex to relatively simple options to obtain network paths between clients. One of the simple solutions is to give all wireless devices static IP

addresses and limit the network to a single subnet. On a more complex network, this is clearly not an option, so more involved solutions such as registration servers that can keep track of the network address of clients are necessary. For the purpose of this paper, the simple static IP address model will be assumed.

13.15.2.2.1 Distributed-Pull Architecture

In the distributed-pull architecture there is a requestor client and a requestee client. The requestor must first send a message to the requestee inquiring its location. Lacking a server to moderate this client to client communication can require an initial handshake stage to verify authenticity and determine what if any location information the requestee should return to the requestor. Once the proper security is cleared the requestee then sends a final message back that includes the location information to the requestor.

13.15.2.2.2 Distributed-Push Architecture

The distributed-push architecture can be a complex system to implement. Each requestee must maintain a registration table to keep track of all the requestors that would like to be updated whenever the requestee moves or some other event occurs. Unlike the centralized-push model where the requestee sends a single message to a server at a set location, in the distributed-push model the requestee needs to send messages to every requestor client. The more requestors that have subscribed, the more messages that must be sent. A rapidly moving requestee with many requestors could be sending out a large stream of data. The requestors themselves are also wireless clients and potentially moving. The requestee thus must keep track of the changing network paths and addresses of the requestors. This process may involve the necessity of administrative messages or potentially a third party server. The distributed-push model can allow the requestors to have options over the frequency of their updates (i.e. User a wants updates every 5 minutes and User B wants updates every 3 minutes).

13.15.3 Data Flow Analysis

Comparing the efficiency of the data flow in each of the architectures helps to quantitatively understand their differences. To do the analysis we select a common context aware application. This instant messaging application is described below in Section 13.15.3.1.

13.15.3.1 *Context-Aware Application*

To study these prospective wireless location architectures it is necessary to select applications that will utilize them. A suite of location aware messaging applications has been implemented at Carnegie Mellon using the Locator@ CMU service. One of these applications is a common instant messaging application. This instant messaging program not only displays the usual user status information (online, offline, away, etc), but also the user's current or last known location. In the centralized architectures the application requests location information from the server for all of its current users at set time intervals.

Using the information obtained from the server the location of each user is updated in the application GUI. For the distributed architectures the GUI is also updated with new location information. However, instead of having a server to retrieve information from, the application must rely on a location aware client that functions in the distributed architecture. Other location-based applications that have been developed allow messages to be sent to all users in a specified location (building, floor, area of campus) and also allow virtual location message bulletin boards to be created.

The instant messaging application is particularly interesting to study in conjunction with the four architecture types for a number of reasons. First, it is a widely used and common application. Second, the instant messaging application involves more than simple one to one user relationships. Users typically have long lists of "buddies". This requires location information to be updated for multiple users spread across many locations. Thirdly, the instant messaging application allows for basic privacy concepts to be explored and tested in a common and familiar user environment.

13.15.3.2 *Architecture Performance*

The location-aware instant messaging application provides a mechanism to compare how architectures will perform. The formulas demonstrating the amount of data required to be sent are shown for an average one hour period. Findings from past research are referenced to provide information on factors that affect the architecture performance, such as number of concurrent users and number of buddies per user. The number of access points on the Wireless Andrew network is a constantly growing number. For the purpose of these calculations we utilize the figure of 640 access points that were present when the application was implemented.

13.15.3.2.1 Centralized-Pull Architecture

In the centralized-pull architecture a server computer is used to poll each of the access points. The amount of information retrieved from a single access point is dependent on the total number of users that are connected to it. The data flow of both the requests and return information to the server from the access points is modeled by Equation 13.1:

$$TB = \%.T*F*\ [(N_{AP} * R_{AP})+(N_{AP} * UAvg * R_{USR})] \qquad ...(13.1)$$

TB : Total bytes sent

T : Length of time (mins)

F : Number of polls per minute

NAP : Number of access points in network

RAP : Size of request message to AP

UAvg : Average number of users per AP

RUsr : Size of single user location information

Taking the Locator@CMU system and Wireless Andrew network as representative we can solve the equation with known data. For time of one hour T = 60 minutes. The frequency of the polls F are every minute and the number of access points NAP = 640. The size of the request to each access point RAP is nearly 2 kilobytes. From past research on the Locator@CMU network we know that the average number of users per AP UAvg is 5.1 and the RUsr is close to 3 bytes. These numbers yield total bytes sent of 77,152 KB per hour which yields an average data rate of 171.4 Kbps, an easy amount for any modern network to handle.

The formula above gave the data flow required to get the information into the central server. Next, we need to calculate the data flow of getting the location information to the context-aware instant messaging application from the central location server. Equation 13.2 models the amount of data required by instant messaging clients from the location server.

$$TB = T*F*[(N_{IM} * R_S) + (N_{IM} * B_{AVg} * R_B)] \quad ...(13.2)$$

TB : Total bytes sent
T : Length of time (mins)
F : Number of polls per minute
NIM : Number of IM Clients currently running
RS : Size of request to location server for buddy locations
BAvg : Average number of buddies per user
RB : Size of location information per buddy

Once again we can solve the equation to determine the amount of information sent to and from all application clients over one hour. The number of instant messaging clients NIM is assumed to be the average wireless users concurrent at a single time: -3,000. The size of the request to the location server RS is 3 kilobytes. The average number of buddies per user BAvg is 48 from our research and the size of location for each buddy RB returned by the server is 3 bytes. Solving Equation 4.2 we get a total bytes sent to the application of 1,182,464 KB or nearly 1 gigabyte of dataflow at a rate of 2.56 Mbps. The total dataflow for the centralized-pull architecture is thus the amount from the access points to the location server (77,152 KB) combined with the amount from the server to the instant messaging application (1,182,464 KB) which is a total of 1,259,616 KB per hour or a rate of 2.73 Mbps.

13.15.3.2.2 *Centralized-Push Architecture*

In the centralized-push architecture each client sends the server its location at set events *(i.e.* change of location or at set time intervals). In the centralized architectures there are two parts to the total data flow: location information to the central server and central server to the application. For both the push and pull model we assume that the communication between the server and the application is the same. Once the information is in the server, the centralized-push and pull have the same dataflow path between the server and the context-

aware application. We will calculate the effect of using both set time intervals as well as an event (change of location) for sending location information to server. The simple formula for set time intervals is given in Equation 13.3.

$$TB = T*F*[(N_{IM}* M_S)] \quad ...(13.3)$$

TB : Total bytes sent
T : Length of time (mins)
F : Number of pushes per minute
NIM : Number of IM Clients currently running
MS : Size of location message to server

Next we develop the formula for event based push. Equation 13.4 calculates the total bytes sent for a simple event based push where the event is any change of access point that the client is currently connected too.

$$TB = T*E*[(N_{IM}*M_S)] \quad \text{(Equation 13.4)}$$

TB : Total bytes sent
T : Length of time (mins)
E : Average number of events per user per minute (event = switching access points)
NIM : Number of IM Clients currently running
MS : Size of location message to server

Inserting information on user movement (1 move per 85 minutes, E = 0.0177 moves per minute). With a time interval of one minute, as in the pull architecture, the total traffic required for reporting client location on a time based push is given by TB = 60 mins * 1 per min * (3,000 * 2 KB) = 360,000 KB at a rate of 800 Kbps. The event based data flow is given by TB = 60 mins * 0.0117 per min * (3,000 * 2KB) = 4,212 KB at a rate of 9.36 Kbps.

13.15.3.2.3 ***Distributed-Pull Architecture***

The same data calculations are now performed on distributed-pull architecture. In the distributed architecture Equation 13.5 can be used to describe the requests for location information from the application computer to each buddy computer (the dotted lines in Figure 13.5 (c))

$$TB = T*F*[N_{IM}*R_{IM}*B_{Avg}] \quad ... (13.5)$$

TB : Total bytes sent
T : Length of time (mins)
F : Number of polls per minute
NIM : Number of IM Clients currently running
RIM : Size of request each client
BAvg : Average number of buddies per user

The only new variable we introduce is RIM which represents the size of the request to each client. Since clients need additional information for authentication

and trust purposes the size of these messages are larger than those to the server and are 5 KB. Using all of the previous information Equation 13.5 yields total bytes per hour of location requests of 43,200,000 KB or-42 GB at a rate of 93.75 Mbps.

Equation 13.6 represents the location information sent to the clients (solid lines in Figure 13.5(c)).

$$TB = T*F*[N_{IM}*R_B*B_{AVG}] \quad \text{... (13.6)}$$

TB : Total bytes sent
T : Length of time (mins)
F : Number of polls per minute
NIM : Number of IM Clients currently running
RB : Size of location information per buddy
BAvg : Average number of buddies per user

The location information is contained in the packet RB of only 3 bytes. The total bytes returned to the clients in a single hour are 25,313 KB at a rate of 56.25 Kbps. The total data flow for the distributed-pull is thus the combination of the requests (43,200,000 KB) plus the return information (25,313 KB) for a total data flow of 43,225,313 KB and a combined rate of 93.8 Mbps.

13.15.3.2.4 ***Distributed-Push Architecture***

Like the centralized-push architecture the distributed-push architecture can be done on a time or an event base. In a distributed push system requestors will need to register and deregister with each requestee that they desire location information to be sent from. The registration and deregistration mechanism is heavily dependent on the implementation. In addition to the registration messages, basic administration message also need to be sent to allow the requestor and requestee to be able to reach each other. To model this administration overhead, we assume that one administration message is necessary per requestor/requestee relationship every twenty minutes and use Equation 13.7.

$$TB = T*A*[N_{IM}* A_M*B_{Avg}] \quad \text{... (13.7)}$$

TB : Total bytes sent
T : Length of time (mins)
A : Number of administrative messages per minute
NIM : Number of IM Clients currently running
AM : Size of administrative message
BAvg : Average number of buddies per user

Equation 13.7 can be solved with the information that we know already plus the size of the administrative messages being 3 KB. This yields a total data flow of 1,296,000 KB for the hour at a data rate of 2.81 Mbps.

For the location push messages that need to be sent to the requestors, the time based solution is given by Equation 13.8.

$$TB = T*F*[N_{IM}*M_R*B_{AVG}] \quad \text{... (13.8)}$$

TB : Total bytes sent
T : Length of time (mins)
F : Number of pushes per minute
NIM : Number of IM Clients currently running
MR : Size of location message to requestor
BAvg : Average number of buddies per user

We can solve this formula knowing that MR is 3 KB. This yields TB = 60 mins * 1 minute * (3000 clients * 3 KB * 48) = 5,529,600 KB at a rate of 12 Mbps. The combined value for the time based push architecture is thus 14.81 Mbps.

The event based formula requires a slight change over the time based formula. Once again we can utilize our study of user movement on Wireless Andrew to solve on a network wide basis with known constants. Equation 13.9 gives the total bytes required.

$$TB = T*E*[N_{IM}*M_R*B_{AVG}] \quad \text{... (13.9)}$$

TB : Total bytes sent
T : Length of time (mins)
E : Average number of events per user per minute (event = switching access points)
NIM : Number of IM Clients currently running
MR : Size of location message to requestor
BAvg : Average number of buddies per user

Solving the equation for total bytes we get TB = 60 mins * 0.0117 per minute * 3,000 * 3 KB * 48 = 303,264 KB at a rate of 673.29 Kbps. The combined value for total data flow for the event based solution is thus 3.47 Mbps.

13.15.4 Comparison of Privacy and Security in Architectures

Sections 13.15.4.1 - 13.15.4.2 discuss each of the architectures in more depth.

13.15.4.1 *Centralized Architectures*

Both the centralized-pull and centralized-push architectures rely on a central server computer. This server is a major source of privacy and security vulnerability. If this server is compromised then all of the users' location data, preferences, and setup information could be obtained by an intruder. It is important that extensive security techniques be used to protect the integrity of the central server. For a centralized architecture to succeed users must have trust in the system. Tens of millions of internet user send their credit card and financial information to central server computers everyday. These merchants and financial institutions have cultivated a large amount of trust in the system through independent certifications and extensive methods for system security. Centralized architecture developers must also employ similar techniques. In a central server privacy presents a major hurdle. Many users are resistant to the

idea of "big brother" knowing where they are and keeping track of their location and past history. It is important for the central server to explicitly and clearly inform the user of what information is recorded and how it is stored. Users should be given easy instructions to understand and clear options to control how their information is stored and who has access to it.

13.15.4.1.1 ***Centralized-Pull Privacy and Security***

In the centralized-pull model the location is obtained in a completely passive methodology. The network infrastructure contains the necessary location information and this information is gleaned by the central server through polling of the access points. The only action that is required by the user is a one time registration to make the logical connection between the wireless MAC Address and the user. It is important that during this step the user is given all of the privacy information and options to choose from.

A major advantage that the centralized-pull model has is that no software needs to be installed on the wireless devices in order to locate them. This also offers a major security benefit in that there is no need to implement security on a client device. Client devices with their widely varying operating systems, applications, and programs can be a major source of security breaches that endanger information on a specific user or set of users. Security can be controlled and implemented entirely as part of the central server in the centralized-pull model.

13.15.4.1.2 Centralized-Push Privacy and Security

The centralized-push model is an active method in which the wireless device sends its location information to the central server. Greater control of privacy is given to users over the centralized-pull method. For example, if a user does not want to be located they can simply chose not to run their location software application.

It is important that secure communication channels be used between the client and server All location and personnel data that is sent must be encrypted to ensure that it is not "sniffed" off of a network connection. Additionally, the client software itself should be secured. If the device is hacked or stolen a user's information can be compromised.

13.15.4.2 Distributed Architectures

The distributed architectures lack a server computer and thus, all of the security and privacy implementations must be done on the wireless devices. One of the major advantages often mentioned is that there is no single point where all information can be compromised. In the distributed model if one or two wireless devices are infiltrated this breach can be limited to the information on those devices. This property of the data diffuseness makes it extremely difficult for an intruder to ever fully get all system information on every user. However, the lack of a central server also makes it harder to detect any break-ins and can increase the vulnerability of the individual nodes. In the centralized

architecture wireless devices never communicated with each other but always with the trusted server. In the distributed architecture all communication is directly with other wireless devices. Each wireless device must independently be trusted. Users can implement a variety of different device specific security techniques ranging from nothing to extensive precautions.

13.15.5 Wireless Architecture Types Summary

In choosing which architecture is best for imremeriting a wireless location service the designer must balance the performance characteristics, ease of implementation, and users' privacy and security concerns. The centralized architectures provide an enormous amount of convenience for application developers and even users. However, the systems must be carefully designed so that users are able to adequately trust the system to maintain their information. This is certainly a Hurdle that can be overcome. Hundreds of millions of people trust and rely on systems that they have far less control of or insight into. For example, supermarkets track our shopping history through discount cards and develop extensive profiles of our likes and dislikes for targeted marketing. We 'google' website merchants and have little hesitation to send even unknown companies our credit card and contact information (address, phone, email). There are many examples where a central server is used to store and track information on users. The key factor in acceptance of this is the trust in the entity running the service as well as the perceived value gained by divulging information about oneself.

For environments where the user must have total control over how information is disseminated, the distributed model is best suited. However, in the distributed model the user must take greater responsibility for ensuring that their system is secure as well as those people that they communicate with. There is no server to enforce standards or verify requests. The centralized-push model offers some of the best qualities of a centralized service while still allowing the user considerable control. In this model requests for location can still be aggregated and verified at the server computer. This significantly reduces the load placed on the individual client. Further, the user maintains control over when the server can locate them. They can choose not to send their information to the server.

For large scale wireless networks with thousands of users, a centralized architecture offers the best performance and implementation capabilities. Distributed models can be ideal for ad-hoc networks or limited research-based implementations, but are very difficult to fully implement across an entire organization. The centralized-pull model is by far the easiest model to implement. The main drawback, however, is that its precision in locating wireless users is limited to the access point signal range. For many applications this location precision might not be sufficient, and a solution where location software is installed on the client is necessary. It is necessary to carefully weigh the needs of applications and users when selecting architecture for use.

In both the push and pull distributed models, all of the logic for request verification must be implemented on the wireless device. The lack of a server with a powerful processor can sometimes limit the level of processing and memory available for security techniques. This is especially true when using PDAs or even smaller wireless devices. It is extremely important for the devices to be able to create a secure communication channel with each other. The sending devices must also trust the receiving device not to relay location information to a third party.

HIGHLIGTS

1. Wireless USB is a short-range, high-bandwidth wireless radio communication protocol created by the Wireless USB Promoter Group. Wireless USB is based on the WiMedia Alliance's Ultra-WideBand (UWB) common radio platform, which is capable of sending 480 Mbit/s at distances up to 3 meters and 110 Mbit/s at up to 10 meters. It was designed to operate in the 3.1 to 10.6 GHz frequency range.
2. Wireless USB is used in game controllers, printers, scanners, digital cameras, MP3 players, hard disks and flash drives. It is also suitable for transferring parallel video streams.
3. One of the imperatives of the modern combat leader is enabling his forces to exchange information easily between all services, at all echelons in the chain of command, in all forms of information to include voice, data, and video.
4. The IEEE 802.11b is now the most common wireless technology standard used in offices, airports, and coffee shops. The IEEE established the 802.11b standard in 1997; it can provide data rates of 1-11 Mbps from 600 meters to 25 miles. 802.11b operates in the 2.4-2.83 GHz range. 802.11b supports multi-hop wireless network devices meaning that each radio supports more than one peer-to-peer link and acts as a repeater or bridge between the established links.
5. The IEEE ratified the 802.11a standard in 1999. The 802.11a standard can provide data rates of 1-54 (72 in the Turbo Mode) Mbps from 300 meters to 25 miles; it operates in the 5.15-5.35 GHz range. 802.11b and 802.11a use identical framing; however, 802.11a uses an entirely different encoding scheme called Orthogonal Frequency Division Multiplexing or OFDM. The 802.11 a also supports multi-hop wireless devices and during heavy congestion it can divide the frequency spectrum into 10 different channels.
6. The Wireless Network design for information systems must support a maneuver Brigade's existing ground and signal platforms so that connectivity in tactical situations could be considered. One solution for a Brigade wireless network design is the Cisco AIRONET 1200 Series. The Cisco AIRONET 1200 Series products are compliant with IEEE 802.11a and 802.11b standards.
7. GPRS (General Packet Radio Service) is a step between GSM and 3G cellular networks. GPRS offers faster data transmission via a GSM network within a range 9.6 Kbits to 115 K bits. This new technology makes it possible for users to make telephone calls and transmit data at the same time.
8. When packet-switched data leaves the GPRS/GSM network, it is transferred to TCP-IP networks such as the Internet or X.25.Thus, GPRS includes new transmission and signaling procedures as well as new protocols for interworking with the IP world and other standard packet networks.

9. The GPRS mobile phones can be classified into the following three classes in terms of the possibility of simultaneous calls (via GSM) and data transmission (via GPRS).
   (*i*) Class A - Simultaneous calls (via GSM) and data transmission (via GPRS)
   (*ii*) Class B - Automatic switching between the GSM and the GPRS mode is possible according to telephone settings.
   (*iii*) Class C - Hand operated switching between the GSM and the GPRS mode
10. In the core network, the existing MSCs (Mobile Services Switching Centers) are based upon circuit-switched technology, and they cannot handle the GPRS style packet traffic. Thus two new and very important components, called GPRS Support Nodes, are added:
   (*i*) Serving GPRS Support Node (SGSN)
   (*ii*) Gateway GPRS Support Node (GGSN)
11. GPRS could possibly be the technology that will allow consumers to really begin to pursue the mobile Internet. GPRS is considered one step ahead of HSCSD (High Speed Circuit Switched Data) and a step towards 3 G (Third-generation) networks. It is the step to 2.5 G for GSM and TDMA (Time Division Multiple Access) service providers.
12. The Wireless Application Protocol has become the standard for communication between server applications and its clients. That is exactly what the Wireless Application Protocol is; for example, WAP is used as a standardized method so that a cellular phone can talk to a server among the cellular network that it belongs to.
13. The Special Interest Group, also known as its acronym SIG, includes companies such as: 3 Com, Ericsson, IBM, Intel, Lucent, Microsoft, Motorola, Nokia and Toshiba, among the few big companies. SIG contains many members of smaller companies as well.
14. IEEE 802.11 is a set of standards for wireless local area network (WLAN) computer communication, developed by the IEEE LAN/MAN Standards Committee (IEEE 802) in the 5 GHz and 2.4 GHz public spectrum bands.
15. The original version of the standard IEEE 802.11 was released in 1997 and clarified in 1999, but is today obsolete. It specified two net bit rates of 1 or 2 megabits per second (Mbit/s), plus forward error correction code. It specified three alternative physical layer technologies: diffuse infrared operating at 1 Mbit/s; frequency-hopping spread spectrum operating at 1 Mbit/s or 2 Mbit/s; and direct-sequence spread spectrum operating at 1 Mbit/s or 2 Mbit/s.
16. The 802.11a standard uses the same data link layer protocol and frame format as the original standard, but an OFDM based air interface (physical layer). It operates in the 5 GHz band with a maximum net data rate of 54 Mbit/s, plus error correction code, which yields realistic net achievable throughput in the mid-20 Mbit/s.
17. Like 802.11b, 802.11 g devices suffer interference from other products operating in the 2.4 GHz band. Devices operating in the 2.4 GHz range include: microwave ovens, Bluetooth devices, baby monitors and cordless telephones.
18. Connectionless protocols are usually described as stateless because the endpoints have no protocol-defined way to remember where they are in a "conversation"

of message exchanges. The alternative to the connectionless approach uses connection-oriented protocols, which are sometimes described as stateful because they can keep track of a conversation.

19. A connection-oriented networking protocol is one that delivers a stream of data in the same order as it was sent, after first establishing a communication session. It may be a circuit switched connection, or a virtual circuit connection in a packet switched network.

20. The Internet Protocol (IP) is a protocol used for communicating data across a packet-switched internetwork using the Internet Protocol Suite, also referred to as TCP/IP.

21. The Open Systems Interconnection Basic Reference Model (OSI Reference Model or OSI Model) is an abstract description for layered communications and computer network protocol design. It was developed as part of the Open Systems Interconnection (OSI) initiative. In its most basic form, it divides network architecture into seven layers which, from top to bottom, are the Application, Presentation, Session, Transport, Network, Data-Link, and Physical Layers. It is therefore often referred to as the OSI Seven Layer Model.

22. TCP is used extensively by many of the Internet's most popular application protocols and resulting applications, including the World Wide Web, E-mail, File Transfer Protocol, Secure Shell, and some streaming media applications.

23. The User Datagram Protocol (UDP) is one of the core members of the Internet Protocol Suite, the set of network protocols used for the Internet. With UDP, computer applications can send messages, sometimes known as datagrams, to other hosts on an Internet Protocol (IP) network without requiring prior communications to set up special transmission channels or data paths. UDP is sometimes called the Universal Datagram Protocol.

24. TCP ("Transmission Control Protocol") is a connection-oriented protocol, which means that upon communication it requires handshaking to set up end-to-end connection. UDP is a simpler message-based connectionless protocol. In connectionless protocols, there is no effort made to set up a dedicated end-to-end connection.

25. Wireless location service architectures can be classified according to at least two main properties. The first property is whether a central server is used to store the location information and respond to client queries. Architectures that have this property are "centralized", while architectures in which there is no server to store location data and these clients directly communicate with each other is said to be "distributed.

26. CDPD devices now take voice channels over an analog cellular network such as AMPS described above. These channels have the ability to transmit relatively short data packets and use these data packets to start a packet-switching session. Like Bluetooth, CDPD uses a channel hopping service among the cellular frequency it uses.

REVIEW QUESTIONS

13.1. What do you mean by USB? Explain its uses and development.

13.2. Define the following terms" WIRELESS NETWORK DESIGN ANALYSIS ".

13.3. Explain the General Packet Radio Service with GPRS telephone and Data Transmission Speeds.

13.4. Explain the following terms with respect to GPRS
(*i*) Roots of GPRS
(*ii*) The Network
(*iii*) GPRS History

13.5. Define the GPRS Architecture with suitable diagram.

13.6. Write a short note on Wireless Application Protocol (WAP) and also explain Technical challenges of WAP.

13.7. What do you mean by term Bluetooth and also explain The Technology behind Bluetooth.

13.8. Define the Cellular Digital Packet Data (CDPD) architecture.

13.9. Explain the "IEEE 802.11".

13.10. Explain the following terms with respect to "IEEE 802.11"
(*i*) IEEE 802.11a
(*ii*) IEEE 802.11b
(*iii*) IEEE 802.11g
(*iv*) IEEE 802.11n

13.11. Explain the Connectionless protocol and Connection-oriented protocol and also write down the list of Connection-oriented protocol and Connectionless protocol.

13.12. What do you mean by Internet Protocol. Explain in brief.

13.13. Explain the following terms with respect to Internet Protocol-
(*i*) IP Encapsulation
(*ii*) Services provided by IP
(*iii*) Reliability
(*iv*) Version history

13.14. What do you mean by IP addressing.

13.15. Explain in brief. 13.5. Explain the following terms with respect to IP addresses-
(*i*) Static IP addresses
(*ii*) Dynamic IP addresses
(*iii*) IP version 4 addresses
(*iv*) IP version 6 addresses
(*v*) IPv6 private addresses

13.16. Explain the OSI reference model with brief description of all layers.

13.17. Write a short note on Transmission Control Protocol and also define the Significance and Applicability of TCP.

13.18. Write a short note on User Datagram Protocol and also define the Ports and Packet structure of UDP.

13.19. Define the difference between TCP and UDP.

13.20. Explain the Wireless Location Service Architectures with suitable diagram.

APPENDIX

APPENDIX–A

Table 1: Analog Cellular System (An Overview Worldwide)

Standard	Mobile Tx / Base Tx (MHz)	Channel Spacing (kHz)	No. of channel	Region
1. AMPS	824-849/869-894	30	832	USA
2. TACS	890-915/935-960	25	1000	Europe
3. ETACS	872-905/917-950	25	1240	U.K.
4. NMT 450	453-457.5/463-467.5	25	180	Europe
5. NMT 900	890-915/935-960	12.5	1999	Europe
6. C- 450	450-455.74/465.74	10	573	Germany
7. RTMS	450-455/460-465	25	200	Italy
8. Radiocom	192.5-199.5/200.5-207.5		560	
215.5-233.5/207.5-215.5		640		
165.2-168.4/169.8-173		256		
414.8-418/424.8-428	12.5	256	France	
9. NTT	925-940/870-885	25/6.25	600/2400	
915-918.5/860-863.5	6.25	560		
922-925/867-870	6.25	480	Japan	
10. JTACS/NTACS	915-925/860-870	25/12.5	400/800	
898-901/843-846	25/12.5	120/240		
918.5-922/863.5-867	12.5	280	Japan	

Table 2: Digital Cordless Air Interface Parameter Summary.

Region	CT2 Europe	CT2 Canada	DECT Europe	PHS Japan	PACS US
Duplexing	TDD		TDD	TDD	FDD or TDD
Frequency band (MHz)	864-868	944-948	1880 -1900	1895-1918	1850-1910/ 1930-1990 or 1920-1930
Carrier spacing (kHz)	100		1728	300	300/300
Number of carriers	40		10	77	400 or 32
Channels/carrier	1		12	4	8 or 4
Channel bit rate kbps	72		1152	384	384
Modulation	GFSK		GFSK	$\pi/4$ DQPSK	$\pi/4$ QPSK
Speech Coding	32 kb/s		32 kb/s	32 kb/s	32 kb/s
Average handset TX power (mW)	5		10	10	25
Peak handset TX power (mW)	10		250	80	100
Frame duration (ms)	2		10	5	2.5 or 2.0

Table 3: Unlicensed/Short Range PCS Standards.

	PACS-UB	DCS1800	PHS
Year of Introduction	1992	1993	1993
Frequencies	1920-1930 MHz	1710 -1785 MHz (R) 1805-1880 MHz (F)	1895-1907 MHz
Multiple Accees	TDMA/FDMA/TDD	TDMA/FDMA/TDD	TDMA/FDMA/TDD
Modulation	$\pi/4$ QPSK	GMSK	$\pi/4$ QPSK
Carrier Separation	300	200 kHz	300 kHz
Data rate	384 kbps	270.83 kbps	384 kbps
Voice Ch./R.F. Ch.	4	8	4
Speech Coding	ADPCM @ 32 kbps	RELP-LTP@ 13 kbps	ADPCM @ 32 kbps
Channel coding	CRC	Conv. r = 1/2	CRC
Receiver	Coherent	Coherent	Coherent
Portable Tx. Power Max./Avg.	200 mW/20 mW	1W/125 mW	80 mW/10 mW

Table 4: GSM Air Interface Specification Summary.

Parameter	Specification
Reverse Channel Frequency	980-915 MHz
Forward Channel Frequency	935-960 MHz
ARFCN Number	0 to 124 and 975 to 1023
Tx/Rx Frequency Spacing	45 MHz
Tx/Rx Time Slot Spacing	3 Time slots
Modulation Data Rate	270.833333 kbps
Frame period	4.615 ms
Users per frame (Full rate)	8
Time slot period	576.9 μs
Bit Period	3.692 μs
Modulation	0.3 GMSK
ARFCN Channel Spacing	200 kHz
Interleaving (max. delay)	40
Voice Coder Bit Rate	13.4 kbps

Table 5: Second Generation Digital Cellular Standard Summary.

	GSM	IS-136	PDC
Year of Introduction Frequencies	1990 890-915 MHz (R) 935-960 MHz (F)	1991 869-894 MHz (F)	1993 810-830 and 1429-1453 MHz (R) 824-849 MHz (R) 940-960 and 1477-1501 MHz (F)
Multiple Access	TDMA/FDMA/FDD	TDMA/FDMA/FDD	TDMA/FDMA/FDD
Modulation	GMSK (BT = 0.3)	$\pi/4$ DQPSK	$\pi/4$ DQPSK
Carrier Separation	200 kHz	300 kHz	25 kHz
Channel Data Rate	270.833 kbps	48.6 kbps	42 kbps
Number of Voice Channels	1000	2500	3000
Spectrum Efficiency	1.35 bps/Hz	1.62 bps/Hz	1.68 bps/Hz
Speech Coding	RELP-LTP @ 13 kbps	VSELP @ 7.95 kbps	VSELP @ 6.7 kbps
Channel coding	CRC with r = 1/2; L = 5 conv.	7 bit CRC with r = 1/2; L = 6 conv.	CRC with Conv.
Equalizers	Adaptive	Adaptive	Adaptive
Portable Tx. Power Max./Avg.	1 W/125 mW	600 mW/200 mW	

Table 6: DECT Radio Specification Summary.

Parameter	Specification
Frequency band	1880-1990 MHz
Number of Carriers	10
RF Channel Bandwidth	1.728 MHz
Multiplexing	FDMA/TDMA 24 slots per frame
Duplex	TDD
Spectral Efficiency	500 Erlangs/km^2/MHz
Speech Coder	32 kbps ADPCM
Average Transmitted Power	10 mW
Frame Length	10 ms
Channel Bit Rate	1152 kbps
Data Rate	32 kbps Trasffic Channel
	6.2 kbps Control Channel
Channel Coding	CRC 16
Dynamic Channel Allocation	Yes
Modulation	GFSK (BT = 0.3)
Speech Channel RF/Channel	12

Table 7: IEEE 802. 11b channels for both DSSS and FH-SS WLAN Standards.

Country	Frequency range available	DSSS channels available	FHSS channels available
United states	2.4 to 2.4835 Gz	1 through 11	2 through 80
Canada	2.4 to 2.4835 GHz	1 through 11	1 through 80
Japan	2.4 to 2.497 GHz	1 through 14	2 through 95
France	2.4465 to 2.4835 GHz	10 through 13	48 through 82
Spain	2.445 to 2.4835 GHz	10 through 11	47 through 73
Remainder	2.4 to 2.4835	1 through 13	2 through 80

Table 8: Comparison of Mobile Communication System—Mobile Station.

Service	Coverage range	Required infra-Structure	Complexity	Hardware cost	Carrier frequency	Functionality
TV remote control	Low	Low	Low	Low	Infrared	Transmitter
Garage door opener	Low	Low	Low	Low	< 100 MHz	Transmitter
Paging system	High	High	Low	Low	<1 GHz	Receiver
Cordless phone	Low	Low	Moderate	Low	1-3 GHz	Transceiver
Cellular phone	High	High	High	Moderate	<2 GHz	Transceiver

Table 9: Comparison of Mobile Communication System—Base Station.

Service	Coverage range	Required infra-structure	Complexity	Hardware cost	Carrier frequency	Functionality
TV remote control	Low	Low	Low	Low	Infrared	Receiver
Garage door opener	Low	Low	Low	Low	< 100 MHz	Transmitter
Paging system	High	High	Low	Low	<1 GHz	Transmitter
Cordless phone	Low	Low	Low	Moderate	1-3 GHz	Transceiver
Cellular phone	High	High	High	High	<2 GHz	Transceiver

Table 10: Microwave Bands.

New Designation	Frequency (GHz)	Old Designation	Frequency (GHz)
C	0.5-1.0	VHF	0.5 -1
D	1.0 -2.0	L	1-2
E	2-3	S	2- 4
G	5-6	X	8 -12.4
H	6-8	Ku	12.4-18
I	8-10	K	18-26.5
J	10-20	Ka	26.5-40
K	20-40	v	
L	40-60		40 -75
M	60 -100		

Table 11: Antenna Gains and Effective Areas.

Type of Antenna	Effective Area Ae (m^2)	Power Gain (relative) to isotropic
Isotropic	$\frac{\lambda^2}{4\pi}$	1
Infinitesimal dipole or loop	$\frac{1.5\lambda^2}{4\pi}$	1.5
Half wave dipole	$\frac{1.64\lambda^2}{4\pi}$	1.64
Horn/mouth area A	0.81 A	$\frac{10\,A}{\lambda^2}$
Parablolic, face area	0.56 A	$\frac{2A}{\lambda^2}$
Turnstile (two crossed, perpendicular dipoles)	$\frac{1.15\lambda^2}{4\pi}$	1.15
G (Anenna gain) = $\frac{4\pi f^2 Ae}{c^2}$ $f \rightarrow$ Carrier frequency $f \rightarrow$ Speed of light $A_e \rightarrow$ Effective area		

Table 12: Antenna Beamwidth for Various Diameter for Parabolic Reflective Antenna at f = 12 GHz.

Antenna diameter	Beamwidth (Degrees)
0.5	3.5
0.75	2.33
1.0	1.75
1.5	1.166
2.0	0.875
2.5	0.7

Table 13: Frequency band and their application

Band	Frequency range	Free space wavelength range	Propagation	Typical use character
ELP extremely	30-300	10000 to 1000 km	GW (Ground wave)	Power line frequencies; used by home control system
VF Voice frequency	300-3000 Hz	1000 to 100 km	GW	Used by Telephone system (Analog subscriber line)
VLF Very low	3 to 30 kHz	100 to 10 km	GW; low attenuation in day and night High atm. noise level	Low range navigation; submarrine comm.
LF Low Frequency	30-300 kHz	10 to 1 km	GW, slightly less than VLF, absorption in day time	Long range navigation, marriage, radio, beacons
MF Medium freq.	300-3000 kHz	1000 to 100 m	GW and night SW (sort wave) attenuation low in night, high in day, atm. noise	Direction finding, AM broad casting
HF High freq.	3-30 MHz	100 to -10 m	SW, quality varies with time of day, season and frequency.	International broadcasting, millitary comm. ship, aircraft
VHF Very High Frequency	30-300 MHz	10 to 1 metre	LOS, scattering because of temp--erature inversion, cosmic noise	VHF T.V., FM broadcast, two-way radio, AM aircraft com.
UHF Ultra High	300 to 3000 MHz	100 to 10 cm	LOS, cosmic noise	UHF T.V., cellular telephone, radar microwave link, personal comm. system.
SHF super high	3 to 30 GHz	10 to 10 m	LOS, rain fall attenuation above 10 GHz atm. atten. due to oxygen, water vapour	Satellite comm. Radar, Terrestion comm. WLL
EHF extremely high freq.	30-300 GHz	10 to 1 mm	LOS, Atm-noise due to O_2, H_2O water	Experimental; WLL
Infra-red	300 GHz-400 THz	1 mm-770 nm	LOS,	Infrared LAN, consumer electronic application
Visible Light	400 THz to 900 THz	770 nm-330 nm	LOS	Optical communication.

Table 14: Trigonometric Identities.

$$\sin\theta = \frac{\text{Perpendicular}}{\text{Hypotenous}} = \frac{1}{\operatorname{cosec}\theta}$$

$$\cos\theta = \frac{\text{Base}}{\text{Hypotenous}} = \frac{1}{\sec\theta}$$

$$\tan\theta = \frac{\text{Perpendicular}}{\text{Base}} = \frac{1}{\cot\theta} = \frac{\sin\theta}{\cos\theta}$$

$$\sin^2\theta + \cos^2\theta = 1$$

$$\sin 2\theta = 2\sin\theta\cos\theta$$

$$\sin^3\theta = 3\sin\theta - 4\sin^3\theta$$

$$\sin(A \pm B) = \sin A \cos \pm \cos A \sin B$$

$$\cos(A \pm B) = \cos A \cos B \mp \sin A \sin B$$

$$\sin C + \sin D = 2\sin\frac{C+D}{2}\cos\frac{C-D}{2}$$

$$\sin C - \sin D = 2\sin\frac{C-D}{2}\cos\frac{C+D}{2}$$

$$\cos C + \cos D = 2\cos\frac{C+D}{2}\cos\frac{C-D}{2}$$

$$\cos C - \cos D = 2\sin\frac{C+D}{2}\sin\frac{D-C}{2}$$

$$\sin A \sin B = \frac{1}{2}[\cos(A-B) - \cos(A+B)]$$

$$\sin A \cos B = \frac{1}{2}[\sin(A+B) + \sin(A-B)]$$

$$\cos A \cos B = \frac{1}{2}[\cos(A+B) + \cos(A-B)]$$

$$\cos 2\theta = 2\cos^2\theta - 1 = \cos^2\theta - \sin^2\theta = 1 - 2\sin^2\theta$$

$$\sin\frac{\theta}{2} = \sqrt{\frac{1-\cos\theta}{2}}$$

$$\cos\frac{\theta}{2} = \sqrt{\frac{1+\cos\theta}{2}}$$

$$\sin^2\theta = \frac{1-\cos 2\theta}{2}, \quad \cos^2\theta = \frac{1+\cos 2\theta}{2}$$

$$\sin\theta = \frac{e^{j\theta} - e^{-j\theta}}{2}$$

$$\cos\theta = \frac{e^{j\theta} + e^{-j\theta}}{2}, \quad e^{j\theta} = \cos\theta + j\sin\theta$$

Table 15: Some Important Integral Result.

$\int \sin(at)\,dt = -\frac{1}{\alpha}\cos(ta)$	$\int \cos(at)\,dt = \frac{1}{a} = \sin(at)$
$\int e^{at}dt = \frac{e^{at}}{a}$	$\int \frac{1}{t}dt = \ln t$
$\int t\cos at\,dt = \left(\frac{1}{a^2}\right)(\cos at + at\sin at)$	$\int \sec t\,dt = \sec t\,\tan t$
$\int \operatorname{cosec} t\,dt = -\operatorname{cosec} t \cot t$	$\int_0^\infty \frac{\sin at}{t}dt = \frac{\pi}{2}, 0, -\frac{\pi}{2}$ for $a > 0, a < 0$
$\int_0^\infty \frac{\sin^2 t}{t}dt = \frac{\pi}{2}$	$\int_0^\infty t^n e^{-at}dt = \frac{n!}{a^{n+1}}$
$\int_0^\infty e^{-a^2-t^2}dt = \frac{\sqrt{\pi}}{2t}$	$\int_0^\infty t\,e^{-a^2t^2}dt = \frac{1}{2a^2}$

APPENDIX–B

ACRONYMS

A

AC	Access Channel
ACA	Adaptive Channel Allocation
ACF	Autocorrelation function
ACI	Adjacent Channel interference
ACK	Acknowledge
ADB/LIDB	Alternate Billing Service and Line Information Database
ADM	Adaptive Delta Modulation
ADPCM	Adaptive Digital Pulse Code Modulation
ADSL	Asynchronous Digital Subscriber Line
AGCH	Access Grant Channel
AIN	Advanced Intelligent Network
AM	Amplitude Modulation
AMPS	Advanced Mobile Phone System
ANSI	American National Standards Institute
APC	Adaptive Predictive Coding
ARDIS	Advance Radio Data Information Systems
ARFCN	Absolute Radio Frequency Channel Numbers
ARPU	Average Revenue Per User
ARQ	Automatic Repeat Request
ATC	Adaptive Transform Coding
ATM	Asynchronous Transfer Mode
AUC	Authentication Center
AWGN	Additive White Gaussian Noise

B

BER	Bit Error Rate
BERSIM	Bit Error Rate Simulator
BCH	Bose-Chaudhuri-Hocquenghem, *also* Broadcast Channel
BCCH	Broadcast Control Channel
BFSK	Binary Frequency Shift Keying
BISDN	Broadband Integrated Services Digital Network
BIU	Base Station Interface Unit
BOC	Bell Operating Company
BPSK	Binary Phase Shift Keying
BRAN	Broadband Radio Access Networks
BRI	Basic Rate Interface

BSC Base Station Controller
BSIC Base Station Identity Code
BSS Base Station Subsystem
BT 3 dB bandwidth-bit-duration product for GMSK
BTA Basic Trading Area
BTS Base Transceiver Station

C

CAD Computer-Aided Design
CAI Common Air Interface
CATT China Academy of Telecommunication Technology
CB Citizens Band
CCCH Common Control Channel
CCH Control Channel
CCI Co-channel Interference
CCIR Consultative Committee for International Radiocommunications
CCITT International Telgraph and Telephone Consultative Committee
CCS Common Channel Signaling
CD Collision Detection
CDF Cumulative Distribution Function
CDMA Code Division Multiple Access
CDPD Cellular Digital Packet Data
CDVCC Coded Digital Verification Colour Code
CELP Code Excited Linear Predictor
CFP Cordless Fixed Part
C / I Carrier-to-Interference Ratio
CIC Circuit Identification Code
CIR Carrier-to-interference Ratio
CIU Cellular Controller Interface Unit
CLNP Connectionless Protocol (Open System Interconnect)
CMA Constant modulus algorithm
CNR Carrier-to-Noise Ratio
CNS Comfort Noise Subsystem
CO Central Office
Codec Coder/decoder
COST Cooperative for Scientific and Technical Research
CPE Customer Premises Equipment
CPFSK Continuous Phase Frequency Shift Keying
CPP Cordless Portable Part
CRC Cyclic Redundancy Code
CSMA Carrier Sense Multiple Access
CT2 Cordless Telephone-2
CVSDM Continuously Variable Slope Delta Modulation
CW Continuous Wave

D

DAM	Diagnostic Acceptability Measure
DAS	Distributed Antenna System
DBAS	Database Service Mangement System
dBi	dB gain with respect to an isotropic antenna
dBd	dB gain with respect to a half-wave dipole
DCA	Dynamic Channel Allocation
DCCH	Dedicated Control Channel
DCE	Data Circuit Terminating Equipment
DCS	Digital Communication System
DCS1800	Digital Communication System—1800
DCT	Discrete Cosine Transform
DECT	Digital European Cordless Telephone
DEM	Digital Elevation Model
DFE	Decision Feedback Equalization
DFT	Discrete Fourier Transform
DLC	Data Link Control
DM	Delta Modulation
DPCM	Differential Pulse Code Modulation
DQPSK	Differential Quadrature Phase Shift Keying
DRT	Diagnostic Rhyme Test
DS	Direct Sequence
DSAT	Digital Supervisory Audio Tone
DSE	Data Switching Exchange
DS/FHMA	Hybrid Direct Sequence/Frequency Hopped Multiple Access
DSL	Digital Subscriber Line
DSP	Digital Signal Processing
DSSS	Direct Sequence Spread Spectrum
DST	Digital Signaling Tone
DTC	Digital Traffic Channel
DTE	Data Terminal Equipment
DTMF	Dual Tone Multiple Frequency
DTX	Discontinuous Transmission Mode
DUP	Data User Part

E

EDGE	Enhanced Data Rates for GSM Evolution
EIA	Electronic Industry Association
EIR	Equipment Identity Register
EIRP	Effective Isotropic Radiated Power
E_b / N_0	Bit Energy-to-noise Density

EOC Embedded Operations Channel
erf Error Function
erfc Complementary Error Function
ERMES European Radio Message System
ERP Effective Radiated Power
E-SMR Extended — Specialized Mobile Radio
ESN Electronic Serial Number
ETACS Extended Total Access Communication System *also* European Total Access Cellular System
ETSI European Telecommunications Standard Institute

F

FACCH Fast Associated Control Channel
FAF Floor Attenuation Factor
FBF Feedback Filter
FC Fast Channel
FCC Federal Communications Commission, Inc., *also* Forward Control Channel
FCCH Frequency Correction Channel
FCDMA Hybrid FDMA/CDMA
FDD Frequency Division Duplex
FDMA Frequency Division Multiple Access
FDTC Forward Data Traffic Channel
FEC Forward Error Correction
FFF Feed Forward Filter
FFSR Feed Forward Signal Regeneration
FH Frequency Hopping
FHMA Frequency Hopped Multiple Access
FHSS Frequency Hopped Spread Spectrum *also* Frequency Hopping Spread Spectrum
FLEX 4-level FSK-based paging standard developed by Motorola
FM Frequency Modulation
FN Frame Number
FPLMTS Future Public Land Mobile Telephone System
FSE Fractionally Spaced Equalizer
FSK Frequency Shift Keying
FTF Fast Transversal Filter
FVC Forward Voice Channel

G

GIS Graphical Information System
GIU Gateway Interface Unit

GMSK Gaussian Minimum Shift Keying
GOS Grade of Service
GPRS General Packet Radio Service
GSC Golay Sequential Coding (a 600 bps paging standard)
GSM Global System for Mobile Communication *also* Global System Mobile

H

HAAT Height Above Average Terrain
HSCSD High Speed Circuit Switched Data
HDB Home Database
HDR High Data Rate
HIPERLAN High Performance Radio Local Area Network
HLR Home Location Register

I

IDCT Inverse Discrete Cosine Transform
iDen Integrated Digital Enhanced Network
IDFT Inverse Discrete Fourier Transform
IEEE Institute of Electrical and Electronics Engineers
IF Intermediate Frequency
IFFT Inverse Fast Fourier Transform
IGA Improved Gaussian Approximation
IIR Infinite Impulse Response
IM Intermodulation
IMSI International Mobile Subscriber Identity
IMT2000 International Mobile Telecommuncation 2000
IMTS Improved Mobile Telephone Service
IP Internet Protocol
IS54 EIA Interim Standard for U.S. Digital Cellular with Analog Control Channels
IS95 EIA Interim Standard for U.S. Code Division Multiple Access
IS-136 EIA Interim Standard 136—USDC with Digital Control Channels
ISDN Integrated Services Digital Network
ISI Intersymbol Interference
ISM Industrial, Scientific, and Medical
ISUP ISDN User Part
ITFS Instructional Television Fixed Service
ITU International Telecommunications Union
ITU-R ITU's Radiocommunications Secto
IXC Interexchange Carrier

J

JDC Japanese Digital Cellular (later called Pacific Digital Cellular)
JRC Joint Radio Committee
JTACS Japanese Total Access Communication System
JTC Joint Technical Committee

L

LAN Local Area Network
LAR Log-area Ratio
LATA Local Access and Transport Area
LBT Listen-before-talk
LCC Lost Call Cleared
LCD Lost Call Delayed
LCR Level Crossing Rate
LEC Local Exchange Carrier
LEO Low Earth Orbit
LMDS Local Multipoint Distribution Systems
LMS Least Mean Square
LOS Line-of-sight
LPC Linear Predictive Coding
LSSB Lower Single Side Band
LTE Linear Transversal Equalizer
LTP Long Term Prediction

M

MAC Medium Access Control
MAHO Mobile Assisted Handoff
MAI Multiple Access Interference
MAN Metropolitan Area Network
MARY Multiple Level Modulation
MC Multicarrier
MCS Multiple Modulation and Coding Schemes
MDBS Mobile Database Stations
MDLP Mobile Data Link Protocol
MDS Multipoint Distribution Service
MDR Medium Data Rate
MFJ Modified Final Judgement
MFSK Minimum Frequency Shift Keying
MIN Mobile Identification Number
MIRS Motorola Integrated Radio System (for SMR use)
ML Maximal Length

MLSE Maximum Likelihood Sequence Estimation
MMAC Multimedia Mobile Access Communication System
MMDS Multichannel Multipoint Distribution Service
MMSE Minimum Mean Square Error
MOS Mean Opinion Score
MoU Memorandum of Understanding
MPE Multipulse Excited
MPSK Minimum Phase Shift Keying
MS Mobile Station
MSB Most Significant Bit
MSC Mobile Switching Center
MSCID MSC Identification
MSE Mean Square Error
MSK Minimum Shift Keying
MSU Message Signal Unit
MTA Major Trading Area
MTP Message Transfer Part
MTSO Mobile Telephone Switching Office
MUX Multiplexer

N

NACK Negative Acknowledge
NADC North American Digital Cellular
NAMPS Narrowband Advanced Mobile Phone System
NBFM Narrowband Frequency Modulation
NEC National
NISDN Narrowband Integrated Service Digital Network
NMT-450 Nordic Mobile Telephone—450
NRZ Non-return to Zero
NSP Network Service Part
NSS Network and Switching Subsystem
NTACS Narrowband Total Access Communication System
NTT Nippon Telephone and Telegraph

O

OBS Obstructed
OFDM Orthogonal Frequency Division Multiplexing
OMAP Operations Maintenance and Administration Part
OMC Operation Maintenance Center
OQPSK Offset Quadrature Phase Shift Keying
OSI Open System Interconnect
OSS Operation Support Subsystem

PABX Private Automatic Branch Exchange
PACS Personal Access Communication System
PAD Packet Assembler Disassembler
PAF Partition Attenuation Factor
PAN Personal Area Network
PBX Private Branch Exchange
PCH Paging Channel
PCM Pulse Code Modulation
PCN Personal Communication Network
PCS Personal Communication System
PDC Pacific Digital Cellular
pdf probability density function
PG Processing Gain
PH Portable Handset
PHP Personal Handyphone
PHS Personal Handyphone System
PL Path Loss
PLL Phase Locked Loop
PLMR Public Land Mobile Radio
PN Pseudo-noise
POCSAG Post Office Code Standard Advisory Group
POTS Plain Old Telephone Service
PR Packet Radio
PRI Primary Rate Interface
PRMA Packet Reservation Multiple Access
PSD Power Spectral Density
PSK Phase Shift Keying
PSTN Public Switched Telephone Network
PTI Permanent Terminal Identifier

Q

QAM Quadrature Amplitude Modulation
QCELP Qualcomm Code Excited Linear Predictive Coder
QMF Quadrature Mirror Filter
QPSK Quadrature Phase Shift Keying

R

RACE Research on Advanced Communications in Europe
RACH Random Access Channel
RCC Reverse Control Channel
RCS Radar Cross Section
RDLAP Radio Data Link Access Protocol

RDTC Reverse Data Traffic Channel
RELP Residual Excited Linear Predictor
RF Radio Frequency
RFP Radio Fixed Part
RLC Resistor Inductor Capacitor
RLS Recursive Least Square
RMD RAM Mobile Data
RPCU Radio Port Control Unit
RPE-LTP Regular Pulse Excited Long-Term Prediction
RRMS Radio Resource Management Protocol
RS Reed-Solomon
RSSI Received Signal Strength Indication
RTT Radio Transmission Technology
RVC Reverse Voice Channel
Rx Receiver
RZ Return to Zero

S

SACCH Slow Associated Control Channel
SAT Supervisory Audio Tone
SBC System Broadcasting Channel *also* Sub-band Coding; Southwestern Bell Corporation
SC Slow Channel
SCCP Signaling Connection Control Part
SCH Synchronization Channel
SCM Station Class Mark
SCORE Spectral Coherence Restoral Algorithm
SCP Service Control Point
SDCCH Stand-alone Dedicated Control Channel
SDMA Space Division Multiple Access
SEIGA Simplified Expression for the Improved Gaussian Approximation
SELP Stochastically Excited Linear Predictive Coder
SEP Switching End Points
SFM Spectral Flatness Measure
SI *see* SIR
SID Station Identity
SIM Subscriber Identity Module
SIR Signal-to-Interference Ratio
SIRCIM Simulation of Indoor Radio Channel Impulse Response Models
SISP Site Specific Propagation
SMR Specialized Mobile Radio
SMRCIM Simulation of Mobile Radio Channel Impulse Response Models

SMS Short Messaging Service *also* Service Management System
S/N *see* SNR
SNR Signal-to-Noise Ratio
SOHO Small Office/Home Office
SONET Synchronous Optical Network
SP Signaling Point
SQNR Signal-to-Quantization Noise Ratio
SS Spread Spectrum
SSB Single Side Band
SSMA Spread Spectrum Multiple Access
SS7 Signaling System No. 7
ST Signaling Tone
STP Short Term Prediction, *also* Signaling Transfer Point
SYN Synchronization channel

T

TACS Total Access Communications System
TCAP Transaction Capabilities Application Part
TCDMA Time Division CDMA
TCH Traffic Channel
TCM Trellis Coded Modulation
TDD Time Division Duplex
TDFH Time Division Frequency Hopping
TDMA Time Division Multiple
TDN Temporary Directory Number
TD-SCDMA Time Division-Synchronous Code Division Multiple Access
TIA Telecommunications Industry Association
TIU Trunk Interface Unit
TTIB Transparent Tone-in-Band
TUP Telephone User Part
Tx Transmitter

U

UF Urban Factor
UMTS Universal Mobile Telecommunications System
UNII Unlicensed National Information Infrastructure
US United States of America
USDC United States Digital Cellular, *see* IS- 136 and NADC
USGS United States Geological Survey
USSB Upper Single Side Band
UTRA UMTS Terrestrial Radio Access

V

VAD	Voice Activity Detector
VCI	Virtual Circuit Identifier
VCO	Voltage Controlled Oscillator
VDB	Visitor Database
VHE	Virtual Home Entertainment
VIU	Visitor Interface Unit
VLR	Visitor Location Register
VLSI	Very Large-Scale Integration
VMAC	Voice Mobile Attenuation Code
VoIP	Voice over Internet Protocol
VQ	Vector Quantization
VSELP	Vector Sum Excited Linear Predictor

W

WACS	Wireless Access Communication System (later called PACS)
WAN	Wide Area Network
WAP	Wireless Applications Protocol
WARC	World Administrative Radio Conference
W-CDMA	Wideband CDMA
WECA	Wireless Ethernet Compatability Alliance
WIN	Wireless Information Network
WIU	Wireless Interface Unit
WLAN	Wireless Local Area Network
WLL	Wireless Local Loop
WRC2000	ITU World Radio Conference
WUPE	Wireless User Premises Equipment

Z

ZF	Zero Forcing

APPENDIX–C

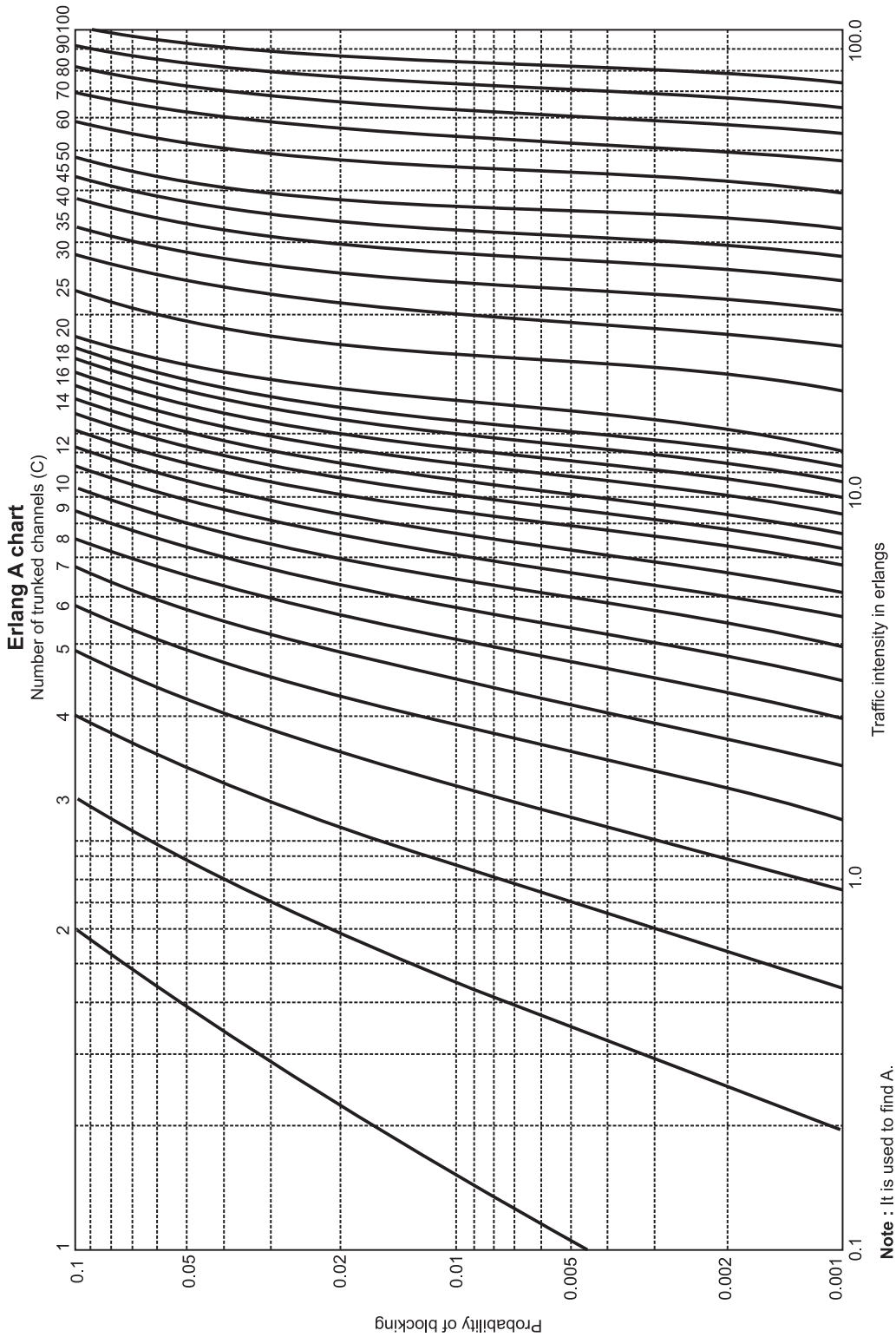
Some Additional Acronyms used in Wireless Communication

ADC	American Digital Cellular
AMPS	Advanced mobile phone service
AoC	Advice of charge
AUC	Authentication center
bps	Bits per second
BSC	Base station controller
BSS	Base station system
BTS	Base transceiver station
CGI	Cell global identity
CUG	Closed user group
DCS	Digital cellular system
DTMF	Dual-tone multifrequency
EIR	Equipment identity register
GIWU	GSM interworking unit
GMSC	Gateway mobile services switching center
GMSK	Gaussian minimum shift keying
GSM	Global system for mobile communication
HLR	Home location register
Hz	Hertz
ISDN	Integrated services digital network
k	Kilo
kbps	Kilobits per second
LA	Location area
LM	Location-area identity
LPC	Linear predictive coding
MHz	Megahertz
MSC	Mobile services switching center
MSN	Mobile service node
MXE	Message center
NMT	Nordic Mobile Telephone
OMC	Operations and maintenance center
OSS	Operation and support system
PCS	Personal communications services
PDC	Personal digital cellular
PLMN	Public land mobile network
SS	Switching system
TAGS	Total access communication system
TOMA	Time division multiple access
VLR	Visitor location register

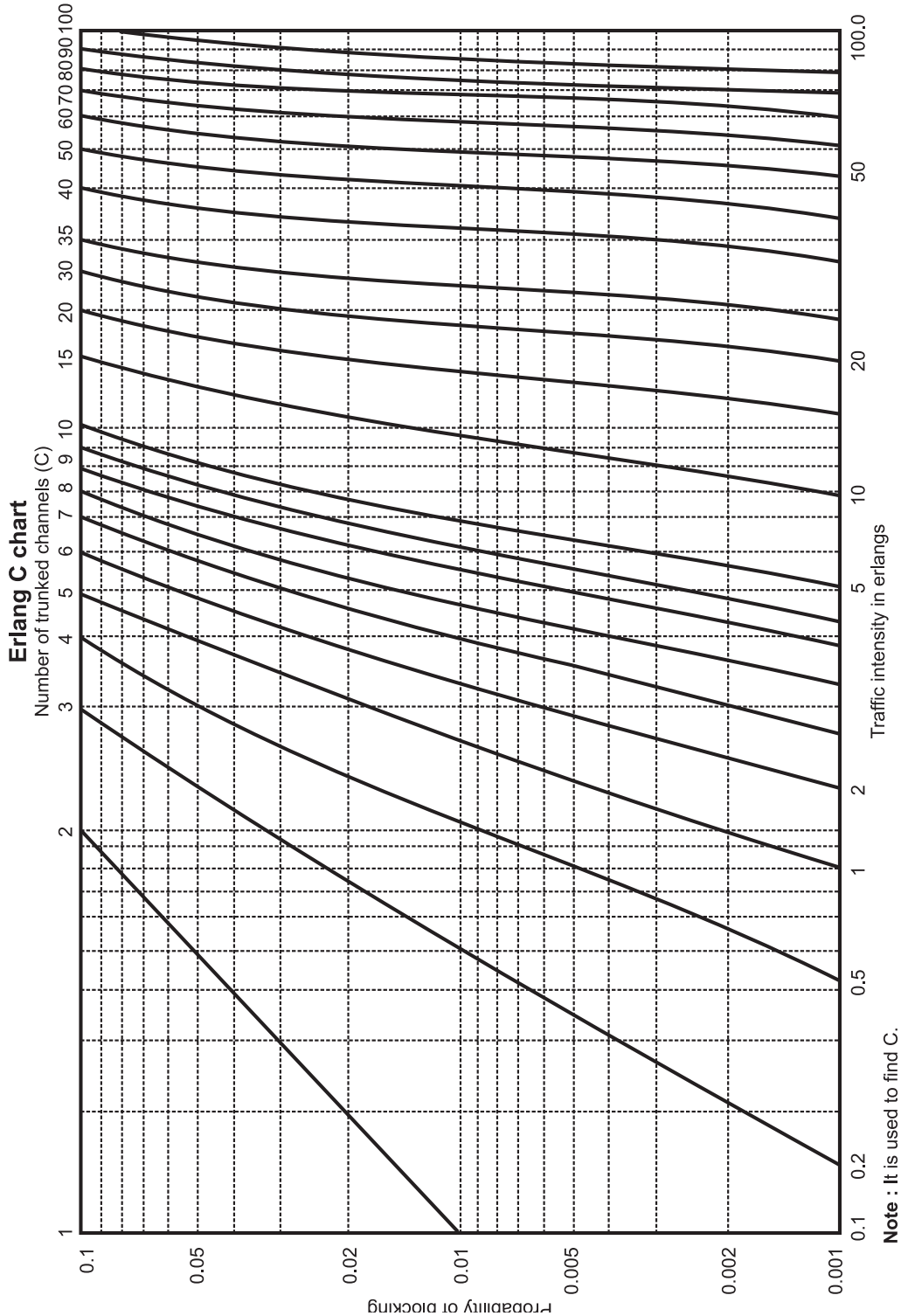
Numerics

1 Base S	The IEEE 802.3 standard for low data rate ethernet using twisted pair cable and daisy chained connections.
10 Base T	IEEE 802.3 standard for twisted pair ethernet.
10 Base 2	IEEE 802.3 standard for ethernet using thin co-axial cable.
10 Base 5	IEEE 802.3 standard for ethernet using thick co-axial cable.
100 base FX	A version of IEEE 802.3 standard for Fast ethernet using 2 optical fibers.
100 base T	A version of IEEE 802.3 standard for Fast ethernet using 2 twisted pair cable.
56 K Modem	A modern technology using two different data rates, one for uploading and one for downloading from the internet.
1XDV	3G extension of IS-95B (shared data and voice)
1XD0	3G extension of IS-95B data only
1XEV	3G extension of IS-95B data with circuit switched voice
1XRTT	3G extension of IS-95B one RF channel
3GPP	3G partnership project for wideband CDMA standards base on backward compatibility with GSM and IS-136/PDC
3GPP2	3G partnership project for CDMA 2000 standards based on backward compatibility with IS-95.
3XRTT	3G extention of IS-95B three RF channels.

APPENDIX–D



Note : It is used to find A.



Note : It is used to find C.

Model Question Papers

MODEL QUESTION PAPER–I

WIRLESS COMMUNICATIONS

Time : 3 Hours **Total Marks : 100**

Note: **(*i*)** ***Attempt All Questions***
(*ii*) ***All questions carry equal marks.***

1. Attempt any four from the following : (5 × 4 = 20)

(*a*) Exaplain impulse response model of a multipath chanel.
(*b*) What do you understand by Pseudo-Noise (PN) sequence?
(*c*) Compare and contract the features of second-generation digital cellular standards-GSM and CDM technologies.
(*d*) What are the advantages of cellular mobile communication systems over conventional mobile communication systems and conventional mobile telephone systems?
(*e*) Define and explain TDMA systems
(*f*) Draw the GSM network architecture. Identify various interfaces used in its different entities.

2. Attempt any four from the following: (5 × 4 = 20)

(*a*) Explain the difference between equalization and diversity.
(*b*) Disinguish Speech coders and explain with suitable diagram each type in brief.
(*d*) Describe paging system in detail.
(*e*) Write a short note on UMTS.
(*f*) Explain frequency allocation.

3. Attempt any two from the following: (10 × 2 = 20)

(*a*) Derive probability of error for Coherent binary PSK modulation in slow flat fading channels.
(*b*) Explain IEEE 802.11 family and mention its special features.
(*c*) Explain the term frequency Reuse in a cellular networks

OR

What are various modulation techniques used for mobile networks?

4. Attempt any two from the following: (10 × 2 = 20)

(*a*) Describe in details the wirless local loops.

(*b*) Compare the similarities and difference in the fundmental concepts of a DSSS system versus FHSS system.

(*c*) The 'near-far interference' is a serious problems in a wireless cellular CDMA network. Explain its reason giving suitable example data.

5. Write short notes on any two from the following: (10 × 2 = 20)

(*a*) GPS

(*b*) MASP

(*c*) HSCSD for 2.5 GSM

OR

(*a*) Explain the OSI reference model with brief description of all layers.

(*b*) Give block diagram and description of earth station and transpander used for satellite communication.

MODEL QUESTION PAPER–II

WIRLESS COMMUNICATIONS

Time : 3 Hours **Total Marks : 100**

Note: **(*i*)** ***Attempt All Questions***
(*ii*) ***All questions carry equal marks.***

1. Attempt any four from the following: (5 × 4 = 20)

(*a*) Define far-field (or Fraunhofer region) of a transmitting antenna. Find out the receiver power in dBm at a free space distance of 1k. if the power transmitted from a unity gain antenna is 50 watts.

(*b*) Why can free-space propagation model not be applied in a mobile radio environment?

(*c*) How IMTS is different from MTS?

(*d*) Do practical link budget design using Log-distance path loss model.

(*e*) What are the various accessing techniques.

(*f*) What is Commn Channel Signalling? Describe the common channel signalling architecutre for wireless networks.

2. Attempt any four from the following: (5 × 4 = 20)

(*a*) Explain the Salient features of USDC/AMPS interim stadara (IS-54).

(*b*) Explain four parameters which are considered for quality speech transmission and reception.

(*c*) Discuss survey of equalization techniques.

(*d*) Give a point-by point brief note on channel planning for wireless systems.

(*e*) A flat Rayleigh fading signal at 6 GHz is received by a mobile travelling at 60 km/hr Determine the average duration of fade below the rms level.

(*f*) Discuss a method of locating co-channel cells in a cellular systems. What are the elements of a or satellite communication? Explain each of them with suitable diagram.

3. Attempt any two from the following: (10 × 2 = 20)

(*a*) What is the fundamental difference in concepts for calculating the capacity of CDMA cellular system over FDMA cellular system? Find out an expression for calculating the member of users in a CDMA cellular system.

(*b*) Which multiple access techniques are used in mobile communication? Explain with the help of example.

(*c*) What is the architecture of WLAN? Explain in brief

OR

Write a short note on Transmission control Protocol and alos define the significance and applicability of TCP

4. Attempt any Two from the following: (10 × 2 = 20)

(*a*) What do you understand by Rake Receiver? Explain the working of a M-branch Rake receiver?

(*b*) Discuss about Channel Assignment Strategies.

(*c*) What is handoff in cellular systems? list the peformance matrices used to make handoff decisions and give the various handoff strategies

OR

How does Gaussian MSM differ from MSK? Discuss GMSK in brief

6. Write short notes on any Two from the following: (10 × 2 = 20)

(*a*) Basic requirements in Wireless communication

(*b*) UMTS

(*c*) Wireless fidelity.